ÉCOLE DOCTORALE SEPT

LABORATOIRE THÉMA - UMR 6049 CNRS

BIOGÉOSCIENCES - CENTRE DE RECHERCHES DE CLIMATOLOGIE - UMR 6282 CNRS

**THÈSE**

Pour obtenir le grade de

Docteur de l'Université de Bourgogne-Franche-Comté

**Discipline : Géographie**

par

**JUSTIN EMERY**

Présentée le 17 Novembre 2016

# LA VILLE SOUS ÉLECTRODES :

## DE LA MESURE À L'ÉVALUATION DE LA POLLUTION ATMOSPHÉRIQUE AUTOMOBILE

---

## VERS UNE SIMULATION MULTI-AGENTS DU TRAFIC ROUTIER EN MILIEU URBAIN

**Sous la direction de**
**Thomas Thévenin (Directeur) et Nadège Martiny (Co-directrice)**

**JURY :**

**BADARIOTTI Dominique**, Professeur, *Université de Strasbourg* ..... *Rapporteur*
**JOSSELIN Didier**, Directeur de recherche CNRS, *ESPACE* ..... *Examinateur*
**MARILLEAU Nicolas,** Ingénieur de recherche IRD-*UMMISCO, UPMC* ..... *Co-encadrant*
**MARTINY Nadège**, Maître de conférences, *Université de Bourgogne-Franche-Comté* ..... *Co-directrice*
**SANDERS Lena**, Directrice de recherche CNRS, *Géographie-Cités* ..... *Rapporteur*
**THÉVENIN Thomas**, Professeur, *Université de Bourgogne-Franche-Comté* ..... *Directeur*

*« DO OR DO NOT, THERE IS NO TRY »*
*(MAÎTRE YODA À LUKE SKYLWALKER, EPISODE V)*

# REMERCIEMENTS

*Je tiens à adresser mes plus chaleureux remerciements à mes deux directeurs, Thomas et Nadège, qui m'ont accompagné tout au long de ces années. Depuis mon master, j'ai bénéficié de votre attention et de vos précieux conseils : je vous remercie pour la confiance et la liberté que vous m'avez accordée depuis le master. Auprès de vous, j'ai pu découvrir et apprendre chaque jour ce qu'est le métier d'enseignant-chercheur. Merci pour tous ces moments conviviaux et chaleureux que nous avons passés ensemble. À mes deux « maîtres Jedi », j'associe naturellement Nicolas qui, depuis quatre ans, a accepté de suivre ce travail et de m'initier à l'art de la simulation multi-agents. À vous trois, je vous exprime ma plus profonde gratitude pour vos nombreux conseils, votre patience, votre intérêt et votre disponibilité tout au long de ces années. Grâce à votre soutien, j'ai pu atteindre cet objectif qui au départ me paraissait si inaccessible et lointain. C'est pourquoi, ma géocomputation c'est aussi vous trois, et cette thèse est aussi en partie votre thèse. Encore merci pour l'investissement et la patience dont vous avez fait preuve ces quatre dernières années et surtout ces derniers mois.*

*Je tiens à remercier les membres du jury qui ont accepté la lourde tâche que peut être l'évaluation d'une thèse. Merci à Lena Sanders et Dominique Badariotti qui, en tant que rapporteurs, ont accepté de juger la qualité de ce travail. Je remercie aussi Didier Josselin d'avoir accepté de participer à cette évaluation et d'apporter son regard sur ce travail.*

*J'adresse mes remerciements au Conseil Régional de Bourgogne ainsi qu'au FEDER (Fond Européen de développement régional) pour leur soutien financier qui a été essentiel au bon déroulement de cette thèse et pour m'avoir permis d'élargir mes compétences à travers le dispositif Jeune Chercheur Entrepreneur. Merci à l'ensemble des personnels de la cellule de régulation de la circulation de la ville de Dijon pour leur accueil. Je remercie plus particulièrement Florent Bouvier pour sa disponibilité et le temps qu'il a pu m'accorder afin d'accéder aux données de comptages, essentielles à la réalisation de ce travail. Je souhaite aussi remercier Sandrine Monteiro et l'ensemble des membres de l'association de qualité de l'air, parrain de cette thèse, ATMOSF'AIR BOURGOGNE. Un grand merci à Julien Villery pour sa considération, sa disponibilité et son aide précieuse afin de mener à bien les calculs CIRCUL'AIR.*

*Mon positionnement au sein des deux laboratoires de géographie dijonnais a été enrichissant, autant sur le point professionnel que personnel, me permettant de côtoyer un grand nombre de personnes que je souhaiterais mentionner ici. Je remercie l'ensemble des membres du laboratoire ThéMA (actuels et anciens), Anne, Cécile, Cécilia, Dany, Françoise, Gilles, Hélène, Isabelle, Jean-Christophe, Marie-Hélène, Valérie et tous les autres. J'ai une pensée toute particulière pour Agnès et Patrick qui sont toujours présents pour nous : merci pour vos attentions de tous les jours. Au cours de ces derniers mois, plusieurs personnes ont accepté de porter leurs regards avisés sur ce travail. Je tiens à remercier Jean-Philippe, Julie et Robert pour les nombreux conseils et commentaires qu'ils ont pu m'apporter lors de la rédaction du manuscrit. Je souhaite également remercier l'ensemble des membres du CRC Albin, Benjamin B., Pascal, Pierre, Thierry, Yves et tous les autres. J'ai une pensée très particulière pour Benjamin P., dit Master P., auprès de qui j'ai pu partager bien plus qu'une passion et qui m'a permis de me perfectionner dans mon apprentissage de « padawan », en m'initiant à la climatologie et au diagramme de Taylor (couperas-tu ma tresse master ?). À ces remerciements, je tiens à associer Joseph qui, dès mon master, m'a initié aux statistiques et à Matlab et sans qui je n'aurais pu mettre en œuvre l'analyse des résultats en autonomie. Un grand merci à vous deux pour tout ce que vous avez pu m'apporter autant sur le plan scientifique qu'humain.*

*Que serait une thèse sans partenaires de route partageant le même fardeau ? J'ai une pensée pour les doctorants des deux laboratoires, que ce soit les anciens ou les actuels doctorants : merci à Basile, Christophe, Clémence, Clémentine, Damien, Émilie, Florian, Guillaume, Joanne, Joseph, Larissa, Marc, Medhi, Nicolas, Rébecca, Samuel, Sophie, Valentine, Vincent, Yohan… Aux anciens, auprès de qui j'ai pu apprendre et me former, à Guillaume, un ami dont les conseils, l'écoute, et le calme ont été importants pour moi, à Rébecca pour son dynamisme, sa bonne humeur et son attention. Aux actuels, particulièrement à Christophe, compagnon de route depuis la 1ere année de Licence et avec qui j'ai partagé bien plus qu'un bureau ! Obrigado à Larissa pour sa joie de vivre et le soleil brésilien qu'elle nous apporte, même lorsque la grisaille*

*dijonnaise devient persistante. À notre dijonnaise d'adoption, Émilie, pour ses encouragements et le temps qu'elle a accordée à la relecture de mes pépites. À Florian, avec qui, malgré la distance entre nos bâtiments, j'ai pu partager d'agréables moments lors de nos pauses et de nos sorties vélo. À tous les autres, stagiaires ou doctorants de passage, et avec qui nous avons pu parcourir un bout de chemin ensemble, Daniele, Natalie, Séverine, ainsi que Maud et Thibaut qui vont bientôt débuter leurs thèses. J'ai une pensée pour tous les doctorants JCE, Alexandre, Caroline, Elodie, Fam-Ky, Marion, Marouane, Pierre-Emmanuel et tous les autres. Mais surtout une pensée toute particulière pour Marion, amie et partenaire de route depuis le début de nos thèses respectives, et sans qui le master AE n'aurait pas du tout eu le même goût.*

*Au terme de ce parcours, je remercie enfin celles et ceux qui me sont le plus proches et que j'ai parfois délaissé au profit de mon travail. À toute ma famille, en particulier mes parents et mes beaux-parents, ma sœur et mon frère, les guedons et tous les autres. À tous mes amis proches, Jo, Simon S. et Antoine, Roukin, Marine et bébé Martinou, Bichon, Bastien M. (pour tous nos vendredis soir et ton aide indéfectible même depuis Perpignan), Stéph (pour tous ces moments passés ensemble depuis nos débuts, maintenant c'est à ton tour !), Vincent B, la troupe de l'AGS (Chacha, Luka, Nono, Lucile, Jem, Lisou, Simon F., Sophian, Pierre,…), les anciens collègues du Master TMEC (Bastien C., Batiste, Céline, Claire, Louise, Pti'Blond, Touf, Vianney …), Fanny, Émilie, Mika et Mélyne, Vincent N. (merci pour les math)… Je tiens à adresser un remerciement très spécial à Norbert Bousigue, qui fut le premier à m'enseigner la géographie il y a maintenant plus de 10 ans, et qui a bien voulu porter son regard sur le manuscrit pour me conseiller et y rechercher les pépites restantes.*

*Enfin, à mon infirmière qui, avec ses électrodes, scope ses patients. À toi, ma Pascaline qui m'a toujours donné l'équilibre et les moyens d'y arriver. Ta présence, ta patience, ton dynamisme et ta joie de vivre ont contribué, bien plus que tu ne le crois, à la réussite de ce travail. Merci pour tout !*

*Je ne voudrais pas oublier d'exprimer ma gratitude à ceux qui n'ont pas été mentionnés ici et que j'aurais pu oublier.*

***Merci à vous tous sans distinction et …***

***… que la force soit avec vous !***

# AVANT-PROPOS

Le travail que nous présentons a été réalisé grâce à l'utilisation de la plate-forme de simulation multi-agents GAMA-Platform (v.1.6.1 et v1.7 disponibles à l'adresse suivante : http://gama-platform.org/). Le modèle SCAUP (*Simulation multi-agents à partir de CApteurs Urbains pour la Pollution atmosphérique automobile*) développé dans le cadre de cette thèse et appliqué sur la ville de Dijon est disponible en accès libre[1] à l'adresse suivante (format *.rar*) :
https://je208836.github.io/TheseEmery/.

L'ensemble du plan d'expérience a été effectué grâce à OpenMole (Reuillon *et al.*, 2013) et son exécution ainsi que les analyses ont été effectuées grâce aux ressources HPC de la DSI-CCUB de l'Université de Bourgogne[2].

[1] Le modèle est disponible sous licence GNU GPL3, qui garantit aux utilisateurs les libertés d'exécuter le logiciel, d'étudier son fonctionnement, redistribuer des copies : https://www.gnu.org/copyleft/gpl.html.
[2] Calculations were performed using HPC resources from DSI-CCUB (Université de Bourgogne).

# SOMMAIRE

**INTRODUCTION GÉNÉRALE ..... 9**

*Une approche de construction de la connaissance : du comptage routier à la PAA à l'échelle urbaine ..... 12*
*Le géographe et la modélisation : de la donnée à l'expérimentation par simulation multi-agents ..... 13*
*La ville de Dijon comme terrain expérimental de la thèse ..... 14*
*Cheminement de la thèse : de la quantification à l'évaluation ..... 15*

**PARTIE 1 - LES MESURES DU TRAFIC ROUTIER EN MILIEU URBAIN : L'APPORT DES DISPOSITIFS DE COMPTAGE PAR CAPTEUR ..... 17**

INTRODUCTION DE LA PREMIÈRE PARTIE ..... 19
CHAPITRE 1. SIMULER L'ESPACE URBAIN EN MOUVEMENT ..... 23

*1.1. Quelles échelles spatio-temporelles pour étudier les dynamiques urbaines ? ..... 24*
*1.2. Entre rationalisme et empirisme ..... 35*
*1.3. La Géocomputation : un arsenal d'outils pour la constitution et l'étude de l'information spatio-temporelle ..... 45*
*Conclusion ..... 53*

CHAPITRE 2. L'ÉTUDE DU TRAFIC ROUTIER EN MILIEU URBAIN, ENTRE DISPOSITIF DE MESURES ET DISPOSITIF DE CONSTRUCTION ..... 55

*2.1. Capter le trafic routier ..... 56*
*2.2. Les systèmes de capteurs mobiles pour capter la trajectoire des véhicules ..... 64*
*2.3. Les données issues des capteurs fixes : une mesure collective ..... 71*
*Conclusion : ..... 78*

CHAPITRE 3. DU COMPTAGE ROUTIER AU DISPOSITIF DE QUANTIFICATION DU TRAFIC ROUTIER : LE CAS DE LA VILLE DE DIJON ..... 81

*3.1. Un rôle pivot des données de comptages routier pour la régulation de la circulation routière en milieu urbain ..... 83*
*3.2. Présentation du réseau de mesure de la ville de Dijon ..... 89*
*3.3. Que nous révèlent les capteurs de la ville de Dijon pour la visualisation du trafic routier en milieu urbain ? ..... 102*

CONCLUSION DE LA PREMIÈRE PARTIE ..... 117

**PARTIE 2 - DE LA MODÉLISATION À LA SIMULATION MULTI-AGENTS DU TRAFIC ROUTIER POUR LA PAA : UNE APPROCHE ORIENTÉE SUR LES DONNÉES ..... 119**

INTRODUCTION DE LA DEUXIÈME PARTIE ..... 121
CHAPITRE 4. MODÉLISER ET SIMULER LE TRAFIC ROUTIER EN MILIEU URBAIN POUR LA QUALITÉ DE L'AIR : DÉFIS ET ENJEUX ..... 125

*4.1. Un retour sur les concepts de modèle(s) et de simulation en sciences humaines ..... 127*
*4.2. Quel modèle de trafic retenir pour l'étude de la PAA ? ..... 133*
*4.3. Les enjeux liés à une approche de simulation du trafic routier à partir des comptages routiers ..... 141*
*Conclusion : ..... 150*

CHAPITRE 5. LES APPORTS DE LA SIMULATION MULTI-AGENTS POUR L'ÉTUDE DU TRAFIC ROUTIER ..... 151

*5.1. Le renouveau de la modélisation du trafic routier par la simulation-multi-agents ..... 152*
*5.2. Les approches de géosimulation pour la représentation du trafic routier : quelle méthode privilégier ? ..... 159*
*5.3. La SMA pour la modélisation du trafic routier à partir des capteurs de trafic routier : le modèle SCAUP ..... 165*
*Conclusion : ..... 182*

CHAPITRE 6. EXPÉRIMENTATION ET MISE EN MOUVEMENT DU TRAFIC ROUTIER PAR LA SIMULATION MULTI-AGENTS ..... 183

*6.1. Les plates-formes de simulation multi-agents : ..... 184*
*6.2. Formalisation du modèle de simulation multi-agents : du projet à l'expérimentation de SCAUP ..... 190*
*6.3. Expérimentation et plan d'expérience du modèle SCAUP : ..... 205*

CONCLUSION DE LA DEUXIÈME PARTIE ..... 217

**PARTIE 3 - DU TRAFIC ROUTIER À LA POLLUTION ATMOSPHÉRIQUE AUTOMOBILE : VALIDATION, ÉVALUATION ET PERSPECTIVES ..... 219**

INTRODUCTION DE LA TROISIÈME PARTIE ..... 221
CHAPITRE 7. SIMULATION DU TRAFIC ROUTIER SUR LA VILLE DE DIJON : DE LA SIMULATION À L'ÉVALUATION ..... 223

*7.1. Exploration du modèle SCAUP à travers le cas de la ville de Dijon ..... 225*
*7.2. Analyse préliminaire de l'approche expérimentale de simulation du trafic routier ..... 235*
*7.3. Comment choisir la meilleure simulation : de la simulation à la réalité ..... 246*
*Conclusion : ..... 260*

CHAPITRE 8. UN DISPOSITIF D'ÉVALUATION ENVIRONNEMENTALE : DU TRAFIC ROUTIER AUX ÉMISSIONS DE POLLUANTS ATMOSPHÉRIQUES AUTOMOBILES ..... 263

*8.1. Les enjeux de la spatialisation des émissions de polluants atmosphériques automobiles ..... 265*
*8.2. Données d'entrée et méthodologie de calcul des émissions avec CIRCUL'AIR ..... 271*
*8.3. Spatialisation de la pollution atmosphérique automobile sur la ville de Dijon ..... 281*
*Conclusion : ..... 298*

CHAPITRE 9. LES PERSPECTIVES DE DÉVELOPPEMENT AUTOUR DE LA MATRICE TECHNIQUE : DE LA QUANTIFICATION À L'ÉVALUATION ..... 301

*9.1. Les réflexions autour du dispositif de quantification pour alimenter et façonner le dispositif de modélisation ..... 303*
*9.2. Une approche de modélisation SCAUP à affiner ..... 311*
*9.3. S'inscrire dans une démarche intégrée d'évaluation de la PAA en milieu urbain : de l'émission à l'imission ..... 320*

CONCLUSION DE LA TROISIÈME PARTIE ..... 333

**CONCLUSION GÉNÉRALE ..... 335**

*Une approche à l'intersection de deux géographies ..... 338*
*La matrice technique : une colonne vertébrale pour la construction d'une connaissance scientifique ? ..... 339*
*Le géographe, l'informaticien et le statisticien : un ancrage dans l'interdisciplinarité ..... 339*
*De l'homme moyen à l'homme capteur : vers une "gouvernance" du capteur ? ..... 341*

**BIBLIOGRAPHIE ..... 345**

**TABLE DES MATIÈRES ..... 361**

**TABLE DES FIGURES ..... 365**

**TABLE DES TABLEAUX ..... 371**

**ANNEXES ..... I**

**ANNEXES CARTOGRAPHIQUES ..... VI**

# INTRODUCTION GÉNÉRALE

*Circuler, réguler, rythmer, perturber, pulser* sont autant de verbes qui caractérisent le trafic routier. Ils font implicitement référence à un fonctionnement du trafic routier qui peut s'apparenter à un métabolisme physique (Barles, 2008). En milieu urbain, cette métaphore prend encore plus de sens, par exemple, lorsque l'on caractérise les grandes infrastructures de transport de l'espace urbain comme des *artères urbaines* qui font intuitivement penser à notre système sanguin. Avant tout, pour qu'il y ait trafic routier, il faut deux éléments (Bavoux *et al.*, 2005) qui sont, d'une part, les véhicules routiers en déplacement, et d'autre part, les infrastructures de transport qui supportent ces mouvements. En d'autres termes, sans route, et inversement sans véhicules, il ne peut y avoir de trafic routier. Dans cette acception, le trafic routier peut s'apparenter à un flux sanguin, où le déplacement des véhicules est assuré et porté, au même titre que par les *artères*, par les infrastructures de transport, tandis que les véhicules peuvent être assimilés à des *globules rouges* permettant d'alimenter nos *organes* (les lieux d'activités et de loisirs) en substances qui peuvent alors être des individus, des marchandises ou des services. Enfin, à cette machine cardiaque, il manque un élément essentiel qui est son *cœur,* c'est-à-dire la machine qui permet de faire circuler, rythmer et pulser les véhicules autant dans l'espace que dans le temps. Ce dernier élément, qui assure la *circulation routière* ou le *trafic routier* (Lévy et Lussault, 2003), est en très grande partie supporté par nos activités et nos déplacements entre différents espaces (domicile, travail, loisir…), qui sont à l'origine des mouvements intraurbains.

Ces mouvements s'établissent à différentes échelles de temps, allant des plus lentes (les siècles) aux plus immédiates (la minute voire la seconde) (Wegener, 1994, 2004). À cette dernière échelle, ce sont d'ailleurs les transports de marchandises et les déplacements des personnes qui rythment et affectent l'espace urbain. De plus, on note qu'à cette échelle près de 65 % des déplacements locaux sont assurés par les véhicules routiers (CGDD, 2010). À ce niveau, le trafic routier joue un rôle important sur l'environnement urbain en affectant, notamment la qualité de l'air. Les impacts environnementaux du trafic routier en milieu urbain constituent aujourd'hui des enjeux socio-politiques de premier ordre (CEREMA, 2008 ; CGDD, 2016). On notera d'ailleurs que, depuis 2013, la pollution atmosphérique a été placée par l'Organisation Mondiale de la Santé (OMS) comme l'un des principaux risques environnementaux pour la santé dans le monde. Par ailleurs, à l'échelle urbaine, le trafic routier est considéré avec l'industrie comme le principal contributeur à une mauvaise qualité de l'air (OMS, 2006 ; Fenger, 2009 ; CITEPA, 2015). C'est pourquoi la mesure et la quantification des émissions constituant la Pollution Atmosphérique Automobile (PAA), autant dans l'espace que dans le temps, apparaissent comme une préoccupation autant politique et sociale qu'environnementale. Dès lors, la première étape pour la mise en œuvre de politiques locales efficaces pour appréhender les émissions de PAA nécessite d'estimer le flux de trafic routier avec précision, ce qui s'opère le plus généralement par l'intermédiaire d'un modèle de trafic routier en entrée du calcul des émissions de PAA (Fouillé *et al.*, 2012 ; PCIT, 2012 ; Primequal-Predit, 2012).

L'évaluation de la PAA constitue l'objectif final de ce travail. Dès lors, il faut d'abord être en mesure de relever et de quantifier avec finesse le trafic routier, que ce soit dans l'espace et dans le temps. À ce niveau, la donnée et le système de mesure du trafic routier prennent une place décisive et considérable dans les calculs d'émissions (Fouillé *et al.*, 2012). Alors que le comptage a été la première mesure mise en place pour le suivi du trafic routier, on remarque aujourd'hui que ce sont plutôt les enquêtes qui sont privilégiées en entrée des modèles afin de mesurer le trafic

routier sur un espace donné (Bonnel, 2002 ; Ortúzar et Willumsen, 2011). Or, à travers les enquêtes, le trafic n'est pas donné, il est plutôt construit par l'intermédiaire d'un modèle dit *classique* et dont la forme la plus connue est le *modèle à quatre étapes* (Bonnel, 2002 ; Ortúzar et Willumsen, 2011 ; Commenges, 2013). De plus les enquêtes ne permettent pas d'appréhender avec finesse les temporalités du trafic routier. Tandis que les enquêtes continuent à présenter leurs avantages et leurs inconvénients pour la modélisation des dynamiques urbaines, d'autres systèmes de mesure, comme les systèmes de capteurs, présentent l'opportunité de relever avec plus de finesse les flux de trafic routier.

C'est particulièrement le cas des capteurs électromagnétiques, qui à notre sens remettent au goût du jour les comptages. Effectivement, au contraire des enquêtes qui se caractérisent par un aspect sémantique très riche et par une résolution temporelle pauvre, le comptage se distingue par son caractère temporel très riche, tandis que sa dimension sémantique en demeure relativement pauvre (Commenges, 2013). À l'inverse, l'enquête apparait qualitativement très riche, tandis que le comptage apparaît quantitativement très riche. En outre, en se situant au cœur de la mesure et en se structurant sous la forme d'un réseau (Akyildiz *et al.*, 2002 ; Servigne *et al.*, 2009), les capteurs fournissent l'opportunité de se rapprocher de la mesure en tant que telle. Par exemple, le suivi d'un patient sous électrodes permet de surveiller ainsi que d'observer son rythme cardiaque, qu'il soit normal ou anormal. Les électrodes permettent alors d'enregistrer l'activité cardiaque d'une personne et ne peuvent être employées à d'autres fins que celui-ci. Les données issues des capteurs électromagnétiques exploités au cours de cette thèse sont des données automatiques dédiées à la gestion du trafic routier en milieu urbain présentant des caractéristiques relativement similaires aux mesures réalisées par les électrodes. Formant un réseau de près de 200 capteurs, ils fournissent, par rapport aux données d'enquêtes, un flot d'informations considérable pour l'observation fine du trafic routier en milieu urbain.

En reprenant la métaphore du métabolisme urbain, on observe alors les pulsations cardiaques d'un patient (*la ville*) par l'intermédiaire des électrodes (*capteurs*). Les systèmes de capteurs nous fournissent l'opportunité de relever, au même titre que des *électrodes*, les pulsations (les rythmes) du trafic routier, autant dans le temps très court (la minute) que dans le temps long (les années). En ce sens, *la ville sous électrodes* portée par les réseaux de capteurs et leurs comptages nous donne aujourd'hui l'opportunité d'appréhender plus finement la dynamique du trafic routier.

### Une approche de construction de la connaissance : du comptage routier à la PAA à l'échelle urbaine

À mesure que l'homme s'équipe de capteurs en tout genre (*smartphone*, GPS, montre connectée…), les villes tendent également à investir dans la construction de réseaux de capteurs de plus en plus intelligents en s'appuyant sur une multitude de capteurs urbains. Dans le domaine des transports, les capteurs permettent, par exemple, d'observer la fréquentation d'un espace, de compter le passage des véhicules, où de s'informer sur le nombre de places de stationnement disponibles. Le capteur et sa mesure se situent aujourd'hui dans des tendances émergentes que l'on peut autant relier aux phénomènes du *big data* qu'à la ville intelligente (*smart city*), dont la géographie est partie prenante (Batty, 2012, 2013, Kitchin, 2013, 2014). Denise Pumain (2014a) relève d'ailleurs que les « *big data renferment certes des indices de notre présence, mais elles sont loin de parler seules : que révèlent-elles de nos préférences et de nos mouvements, quelle part de vérité contiennent-elles, peut-on s'y*

*fier vraiment ? Pour en tirer parti, des recherches d'un nouveau genre sont requises à tous les niveaux : observation soigneuse des technologies et des services qui produisent les données massives,* […], *observation des apports de leur analyse aux savoirs préexistants, obtenus avec d'autres méthodes et d'autres indicateurs. Il serait en effet désolant de ne redécouvrir avec elles que la pertinence du modèle gravitaire, alors qu'elles nous informent bien plus subtilement quant au façonnement réciproque des structures territoriales et des pratiques et des représentations dans des sociétés inégalement hyper-mobiles* ».

Par conséquent ces derniers propos invitent les géographes à s'inscrire dans une démarche expérimentale, en positionnant non plus la donnée en fin du processus, mais plutôt dans une vision procédurale de construction de la connaissance (Joliveau, 2004), positionnant alors la donnée au premier plan de la démarche sans y apporter d'hypothèses préalables. Pour le trafic routier, la donnée de comptage constitue la première porte d'entrée de ce travail, en vue d'en extraire une information sur la circulation routière pour ensuite en construire une connaissance sur les émissions de PAA. Dans ce sens, en quoi les données de comptages permettraient-elles d'améliorer l'estimation et la spatialisation de la PAA en milieu urbain ? Cette démarche de modélisation nécessite la mise en place de trois phases de travail. La première vise à *entitiser* les données de comptage routier mesurées à travers un réseau de capteurs électromagnétiques : comment qualifier ces données ? Que nous révèlent les comptages routiers ? La deuxième phase, quant à elle, vise à passer de ce comptage ponctuel à une information linéaire sur un plus grand ensemble d'infrastructures de transport : comment générer un trafic routier à partir des comptages routiers ? Enfin la dernière phase vise à transformer cette information sur la circulation routière en une connaissance sur les émissions de polluants atmosphériques à l'échelle de l'espace urbain : comment à partir des informations sur la circulation routière en estimer ensuite les émissions de polluants atmosphériques qui lui sont rattachées ?

## Le géographe et la modélisation : de la donnée à l'expérimentation par simulation multi-agents

Toutefois, il faut noter que le passage de la donnée de comptage aux émissions de polluants atmosphériques n'est pas aisé et nécessite l'application et la mise en œuvre d'une démarche de modélisation (Chorley, 1964) où les données sont positionnées au premier plan (*data driven*). Le concept de *data driven* en géographie a, plus particulièrement, été initié en réponse aux flux massifs de données issues des systèmes de capteurs et des individus (Miller et Goodchild, 2014). Dans le champ des méthodes axées sur les données, le cadre méthodologique dans lequel s'inscrit ce travail se rattache à la *géocomputation* (Openshaw et Abrahart, 2000 ; Brunsdon et Singleton, 2015). Cet ancrage apparait intéressant car il fournit au géographe un arsenal d'outils axés sur les données pour la représentation des informations spatio-temporelles (Banos, 2013). C'est particulièrement le cas des outils de géosimulation, dont la Simulation Multi-agents (SMA) est aujourd'hui un outil largement investigué et exploité par les géographes pour la représentation des dynamiques spatio-temporelles (Daudé, 2005 ; Amblard et Phan, 2006 ; Daudé et Langlois, 2006 ; Sanders, 2007 ; Banos, 2013).

C'est l'approche que nous avons privilégiée pour le passage de la donnée de comptage à une information simulée sur les infrastructures de transport. Par nature générative (Epstein, 1999), la SMA fournit un cadre expérimental de simulation du trafic routier d'autant plus intéressant puisqu'il s'inscrit dans le champ des systèmes complexes (Batty, 2007). De plus, dans une logique

de SMA, l'approche guidée par les données vise à obtenir des résultats de simulation qui sont plus proches de la donnée (Hassan *et al.*, 2010a, 2010b). L'interdépendance entre donnée et SMA fournit un cadre expérimental de modélisation qui s'insère autour de plusieurs éléments clés caractérisant la modélisation en géographie qui, pour Arnaud Banos (2013), vise à :

- **Étudier un système réel** qui sera ici appliqué sur la ville de Dijon
- **S'enraciner dans la donnée** en exploitant exclusivement les données de comptage routier issues d'un système de capteurs *in situ*
- S'inscrire dans une **démarche de simplification et de parcimonie** ; en nous entourant de la théorie des systèmes complexes nous envisagerons une formalisation et une conceptualisation du trafic routier autour de deux éléments qui définissent le trafic routier : les véhicules et la route.
- Construire un modèle pour **réaliser des expérimentations**, l'approche de simulation à travers les données de capteurs sera expérimentée pour améliorer les calculs d'émissions de la PAA en milieu urbain en s'appuyant sur les outils utilisés par les Associations Agréées de la Surveillance de la Qualité de l'Air (AASQA).

La démarche de modélisation entreprise au cours de cette thèse vise à construire un modèle de simulation du trafic routier à partir des données issues des capteurs urbains pour l'évaluation de la PAA. En s'appuyant sur les comptages routiers, cette démarche de modélisation vise à apporter une autre réponse que les approches dites *classiques* pour l'estimation des niveaux de polluants automobiles en milieu urbain. Au même titre que les personnels infirmiers « *scopent* » leurs patients, c'est-à-dire, examinent et observent les mesures cardiaques d'un patient par l'intermédiaire d'un moniteur[3], la démarche de modélisation SCAUP (*Simulation multi-agents à partir de CApteurs Urbains pour la Pollution atmosphérique automobile*) proposée dans cette thèse vise à assurer un processus d'observation du trafic routier qui s'en rapproche. En poussant la métaphore du métabolisme urbain, l'approche que nous proposons vise ainsi à *SCAUPER* le trafic routier pour l'étude de la PAA en en exploitant les électrodes urbaines dédiées à la mesure et au suivi du trafic routier.

## La ville de Dijon comme terrain expérimental de la thèse

Ce travail s'inscrit dans une optique de recherche appliquée à un territoire urbain de la région (financement par la région Bourgogne-Franche-Comté). C'est pourquoi l'application du cadre expérimental de cette thèse a été réalisée en s'appuyant sur le territoire de la ville de Dijon, où près de 200 capteurs du trafic routier sont disséminés sur le réseau routier depuis plus de dix ans. Ces données de comptage, dont la vocation la régulation du trafic routier, sont aussi exploitées par les Associations Agréées de la Surveillance de la Qualité de l'Air (AASQA) pour la réalisation des cadastres d'inventaire des émissions du territoire (PCIT, 2012).

Les comptages font alors le pont entre deux acteurs du territoire dijonnais qui sont, d'une part, le régulateur de la circulation routière en charge de la collecte et de la gestion des comptages routiers du territoire, et d'autre part, les AASQA qui exploitent ces comptages pour la réalisation des inventaires territoriaux de la pollution atmosphérique, qui sont des documents d'entrée en matière de gestion des impacts de la pollution atmosphérique à l'échelle des territoires. Enfin, il

[3] Définition de « **scoper** » issues du Petit Robert 2006.

faut noter que cette thèse est parrainée par l'AASQA locale dijonnaise : ATMOSF'AIR BOURGOGNE. C'est pourquoi ce travail s'inscrit dans une recherche appliquée en support des AASQA en utilisant leur expertise en matière de pollution atmosphérique. Le contexte local dijonnais fournit l'opportunité de tester notre approche sur un système urbain *réel* en s'appuyant sur les compétences du gestionnaire de trafic routier pour le recueil et la localisation des comptages routiers, et sur les compétences d'ATMOSF'AIR BOURGOGNE quant au passage des informations sur la circulation routière au calcul des émissions de PAA.

## Cheminement de la thèse : de la quantification à l'évaluation

L'ensemble de ces éléments montre tout l'intérêt de notre cadre expérimental, qui vise à rapprocher la donnée de ce qui est simulé, dans une optique de recherche autant appliquée à travers le cadre dijonnais, que de recherche fondamentale lorsque nous abordons les concepts des approches guidées par les données. D'une manière générale, en nous appuyant sur la richesse des données de capteurs et sur la Simulation Multi-Agents (SMA), sommes-nous en capacité de générer un trafic routier émergent et représentatif de la circulation routière observée sur la ville de Dijon ? Cependant, ce postulat de départ soulève de nombreux écueils et enjeux qui seront soulevés tout au long de cet exposé :

- En quoi les systèmes de capteurs permettent de renouveler, voire d'actualiser, la représentation du trafic routier en milieu urbain ? Quel rôle jouent les données dans la démarche de construction de la connaissance en géographie ?
- Comment à partir de ces comptages construire un modèle de simulation du trafic routier en vue d'obtenir un instrument d'observation de la circulation routière en milieu urbain plus précis ? Dans quel(s) cadre(s) conceptuel(s) et méthodologique(s) valoriser les données pour la simulation du trafic routier ?
- Dans quel cadre la simulation multi-agents du trafic routier permet d'alimenter les outils de spatialisation de la pollution atmosphérique en milieu urbain ? Quels en sont les apports, ainsi que les limites, pour la spatialisation des émissions de PAA à l'échelle urbaine ? Quelles en sont les réponses opérationnelles pour appuyer l'expertise des AASQA en matière de qualité de l'air ?

Ces questionnements s'intègrent aisément dans un triptyque (re)mis en avant par Hadrien Commenges (2013) à travers la notion de *matrice technique*. Ce terme est employé pour définir « *un ensemble de dispositifs techniques qui se tient, c'est-à-dire dont chaque dispositif est relié aux autres par des relations durables qui assurent la stabilité de l'ensemble* » (Commenges, 2013, p. 75). Trois dispositifs sont intégrés dans cette matrice : *quantification*, *modélisation* et *évaluation*. Ces trois dispositifs ne sont pas à considérer comme des isolats mais bien comme des éléments en relation et en interdépendance entre eux, sans quoi la matrice technique ne peut tenir. Nous avons pour notre part souhaité exploiter ce cadre pour construire le cheminement de ce travail, qui à notre sens s'intègre aisément à notre problématique de travail. Par ailleurs, on peut faire un parallèle avec la démarche de construction de la connaissance qui se constitue de trois phases : le traitement de la donnée (*la quantification des véhicules routiers*), vers l'extraction d'une information (*la simulation multi-agents du trafic routier*), pour aboutir à la constitution d'une connaissance (*l'évaluation de la PAA*). C'est pourquoi la mise en œuvre d'une démarche de modélisation quelle qu'elle soit nécessite de s'arrêter sur :

- 1. la *quantification* du phénomène que l'on souhaite étudier,
- 2. la *modélisation* qui en sera mise en place à partir des données à disposition,
- 3. l'objectif visé par le modèle pour l'*évaluation* d'un phénomène quelconque.

Le cheminement est structuré autour des trois composantes de la matrice technique pour présenter la démarche de modélisation SCAUP. La première partie a pour objet de cerner les principaux enjeux de la *quantification* du trafic routier pour l'étude spatio-temporelle du trafic routier. La première étape de cette partie consistera à mettre en avant le contexte conceptuel et méthodologique de ce travail. Ensuite, nous nous focaliserons plus spécifiquement sur la mesure du trafic routier en nous intéressant plus particulièrement aux données. Il s'agit ici de différencier les systèmes de capteurs pour la mesure du trafic routier. Enfin, la dernière étape de cette partie vise à présenter et à exploiter les données de comptage recueillies sur la ville de Dijon et à en évaluer leurs atouts et limites pour l'observation du trafic routier en milieu urbain. Cette première partie permettra, en outre, de mettre en avant les défis conceptuels et méthodologiques de la manipulation des capteurs de trafic routier, dans l'optique d'une simulation de ces données au sein d'un modèle de trafic routier.

Consacrée à la *modélisation*, la seconde partie présentera les défis et enjeux qu'impliquent une approche de simulation du trafic routier non seulement initialisée par l'intermédiaire des données de comptages routiers, mais aussi dans l'optique d'une évaluation des émissions de PAA. Ensuite, nous nous focaliserons sur les apports de la simulation multi-agents dans le domaine de la modélisation du trafic routier en milieu urbain, afin de présenter ensuite la démarche de modélisation SCAUP initiée dans le cadre de ce travail. Enfin, nous passerons à l'expérimentation de ce modèle sur un espace *théorique* pour juger du fonctionnement et des capacités de l'approche proposée pour la simulation du trafic routier en milieu urbain.

La dernière partie sera réservée à l'application expérimentale du modèle SCAUP sur la ville de Dijon. Un premier chapitre se focalisera sur une démarche d'*évaluation* du modèle en vue de sélectionner les données de simulation les plus caractéristiques de ce qui est observé à travers les comptages *in situ*. Cette première phase d'évaluation est importante car elle permet, à la suite, de mettre en avant le processus de calcul des émissions réalisées en collaboration avec ATMOSF'AIR BOURGOGNE qui sera entrepris dans un second temps. Enfin, les perspectives de recherches seront proposées à l'issue de cette dernière partie.

# PARTIE 1 -

# LES MESURES DU TRAFIC ROUTIER EN MILIEU URBAIN : L'APPORT DES DISPOSITIFS DE COMPTAGE PAR CAPTEUR


INTRODUCTION DE LA PREMIÈRE PARTIE....19
CHAPITRE 1. SIMULER L'ESPACE URBAIN EN MOUVEMENT....23
*1.1. Quelles échelles spatio-temporelles pour étudier les dynamiques urbaines ?....24*
*1.2. Entre rationalisme et empirisme....35*
*1.3. La Géocomputation : un arsenal d'outils pour la constitution et l'étude de l'information spatio-temporelle....45*
*Conclusion....53*
CHAPITRE 2. L'ÉTUDE DU TRAFIC ROUTIER EN MILIEU URBAIN, ENTRE DISPOSITIF DE MESURES ET DISPOSITIF DE CONSTRUCTION....55
*2.1. Capter le trafic routier....56*
*2.2. Les systèmes de capteurs mobiles pour capter la trajectoire des véhicules....64*
*2.3. Les données issues des capteurs fixes : une mesure collective....71*
*Conclusion :....78*
CHAPITRE 3. DU COMPTAGE ROUTIER AU DISPOSITIF DE QUANTIFICATION DU TRAFIC ROUTIER : LE CAS DE LA VILLE DE DIJON....81
*3.1. Un rôle pivot des données de comptages routier pour la régulation de la circulation routière en milieu urbain....83*
*3.2. Présentation du réseau de mesure de la ville de Dijon....89*
*3.3. Que nous révèlent les capteurs de la ville de Dijon pour la visualisation du trafic routier en milieu urbain ?....102*
CONCLUSION DE LA PREMIÈRE PARTIE....117

# INTRODUCTION DE LA PREMIÈRE PARTIE

Depuis 2013, l'OMS alerte les décideurs publics sur les effets néfastes de la pollution atmosphérique pour la santé humaine[4]. Au-delà de cet aspect sanitaire, la pollution de l'air s'avance non seulement comme une véritable préoccupation socio-environnementale, mais aussi politique afin d'en *identifier* les sources, d'en *mesurer* ses impacts autant sur les populations que sur la structure atmosphérique, en vue de mettre en œuvre des actions pour les *diminuer* (Primequal-Predit, 2012). La quantification, la surveillance et la prévention de la pollution atmosphérique présentent donc un réel enjeu de santé publique autant à l'échelle planétaire que locale, pour en identifier les facteurs naturels et anthropiques rentrant en jeu. À l'échelle locale, la pollution atmosphérique est à relier directement à l'intensité des sources d'émission qui accentuent, ou pas, les concentrations de polluants dans l'atmosphère (Roussel, 2006). À cette échelle, l'Agence Européenne pour l'Environnement (AEE) considère que dans les villes européennes, les épisodes de pollution atmosphérique d'origine industrielle appartiennent à un autre temps et, à l'heure actuelle, que la principale source d'une mauvaise qualité de l'air est à relier majoritairement au trafic routier (AEE, 2014). Tout au long de ce travail, nous nous focaliserons sur cette échelle locale où les concentrations de polluants atmosphériques sont plus directement associées aux sources primaires (Dab et Roussel, 2001), au premier rang desquelles le trafic routier constitue la principale cause d'une mauvaise qualité de l'air (Primequal-Predit, 2012 ; AEE, 2014 ; CITEPA, 2015).

Comme il est mentionné dans un ouvrage de synthèse consacré à la recherche sur la pollution de l'air et des transports terrestres (Primequal-Predit, 2012, p. 21): « *la première étape pour mettre en œuvre des politiques publiques pertinentes et efficaces est de pouvoir appréhender et définir précisément la pollution de l'air dans son ensemble* ». De ce point de vue, l'instrument de mesure occupe une place centrale et déterminante pour évaluer les impacts de la pollution atmosphérique autant dans le temps que dans l'espace. C'est particulièrement le cas des études relatives à la pollution de proximité automobile qui se caractérise par l'émission de polluants primaires dont les caractéristiques présentent une forte volatilité spatio-temporelle (Dab et Roussel, 2001 ; Roussel, 2006 ; Elichegaray *et al.*, 2008). Plus directement reliée à leurs sources de polluants (les véhicules), l'évaluation de la Pollution Atmosphérique d'origine Automobile (PAA) implique, tout d'abord, de disposer d'outils de mesure adéquats pour spatialiser les volumes de trafic routier sur l'ensemble des tronçons routiers d'une agglomération (PCIT, 2012).

Dans cette acception, préalablement à la mise en place d'outils d'évaluation de la PAA, il apparaît important d'obtenir une bonne image des flux de trafic routier à une échelle spatiale fine, mais également à une échelle temporelle fine. Pour y parvenir, la piste privilégiée est de mobiliser les modèles de trafic et leurs résultats (les trafics modélisés) pour alimenter un modèle d'émission et de dispersion des polluants de l'air (PCIT, 2012). Cependant, les relations entre trafic routier et pollution atmosphérique s'observent à différents pas de temps dont les plus caractéristiques s'opèrent à des échelles spatio-temporelles relativement fines (Coria *et al.*, 2015). À cette échelle, les modèles de trafic utilisés pour l'étude de la PAA ne permettent pas de répondre finement à

[4] Communiqué de presse du 17 Octobre 2013 de l'OMS et de l'agence internationale de recherche sur le cancer : « *Pollution de l'air extérieur : une cause environnementale des décès associés au cancer* » : http://www.iarc.fr/en/media-centre/iarcnews/pdf/pr221_E.pdf.

ces questions. De plus, au sein de ces modèles, le trafic routier est le résultat d'un processus de construction et de modélisation basé sur la demande et l'offre de déplacement (Bonnel, 2002) et la question du système de mesure utilisé en aval n'est que peu souvent mise en avant, alors qu'il est décisif pour l'estimation des PAA (Fouillé *et al.*, 2012). De ce point de vue, nous proposons plutôt d'inverser la tendance, via une approche *bottom-up*, en se rapprochant du véhicule et de sa mesure en tant que tel dans une perspective d'évaluation de la PAA. C'est pourquoi, dans cette partie, nous nous focaliserons plus spécifiquement sur les données issues des capteurs disséminés dans l'espace urbain et permettant de quantifier les véhicules routiers dans le temps. L'intérêt de mobiliser les capteurs réside dans leur capacité à se situer au cœur du phénomène à mesurer, se distinguant dès lors des données issues des enquêtes (Akyildiz *et al.*, 2002 ; Servigne *et al.*, 2009). C'est pourquoi, dans un premier temps, il est nécessaire de s'interroger sur les dispositifs de quantification du trafic routier. La quantification se situe en amont de la démarche de modélisation et inclut un double travail : « *convenir, c'est-à-dire s'accorder sur la définition d'objets conceptuels ; puis mesurer les objets ainsi définis, les exprimer sous forme numérique* » (Commenges, 2013, p. 18). Prérequis à notre approche, le dispositif de quantification représente la donnée et les objets qui lui sont associés, permettant ensuite d'alimenter le modèle de représentation du trafic routier. Cette étape de dénombrement constitue la première phase du travail, dont l'objectif est de disposer de mesures relatives au trafic routier en milieu urbain. Le trafic routier est un objet éminemment dynamique qui s'inscrit autant dans le temps que dans l'espace. Il se définit avant tout comme étant un flux de déplacements de véhicules sur une infrastructure routière (Bavoux *et al.*, 2005).

L'objectif de cette partie vise à montrer comment les données de capteurs pourraient permettre de renouveler l'analyse du trafic routier en milieu urbain pour l'évaluation de la PAA. Ce propos constitue l'un des points de départ de cette thèse. En effet, le parti pris de ce travail réside dans l'application d'une approche guidée par des données spatio-temporelles issues de capteurs implantés dans l'espace urbain. Néanmoins, l'utilisation des données issues des capteurs n'est pas directe. Il faut donc, au préalable de la démarche initiée dans cette thèse, identifier les objets et dispositifs pertinents à notre problématique guidée par les données : quel processus « *d'entitisation* [5]» envisager pour l'étude de la dynamique du trafic routier à travers un système de mesure ? Comment disposer de mesures de trafic fiables et continues dans la complexité d'un espace urbain ? En quoi les capteurs permettraient de renouveler, voire d'actualiser, la représentation spatio-temporelle du trafic routier ?

La première partie de la thèse s'appuie sur la construction d'une démarche géographique spatio-temporelle depuis un corpus de données issues d'un système de capteurs placés dans l'espace urbain. Elle se construit à travers trois cadres de développement : le premier chapitre vise à passer en revue les défis thématiques et conceptuels d'une démarche de simulation des dynamiques du trafic routier. Ce premier travail permettra d'inscrire la problématique de la thèse dans une optique spatio-temporelle, mais également dans un contexte d'évaluation de la PAA. Il s'agira aussi de s'interroger sur la place des données dans le processus de construction d'une information spatio-temporelle et des outils y étant associés en se focalisant plus particulièrement sur les

[5] MSH de Dijon, 24 Octobre 2014, « *En quel sens parlons-nous de sciences « sociales »* » : Conférence de Jean-Louis Fabian et Lena Sanders sur les processus d'entitisation : http://msh-dijon.u-bourgogne.fr/toute-lactualite/actualites-externes/677-ecoutez-ou-reecoutez-la-matinee-de-reflexion-l-avenir-des-shs-enjeux-et-perspectives-pour-la-recherche-du-24-octobre-dernier.html.

concepts issus de la *géocomputation* (Openshaw *et al.*, 1987 ; Openshaw et Abrahart, 2000). Les deuxième et troisième chapitres analysent de plus près les mesures issues des systèmes de capteurs pour appréhender le trafic routier en milieu urbain. Le corpus de données mobilisé dans ce travail sera présenté à la suite, afin d'en expliciter les défis méthodologiques pour la simulation du trafic routier. À travers cette partie, ce sont les enjeux conceptuels, empiriques et heuristiques de la thèse qui seront présentés. Ils permettront d'introduire le travail de simulation du trafic routier présenté en partie 2.

# Chapitre 1. SIMULER L'ESPACE URBAIN EN MOUVEMENT

Les mouvements de la ville se manifestent à différentes échelles, allant du niveau le plus global (les réseaux de villes) au plus local (les individus). À l'échelle locale, ce sont les mouvements des individus pour se déplacer vers leurs lieux de travail et de loisirs qui constituent la majorité des *palpitations urbaines* (Banos et Thévenin, 2005). À ce niveau, on parle plutôt de mouvement intraurbain par opposition aux mouvements interurbains qui rendent compte des échanges et relations qu'entretiennent les villes et les espaces qui les entourent (Pumain, 2014b). Au cours de ce travail, nous nous intéresserons aux mouvements intraurbains, ces mouvements individuels constituant ce que Arnaud Banos (2013, p. 38), appelle la « *fourmilière urbaine* » ; pour lui, au-delà d'une métaphore naturaliste, « *une telle mise en parallèle ouvre la voie à des horizons méthodologiques renouvelés, faisant de la modélisation et de la simulation des méthodes d'investigation scientifique à part entière, à même de nous aider à mieux explorer les phénomènes dans leur complexité* ».

Ici, le dispositif de quantification à mobiliser est primordial en vue de quantifier les mouvements de l'espace urbain et ensuite de pouvoir en évaluer ses impacts sur les niveaux de pollution atmosphérique (PAA). La question transversale de ces enjeux est de fournir aux décideurs locaux une mesure de ces phénomènes, qu'ils soient humains ou physiques. Si ultérieurement la pauvreté des données a été largement mise en avant, l'avènement des *big data* ouvre la voie à de nouvelles méthodologies pour l'étude et la représentation des mouvements intraurbains. Néanmoins, ce type de données se réfère à des données qui dépassent notre capacité à les analyser (Miller et Goodchild, 2014) et donc à en extraire une connaissance. Le flot de données présent nécessite d'avoir une attention particulière sur la méthodologie et les outils employés pour les faire parler et en extraire une information géographique pertinente (Kitchin, 2014). Si certains voient à travers les *big data* l'apparition d'un nouveau paradigme scientifique se faisant au détriment des approches rationalistes (Hey *et al.*, 2009), l'émergence d'un nouveau paradigme scientifique ne peut être entrevue que par une remise en cause des outils actuellement utilisés (Kuhn, 1970). Cependant, force est de constater que les démarches guidées par les données ne remettent pas totalement en cause les outils traditionnels (Miller et Goodchild, 2014). Il s'agit plutôt d'un ensemble de concepts, de méthodes et d'outils, qui à travers le concept interdisciplinaire de la « *géocomputation* », ouvre la voie à la mise en place de diverses approches axées sur les données pour l'étude des dynamiques urbaines (Openshaw et Abrahart, 2000 ; Brunsdon et Singleton, 2015).

Ce chapitre vise à apporter des éléments de réflexion sur l'insertion des données spatio-temporelles en géographie. Il permettra, en outre, d'apporter des positionnements scientifiques sur les *objets,* les approches et les outils utilisés dans ce travail. En effet, si l'approche guidée par les données vise à faire parler les données d'elles-mêmes (Miller et Goodchild, 2014), cela nous amène à questionner les opportunités et les limites thématiques et épistémologiques de l'approche entreprise. Ainsi, ce sont les positionnements thématiques, épistémologiques et scientifiques que nous souhaitons explorer : quelle(s) représentation(s) de la ville envisager et quelles données collecter pour la représenter ? Quel rôle joue la donnée dans les démarches de construction de la connaissance scientifique ? Quelle stratégie adopter pour modéliser puis étudier les *big data* ? Quels sont les outils sur lesquels s'appuyer pour manipuler une information spatio-temporelle dans l'optique d'un positionnement guidé par les données ? L'ensemble de ces questions est traité tour à tour à travers les différentes sections de ce chapitre.

## 1.1. Quelles échelles spatio-temporelles pour étudier les dynamiques urbaines ?

### 1.1.1. Dynamiques urbaines et système urbain

L'étude des dynamiques spatiales dans un sens large renvoie « *à tout changement impliquant la dimension spatiale* » (Lévy et Lussault, 2003, p. 304). Dans une perspective intraurbaine, cette dynamique se rattache à des changements de localisation des individus, des équipements, ou des activités qui vont impliquer une mise en mouvement de la ville au cours d'une journée. Ces mouvements couvrent des processus fonctionnant sur des échelles de temps différentes, allant des années pour les processus les plus lents, à quelques minutes voire secondes pour les processus les plus rapides (Batty, 2007). Sur le plan spatial, il s'agit aussi d'articuler à la fois les niveaux les plus fins (les populations, les ménages) avec des niveaux plus globaux (les réseaux routiers, la ville). Ces questionnements s'avèrent de plus en plus féconds dans les contextes socio-économiques et environnementaux actuels. Le besoin d'une meilleure compréhension des systèmes intraurbains s'inscrit dans les outils non seulement d'aide à la décision, mais aussi d'évaluation des actions d'aménagement du territoire, pour mieux guider les collectivités territoriales dans leurs prises de décisions territoriales (Thériault et Rosiers, 2008). Les thématiques abordées dans les études des dynamiques intraurbaines sont nombreuses, allant de l'étude des mobilités quotidiennes (Banos *et al.*, 2010), aux impacts des infrastructures de transport en milieu urbain (Foltête *et al.*, 2008), jusqu'aux mobilités résidentielles des individus (Hirtzel, 2015).

Si ces travaux relèvent de thématiques urbaines très différentes, elles se rejoignent lorsqu'il s'agit d'introduire le *temps* dans l'analyse spatiale d'un système intraurbain. Ce questionnement de fond soulève deux interrogations (Lévy et Lussault, 2003) : d'une part, comment conceptualiser le changement, et d'autre part, comment appréhender les processus de changement ? Il faut alors s'interroger sur ce qu'il est possible d'observer (de quantifier) dans le temps, puis ensuite, de comprendre les dynamiques locales sous-jacentes à la mise en mouvement de l'espace urbain.

#### a. *Observer l'espace urbain à des niveaux plus fins : vers une approche objet-centrée ?*

L'*objet* se réfère à des *choses* ou à des *réalités* matérielles ou immatérielles (Lévy et Lussault, 2003). En géographie, l'*objet géographique* s'illustre par sa dimension spatiale, mettant en jeu des lieux, consistant à « *rendre perceptible une idée, une image ou un phénomène* » (Brunet *et al.*, 2005, p. 323). Les *objets* sont construits et conceptualisés en vue de répondre à une problématique précise. Généralement, ils correspondent à des entités physiques (l'individu, le bâtiment, la rue…) ou à des concepts plus abstraits se référant à des constructions plus générales comme les ménages ou la ville (Mathian et Sanders, 2014). Les *objets* sont dans notre monde et peuvent être humains (les individus) ou pas (le bâtiment), animés (l'automobile) ou pas (le bâtiment). Dans une optique spatio-temporelle, les *objets* doivent se conceptualiser dans un contexte dynamique, en s'intéressant à l'évolution de ces *objets* dans le temps (Mathian et Sanders, 2014). Essentiellement influencées par des phénomènes plaçant l'individu au centre du phénomène à étudier, les dynamiques intraurbaines peuvent être observées à différentes granularités (Moulin et Ali, 2008). Lorsque l'on s'intéresse aux dynamiques intraurbaines, ce sont donc bien les *objets* mettant en

mouvement la ville qui nous intéressent, par exemple en rendant compte du déplacement d'une personne pour relier son lieu de travail depuis son domicile. Dans ces conditions, on s'intéresse à des *objets mobiles* qui se déplacent dans l'espace et dont les caractéristiques apparaissent éminemment dynamiques. Cette approche, centrée sur les *objets*, est relativement récente et constitue le fruit d'un double développement : d'une part, un essor des outils informatiques et de simulation ; et d'autre part, une focalisation centrée sur les comportements locaux pour expliquer les structures à un niveau plus global (Batty, 2007). C'est à travers les outils de microsimulation des déplacements urbains que ces approches se sont tout d'abord développées. À l'inverse des méthodes classiques, l'approche *individu-centrée* semble offrir un cadre de modélisation intéressant permettant d'appréhender les déplacements à partir des comportements individuels sur trois dimensions : les deux premières représentent l'espace et la troisième le temps (Banos *et al.*, 2010).

Cette conception nécessite pour l'observation des dynamiques urbaines de restituer une photographie de la ville en reconstituant virtuellement, d'une part, un environnement de la ville (le bâti, le réseau routier), et d'autre part, les usagers (la population) utilisant cet environnement (Batty *et al.*, 1998). Si la reconstitution de l'environnement virtuel est relativement aisée à reproduire grâce aux Systèmes d'Information Géographique (SIG) et aux données mises à disposition comme les bases de données de l'IGN ou OpenStreetMap, ce sont bien les données relatives aux usagers et à leurs déplacements qui posent le plus de contraintes pour une approche spatio-temporelle de l'espace urbain (Banos et Thévenin, 2008). Traditionnellement, l'observation et l'étude des dynamiques intraurbaines, contraintes par la rareté des données, sont réalisées avec des données agrégées et rattachées à des zonages spatiaux plus ou moins fins (Thériault et Rosiers, 2008). L'exemple le plus caractéristique s'illustre lors de la mobilisation des Enquêtes Ménages Déplacements (EMD), qui offrent l'opportunité de saisir les temporalités des trajets quotidiens sur des zonages relativement fins (Thévenin *et al.*, 2007).

Or, la mobilisation de ces zonages dans des problématiques intraurbaines ne convient pas toujours à l'observation des dynamiques urbaines. En effet, « *la connaissance des flux à un certain niveau de l'échelle spatiale tend à limiter les analyses des transports en ce qu'elle s'éloigne de l'échelle à laquelle se composent les réseaux de transport* » (Robert, 2001, p. 45). Par ailleurs, si les données ont longtemps été jugées trop pauvres, l'essor des données individuelles issues des nouvelles technologies et des capteurs rend possible l'étude des processus urbains à de nouvelles échelles autant spatiales que temporelles, amenant les chercheurs à repenser les concepts et outils actuels (Pumain, 2014a). Tout comme l'homme avec les nouvelles technologies, la ville s'enrichit de ces nouveaux équipements permettant de relever à des rythmes réguliers et à de multiples endroits les *pulsations* de l'espace urbain à des niveaux relativement fins. Par conséquent, ces nouvelles données ouvrent la voie à de nouvelles observations et analyses, allant du niveau le plus fin (l'individu) à des niveaux plus agrégés (les groupes d'individus). Accompagnés par le développement des outils informatiques et de ces nouvelles données, Marius Thériault et François Des Rosiers (2008) y voient la possibilité de s'affranchir des contraintes d'agrégation pour l'étude des dynamiques urbaines. Cette nouvelle impulsion pour l'étude des dynamiques urbaines à une échelle locale s'opère dans trois orientations (Batty, 2007) :

- Une **prise de conscience des problèmes spatiaux locaux** liés aux déplacements des individus dans un contexte urbain global
- Une **disponibilité de plus en plus accrue des données individuelles** permettant la

mesure et la collecte de phénomènes à des échelles plus fines

- Un **essor des démarches *bottom-up*** en vue d'observer à partir d'actions locales des structures émergentes à des niveaux supérieurs.

### b. *La mise en mouvement de l'espace urbain : une recherche dans la complexité ?*

Ces trois orientations s'appuient sur le cadre conceptuel de la théorie des systèmes et de la complexité (Batty, 2007). Dans cette conception, l'espace urbain se formalise selon un grand système constitué de différents sous-systèmes emboîtés les uns dans les autres (Antoni, 2010). Cette démarche systémique permet de concevoir tout problème complexe comme étant le résultat d'un ensemble de sous-systèmes en interaction (Cambien, 2010). En d'autres termes, les changements de localisation, nos déplacements d'un lieu vers un autre, affectent, transforment et modifient la ville. Ainsi, ils sont le produit d'un ensemble d'interactions et d'interrelations au sein de différents sous-systèmes caractérisant l'espace urbain comme, par exemple, les individus, les logements, les lieux d'activité, l'occupation du sol et les réseaux de transport urbain. Ces différentes interactions sont à l'origine du mouvement intraurbain qui s'établit sur différentes échelles de temps (Wegener, 1994 ; 2004). Michael Wegener (1994 ; 2004) distingue neuf types de sous-systèmes urbains, qu'il classifie en fonction de leur vitesse de changement, allant des plus lents vers les plus rapides, en passant par les changements les plus complexes présentés en tableau 1.1 :

| | Temporalités | Sous-sytèmes |
|---|---|---|
| SYSTÈME URBAIN | Très lentes | Occupation du sol |
| | | Réseaux |
| | Lentes | Lieux de travail |
| | | Logements |
| | Rapides | L'emploi |
| | | Les populations |
| | Immédiates | Les transports de marchandises |
| | | Les déplacements |
| | Complexes | L'environnement urbain |

**Tableau 1.1 : Les temporalités du système urbain (d'après Wegener, 1994, 2004)**

Tout d'abord, les **sous-systèmes les plus lents** représentent les éléments physiques les plus stables et permanents de la ville. Les réseaux, par exemple, représentent les structures physiques les plus pérennes de l'espace urbain. Plus bas, se situent les **sous-systèmes lents :** ils se rattachent aux bâtiments dont la durée de vie peut facilement dépasser le siècle. Par exemple, les lieux de travail ont une durée de vie supérieure à celle des entreprises qui les occupent. Les **processus rapides** se rapportent aux cycles de vie des populations : « *les emplois s'établissent, ferment et se délocalisent […] les ménages se fondent, s'agrandissent, ajustent leurs localisations lors de leurs cycles de vie modifiant la distribution des populations* » (Wegener, 2004, p. 2). Les phénomènes les plus flexibles s'opèrent à des **temporalités très rapides/immédiates** s'ajustant de la minute à l'heure. Les transports de marchandises ou des individus donnent lieu à des interactions multiples entre les

lieux d'activités et les réseaux de transport qui s'ajustent, eux-mêmes, en fonction des conditions de déplacement (fluide ou congestionné). Enfin, il reste une dernière catégorie qui s'opère sur des **temporalités complexes**. Ce dernier sous-système se rattache à l'environnement urbain et aux impacts des huit autres sous-systèmes sur l'environnement urbain, comme la pollution atmosphérique, le bruit ou la consommation d'espace naturel. Plus complexes, ces phénomènes s'accumulent progressivement dans le temps, comme par exemple, les impacts et les processus des émissions de $CO_2$ sur le climat, qui sont plus difficilement observables autant dans le temps que dans l'espace.

La grille de lecture systémique proposée par Michael Wegener (1994 ; 2004) permet d'illustrer la complexité des dynamiques intraurbaines. L'ensemble des sous-systèmes interagit à des échelles temporelles foncièrement différentes, mais également à différentes échelles spatiales. Ainsi, il ne faut pas voir le système urbain comme un ensemble d'éléments isolés mais plutôt comme interdépendants[6]. Toutefois, on comprend vite la difficulté d'accéder à l'ensemble des données pour reconstituer l'ensemble du *complexe* urbain. C'est pour cela que, selon la problématique considérée, le chercheur aura tendance à mettre l'accent sur quelques sous-systèmes, et non sur l'ensemble du système urbain (Cambien, 2010). Le travail présenté dans cette thèse s'intéresse à la dynamique du trafic routier en milieu urbain. C'est pourquoi, comme illustré en tableau 1.1, nous nous inscrirons plutôt dans le sous-système des déplacements et donc dans des temporalités très rapides/immédiates. Traditionnellement, appréhender la ville sous cet angle permet de réduire la formalisation du système urbain en trois sous-systèmes imbriqués (Masson, 2000) illustrés en figure 1.1, sur les caractéristiques desquels nous reviendrons plus spécifiquement plus bas (*cf.* section 5.3.1, p. 165) :

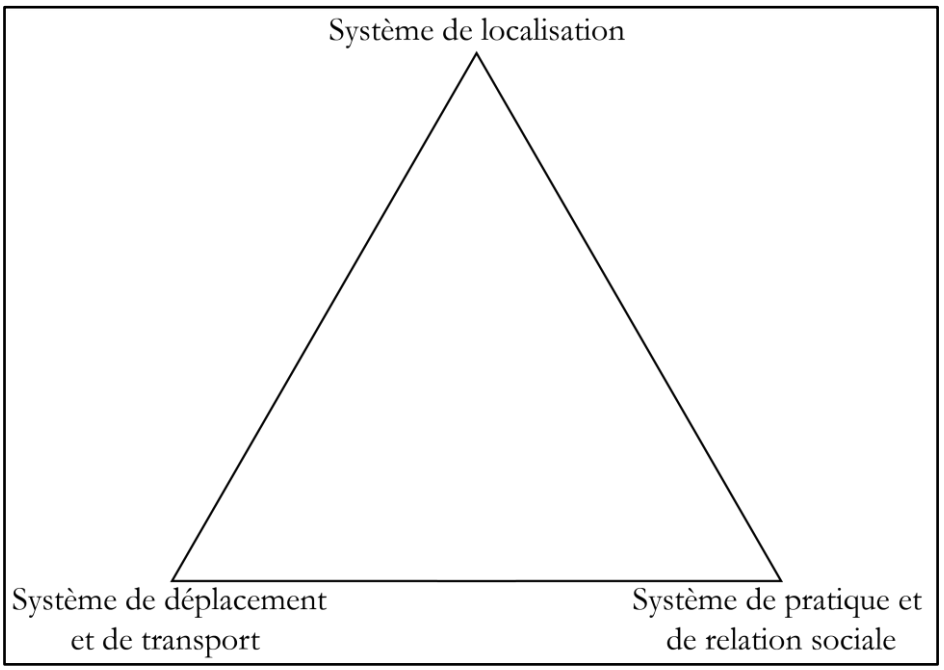



**Figure 1.1 : L'approche traditionnelle en trois sous-systèmes urbains pour la modélisation du trafic routier (d'après Masson, 2000)**

[6] On peut signaler au passage que cette formalisation a permis une prise de conscience globale des modèles urbains au sein d'une boucle de rétroaction entre système de déplacement quotidien, résidentiel et occupation du sol par les applications LUTI (Land-Use Transport Interaction) ; pour plus de détails, nous pouvons renvoyer le lecteur vers l'article de Michael Wegener (2004, p. 3) ou vers l'ouvrage de Jean-Philippe Antoni (2010).

Cette formalisation permet de décomposer l'espace urbain en deux sous-systèmes (les localisations et les déplacements) renvoyant plutôt aux contraintes spatiales de la ville ; tandis que le sous-système de pratiques et de relations sociales s'attache plutôt aux temporalités de l'activité humaine. Ainsi, la démarche mise en œuvre permet de prendre en compte les interactions entre l'espace et les individus, sous la forme d'allers et retours entre un système spatial qui influence les comportements de déplacement et inversement. En appréhendant les dynamiques urbaines locales par la systémique au niveau des *objets*, on touche l'idée que l'ensemble de ces *objets* fait émerger un phénomène urbain. En effet, les dynamiques urbaines se déroulent à des échelles spatiales et temporelles différentes (Wegener, 2004) ; il devrait alors être possible par la somme des comportements individuels de faire émerger de nouvelles structures (Batty, 2007).

Le concept d'émergence apparait « *quand les interactions entre des objets ou d'autres éléments à un niveau influencent d'autres objets ou d'autres éléments à un autre niveau. Plus précisément, un phénomène est émergent s'il faut créer une nouvelle catégorie pour le décrire, catégorie qui n'est pas nécessaire pour caractériser les comportements des objets qui ont mené à son émergence* » (Antoni, 2010, p. 5). Toutefois, ce postulat peut aussi se confronter à une vision paradoxale où « *le tout est plus que la somme des parties* », cette conséquence renvoyant la plupart du temps à la formalisation du modèle mis en place, qui est nécessairement incomplet et/ou partiel (Batty, 2007). Dès lors, pour mieux répondre à cette problématique intrinsèque à la théorie de la complexité, l'une des réponses apportées est de s'inscrire dans une démarche la plus parcimonieuse possible. Ce positionnement conventionnel, qui a tendance à devenir *hégémonique* (Banos, 2013), cherche à limiter autant que possible le nombre de paramètres à inclure dans le modèle. L'étude des dynamiques intraurbaines n'échappe pas à ces questionnements entre parcimonie et/ou exhaustivité des dynamiques. Nous avons vu plus haut que le système urbain peut être appréhendé en un ensemble de sous-systèmes. De plus la multitude des thématiques pour l'étude des dynamiques intraurbaines couvre un ensemble de concepts portant autant sur la mobilité quotidienne, que sur la mobilité résidentielle et l'occupation du sol.

Dans cette acception, face à la complexité des phénomènes urbains, l'approche réductionniste présente l'avantage de décomposer « *le tout en un ensemble de sous-système plus simples (les parties), jusqu'à ce que le cœur même du ou des phénomènes étudiés apparaisse, tel un diamant débarrassé de sa gangue* » (Banos, 2013, p. 39). Ici, l'approche présentée s'inscrit dans ce sens en se focalisant plus particulièrement sur la richesse des données de capteurs, en vue de s'inscrire au sein d'un seul sous-système urbain, qui est celui de *déplacement et de transport*. En effet, comme nous le verrons au cours du Chapitre 2 (p. 55), les capteurs se situent au cœur du phénomène qu'ils mesurent (Servigne *et al.*, 2009) ; ils restreignent nécessairement les champs d'application dans lequel ils s'inscrivent. En mobilisant les informations issues des comptages routiers, ceci nous amène, tout d'abord, à aborder une approche réductionniste du système urbain en se centrant non pas sur un ensemble de systèmes, mais bien sur un sous-système et un seul, tout en nous inscrivant dans une approche *individu-centrée*, en vue d'appréhender les dynamiques inhérentes au sous-système de transport et de déplacement.

### 1.1.2. **Étudier les dynamiques urbaines par la géographie des transports**

Dans le champ de la géographie des transports, on parle habituellement de *déplacement* pour rendre compte de la complexité des mouvements des individus. Dans leur acception généraliste, ces deux termes renvoient au sous-système des déplacements (Bavoux *et al.*, 2005). On peut donc concevoir, intuitivement, qu'entre dynamiques urbaines et déplacements des individus, la frontière est mince, et probablement que la meilleure photographie de ces mouvements s'illustre en partie lors de l'usage des modes de transport autant individuels que collectifs (marche à pied, transports en commun, véhicules particuliers,…).

La géographie des transports est une des branches de la géographie humaine (Cuyala, 2014). Elle vise à étudier les « *rapports complexes entre transport et territoire* », dans le sens où les transports (urbains ou non) sont « *un bon révélateur de l'état de notre planète et de ses dynamiques, tant politiques qu'économiques ou sociales »* (Bavoux *et al.*, 2005, p. 6). À ces trois thématiques, il est important d'ajouter la dimension environnementale du transport, tant les externalités actuelles qui lui sont reliées prennent une part importante dans le débat sociétal actuel (*cf.* section 1.1.3, p. 31). La focalisation de ce travail sur le sous-système de déplacement (urbain) s'inscrit dans des temporalités urbaines très rapides (*cf.* tableau 1.1). Par nature, le fonctionnement urbain implique nécessairement des déplacements d'un lieu vers un autre (Dupuy et Benguigui, 2015). Ainsi, ils sont un des composants essentiels du système urbain, s'opérant par de multiples interactions et agissant autant sur les rythmes de l'espace urbain, l'occupation du sol que sur les localisations des individus.

Le terme de transports (au sens large) regroupe l'ensemble des « *dispositifs, modes et moyens permettant l'acheminement de personnes ou d'objets matériels d'un lieu vers un autre. Par extension, ils regroupent l'ensemble des moyens de la mobilité* » (Lévy et Lussault, 2003, p. 1027). Dans ce cadre, les transports urbains apparaissent comme des systèmes techniques (Dupuy, 1987) dédiés aux déplacements des individus : « *sans route il n'y a pas de déplacement !* ». À cette définition s'ajoute le rôle territorial des transports qui est un bon indicateur de l'intégration d'un territoire, par exemple grâce aux mesures d'accessibilité (Dupuy, 1984) ou de vulnérabilité (Gleyze, 2005). Les transports urbains s'inscrivent alors dans un double schéma entre territoire et réseaux de transport (Dupuy, 1992) : le lien entre urbanisation et réseaux conduit, par exemple, à accroître l'extension des espaces urbains. Les réseaux de transports se positionnent comme « *un système complexe utilisant des infrastructures et des équipements fixes ou mobiles, destinés à offrir, grâce à un mode d'organisation et de gestion adapté, les services que réclame une formation sociale* » (Plassard, 1992, p. 540). Observer les dynamiques urbaines par l'étude des transports urbains s'avère donc être un bon indicateur du rythme de la ville et de ses dynamiques sur trois dimensions : sociale, économique et environnementale. L'étude des transports revêt une forte inscription géographique et ses domaines d'intérêts (pour les géographes) se constituent autour de trois grands domaines (Mérenne, 2008) :

- Les transports contribuent à l'**organisation et à l'aménagement du territoire** : les points d'accès et les infrastructures y étant reliées structurent, hiérarchisent et transforment les milieux géographiques dans lesquels ils s'inscrivent. En milieu urbain, le contrôle et/ou la régulation des flux de déplacements passent, par exemple, par la mise en place des Plans Locaux d'Urbanisme (PLU), des Plans de Déplacement Urbain (PDU) ou des Schémas de Cohérence Territoriale (SCOT) établis par les collectivités territoriales.

- Les transports jouent un **rôle fondamental sur la localisation des lieux** d'activités et/ou de résidence : ils influencent les localisations en facilitant ou freinant le développement des lieux. Par exemple, les zones d'activités situées en périphérie des villes dans lesquelles s'implantent des commerces situés le plus habituellement à proximité des grandes infrastructures de transports (Boquet *et al.*, 2010).
- Les transports constituent **un secteur d'activité à part entière** : le secteur des transports suit ses propres logiques d'organisation, ses contraintes de localisation et ses impacts sur les autres activités (Mérenne, 2008).

Les déplacements des populations dans un sens plus large sont à relier au déplacement des individus, mais comme nous l'avons vu, ce déplacement n'est pas seulement physique, il s'inscrit aussi dans une dimension sociale et économique. Ainsi, la spécificité de l'approche géographique réside dans cette inscription des transports au sein du territoire autant sous sa forme physique (le trafic routier) qu'en tant que mobile en vue de rechercher les interactions avec le territoire en termes d'impacts et aussi de contraintes (Mérenne, 2008). Au niveau des thématiques traitées en géographie des transports, la conclusion d'un ensemble de corpus dédiés à la géographie des transports et intitulé « *Villes et grands équipements de transport* [7]» regroupe les travaux en géographie des transports autour de trois grandes thématiques[8] (Dobruszkes et Lanneaux, 2010) :

- Les questions **relatives à la grande vitesse ferroviaire**
- Les questions d'**organisation à l'échelle régionale**
- Les questions **relatives aux transports urbains** autour non seulement de la place de la voiture en ville, mais aussi autour des questions relatives au développement des transports en commun en milieu urbain.

Ainsi, la géographie des transports est un domaine d'étude multiscalaire allant des plus grandes infrastructures de transport au niveau national (les lignes à grandes vitesses ou les autoroutes), en passant par des enjeux à une échelle régionale, jusqu'aux enjeux locaux des transports. Le travail présenté dans cette thèse se situe dans cette dernière thématique. En effet, en milieu urbain, être capable de mieux connaitre, réguler, et planifier les transports dans une perspective spatio-temporelle revêt des enjeux politiques et sociétaux très larges suscitant un vif intérêt pour non seulement la géographie, mais aussi pour d'autres disciplines comme l'économie ou la sociologie (Mérenne, 2008). Dans leurs conclusions, Philippe Dobruszkes et Marie-Agnès Lanneaux (2010, p. 245) relèvent aussi une rareté des méthodes quantitatives, alors que l'on peut y voir une « *façon de concevoir la géographie des transports* ». Entrer par celle-ci pour l'étude des dynamiques urbaines permet d'utiliser les outils de traitement des réseaux en mobilisant la théorie des graphes permise par l'analyse spatiale et sa modélisation (Appert et Chapelon, 2003 ; Mathis, 2003). L'intérêt de ces méthodes réside, en outre, autant dans leur capacité à intégrer espace et temps pour l'analyse des réseaux de transports, que dans leur vocation opérationnelle (Mathis, 2003). Enfin, l'un des enjeux de l'étude spatio-temporelle du trafic routier se situe aujourd'hui autour de ces impacts environnementaux sur les populations locales (bruit, qualité de l'air, paysages…).

[7] L'ensemble des articles est accessible à l'adresse suivante : https://belgeo.revues.org/6629 : 2010, « *Villes et grands équipements de transport : Compétition, tensions, recompositions* », Belgeo : Revue Belge de Géographie, n°1-2.
[8] À noter que les équipements de transports rattachés au secteur aérien ne sont pas inclus lors de cette synthèse. Pourtant, le « *couple ville-aéroport* » (Bavoux *et al.*, 2005, p. 177) présente des enjeux essentiels, par exemple, pour relier ces grands équipements de transports (souvent positionnés en périphérie) aux autres équipements urbains.

### 1.1.3. Une approche spatio-temporelle pour l'évaluation de la pollution atmosphérique automobile (PAA) en milieu urbain

Historiquement, les pollutions atmosphériques ont plutôt été d'origine industrielle, mais aujourd'hui les émissions des polluants « majeurs » comme les oxydes d'azote ($NO_x$), les composés organiques volatils (COV) ou les particules ($PM_{2.5}$ et $PM_{10}$) sont reliées à l'usage de l'automobile (Fenger, 2009 ; AEE, 2014). L'approche environnementale du trafic routier revêt de multiples enjeux autant pour les collectivités territoriales que pour les individus. Les phénomènes de pollution atmosphérique sont une bonne illustration des principales externalités environnementales du trafic routier. D'après l'Organisation Mondiale de la Santé (OMS), la pollution de l'air est le principal risque environnemental pour la santé dans le monde[9]. En milieu urbain, le trafic routier apparaît comme l'un des principaux contributeurs à une mauvaise qualité de l'air (OMS, 2006 ; Fenger, 2009). La pollution atmosphérique induite par l'automobile (PAA) apparaît comme une préoccupation de plus en plus locale du fait de sa forte variabilité spatio-temporelle liée aux déplacements des individus et aux conditions de trafic (Keuken *et al.*, 2010 ; Uherek *et al.*, 2010). Ce sont les outils d'action locaux actuels qui apparaissent le mieux se saisir de cette problématique. Par exemple, les PDU visent à mieux réguler la circulation en ville en favorisant de nouveaux comportements de déplacement, ou en menant des actions restrictives à l'usage de l'automobile (CEREMA, 2008).

Cependant, d'un jour à l'autre, la pollution peut varier énormément ; cela s'explique en partie par la variabilité des émissions liées à l'utilisation de l'automobile en ville, mais également par les conditions météorologiques (Coria *et al.*, 2015). Deux facteurs apparaissent déterminants (Dab et Roussel, 2001) : les **facteurs naturels** liés aux conditions climatiques et jouant un rôle sur les concentrations de polluants et les **facteurs anthropiques** liés à l'utilisation de l'automobile en jouant un rôle sur l'intensité des émissions de polluants. Ces facteurs vont par conséquent conditionner la contribution directe du transport routier à la pollution locale qui n'est pas figée dans le temps et dans l'espace, mais régie par un fonctionnement de cycle : annuel, hebdomadaire, et journalier (Elichegaray *et al.*, 2008). La pollution atmosphérique se manifeste à plusieurs échelles en fonction de différents critères, liés autant à la nature des polluants, qu'aux conditions météorologiques et à l'intensité des sources d'émissions. Par exemple, les impacts des émissions de $CO_2$ s'observent à l'échelle planétaire, alors que les émissions de $NO_x$ s'appréhendent à l'échelle de la ville voire de la rue. Ainsi, la pollution atmosphérique doit être considérée selon différentes échelles spatiales allant de la ville à l'individu (Roussel, 2006).

Ici nous nous focaliserons sur les phénomènes de pollution à l'échelle locale. À cette échelle, celle-ci agit sur quelques kilomètres et se manifeste à proximité des sources de pollution. Elle se caractérise par l'émission de composés toxiques et cancérigènes impactant plus spécifiquement les populations environnantes (OMS, 2006). Ce sont les polluants primaires, comme les oxydes d'azote ($NO_x$), qui participent le plus à une mauvaise qualité de l'air. Il est néanmoins possible de différencier deux types de pollution faisant la spécificité de l'échelle locale : la *pollution urbaine* rendant compte des niveaux de fond et la *pollution de proximité*. En effet, on peut observer que

[9] Communiqué de presse du 17 Octobre 2013 de l'OMS et de l'agence internationale de recherche sur le cancer : « *Pollution de l'air extérieur : une cause environnementale des décès associés au cancer* » : http://www.iarc.fr/en/media-centre/iarcnews/pdf/pr221_E.pdf.

lorsque la circulation routière est plus intense, à la pollution de fond urbaine se superpose une pollution directement induite par le trafic (Fenger, 2009 ; Oxley *et al.*, 2009). Ainsi, la PAA présente un effet sur la qualité de l'air urbain et apparaît être une source de nuisance locale. Les mesures montrent en effet que les concentrations de polluants primaires directement émis par les moteurs thermiques (comme les oxydes d'azote ou les particules) sont de deux à trois fois plus importantes à proximité des axes routiers qu'en fond urbain (Krzyżanowski *et al.*, 2005 ; Uherek *et al.*, 2010).

Comme illustré en figure 1.2, Isabelle Roussel (2010) distingue à l'échelle locale la pollution de proximité de la pollution de fond urbaine. En effet, la pollution de proximité se manifeste à deux échelles du territoire : l'échelle urbaine (niveau de pollution de fond) ; et l'échelle de proximité directement induite par les émissions directes de polluants primaires (impact direct d'une source). Cette dernière étant plutôt déterminée par le voisinage de sources d'émissions, comme l'automobile, et caractérisée par des *hot-spots*. À ce niveau, les polluants ont une durée de vie courte, allant de la minute à la journée, et sont aussi précurseurs de pollution se manifestant à l'échelle régionale (Sportisse, 2007a ; Elichegaray *et al.*, 2008). Dès lors, la pollution de proximité ne résulte pas forcément d'un transfert sur de longues distances (Elichegaray *et al.*, 2008). Du fait de leurs courtes durées de vie (Sportisse, 2007a), l'impact de ces polluants présente des effets directs et à court terme sur la santé humaine (Krzyżanowski *et al.*, 2005). Dans ces conditions, les effets des émissions directes liées à la circulation routière présentent un réel risque sanitaire en milieu urbain (Makri et Stilianakis, 2008).

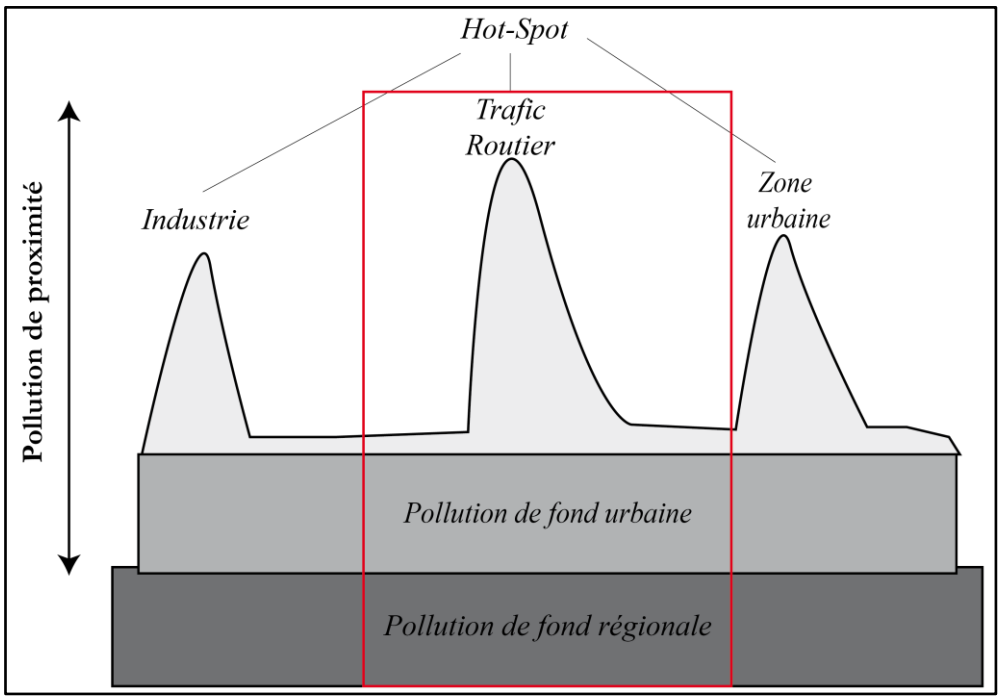



**Figure 1.2 : L'intégration des échelles dans l'étude de la pollution atmosphérique (source : Roussel, 2006, p. 9)**

L'étude de la PAA amène à devoir identifier les *hot-spots* du trafic routier qui vont avoir un impact direct sur les concentrations de polluants à un moment et sur un espace donné (Roussel, 2006). En milieu urbain, les secteurs contribuant le plus significativement à la pollution de proximité sont les installations industrielles, la circulation routière, et le secteur résidentiel (CITEPA, 2015).

En milieu urbain, le trafic routier contribue à hauteur de 65 % des émissions de $NO_x$ et de 85 % des émissions de particules (Primequal-Predit, 2012). La pollution atmosphérique de proximité tient compte de l'intensité des sources d'émissions qui vont accentuer, ou non, les niveaux de pollution de fond (superposition des pollutions illustrée en figure 1.2). Pour le trafic routier, l'intensité des émissions est donc à relier au nombre de véhicules passant en un point donné ; elle s'exprime en quantité de véhicules sur une unité de temps donné : le débit. La pollution de proximité, telle que nous l'appréhenderons au cours de ce travail, se définit comme la combinaison d'une pollution de fond et de l'impact du trafic routier en milieu urbain (*cf.* cadre rouge illustré en figure 1.2). Appréhender la PAA par la pollution de proximité apparaît adapté en vue d'identifier plus facilement les impacts du trafic routier en milieu urbain. Toutefois, ces phénomènes de pollution de proximité demeurent fortement diffus, du fait non seulement de la forte variabilité spatiale et temporelle de la circulation automobile, mais aussi météorologique (Fenger, 1999 ; Keuken *et al.*, 2010). Il se pose de même cette question spatio-temporelle pour connaître et identifier la part des populations exposées sur le court ou le long terme à la PAA. En effet, malgré de nombreux travaux épidémiologiques, la part du trafic routier au sein des concentrations ambiantes de $NO_x$ ou de particules reste encore mal déterminée dans les études d'impacts épidémiologiques (Reungoat *et al.*, 2004 ; Primequal-Predit, 2012).

Historiquement, les épisodes de pollution atmosphérique, comme lors de l'époque industrielle, présentaient des impacts sanitaires beaucoup plus courts et directs entre une concentration importante des polluants et un accroissement de la mortalité. Aujourd'hui, les pollutions s'illustrent plutôt par des concentrations plus faibles, mais les individus y sont exposés de manière chroniques sur une longue période (Blanchard *et al.*, 2008 ; Elichegaray *et al.*, 2008). Pour Jes Fenger (2009), il apparaît relativement facile de détecter et de surveiller l'impact des polluants majeurs. Celui-ci montre, toutefois, que les plus grandes difficultés se font jour lorsque l'on s'intéresse à des composés ayant un temps de résidence court (de la minute à l'heure) et ayant des impacts sanitaires (comme le cancer) à des échelles de temps beaucoup plus longues.

La PAA s'inscrit dans ce contexte. Les effets sanitaires des polluants lui étant reliés, comme les oxydes d'azote ou les particules, vont avoir des effets lors de l'exposition à long terme des individus se situant à proximité des sources de polluants (OMS, 2006). Nous avons vu que la PAA apparaît à une échelle spatiale relativement fine (le bord des infrastructures de transport). Si les problématiques spatiales sont relativement bien prises en compte, la majorité des études portant sur l'exposition des personnes a tendance à ne pas prêter assez attention au rôle du temps et du déplacement des individus (Kwan, 2012, 2013). L'estimation de l'exposition à la PAA dans un contexte où les phénomènes varient autant dans le temps que dans l'espace induit pourtant la mise en œuvre d'un processus spatio-temporel (Gulliver et Briggs, 2005). Le contexte spatio-temporel apparaît intéressant pour l'étude de l'exposition des individus puisque les personnes se déplacent en dehors de leur domicile, arpentent les rues et sont exposées à une PAA qui change constamment au cours de la journée (Kwan, 2013). Néanmoins, avant d'être en capacité d'étudier les impacts de la PAA sur les populations environnantes, il faut être en mesure de quantifier les émissions de polluants issues de la circulation automobile.

L'exploration des relations entre trafic routier et pollution atmosphérique s'observe à différents pas de temps dont les plus caractéristiques s'opèrent à des échelles relativement fines, afin de prendre en compte la variabilité du trafic routier. Jessica Coria *et al.* (2015) montrent que les

concentrations d'oxyde d'azote ($NO_x$) et de particules ($PM_{10}$)[10] varient au cours d'une journée de manière quasiment synchrone par rapport au trafic routier. En analysant les cycles horaires, présentés en figure 1.3, pour plusieurs grandes agglomérations (Stockholm, Londres et Santiago), ils s'intéressent aux relations entre le trafic routier et les concentrations de polluants atmosphériques locaux. Tout en relativisant leurs résultats, qui ne prennent pas en compte les conditions météorologiques locales, ils montrent que le trafic routier impacte la qualité de l'air à un pas de temps relativement fin en distinguant deux périodes : la première concerne les périodes de forts trafics pendant lesquelles les niveaux de polluants sont les plus corrélés au trafic routier ; inversement, la seconde période met en avant les périodes de « creux » où les concentrations de polluants ne sont pas reliées au trafic routier. La figure 1.3 permet d'illustrer la granularité temporelle (horaire) à laquelle il est possible d'observer la contribution du trafic routier sur les émissions de polluants à l'échelle locale.

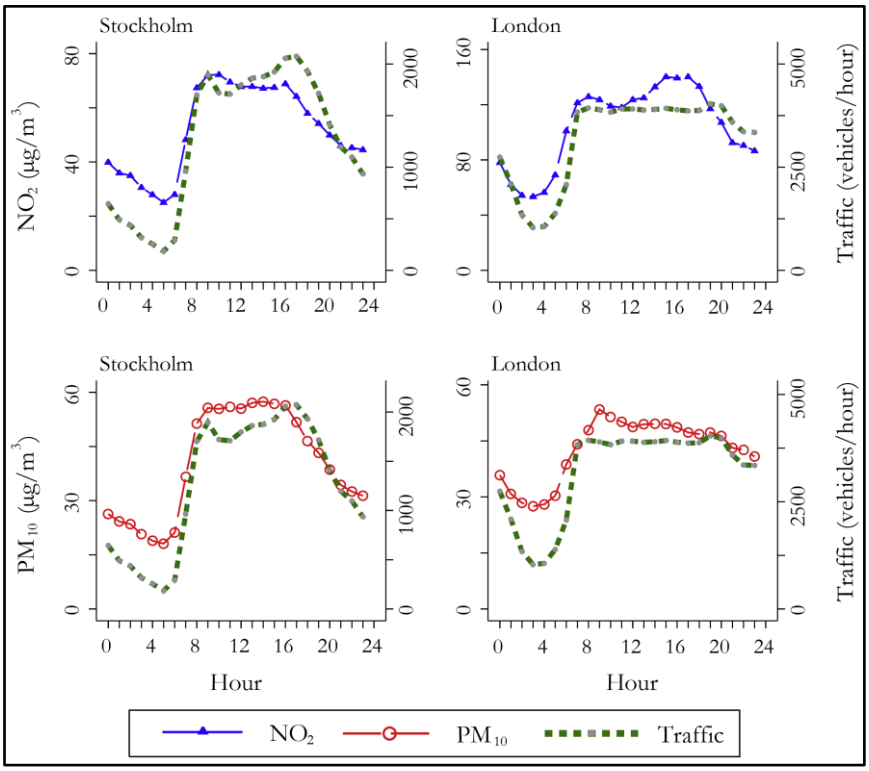



**Figure 1.3 : Les variations horaires de la pollution de l'air pour les villes de Stockholm, Londres en dioxyde d'azotes ($NO_2$) et en particules ($PM_{10}$) (source : Coria *et al.*, 2015, p. 180, figure 1)**

Pour terminer, les contextes géographiques et temporels jouent un rôle primordial pour estimer non seulement les niveaux de PAA mais aussi les niveaux d'exposition des populations environnantes (Gulliver et Briggs, 2005). On se retrouve confronté à une double question lors de l'étude de la PAA : à quel(s) endroit(s) et à quel(s) moment(s) se situent les émissions de polluants les plus importantes ? Et quels sont les niveaux d'exposition des populations environnantes au cours du temps ? Ces problématiques montrent clairement que pour fournir une évaluation plus précise des impacts atmosphériques du trafic routier, il est nécessaire de prendre en compte autant les dynamiques temporelles que spatiales.

[10] Ces deux polluants (NOx et PM10) étant définis par l'OMS (OMS, 2006) comme les plus caractéristiques des impacts sanitaires du trafic routier en milieu urbain (Coria *et al.*, 2015).

Selon l'objectif, le contexte spatio-temporel s'oriente dans deux directions présentées en figure 1.4. La première approche qui constitue l'objet de cette thèse vise à estimer les niveaux d'émissions de polluants atmosphériques rattachés à un secteur donné ; alors que la seconde approche s'attèle plutôt à utiliser les premières estimations pour estimer l'exposition des populations dans un contexte plus épidémiologique. Que ce soit pour l'une ou l'autre approche, dans un contexte urbain, il est possible de positionner les dynamiques urbaines à l'interface des deux approches comme illustré ci-dessous :

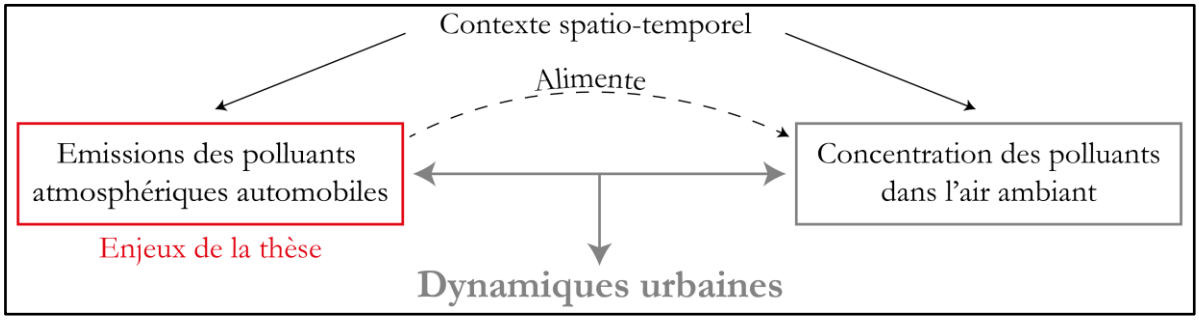



**Figure 1.4 : Les contextes spatio-temporels pour l'étude des pollutions atmosphériques automobiles en milieu urbain**

La figure 1.4 permet d'observer que les deux thématiques sont interconnectées dans le sens où les valeurs des émissions de polluants atmosphériques servent à alimenter en aval les études portant sur les impacts sanitaires de ceux-ci sur la population environnante. À ce niveau, on se retrouve confronté à ce que Laurent Fouillé *et al.*, 2012 (p. 199) identifient comme un phénomène « *GIGO : Garbage In, Garbage Out !* » dans le sens où « *la pollution à la source contamine l'aval* », c'est-à-dire les outils d'évaluation des impacts sanitaires et sociaux de la pollution. Dans ce travail, l'approche que nous avons choisi de privilégier est d'obtenir une bonne représentation de la dynamique du trafic routier pour ensuite en estimer les émissions (*cf.* figure 1.4). Pour cela, nous nous focaliserons sur les concepts de la pollution de proximité en vue de mieux appréhender l'intensité de la circulation routière en milieu urbain. En effet, il apparaît important à notre sens de prendre les choses dans le bon sens, en se focalisant sur une bonne reconstitution du trafic routier pour en estimer les niveaux d'émissions des PAA sans quoi il y a un risque d'étendre les erreurs d'estimation à l'ensemble du processus d'étude de la pollution atmosphérique comme illustré précédemment.

## 1.2. Entre rationalisme et empirisme

Les données et leur structuration prennent une dimension primordiale dans le travail scientifique. Que ce soit pour l'étude des dynamiques urbaines, la géographie des transports ou l'étude de la PAA, les jeux de données à mobiliser revêtent des enjeux primordiaux pour en étudier les processus. Cependant, leur utilisation doit s'inscrire dans un cadre conceptuel permettant, d'une part, d'extraire une information, et d'autre part, de la construire pour en extraire une connaissance en rapport avec cette problématique. Il faut aussi noter que selon l'utilisateur final la donnée peut être une information pour les uns, tandis que pour d'autres elle ne sera qu'une donnée. Les *big data* s'inscrivent dans ce contexte, où dans un sens, on structure la donnée pour en extraire une information définie, alors qu'à l'inverse, on structure la donnée pour en rechercher ensuite une information sans l'avoir préalablement définie. C'est d'ailleurs tout l'enjeu

des *big data* et de leur structuration dans l'optique d'en extraire une information. Cette seconde section permettra aussi d'entrevoir les positionnements scientifiques qui sont en lien avec l'utilisation des données. Elle vise à apporter des éléments de réflexion quant à la manipulation des données et aux démarches suivies par les géographes pour manipuler, construire et développer une connaissance sur les processus spatio-temporels.

#### 1.2.1. **De la donnée à la connaissance d'un phénomène : quel positionnement épistémologique ?**

La constitution des corpus de données se formalise sous la forme d'un Système d'Information (SI) contribuant à faciliter la conceptualisation et l'organisation des données. Pour les gestionnaires, il se définit comme étant « *un ensemble organisé de ressources : matériel, logiciel, personnel, données, procédures… permettant d'acquérir, de traiter, de stocker des informations (sous forme de données, textes, images, sons…) dans et entre des organisations […]. Il est un système d'acteurs sociaux qui mémorise et transforme des représentations via des technologies de l'information et des modes opératoires* » (Reix, 2011, p. 12). Le rôle fondamental du SI est de fournir des informations utiles pour le pilotage d'une problématique. Dans cette conception, il se constitue comme un outil d'aide à la décision fournissant des informations à ses utilisateurs (Reix, 2011). Le SI est le résultat d'un travail de construction et de représentation de la connaissance. Par l'intermédiaire des SI, il s'agit de construire une information à partir des données (et *a fortiori* de l'observation) pour leur donner un sens. Le pilotage d'un SI, géographique ou non, revêt bien cette dimension informationnelle (Joliveau, 2004 ; Reix, 2011) conduisant à fabriquer une image (une représentation) à partir de « choses » observées et de notre connaissance. Ainsi, il ne peut y avoir de communication, d'information et de traitement de données qu'à propos de réalités qui nous paraissent faire sens (Citton, 2012). En d'autres termes, ce processus de construction renvoie à une conception du monde réel et aux choix empiriques qui vont permettre de spécifier les « choses » sous la forme d'une information dont le but est d'en extraire une connaissance. Il permet alors la constitution d'informations qui par la structuration et la conceptualisation des données permettent d'en extraire des connaissances, comme schématisé en figure 1.5.

En France, on a tendance à englober les termes de données et d'information sous la même acception. Toutefois, dans une vision anglo-saxonne ces deux termes ne recouvrent pas la même définition. Pour Donna Peuquet (2002), en science, dériver l'information des données brutes se définit comme étant un processus de séparation du signal et du bruit. Par cette définition, on distingue clairement deux niveaux : le premier concerne la donnée brute issue du système de mesure ; alors que le second est un processus d'interprétation, d'organisation et de filtrage d'un ensemble de données. À ces deux niveaux, Donna Peuquet (2002) rajoute une troisième distinction renvoyant au terme de connaissance définit comme étant un savoir cumulatif de l'information basé sur un « modèle monde ». La structuration des données est donc comme le relève Thomas Thévenin (2011), un processus de construction de la connaissance. L'information est une donnée, mais celle-ci ne devient une information que lorsque que celle-ci est reçue et interprétée. Elle est plus sélective que les données de par sa hiérarchisation et sa contextualisation qui lui donnent un sens (Peuquet, 2002). Dans ce sens, la base de données « *n'est pas directement reliée à des objets réels, mais à l'interprétation de ceux-ci par l'utilisateur final de la base de données* » (Joliveau, 2004, p. 45). Ainsi, la constitution d'un corpus de données renvoie à une conceptualisation de la

réalité issue aussi bien de notre connaissance que de la représentation que nous nous en faisons. Cette structuration peut être envisagée sous une forme procédurale, illustrée en figure 1.5 (Joliveau, 2004) :

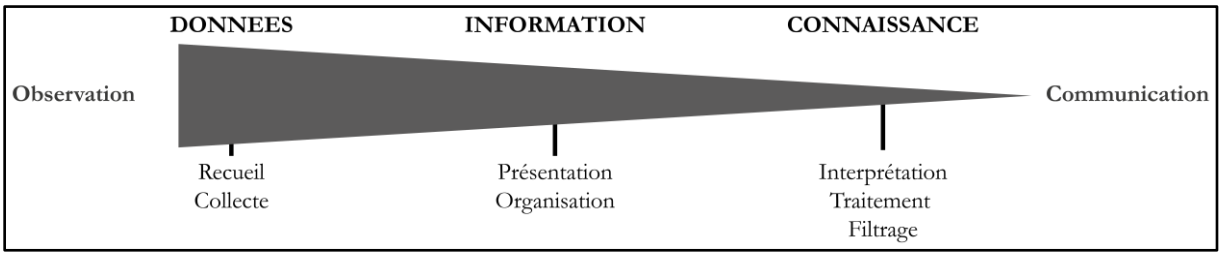



**Figure 1.5 : Le processus de construction de l'information et de la connaissance (d'après Joliveau, 2004)**

En figure 1.5, des exemples de traitement sont rattachés à chacune des étapes. Au départ, les données se caractérisent par l'observation d'un phénomène permettant le recueil et la collecte des données. Ensuite, l'information se caractérise par une étape de présentation et d'organisation des données brutes. Cette étape se traduit par un processus sélectif et d'épuration des données. La construction de l'information se constitue selon deux finalités (Noucher, 2009) :

- i. répondre à un problème donné
- ii. être mobilisable par les utilisateurs concernés.

Pour Thierry Joliveau (2004), étant donné leur mode de recueil et de collecte, la plupart des données sont déconnectées du contexte thématique dans lequel elles s'inscrivent. Elles sont plutôt réinterprétées en vue d'apporter une réponse à un problème donné et deviennent donc des informations sur ce problème. Enfin, la vocation finale de l'information se situe dans sa transmission et sa communication en vue d'en extraire une connaissance sur un problème donné. La connaissance revêt ici la définition que nous en avons faite plus haut, c'est-à-dire un savoir cumulatif entre l'information que nous recevons et nos connaissances autour de ce problème. Ce processus de construction de la connaissance est une démarche cognitive qui « *se traduit par un resserrement continu du cadre de raisonnement qui devient à chaque fois plus personnel et plus subjectif* » (Joliveau, 2004, p. 64).

Cette question de nature épistémologique selon laquelle la connaissance provient à la fois de notre représentation et du raisonnement met l'accent sur deux positionnements épistémologiques en philosophie : le *rationalisme* et l'*empirisme* (Amblard *et al.*, 2006). Il ne s'agit pas ici de décrire ces deux démarches scientifiques qui font l'objet de nombreux travaux[11] mais plutôt d'en expliciter les grandes lignes. Ce paragraphe permet plutôt d'entrevoir le positionnement scientifique qui oriente le travail présenté dans cette thèse. La démarche rationaliste part de la théorie d'un système explicatif qui se réfère à un ensemble de lois (Bailly et Ferras, 2010). Tandis que la démarche *empiriste* recherche la théorie dans les phénomènes rendant plutôt compte de faits observés. Elle se rattache davantage à l'observation et aux données pour en tirer une théorie, des conclusions.

[11] Nous pouvons, par exemple, renvoyer le lecteur vers le premier chapitre de la thèse de Thierno Aw (2010) qui dans la partie 1.1 (p. 36-52) dresse une description détaillée des démarches scientifiques.

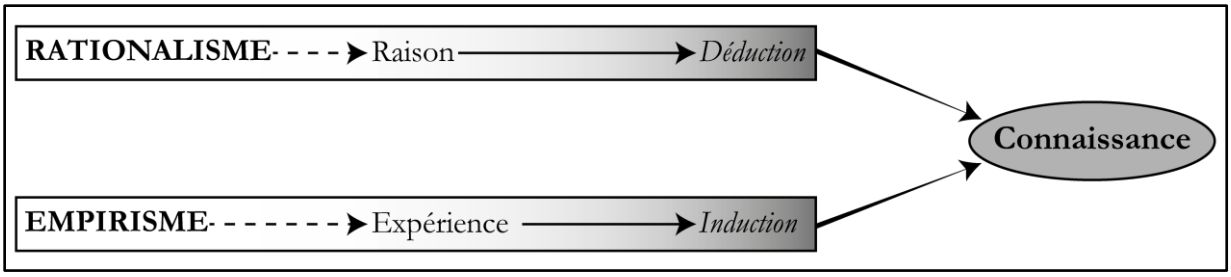



**Figure 1.6 : Schéma et démarche de constitution de la connaissance en épistémologie des sciences (d'après Besnier, 2011)**

Ces deux courants de réflexion scientifique dont nous proposons une synthèse en figure 1.6 assument deux attitudes tranchées (Amblard *et al.*, 2006 ; Aw, 2010 ; Besnier, 2011) :

- le *rationalisme* affiche une démarche introvertie tournée vers les certitudes du chercheur et se rattache à la raison pour la production de connaissances scientifiques (Besnier, 2011). Dans ce courant de pensée, la théorie joue un rôle primordial pour la compréhension d'un phénomène. Pour Descartes (cité dans Amblard *et al.*, 2006), « *la connaissance est d'abord fondée sur la raison* ». Les *rationalistes* font ainsi appel à la raison pour l'acquisition et la production des connaissances scientifiques. La raison recouvre deux aspects qui sont d'une part, la théorie qui est de l'ordre du cognitif et appartient à un domaine, et d'autre part, la pratique qui est de l'ordre du normatif en vue d'en établir des lois (Aw, 2010). À ce courant de pensée, on rattache la démarche déductive qui vise à interpréter notre connaissance à partir d'un point de vue général (Besnier, 2011).
- l'*empirisme* s'appuie plutôt sur l'expérience. Il s'inscrit dans une démarche extravertie selon laquelle toutes nos connaissances dérivent de notre expérience plutôt que de la raison (Besnier, 2011). L'idée défendue par les *empiristes* renvoie à une confrontation à la réalité en s'appuyant sur l'expérience et l'observation. Dans cette conception, l'expérience naît à partir d'observations concrètes permettant ainsi de généraliser certaines théories. On peut, par exemple, rattacher à ce courant de pensée la théorie gravitationnelle de Newton dont les géographes s'inspirent encore aujourd'hui. À l'*empirisme*, on rattache plus usuellement la démarche scientifique d'*induction* qui vise à passer d'un énoncé d'observation à la constitution d'une loi empirique. En d'autres termes, ce sont les expériences et l'observation qui créent nos idées et constituent le socle d'une connaissance scientifique.

Que ce soit suivant la première ou la seconde approche, la même question de fond est soulevée : comment se fonde la connaissance ? Que représente-t-on ? Qui représente et comment représente-on la réalité ? Pour synthétiser ces considérations abstraites, la démarche *rationaliste* se caractérise par la recherche d'une théorie universelle qui permet d'expliquer notre monde. C'est dans un second temps que les données (et l'observation) sont utilisées pour être confrontées à la théorie. Dès lors, théoriser débouche sur une analyse cohérente à partir d'hypothèses préalablement définies et identifiées par *déduction*. À l'inverse, la démarche *empiriste* s'appuie d'abord sur les données, sur leurs spécificités et leur originalité pour tenter d'apporter une explication plus générale. Partir de l'observation induit une démarche plus élémentaire et intuitive (*inductive*) lorsque l'analyse de ces données démarre sans *a priori* ni problématique. On constate assez facilement les limites de cette approche par *induction* qui à partir des « choses » cherche à tendre vers des propositions d'un ordre plus général (Lévy et Lussault, 2003).

Par ailleurs, l'adjectif de l'*empirisme* signifie aussi *pratique* permettant de construire son raisonnement et ses principes en fonction de l'expérience (Brunet *et al.*, 2005). À notre sens, c'est plutôt dans cette dernière considération que la démarche empiriste mérite notre attention. Dans cette conception, elle nous permet d'aborder la mise en place de méthodes expérimentales sans pour autant rejeter et dénigrer les apports de la démarche *rationaliste* qui constitue la pensée dominante dans les démarches actuelles. Si nous avons opposé ces deux positionnements, il faut en revanche souligner que les deux conceptions se complètent : « *Rien n'est dans l'esprit qui ne vienne des sens (empirisme), si ce n'est l'esprit lui-même (rationalisme)* » (Besnier, 2011, p. 15). Ici, c'est la place prise par les méthodes expérimentales et l'observation qui nous intéresse plus particulièrement. En d'autres termes, le rôle de la donnée dans le processus de constitution d'un raisonnement constitue le cadre de réflexion préalable à ce travail. Il s'avère que ce positionnement nécessite un postulat de départ s'inscrivant d'abord dans l'*empirisme*, mais la démarche scientifique se construit par un va et vient entre données et théories, entre expérience et raison et donc *a fortiori* entre *empirisme* et *rationalisme*.

C'est pourquoi, nous privilégierons plutôt une approche dite abductive (Banos, 2013, p. 22) : « *l'abduction renvoie à la capacité du scientifique à se mettre en positionnement d'étonnement, à se laisser guider par la recherche de l'inattendu et plus généralement à laisser libre cours à sa créativité* ». Ainsi, la démarche proposée vise à ne pas se restreindre à la simple rigidité entre les positionnements scientifiques en privilégiant un aller-retour entre les deux positionnements, c'est pourquoi nous en proposons une synthèse sous la forme d'un signe infini comme illustré ci-dessous. Sous ce terme d'abduction, la donnée prend une place non-négligeable dans la construction de la connaissance scientifique et se juxtapose aux deux positionnements traditionnels pour l'expérimentation et l'exploration (Audard *et al.*, 2014). En outre, s'insérer dans cette optique, « *ne peut cependant se concevoir, selon Peirce, que si l'on accepte l'idée d'un système logique complexe, combinant abduction, déduction et induction* » (Banos, 2005, p. 2) :

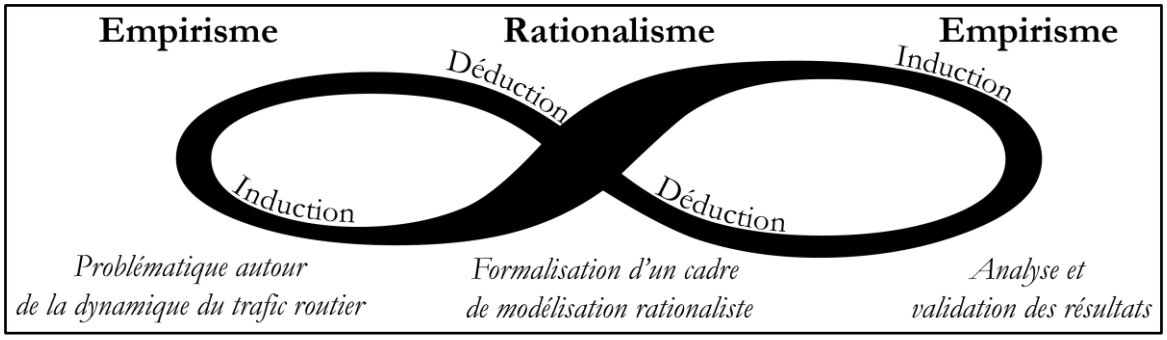



**Figure 1.7 : L'abduction : un double processus de réflexion entre empirisme et rationalisme**

Le choix de cette posture apparait primordial et a des conséquences sur le modèle qui en sera ensuite construit. De plus, il faut souligner que les choix que nous faisons autour du dispositif de quantification à mobiliser restreignent les choix de modélisation et la démarche qui vont guider la suite de ce travail. Comme illustré en figure 1.7, les deux positionnements scientifiques seront employés au sens de la démarche abductive introduite par le philosophe Charles Sanders Peirce[12]

[12] Peirce, C.S., 1993. *À la recherche d'une méthode*, Champ social éditions. ed, Collection Études (Perpignan). Presses universitaires de Perpignan, Perpignan, France, 224p.

et repris par Arnaud Banos (2005, 2013) lorsqu'il décrit les démarches exploratoires des données en géographie. En s'appuyant sur la figure 1.7, nous positionnons tout d'abord notre travail dans le domaine de l'*empirisme* qui constitue le point de départ de la thèse visant à placer les données (*l'observation*) au centre de notre démarche par un processus d'*induction*. De surcroît, la démarche abductive s'inscrit aussi dans ce postulat de départ, tout du moins en partie, « *lorsqu'il s'agit d'explorer ces données, de manière à faire germer des hypothèses* » (Banos, 2005, p. 3). Ensuite, c'est une démarche *rationaliste* qui sera entreprise afin de formaliser et de construire la démarche de modélisation. Par exemple, la mobilisation du cadre conceptuel des systèmes complexes s'inscrit dans le *rationalisme*. Enfin, la dernière étape vise à analyser et interpréter les résultats par *induction* en s'entourant des deux concepts qui par un va et vient entre les deux approches scientifiques, permettra de mettre en œuvre la démarche entreprise dans cette thèse.

### 1.2.2. **Le big data : une question de paradigme scientifique ?**

Les questionnements entre *empirisme/rationalisme* et *déduction/induction* prennent encore plus de sens avec l'essor des *big data*. Pour Rob Kitchin (2014), l'essor et l'analyse des *big data* constituent une rupture dans l'analyse des données modifiant la manière dont l'information scientifique est produite. En reprenant l'ouvrage de Tony Hey *et al.* (2009), il s'interroge sur la constitution d'un nouveau paradigme dans l'analyse des données avec l'essor des *big data*. Le paradigme est défini par Thomas Kuhn (1970, p. 238), qui y apporte une double définition : « *d'une part, il* [le paradigme] *représente tout l'ensemble des croyances, des valeurs reconnues et des techniques qui sont communes aux membres d'un groupe donné. D'autre part, il dénote un élément isolé de cet ensemble : les solutions concrètes d'énigmes, qui sont employées comme modèles ou exemples, peuvent remplacer les règles explicites en tant que bases de solutions pour les énigmes qui subsistent dans la science normale* ». Si la première définition revêt un sens sociologique et un rattachement disciplinaire, c'est plutôt la deuxième définition qui nous intéresse plus spécialement. En effet, le terme d'*énigme* représente les problèmes scientifiques qui donnent à chacun l'occasion de prouver sa capacité à trouver des moyens pour résoudre des problèmes au sein du même champ disciplinaire (Kuhn, 1970). S'appuyant sur une présentation de Jim Gray, Tony Hey *et al.* (2009), repris par Rob Kitchin (2014), tentent de montrer que le développement des techniques et technologies pour l'analyse des *big data* se distingue des autres paradigmes scientifiques. Ils relèvent que les nouvelles technologies et le « déluge » de données modifient non seulement les démarches scientifiques traditionnelles (empiriques, théoriques), mais aussi les plus contemporaines (informatiques).

L'émergence d'un nouveau paradigme implique une discussion autour des outils disponibles qui montrent leurs limites et ne peuvent plus répondre d'une manière efficace aux problématiques posées (Kuhn, 1970). Cette prise de conscience signifie qu'il faut trouver des nouveaux outils pour mieux répondre à ces nouvelles problématiques et à ce que Thomas Kuhn (1970) définit comme une crise[13]. Avec l'essor des *big data,* ce sont bien les outils du chercheur qui présentent leurs limites pour capter et analyser dans leur dimension ces flots de données. Ces modifications amènent les différents auteurs à envisager l'émergence d'un quatrième paradigme scientifique dont l'évolution est présentée en tableau 1.2 (Anderson, 2008 ; Hey *et al.*, 2009 ; Kitchin, 2014 ;

[13] Pour Thomas Kuhn, 1970 (pp. 114–132), les crises sont des conditions préalables et nécessaires à l'apparition de nouvelles méthodes, voire de nouvelles théories impliquant une réaction de la communauté scientifique, soit en définissant de nouveaux outils, soit en rejetant tout simplement un paradigme, car ils se retrouvent en face « *d'anomalies ou de preuves concrètes* ».

Miller et Goodchild, 2014) :

| Paradigme | Nature | Forme | Quand ? |
|---|---|---|---|
| Premier | Science empirique | *Exploration et observation* | Pre-Renaissance |
| Second | Science théorique | *Modélisation et généralisation* | Pre-Ordinateur |
| Troisième | Science informatique | *Simulation des phénomènes complexes* | Pre-Big data |
| Quatrième | Science exploratoire | *Big data ; Data-driven ; Data-mining* | Aujourd'hui |

**Tableau 1.2 : Les quatre paradigmes scientifiques (d'après Kitchin, 2014, p. 3, et Hey *et al.*, 2009)**

Le tableau 1.2 permet de commenter l'évolution des paradigmes au cours du temps. Historiquement, deux paradigmes scientifiques ont dominé la recherche scientifique jusqu'au dernier quart du XX$^{e}$ siècle et la démocratisation des outils informatiques. Les premiers paradigmes furent successivement celui de l'empirisme puis du rationalisme sur lesquels nous nous sommes arrêtés précédemment pour en décrire les principaux fondements (*cf.* section 1.2.1). Ensuite, ce sont les approches informatiques qui ont pris le pas permettant par l'expérimentation informatique de valider ou non des théories. Si l'expérimentation revêt toujours autant d'intérêt pour les géographes avec, par exemple, la modélisation numérique du climat (Pohl, 2015) ou des systèmes urbains (comme le projet MIRO (Banos *et al.*, 2010) ou MOBISIM (Antoni et Vuidel, 2010), l'émergence des *big data* amène certains auteurs à proposer un quatrième paradigme (Hey *et al.*, 2009 ; Kitchin, 2014). Ce nouveau paradigme s'illustre par le positionnement de la donnée dans le processus de construction de la connaissance scientifique. Harvey Miller et Michael Goodchild (2014) caractérisent ces propos en montrant que les données ne sont plus seulement mobilisées pour étalonner, valider et tester des connaissances. Elles jouent un rôle central et les méthodes sont à présent axées et guidées sur la donnée (*data-driven)* et non pas sur la théorie. Ainsi, il ne serait plus nécessaire de partir de la théorie pour constituer la connaissance scientifique mais plutôt de partir des données : « *les méthodes scientifiques sont configurées pour satisfaire les données alors que les données sont configurées pour satisfaire les méthodes* » (Miller et Goodchild, 2014, p. 5).

Allant dans ce sens, on assiste à ce que Chris Anderson (2008) définit avec une certaine provocation « *la fin de la théorie : la renaissance de l'empirisme* (*The end of theory : Empiricism reborn*) », qui est engendrée par les outils d'analyse des *big data*. Néanmoins, comme le souligne Rob Kitchin (2014) plutôt que la *fin des théories*, ce sont plutôt les méthodes axées sur les données qui constituent l'émergence d'un nouveau paradigme scientifique. De plus, plutôt qu'un abandon des théories, les *big data* nécessitent tout d'abord la mise en place de nouvelles théories pour se constituer en tant qu'un paradigme à proprement parler. Sans cela, la connaissance scientifique se limiterait à la constitution de simples corrélations et causalités issues des données (Brunsdon, 2014 ; Brunsdon et Singleton, 2015). Si, pour le moment, les *big data* ne peuvent pas relever d'un nouveau paradigme scientifique, en raison d'un manque de recul et de bases théoriques solides (Kitchin, 2014), ce sont plutôt les approches guidées par les données que nous devons privilégier (Miller et Goodchild, 2014).

En d'autres termes, l'utilisation d'une approche déductive présente une utilité particulière dans la compréhension et l'explication d'un phénomène lorsque le chercheur est confronté à la rareté des données et à la faiblesse de calcul. Mais, inversement dans le cas où les données sont plus abondantes, cette approche peut être plus limitée en amenant plutôt le chercheur à positionner les données au premier plan (Kitchin, 2014). Dans la mesure où une grande partie des données associées au *big data* sont spatialisée (géolocalisées), la géographie se doit aussi de prendre part à ce débat. En effet, elles peuvent présenter un apport non-négligeable pour le traitement de l'information géographique (Pumain, 2014a), de même que pour l'étude des dynamiques intraurbaines ou de la planification urbaine (Batty, 2013), qui ne sont que partiellement ou peu captées par les dispositifs de quantification traditionnellement mobilisés (les enquêtes). Les relations entre *big data* et informations géographiques, voire avec les transports, ouvrent probablement la voie à une lecture originale des phénomènes spatiaux et temporels révélés par ces grands ensembles de données : les *big data.*

### 1.2.3. **Allier temps et espace en géographie : des défis d'exploration à relever**

Depuis plusieurs années, avec l'essor du *big data* et des données individuelles, l'approche spatio-temporelle connait un regain d'intérêt, que ce soit pour son acquisition, sa structuration, son analyse et sa visualisation. Si la donnée reflète un phénomène que nous connaissons, ne pas la modifier, sans inclure d'hypothèses sur la donnée, nécessite tout de même de mettre en place un dispositif de visualisation qui ne peut se faire autrement que par la structuration et la modélisation de la donnée. Ainsi se saisir de ce cadre de réflexion permet, d'une part, d'*identifier les entités* pertinentes à un domaine d'étude, et d'autre part, de se saisir des *opérations* qui sont reliées à ces entités (Mathian et Sanders, 2014). Construire une information spatio-temporelle à partir des données nécessite donc d'en construire les *objets*. Par exemple, il s'agit dans le cas d'une reconstruction de la dynamique urbaine d'identifier les composants élémentaires de mise en mouvement de l'espace urbain. D'une part, ceux-ci (*les objets*) sont spatiaux : les infrastructures de transport, la localisation des résidences, des lieux d'activités… ; d'autre part, ils sont temporels : le programme d'activité, le besoin de déplacement. À partir des données disponibles, la construction empirique des *objets* doit permettre de définir les éléments pertinents, ou pas, de formalisation des dynamiques urbaines. Cette tendance est d'autant plus importante aujourd'hui avec l'émergence des données géolocalisées et des *big data* : Alan MacEachren et Menno-Jan Kraak (2001) estiment que près de 80% de ces données incluent une référence spatiale. Géolocalisées, les données sont aussi le plus régulièrement insérées dans le temps. C'est pourquoi elles présentent un véritable intérêt pour le fonctionnement du territoire, même s'il se pose souvent la question de la représentation conjointe des données autant sur le plan spatial que temporel. Dans ce sens, les phénomènes géographiques s'organisent sous trois dimensions (Miller, 2015) : l'espace, le temps et la thématique. Ces trois dimensions rejoignent le cadre défini par Donna Peuquet (2002) pour la formalisation d'une ontologie spatio-temporelle autour de trois questionnements : quoi ? Où ? Quand ? Ces cadres de réflexion s'inscrivent dans la même acception qui vise à construire les éléments de connaissance scientifique lorsqu'il s'agit de construire une information spatio-temporelle à partir des données disponibles (Peuquet, 2002 ; Thévenin, 2011 ; Mathian et Sanders, 2014)

Les concepts spatiaux sont depuis longtemps abordés par les géographes. Néanmoins, en dehors des travaux autour de la *time geography* (Hägerstrand, 1970) ou de la *géohistoire* (Grataloup, 2015), les

problématiques du temps n'ont été que peu abordées « *frontalement* » par les géographes (Lévy et Lussault, 2003). Pourtant le temps est de plus en plus présent dans notre société et dans son fonctionnement, il rythme autant nos activités que nos déplacements. On peut prendre en exemple cette notion ambivalente que nous entretenons entre le temps et la distance dans notre quotidien : il n'est pas rare que lorsque que l'on nous pose la question de notre distance à un lieu, nous répondions non pas en kilomètres mais plutôt en temps : « Tu habites loin ? » ; « J'habite à ½ heure d'ici ». Pour Jean-Paul Cheylan (2007, p. 3), « *l'exploration des relations entre objets de l'espace d'une part et objets du temps d'autre part laisse percevoir une assez forte analogie* ».

En reprenant deux auteurs[14], Jean-Paul Cheylan (2007) illustre cette analogie par treize formalisations schématiques afin d'illustrer les relations spatiales (déconnexion, intersection, connexion…) et temporelles (avant, rencontre, débute…) entre deux entités. Cet ancrage dans la temporalité apparait, actuellement, d'autant plus important lorsque que l'on s'intéresse aux bases de données spatio-temporelles et aux *big data,* « *une grande partie de ces données est spatialement référencée et estampillée dans le temps. Omniprésents, ces flux de données sont potentiellement révolutionnaires, car ils permettent de saisir la dynamique spatio-temporelle plus directement mais aussi à des échelles spatiales et temporelles multiples* » (Miller, 2015, p. 97). Cette citation de Harvey J. Miller montre toute la complexité d'étudier les phénomènes spatio-temporels pour le géographe. En effet, l'étude conjointe de l'espace et du temps s'opère à différentes échelles de temps et d'espaces qui peuvent se juxtaposer, se chevaucher, se rencontrer, voire se succéder (Cheylan, 2007). Il faut cependant relever que temps et espace ne sont pas symétriques (Lévy et Lussault, 2003, p. 991) : « *le temps, quel qu'il soit, contrairement à l'espace, ne connait pas la commutativité. Il en découle la notion d'irréversibilité du phénomène temporel. Alors que le trajet de A vers B est commutatif, donc réversible, on ne peut parcourir récursivement l'écart qui sépare un temps t+1 d'un temps t* ». Le schéma ci-dessous synthétise cette complexité du temps en distinguant plusieurs temporalités et en décomposant celui-ci par un traitement du *signal/bruit* :

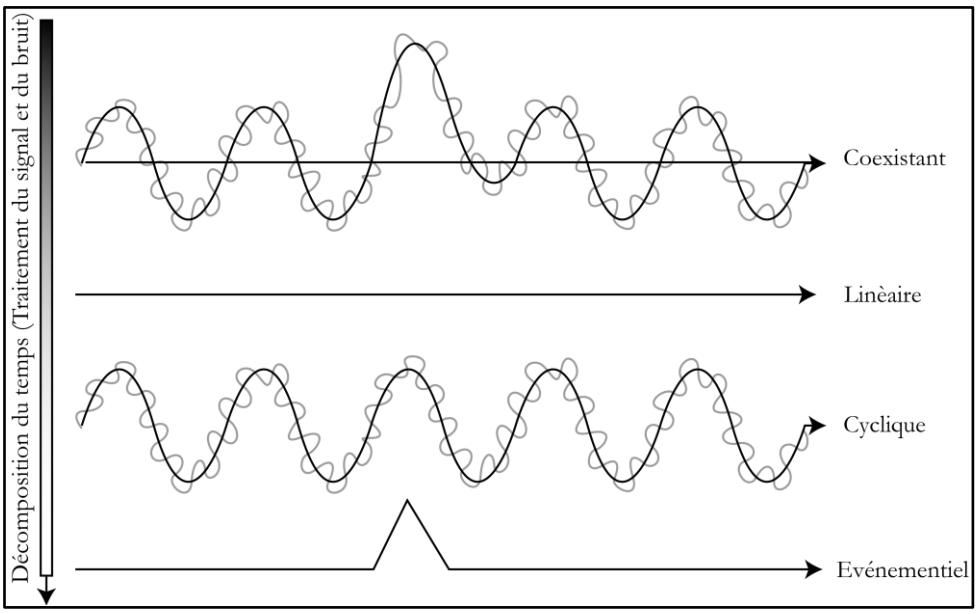



**Figure 1.8 : La conceptualisation des temporalités en géographie (d'après Klein, 2007)**

[14] Max Egenhoffer (1991) pour l'espace et Peter Allen (1983) pour le temps, cités dans : Cheylan, J.-P., 2007. « Les processus spatio-temporels : quelques notions et concepts préalables à leur représentation », *Les dossiers de Mappemonde « Cartographie animée »*, n°87, (3-2007), http://mappemonde.mgm.fr/num15/articles/art07303.html.

Si le temps linéaire revêt cette dimension irréversible et linéaire, il existe une seconde conception de celui-ci reposant sur sa récurrence : le *cycle*. Étudiée sous cet angle, la cyclicité consiste à regarder des processus répétitifs et reproductibles dans le temps ; ce concept est par exemple bien pris en compte en géographie physique lors de l'étude des cycles diurnes ou saisonniers du climat. En géographie humaine, Christian Grataloup (2015, p. 221) introduit ce temps cyclique sous le terme de *temps social* se définissant comme la « *dimension temporelle d'une société, une articulation de la reproduction et du changement* ». Une dernière conception du temps est le temps *événementiel* (Klein, 2007, p. 17) : « *il peut être défini comme une courte période au cours de laquelle apparait une nouveauté. Il dénote un élément temporel localisé qui rompt avec le cycle quotidien. L'événement est généralement localisé dans des espaces limités par rapport à l'ensemble considéré* ». Pour illustrer cette temporalité, Olivier Klein (2007) prend l'exemple d'un match de football débutant à un moment précis (20 h). Si le temps ainsi décomposé montre la diversité des temporalités, il apparait en réalité *coexistant*. En d'autres termes, les temps *linéaire, cyclique et événementiel* entrent en relation complexe et coexistante, par exemple, en se chevauchant, se recouvrant et se rencontrant (Cheylan, 2007). Dans ce dernier, cas, il est d'usage de décomposer le temps (illustré par l'axe en ordonnée en figure 1.8) en vue d'extraire de ce bruit les différents signaux temporels cités précédemment.

Construire les *objets*, dans une approche spatio-temporelle à partir des données, nécessite de prendre en compte cette complexité du temps, tout en étant relié à l'espace. Ces *objets* (*spatio-temporels*) comportent plusieurs dimensions : spatiale, temporelle et attributaire (Peuquet, 2002 ; Klein, 2010). Pour représenter les *objets spatio-temporels* en mouvement, il faut donc être en mesure de les construire mais aussi de les positionner au sein de ces trois dimensions. Les représentations des espaces-temps sont essentielles à une bonne connaissance des structures et processus des dynamiques urbaines (Thériault et Rosiers, 2008). Leurs constructions empiriques, à partir des données, se positionnent au sein d'un triptyque permettant de replacer l'information collectée dans une perspective dynamique permettant de suivre l'évolution de cet *objet*. C'est pourquoi, il faut être capable de prendre en compte le *changement spatial* [15] pour formaliser les dynamiques spatiales et restituer les dynamiques spatio-temporelles. Si les bases conceptuelles apparaissent bien mises en place, il faut tout de même soulever la faiblesse des moyens de représentation spatio-temporelle, du point de vue de leur formulation, structuration et de leur représentation (Cheylan, 2007).

En ce qui concerne les différentes pistes méthodologiques intégrant à la fois espace et temps, Thomas Thévenin (2011) identifie trois approches qui sont la géovisualisation, la géoexploration, et la géosimulation. Ces trois approches reposent sur une interaction forte entre les données spatio-temporelles, les méthodes issues de la statistique et un usage central de l'ordinateur. Dans un sens plus large, les trois approches peuvent être rattachées au concept de la *géocomputation*. L'inscription de ce travail dans le domaine de la *géocomputation* permet d'entrevoir des pistes opérationnelles et originales, permettant de mobiliser des outils issus, d'une part, des outils de l'analyse exploratoire des données (Miller et Han, 2009 ; Miller, 2015), et d'autre part, des outils issus des travaux sur l'intelligence artificielle (Torrens, 2015).

[15] La définition que nous retenons du changement spatial est celle proposée par Bernard Elissalde (2000, p. 235) : « *Le changement spatial n'est pas uniquement synonyme de rupture ou de bifurcation, mais nécessite de s'intéresser aux rythmes et aux échelles de temps variés, c'est-à-dire à la diversité des temporalités dans l'espace géographique. Il intègre l'idée d'un mouvement permanent animé par des processus, eux-mêmes constitués d'enchaînements qui ne sont pas tous linéaires, mais également spiralaires* [cycliques] ».

## 1.3. La *Géocomputation* : un arsenal d'outils pour la constitution et l'étude de l'information spatio-temporelle

Si nous avons discuté des principales thématiques et des enjeux épistémologiques au cours des sous-sections précédentes, il est important d'apporter un éclairage sur les outils permettant une représentation conjointe de l'espace et du temps en géographie. La représentation des données spatio-temporelles est au cœur des nouvelles problématiques autour de l'information géographique (Goodchild, 2013). La représentation conjointe de l'espace et du temps constitue des cas d'étude très concrets pour la géographie, par exemple, pour l'étude des déplacements individuels (Banos *et al.*, 2010) ou de l'évolution des villes (Pumain, 2014b). Pour y aboutir, le principal matériel utilisé par le géographe pour la représentation de l'information géographique se structure autour des SIG.

Toutefois, force est de constater que l'intégration de la dimension temporelle au sein de cet outil s'avère souvent limitée (O'Sullivan, 2005). Michael Goodchild, 2000 (p. 14) soulève ce point : « *le SIG reste une technologie largement bidimensionnelle dominée par la métaphore de la carte papier* ». Sous cet angle, le cadre *stricto-sensu* du SIG vu comme outil n'est pas suffisant pour embrasser l'ensemble des dimensions spatio-temporelles, mais aussi des thématiques portant sur les phénomènes géographiques. Ainsi, l'information géographique doit être replacée dans un contexte plus général de réflexion autour des méthodes, des outils et des thématiques relevant de la manipulation des données spatio-temporelles. C'est dans ce contexte que différents termes ont vu le jour, en vue de constituer une « *véritable science de l'information géographique* » (Joliveau, 2004, p. 48) : selon les auteurs, certains parleront de *GISciences* (Goodchild, 1992 ; Goodchild et Longley, 2014), de *Géocomputation* (Openshaw et Abrahart, 2000 ; Brunsdon et Singleton, 2015) ou de *Géomatique* (Joliveau, 2004) pour le traitement de l'information géographique. L'ensemble de ces termes que nous pouvons respectivement rattacher à une vision nord-américaine, britannique et française reprend le même postulat de départ visant à dépasser le simple usage de l'information géographique à travers les SIG : « *ce mouvement est épistémologique et porte sur des questions de formalisation de l'approche géographique, de méthodologie d'analyse et de modélisation spatiale »* (Joliveau, 2004, p. 455). L'insertion de la réflexion géographique relativement à ces trois termes renvoit à enrichir le SIG par un arsenal de concepts, de méthodes et d'outils (juxtaposés ou intégrés) en vue de répondre aux nouveaux enjeux de traitement et d'étude de l'espace géographique, que nous explorons plus spécifiquement au cours de cette section.

### 1.3.1. Les approches guidées par les données en géographie : un travail s'inscrivant dans l'interdisciplinarité ?

Avec l'émergence des grands ensembles de données estampillées autant dans le temps que dans l'espace, le contexte de la recherche géographique se situe lui aussi dans les questionnements autour d'un quatrième paradigme (Kitchin, 2014 ; Miller et Goodchild, 2014). Il existe un véritable potentiel pour le géographe à mobiliser les *big data* en vue, d'une part, d'en extraire une information géographique, et d'autre part, de construire des modèles d'analyse spatiale (Miller et Goodchild, 2014). En géographie, l'approche guidée par les données relève de l'*analyse spatiale* pour Luc Anselin (1989, p. 1), « *elle vise à laisser les données parler d'elles-mêmes et tenter d'obtenir des informations sur la configuration, les structures et les interactions spatiales sans les contraindre d'une notion théorique préconçue* ». Plus ancien, ce positionnement, guidé par la donnée, cherche plutôt à

maintenir les principes fondamentaux des démarches scientifiques plutôt qu'à se positionner en tant que nouveau paradigme scientifique (Kitchin, 2014). Pour Luc Anselin (1989), les différentes techniques guidées par les données possèdent deux aspects en commun. Tout d'abord, on observe la structure des données (*observed pattern*) recueillies. Ensuite, les structures, modèles, ou formes spatiales à analyser sont exclusivement dérivées des données. Dans ce sens, les méthodes axées sur les données ne s'inscrivent pas seules dans ce quatrième paradigme scientifique et ouvrent une seconde voie à l'exploitation des *big data*. Comme, par exemple, en étant plus ouvert à l'aide de méthodologies « hybrides : *abductives* », combinant autant les approches déductives qu'inductives pour la compréhension de phénomène spatiaux (Kitchin, 2014). C'est pourquoi, pour nous, les démarches guidées par les données ne remettent pas en cause les outils traditionnels et rejoignent plutôt le troisième paradigme autour des sciences informatiques illustré en tableau 1.2 (p. 41).

Le développement de l'*analyse spatiale* s'est fait dans un contexte de fortes turbulences de la science géographique avec l'émergence de la *nouvelle géographie*[16]. Ce mouvement se traduit alors par un tournant *théorique* et *quantitatif* qui va profondément marquer la géographie (alimenter les débats), et l'analyse spatiale qui en découle, avec, d'une part, l'essor des analyses mathématiques et statistiques, et d'autre part, le développement des outils informatiques (Cuyala, 2014). En prenant appui sur ces deux champs disciplinaires, il est possible d'identifier deux apports de la statistique et de l'informatique pour l'*analyse spatiale* (Mathian et Sanders, 2014). Le premier concerne un « raffinement » des aspects méthodologiques et de traitement des données grâce à la statistique. Et le second, relié à l'informatique, concerne les aspects liés au traitement de l'information géographique et à la modélisation des données. Ces développements ont, en outre, permis de développer différentes méthodes reliées aux données spatiales. Les premières approches, les plus caractéristiques, se sont développées avec l'analyse exploratoire des données initiée par John Tukey (1977). Le développement des outils numériques et des différentes approches qui découlent de l'exploration des données permettent d'ouvrir de nouvelles méthodologies s'émancipant des différentes hypothèses sur lesquelles les méthodes traditionnelles s'appuyaient auparavant (Tukey, 1977). Cette tendance s'est considérablement accrue avec l'essor des données de plus en plus massives, sous la forme du *data mining* et du *knowledge discovery* (Miller et Han, 2009, p. 38): « *Les applications de data-mining et knowledge discovery ont été activement développées, y compris pour l'interprétation des cartes et des outils d'extraction de l'information, mais aussi pour l'extraction d'images de télédétection, l'identification de modèles spatio-temporels (patterns), l'analyse des flux et des mouvements au sein de l'espace géographique* ». Ces approches tournées vers les données (*inductives*) permettent par l'enchaînement de méthodes d'analyse variées de détecter plus facilement des *patterns* en vue d'en extraire une information géographique. L'analyse exploratoire des données et les nouveaux outils qui y sont associés (analyse exploratoire des données, *data-mining*…) permettent ainsi d'entrer bien plus rapidement dans une démarche empirique et abductive en s'émancipant de la théorie (Banos, 2013).

Le développement d'un arsenal d'outils et de méthodes pour l'analyse exploratoire des données spatiales nous amène à situer notre travail à l'intersection de trois disciplines, comme illustré en figure 1.9 : *la Géographie, la Statistique et l'Informatique*. Ce positionnement interdisciplinaire met

[16] À l'essor de la nouvelle géographie à partir des années 1970, Sylvain Cuyala (2014) y rattache sous la même acception les expressions de géographie quantitative, géographie théorique et quantitative, et l'analyse spatiale.

l'accent sur les outils de la *géocomputation* dont la fonction principale se structure le plus habituellement autour des démarches inductives (Fotheringham, 1998). De plus, la *géocomputation* se réfère aux « *méthodes d'analyse spatiale dans lesquelles l'ordinateur joue un rôle pivot* » (Fotheringham, 1998, p. 283). Chris Miller (2015), dans son article, l'identifie aussi comme étant un domaine de recherche interdisciplinaire qui cherche à résoudre les problèmes du monde réel en insistant sur leur inscription spatiale et temporelle :

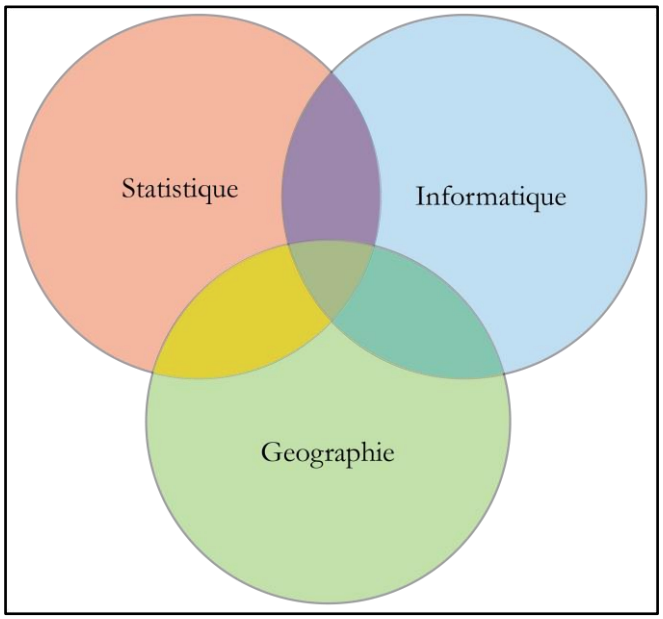



**Figure 1.9 : La *Géocomputation* : un positionnement interdisciplinaire pour l'analyse et la modélisation des données spatiales (source : Alex Singleton, Big data in the real world : *http ://www.alex-singleton.com/talk/2014/12/02/Big-Data-in-the-real-world/*)**

### 1.3.2. Qu'est-ce que la *Géocomputation* ?

Le terme de *Géocomputation* est relativement récent ; il a été introduit en 1996 par Stan Openshaw et Robert J. Abrahart (2000, p. 3)[17] : « *La géocomputation peut être considérée comme l'application du paradigme des sciences informatiques (cf.* section 1.2.2, p. 40*) pour étudier un large éventail de problèmes dans le contexte des systèmes spatiaux ».* Pour les auteurs, la *géocomputation* constitue une seconde révolution qui fait suite au développement des SIG. Plus qu'une discipline, le terme renvoie non seulement à un ensemble de méthodes « *computationnelles* » visant à couvrir l'ensemble des applications liées à la manipulation de données spatiales, mais aussi temporelles (Brunsdon et Singleton, 2015). Dès lors, le terme renvoi à un large éventail d'outils provenant de différentes disciplines et dont l'ancrage commun se situe au niveau de l'information spatiale. Appréhendée dans ce sens la *géocomputation* se distingue des concepts de la *GIScience* introduit par Michael Goodchild (1992, p. 32) comme étant *« la science derrière les systèmes qui sont concernés par l'ensemble des questions fondamentales soulevées par les SIG et les technologies qui leur sont liées »*[18]. La distinction des deux concepts, *GIScience* et *géocomputation*, illustre deux positionnements face aux données spatiales. Les premiers placent le SIG au centre du développement de leur démarche (Goodchild, 1992). Alors que les seconds placent l'ordinateur au centre de la démarche et placent l'outil SIG sur le même plan que d'autres

[17] La première édition de l'ouvrage de Stan Openshaw et Robert J. Abrahart date de 1996.
[18] Michael Goodchild (1992, p. 32) : « *the science behind the systems, concerned with the set of fundamental questions raised by GIS et allied technologies* ».

outils informatiques comme les gestionnaires de bases de données (SGBD) ou les logiciels statistiques. Mark Gahegan (1999, p. 204) illustre bien ce propos vis-à-vis de l'outil SIG : « *la géocomputation représente une tentative consciente de retour en arrière dans l'analyse géographique se faisant avec ou sans les SIG* ». Dès lors, la *géocomputation* se construit en complémentarité et en extension des SIG (Openshaw, 2000 ; Cheng *et al.*, 2012). Pour Stan Openshaw (2000, p. 8) la relation entre les deux disciplines est importante, mais « *la géocomputation a aussi d'autres relations qui peuvent être tout aussi importante, par exemple, avec l'informatique ou des méthodes numériques. Peut-être aussi que dans une perspective géographique, la géocomputation est ce que l'on fait après les SIG dans le sens où on ne cherche pas obligatoirement à utiliser toutes la richesse des données créées par le SIG ».* À cet effet, la *géocomputation* est bien plus qu'un SIG ; d'ailleurs, le terme de *computation* fait explicitement appel aux apports issus des outils de l'intelligence artificielle (AI) et du calcul à haute performance (HPC). Pour Stan Openshaw (2000, p. 5), le recours au calcul numérique permet de s'inscrire dans une démarche empiriste « *le calcul numérique (computation) implique un paradigme très particulier basé sur l'approximation numérique plutôt que sur la précision analytique,* […] *il est basé sur la substitution de vastes calculs comme un substitut au manque de connaissance ou de théorie. Guidé par les données ou non, le calcul numérique est utilisé comme un outil purement théorique pour comprendre comment les systèmes complexes fonctionnent* ». Dans ce sens, le rôle de l'ordinateur joue ici un rôle central et fondamental dans le processus de construction de la connaissance géographique (Fotheringham, 1998). Toutefois, l'ordinateur ne peut être appréhendé seul, et il implique l'usage de trois technologies de pointe dans le domaine de la géographie quantitative (Openshaw et Abrahart, 2000) :

- 1. Les SIG pour produire et créer les données ;
- 2. L'AI pour fournir des outils intelligents ;
- 3. Le HPC pour fournir de la puissance.

Une bonne illustration de l'interfaçage et de l'imbrication de ces éléments se situe au niveau de la « *machine d'analyse géographique (GAM) »*[19] de Stan Openshaw *et al.* (1987, p. 7). La construction de cet outil de géovisualisation était original et novateur dans les années 80 (Banos, 2013). Il se fonde sur une évolution du positionnement du chercheur face aux données, devenu plus massives, mais également face à l'évolution des outils informatiques qui constitue un des points de départ de la *géocomputation* (Gahegan, 1999). C'est aussi un positionnement inductif, voire abductif (Banos, 2005), qui est privilégié : Stan Openshaw (2000, p. 10) justifie cela pour deux raisons : « *(1) la connaissance des données empêche de tester des hypothèses a priori correctes ; (2) les hypothèses qui pourraient être légitimement testées reflètent des connaissances et des théories qui pourraient bien être fausses* ». Ainsi, l'approche initiée par Openshaw a ouvert la voie à la mise en place d'outils tournés vers les données où la *computation* permet de s'affranchir des hypothèses pour visualiser, explorer, et simuler les processus spatiaux, qui n'auraient pu être appréhendés autrement qu'avec l'aide d'un support informatique.

À l'ère du *big data*, la *géocomputation* apparaît d'autant plus comme un domaine de recherche intéressant et original dans lequel la Géographie peut (et doit) s'inscrire. Nous pouvons même aller plus loin en identifiant les prémices d'un nouveau paradigme scientifique qui se fonderait à travers la *géocomputation* à l'intersection, d'une part, des apports issus de la science informatique, et d'autre part, des apports issus des outils de la science exploratoire (*cf.* section 1.2.2, p. 40). Dans

[19] Traduction de « *Geographical Analysis Machine* : GAM ».

cette acception, les outils de la *géocomputation* seraient peut-être l'artefact d'un nouveau paradigme scientifique se construisant à l'interface des deux paradigmes scientifiques précédemment cités, ou comme nous l'avons vu précédemment ce sont plutôt les méthodes axées sur les données, et s'inscrivant dans un contexte pluridisciplinaire, qui constituent l'émergence d'un nouveau paradigme scientifique plutôt que la remise en cause même des outils actuels (Kitchin, 2014). C'est pourquoi, dans un troisième temps, nous porterons un regard plus particulier sur les outils et les concepts rattachés à la *géocomputation*.

### 1.3.3. **Les outils pour la *géocomputation* : des méthodes axées sur les données**

Le site des conférences sur la *géocomputation* (http://www.géocomputation.org/) dresse un bilan des technologies reliées à la *géocomputation*. Actuellement, les développements les plus significatifs autour de la *géocomputation* s'opèrent autour des concepts de la complexité et des données spatio-temporelles (Brunsdon et Singleton, 2015). En outre l'analyse des données spatio-temporelles constitue une bonne illustration des apports technologiques et méthodologiques s'appuyant sur la *géocomputation*. La figure 1.10 illustre ces différents outils à travers deux degrés d'intersection caractérisant, d'une part, les disciplines rattachées à la *géocomputation*, et d'autre part, ses outils. Le premier niveau, plus général, s'établit à l'interface de chacune des grandes disciplines scientifiques caractérisant la *Géocomputation* (Géographie, Informatique et Statistique) et permet d'identifier trois sous-disciplines scientifiques à chaque intersection de la figure 1.10 :

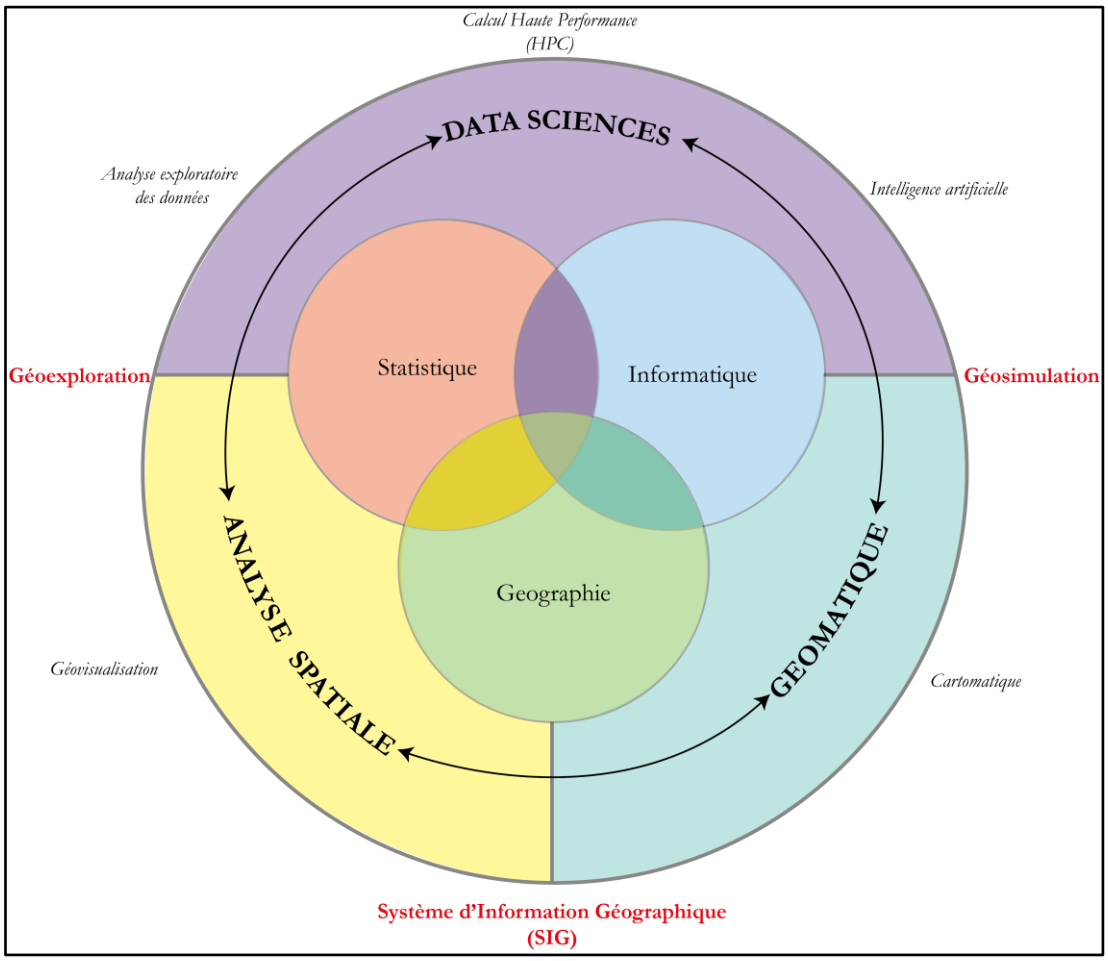



**Figure 1.10 : La *Géocomputation* vue par les outils pour l'analyse des données spatio-temporelles : des approches interdisciplinaires**

- À l'interface de la statistique et de l'informatique, la Data Science : d'un ordre général, la science des données se fonde sur l'extraction des connaissances à partir des données (data driven) et se base sur un processus de découverte de la connaissance à partir des bases de données (KDD : *Knowledge Discovery in databases*). Pour Donna Peuquet (2002), c'est un processus de construction de la connaissance allant de l'observation vers l'analyse, puis vers la théorie. C'est donc bien un processus inductif qui est mis en avant au sein duquel a été développée « *une collection de techniques dans le domaine de la gestion des données, des statistiques, de la visualisation et de l'intelligence artificielle* » (Peuquet, 2002, p. 217).
- À l'interface de la géographie et de l'informatique, la Géomatique: ici, nous prenons délibérément une définition plutôt restrictive de la géomatique comme étant « *l'ensemble des utilisations techniques de l'informatique en géographie* » (Lévy et Lussault, 2003, p. 440) et désignant « *l'activité dont le SIG est le produit* » (Joliveau, 2004, p. 71). Sous ce terme, on entend regrouper l'ensemble des outils liés à l'intersection de la Géographie et de l'Informatique, c'est particulièrement le cas des SIG et des outils de cartographie automatique (cartomatique).
- À l'interface de la statistique et de la géographie, l'Analyse spatiale: elle se définit comme « *l'analyse formalisée de la configuration, et des propriétés de l'espace géographique, tel qu'il est produit et vécu par les sociétés humaines* » (Pumain et Saint-Julien, 2010a, p. 155). Elle se fonde sur les apports des outils statistiques appliqués à des objets localisés dans l'espace : *« l'emploi de la statistique classique conduit à oublier dans un premier temps la localisation des objets […] au contraire, les méthodes géostatistiques et de la morphologie mathématique, regroupées dans la statistique spatiale, analysent directement les informations géocodées* »[20].

Il est important de noter que les liens entre ces trois disciplines ne sont pas hermétiques. Par exemple, la géomatique s'entoure des apports de l'analyse spatiale pour représenter l'information contenue dans une base de données. C'est particulièrement le cas des outils de cartographie qui visent à réduire et résumer l'information géographique par des processus de discrétisation. Ces propos nous ont amené à identifier un second niveau d'intersection entre chacune de ces sous-disciplines (Data Sciences, Analyse spatiale, et Géomatique), où différents outils peuvent être positionnés. En s'appuyant sur la figure 1.10, nous nous focaliserons sur trois outils caractéristiques des intersections (illustrés en rouge) : les SIG, la Géoexploration et la Géosimulation[21].

Le positionnement des SIG à l'intersection de l'analyse spatiale et de la géomatique montre ce que Jacques Lévy et Michel Lussault (2003, p. 83) relèvent, à savoir que « *l'analyse spatiale a été grandement facilitée par le développement des SIG* ». Toutefois, comme nous l'avons précédemment vu, les SIG ne constituent pas les seuls outils de la *géocomputation* (Openshaw, 2000). Il est alors nécessaire de dépasser le cadre de cet outil en le replaçant au même niveau que d'autres outils qui se fondent eux aussi sur une intersection des sous-disciplines, mais aussi des outils. Par exemple, le couplage des SIG avec les outils d'exploration ou de simulation ont permis de développer,

[20] Denise Pumain : définition de « l'**analyse spatiale** » : http://www.hypergeo.eu/spip.php?article38.

[21] Notons ici que le positionnement des outils vis-à-vis de cette grille de lecture illustrée en figure 1.10 n'est pas figé et constitue un essai de positionnement scientifique des outils mobilisés au sein de la *Géocomputation*. En fonction des affinités de chacun, les différents outils peuvent être positionnés différemment, par exemple, la géovisualisation peut être positionnée au même plan que la géoexploration lorsque celle-ci est abordée dans un caractère non-paramétrique (Banos, 2013). Ici, nous avons englobé sous la même acception les deux types de géovisualisation (paramétrique et non paramétriques) permettant de plutôt les positionner au centre de l'intersection entre Statistique et Géographie.

d'une part, des outils dits de géoexploration, et d'autre part, de géosimulation.

Pour Arnaud Banos (2013) la visualisation doit s'allier à l'exploration : « *la visualisation – physiquement passive mais cérébralement active – d'animations spatio-temporelles ne saurait constituer une fin en soi. Il s'agit également de permettre à l'utilisateur d'explorer, de tester et de vérifier toutes les idées qui ont pu émerger au cours de ce processus* » (Banos, 2013, p. 26). La démarche exploratoire se place non seulement au début du travail de visualisation géographique, mais aussi à la fin (MacEachren *et al.*, 2004). Elle a été initiée par les travaux de John Tukey (1977) ; celui-ci préconisait de considérer l'analyse des données à l'aide de différentes méthodes visuelles. Dans un article antérieur (Tukey, 1962, p. 64), il relevait déjà le rôle « *important et vital* » de l'ordinateur pour l'analyse exploratoire des données : « *Il y a des cas où l'ordinateur rend possible ce qui aurait été totalement infaisable* ». Reprenant les méthodes initiées par Tukey pour l'analyse et l'exploration des données spatiales, Andrienko et Andrienko (1999) mettent en avant trois impératifs pour la manipulation de données spatiales : l'utilisateur doit être maître de ses choix dans le mode de représentation ; par la juxtaposition de différents supports graphiques, l'utilisateur doit être en mesure de comparer ces représentations ; la navigation doit être facilitée entre les différents documents graphiques par un simple clic de souris. Ces trois impératifs mettent l'accent sur l'interactivité entre homme et machine.

Pour répondre à ce besoin d'interactivité et d'exploration, l'outil informatique prend une place considérable à travers le développement d'interfaces visuelles pour faciliter les relations entre la machine (l'ordinateur) et l'homme, mais également entre la carte et ses représentations graphiques. L'individu devient donc actif pour explorer les jeux de données spatiaux qu'il a à disposition (Banos, 2005). En d'autres termes, l'analyse exploratoire des données spatio-temporelles s'inscrit au sein de trois objectifs (Klein, 2007) : **identifier des formes spatiales** à un moment donné ; **comparer des formes spatiales** à différents moments pour identifier et quantifier les variations et les évolutions dans le temps ; **identifier des formes spatio-temporelles** pendant une période de temps. Toutefois, si la carte animée et son exploration constituent les principaux supports visuels mobilisés en géographie pour la représentation des dynamiques spatio-temporelles, le développement des plates-formes de géosimulation, via la simulation multi-agents ou les automates cellulaires, offre aujourd'hui de nouvelles opportunités pour l'étude et la reproduction des processus spatio-temporels. Il permet de se positionner à une seconde intersection illustrée dans la figure 1.10. La géosimulation se définit comme une démarche de simulation du monde réel désignant « *un ensemble de concepts, méthodes et techniques, qui combinent les apports de l'intelligence artificielle et des sciences de l'information géographique, pour simuler la dynamique des systèmes spatiaux* »[22].

À travers le cadre de réflexion présenté en figure 1.10, on comprend aisément que la *géocomputation* est un positionnement pluridisciplinaire qui se fonde sur un grand nombre d'applications variées tournées, d'une part, vers les données, et d'autre part, sur la façon dont le monde fonctionne (Goodchild et Longley, 2014). Ce sont aussi les apports des outils informatiques qui constituent un développement non-négligeable des outils reliés à la *géocomputation* (Openshaw, 2000). Cette diversité des approches appliquées dans un cadre géographique sera reprise tout au long de cette thèse qui à notre sens se rattache à la *géocomputation*. En résumé, différents critères nous ont amenés à nous positionner en faveur de la *géocomputation* dans le cadre de ce travail :

[22] Lena Sander : définition « **Géosimulation** » : http://www.hypergeo.eu/spip.php?article311.

- Elle présente des caractéristiques qui lui sont propres en reposant sur la mise en place d'un nouveau paradigme pour la manipulation de l'information géographique
- Elle constitue une des pistes d'investigation pour l'étude des dynamiques urbaines, c'est particulièrement le cas des méthodes de géosimulation
- Ce sont des méthodes et des outils axés sur les données et présentant une réponse à l'analyse des *big data* dans le contexte urbain. Cette dimension se construit dans la continuité des apports issus de l'analyse exploratoire des données et des KDD.
- Les outils qui lui sont liés s'insèrent dans un cadre géographique original visant à dépasser les limites intrinsèques aux SIG en se reposant sur les apports combinés des outils issus autant de l'AI que de la statistique.

## Conclusion

En évoquant les enjeux thématiques, nous avons pu voir que les enjeux liés à l'étude des dynamiques urbaines, de la géographie des transports et de la pollution atmosphérique se rejoignent lorsque l'on aborde les problèmes liés à la quantification de ces phénomènes. Dans les trois cas, les questionnements autour de la quantification du phénomène à observer revêt un enjeu important : quels sont les processus à observer ? À quelles échelles spatiales et temporelles observer ces phénomènes ? Par ailleurs, ce premier chapitre permet de délimiter les champs d'application de ce travail autour de la notion de pollution de proximité automobile. La constitution d'une approche spatio-temporelle du trafic routier nécessite non seulement de s'intéresser au corpus de données à disposition, mais aussi aux outils permettant d'en explorer les processus. Les positionnements épistémologiques abordés ont permis de mettre en lumière les processus de construction pour mieux appréhender un problème vis-à-vis des données. En plaçant la donnée au premier plan de la démarche, c'est d'abord une démarche empirique que nous souhaitons entreprendre en vue d'explorer non seulement les forces mais aussi les faiblesses des données de capteurs. Néanmoins, le parti pris du scientifique vis-à-vis des corpus de données se distingue entre deux grandes approches : le rationalisme et l'empirisme. Dans le cas de la démarche empirique, le rôle de la donnée n'est pas neutre et il joue un rôle central dans la constitution de la démarche entreprise et oriente le travail à mettre en place. En revanche, la réalité de la trajectoire du scientifique montre que les enjeux actuels se situent autour d'un va-et-vient entre *empirisme* et *rationalisme* (Besnier, 2011), c'est pourquoi l'adoption d'un positionnement *abductif* permet de ne pas se restreindre à la simple rigidité des deux positionnements traditionnels (Banos, 2005). D'autant plus que l'inscription d'une démarche *abductive* s'insère elle aussi dans un processus de raffinement et de construction de la connaissance (*cf.* figure 1.5, p. 37).

Enfin, nous avons pu aborder sous un angle plutôt opérationnel les différents outils permettant d'explorer mais aussi de représenter les dynamiques spatiales. Le cadre conceptuel et méthodologique de la *géocomputation*, nous a permis d'identifier différentes approches pour l'étude spatio-temporelle des dynamiques en géographie. Ce domaine de recherche à l'interface de trois disciplines (Géographie, Informatique et Statistique : *cf.* figure 1.9, p. 47) permet, en outre, de mettre l'accent sur la représentation des structures spatiales dans le temps et sur les méthodologies axées sur les données. On a pu voir que ce domaine de recherche interdisciplinaire apparaît être une réponse adéquate à notre travail sur plusieurs points, cités plus haut, nous amenant ainsi à nous positionner en faveur de la *géocomputation*.

Le parti pris de cette thèse se situe au niveau du dispositif de quantification mobilisé dont nous allons présenter les caractéristiques pour l'étude du trafic routier au sein du chapitre suivant. Dans une optique *individu-centrée*, la démarche de constitution du dispositif de quantification implique de s'accorder sur les *objets* conceptuels (le trafic routier) et leur mesure (les capteurs fixes) en vue d'en constituer un corpus de données (Commenges, 2013). Les caractéristiques d'un réseau de capteurs apparaissent foncièrement spatiales du fait de la mesure d'un phénomène en un point donné. À cette dimension spatiale, les réseaux de capteurs présentent aussi une composante temporelle qui est elle-aussi inhérente au dispositif de comptage, du fait d'une mesure en temps réel ou à fréquence régulière (Bouju *et al.*, 2009). Cette dernière composante montre la spécificité du dispositif de mesure utilisé au cours de ce travail. Elle est d'autant plus un prérequis à l'approche présentée dans cette thèse qu'elle doit prendre une part importante dans la

structuration d'une approche de modélisation des données spatio-temporelles. La constitution d'un réseau de capteurs forme un système de comptage qui est centré sur la donnée (Akyildiz *et al.*, 2002). Cette donnée doit, toutefois, être décortiquée pour en extraire une information ; en effet, « *la plupart des données accessibles en géographie sont produites par inadvertance, à d'autres fins que l'emploi qu'en fait la géographie* »[23]. Il est donc souvent nécessaire de structurer la donnée dans ce cadre formel pour en extraire une information géographique (Peuquet, 2002 ; Joliveau, 2004).

[23] Roger Brunet *et al.* (2005, p. 168) : définition de « **donnée** ».

# Chapitre 2. L'ÉTUDE DU TRAFIC ROUTIER EN MILIEU URBAIN, ENTRE DISPOSITIF DE MESURES ET DISPOSITIF DE CONSTRUCTION

D'après le Commissariat Général au Développement Durable (CGDD), la circulation routière concerne près de 65 % des déplacements locaux des individus (CGDD, 2010). Malgré un ralentissement de son usage dans les grands pôles urbains régionaux, la voiture demeure le mode de déplacement dominant pour les déplacements locaux[24]. Véritable vecteur de la liberté de déplacement des individus (Dupuy, 1995), l'usage de l'automobile en milieu urbain présente des caractéristiques propres à la mobilité des personnes sur les infrastructures routières. L'étude du trafic routier relève du domaine de la socio-économie des transports (Bonnel, 2002 ; Lannoy, 2003 ; Commenges, 2013). Cet ancrage n'est pas neutre car celui-ci a conditionné la mesure du trafic routier que nous faisons aujourd'hui par « *l'invention de la notion de déplacement* » (Gallez et Kaufmann, 2009, p. 4) et des concepts de mesure de la *mobilité quotidienne* (Commenges, 2013). Ainsi, en entrant par la théorie des systèmes complexes, la géographie des transports et la *géocomputation* nous obtenons une vision géographique du trafic routier. La mesure d'un phénomène vise à fournir « *des* ***valeurs*** *aux phénomènes auxquels on s'intéresse dans le cadre d'un questionnement géographique. Elle sert à caractériser les* ***attributs des objets*** *que l'on étudie et intervient en amont de tout traitement* […] *Élaborer un système de mesure suppose au préalable d'****identifier des objets et des attributs pertinents*** *par rapport aux questions posées* […] *Suivant les informations disponibles et ou mesurables, il peut y avoir un* ***décalage plus ou moins important entre le phénomène que l'on souhaite mesurer et ce qui est effectivement observable*** »[25]. Cette première définition générale permet de mettre en lumière les objectifs de ce second chapitre. Celui-ci permettra de dégager les enjeux de l'instrument de mesure mobilisé pour la modélisation du trafic routier. Dans le cadre de cette thèse, nous proposons de nous intéresser à une mesure directe du trafic routier. Nous nous attacherons notamment à définir les variables du trafic routier fournies par les dispositifs de mesure de comptage. Ce travail s'intéresse à un certain type de système de mesure se définissant par le jeu de critères suivant :

- Observer et étudier la circulation routière par un dispositif automatique
- Apporter une observation en temps réel ou quasi-réel du trafic routier
- Prendre en compte l'ensemble du trafic routier en milieu urbain (trafic de transit et intraurbain).

Le premier critère permet de définir l'objet d'étude de ce travail. Il délimite d'ores et déjà les éléments essentiels à l'étude du trafic routier, desquels nous verrons les composantes en première partie. Le second critère exclut les dispositifs dits qualitatifs dont le coût de mise en place ne permet pas un suivi des déplacements en temps réel ; il permet de restreindre notre choix à des dispositifs quantitatifs issus des systèmes de capteurs. Enfin le dernier critère permet de délimiter l'espace d'étude de ce travail dans le sens où les données nécessaires à la démarche envisagée doivent permettre de rendre compte du trafic routier en milieu urbain dans sa globalité. De plus, ces critères induisent une distinction entre les mesures individuelles et ceux qui sont collectives.

[24] Les déplacements locaux représentent 98 % des déplacements métropolitains. Ils couvrent l'ensemble des déplacements réalisés dans un rayon de 80 km autour du domicile des individus et concerne essentiellement les déplacements réalisés quotidiennement (CGDD, 2010).
[25] Lena Sanders : définition de « **mesure** » : http://www.hypergeo.eu/spip.php?article61.

## 2.1. Capter le trafic routier

### 2.1.1. **Trafic versus Mobilité**

Le trafic routier relève de l'utilisation de véhicules motorisés sur une infrastructure de transport. Il implique le déplacement des individus ou d'un objet, le véhicule, dans l'espace. Avant de définir la notion de trafic routier, il est important de cerner la notion de *mobilité*. Comme précisé en introduction, l'étude de la circulation routière s'inscrit dans le domaine de la socio-économie du transport (Lannoy, 2003 ; Commenges, 2013). Malgré une évolution parallèle, la définition de la mobilité que ce domaine en fait se distingue de celle des sociologues issus de l'école de Chicago (Lannoy, 2003 ; Lévy et Lussault, 2003 ; Gallez et Kaufmann, 2009), les premiers plaçant l'automobile au centre de leurs recherches. La mobilité est approchée comme étant un flux de transport, renvoyant au franchissement des individus dans « *l'espace des circulations* » (Lévy et Lussault, 2003). Pour Hadrien Commenges (2013), la définition que fait la socio-économie des transports est synonyme de la demande de transport sous la forme de flux de déplacement. Alors que les sociologues considèrent plutôt l'automobile comme un objet secondaire (Lannoy, 2003). Ils étudient la mobilité comme un changement de position, de rôle et de statut. Dans ce sens, la mobilité n'est pas seulement physique, elle est aussi sociale, économique et technique (Orfeuil, 2000, 2004 ; Gallez et Kaufmann, 2009). Ainsi, les déplacements urbains (et *a fortiori* les déplacements motorisés) doivent s'inscrire dans un système plus large où les interactions entre la localisation des hommes et des activités génèrent des différenciations spatiales. Le monde urbain n'est pas seulement vu comme un vecteur de la mobilité, puisque celle-ci forme plutôt un *couple infernal* (Wiel, 2005) où « *la ville conditionne les formes de la mobilité comme les conditions de mobilité influent sur la forme de la ville* » (Wiel, 2005, pp. 12–13).

Au final, la mobilité est un terme polysémique dont il est difficile de donner une seule définition (Lévy et Lussault, 2003 ; Commenges, 2013). Les définitions que nous avons abordées sont complémentaires. La définition apportée par les sociologues s'intéresse principalement aux déterminants de la mobilité en s'interrogeant sur le sens du déplacement, ses motifs, ses projets… À l'inverse, la socio-économie des transports s'intéresse plutôt aux masses de déplacement, à les quantifier physiquement sans réellement y apporter de description sémantique, à part lors de l'analyse des programmes d'activités (Commenges, 2013). Pour ce travail, nous nous arrêterons sur la définition développée en socio-économie des transports. Cette définition est certes restrictive mais permet de positionner l'*objet* de notre étude, le trafic routier, au centre de la démarche. Par ailleurs, le *trafic* est synonyme de *circulation* (Brunet *et al.*, 2005), c'est pourquoi nous utiliserons les notions de *trafic routier* et de *circulation routière* de façon indifférenciée dans cette thèse[26]. Ces deux termes se fondent sur le principe fondamental du fonctionnement d'un système permettant d'en apprécier sa dynamique (Lévy et Lussault, 2003). Pour les auteurs, l'étude de la circulation est une des composantes de la géographie des mobilités et se caractérise spatialement par les trois éléments illustrés en figure 1.11.

Tout d'abord, la circulation s'organise entre des **espaces connectés (ou reliés)**, comme par exemple entre le lieu de résidence et de travail, générant ainsi des flux entre les deux zones. Le flux peut alors être défini comme « *le déplacement de personnes entre des lieux différents sans préjuger du*

[26] Ici, nous pouvons souligner que les Anglo-saxons englobent les concepts de « *traffic* » et de « *circulation* » sous la même acception.

*mode d'acheminement* » (Bavoux *et al.*, 2005, p. 216). La **nature des éléments** qui circulent se rapporte aux modes et aux moyens d'acheminement de ces flux dont le **mouvement** est régi par une vitesse et une capacité à se déplacer (Lévy et Lussault, 2003). Ces deux derniers éléments permettent d'introduire le concept de trafic routier qui s'illustre lorsque « *ses flux sont acheminés par un mode de transport particulier* [l'automobile] *et parcourant les réseaux* » (Bavoux *et al.*, 2005, p. 216). À cet égard, le trafic routier (ou la circulation routière) se définit lors de cette phase de transition. Il peut être considéré comme le flux de déplacement de véhicules sur une infrastructure routière. Il est la traduction d'un mouvement physique entre un lieu d'origine et de destination et revêt une des formes de la mobilité géographique.

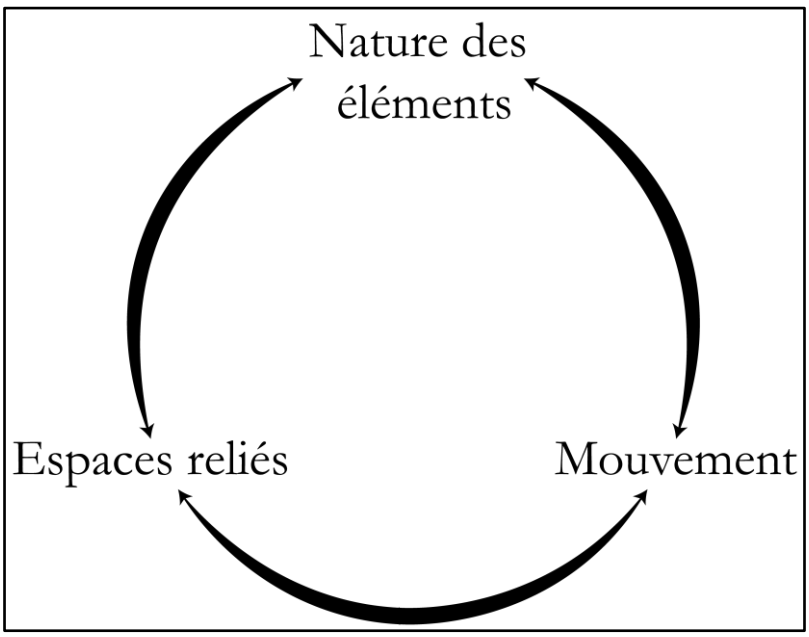



**Figure 1.11 : Les caractéristiques spatiales de la circulation (d'après Lévy et Lussault, 2003)**

Dans cette optique, deux concepts socio-économiques doivent être introduits : le flux comme étant un *besoin* ou une *demande* de déplacement entre deux lieux, et l'infrastructure comme l'*offre* de transport à disposition des usagers pour se mouvoir. À ce niveau, l'objet d'observation n'est plus exclusivement le véhicule mais bien les personnes (Bonnel, 2002). On ne se rattache plus à l'objet vecteur du déplacement, comme le véhicule, mais plutôt aux besoins des usagers pour rejoindre un lieu en fonction de leurs capacités et de la capacité des infrastructures mises à disposition. Le déplacement n'est donc plus approché comme un objet en soi, il est le résultat d'un arbitrage entre la localisation des individus et leurs activités. En d'autres termes, il est le résultat d'un système offre-demande. La socio-économie des transports s'en est inspirée pour estimer les flux de transport. Pour la socio-économie des transports, le fonctionnement du système urbain est le résultat de l'imbrication de trois sous-systèmes (*cf.* figure 1.1, p. 27) :

- Un **système de localisation** correspondant à la localisation des zones d'activités, de résidence ou d'équipements de loisirs, par exemple.
- Un **système de pratiques et de relations sociales** rythmant l'organisation de l'espace urbain par l'activité professionnelle ou de loisirs, par exemple.
- Un **système de déplacements** qui s'articule autour des deux autres sous-systèmes et qui se traduit par un ensemble de flux.

Le trafic routier appartient à ce dernier sous-système et est alors lié à un « *changement de localisation s'exprimant dans le temps* » (Bavoux *et al.*, 2005). Il fait référence au déplacement d'un véhicule d'une origine vers une destination. Ce mouvement est lié à l'organisation de l'espace urbain qui se structure, d'une part, autour des activités des individus (Massot et Orfeuil, 2008), et d'autre part, autour d'infrastructures supportant ces échanges, formant ce que Marcel Roncayolo (1990) ou Sabine Barles (2008) appellent le *métabolisme urbain*. En outre, le réseau routier joue un rôle essentiel dans la structuration du trafic routier. Sans route, il n'y aurait pas de circulation. La vocation première des réseaux routiers est son caractère technique permettant la mise en relation physique des lieux (Dupuy, 1984). Les caractéristiques structurelles et fonctionnelles des réseaux routiers jouent un rôle non-négligeable dans la structuration des conditions de circulation automobile (Gleyze, 2005). Par exemple, les phénomènes de congestion automobile sont le résultat de deux facteurs : le nombre d'individus empruntant le même tronçon et la capacité de l'infrastructure à assimiler ses flux de déplacement automobile. L'infrastructure joue ici un rôle déterminant sur les déplacements routiers et constitue elle-aussi un élément essentiel du trafic routier (Dupuy, 1984) et d'organisation structurelle des déplacements des personnes et des marchandises en milieu urbain (Gleyze, 2005). Le trafic routier se constitue donc de deux composantes qui sont à l'origine du mouvement : **le véhicule ;** et le support permettant ce mouvement : **l'infrastructure**. Il peut alors se définir comme un objet spatio-temporel où seuls les véhicules mobiles sur l'infrastructure sont pris en compte.

### 2.1.2. **Enquêtes versus Comptages : quelle(s) mesure(s), pour quel(s) usage(s) ?**

La mesure du trafic routier vise à connaître le nombre de véhicules présents sur une route et à un moment donné. L'identification de l'objet de cette recherche, le trafic routier, nécessite de s'intéresser aux dispositifs de quantification permettant cette mesure. La revue de littérature mobilisée fait ressortir deux grandes approches pour mesurer du trafic routier : une **approche indirecte** prenant en compte les déterminants de la mobilité et de l'utilisation de véhicules motorisés par la localisation de zones émettrices (les origines) et réceptrices (les destinations) ; et une **approche plus directe** par la mesure du trafic routier ou son suivi au cours d'une période de recueil. Cette distinction permet d'illustrer l'éloignement entre le phénomène que nous observons et celui que nous mesurons. En effet, la mesure vise à donner une connaissance exacte d'un phénomène. Par ailleurs, comme le relève la définition de la *mesure*, « *la nature ne se prête pas immédiatement à être mesurée, mais elle* [la mesure] *se fait au moyen d'une ruse »* (Lévy et Lussault, 2003, p. 655). Il est alors primordial d'identifier la relation entre l'instrument de mesure utilisé et ce qui est effectivement observable. L'objectif de cette section n'est pas de décrire chacun de ces dispositifs[27], mais plutôt d'illustrer cet éloignement entre la mesure et ce que nous souhaitons observer comme illustré en figure 1.12.

Aux deux grandes approches, indirectes et directes, on rattache usuellement deux grands dispositifs de quantification du trafic routier, qui sont respectivement **l'enquête et le comptage**. Comme présenté en bas de la figure 1.12, la principale distinction entre ces deux dispositifs repose sur le degré d'interaction entre le dispositif de comptage et l'individu compté (Commenges, 2013). Le premier, qualitatif et plus descriptif, cherche à replacer les déplacements

[27] Pour plus de détails, nous renvoyons le lecteur à la section 2.3 (p. 64) en ce qui concerne les dispositifs de comptages par capteur et vers les travaux d'Hadrien Commenges (2013) pour les données d'enquête.

des individus dans l'espace urbain en étudiant les déterminants de ces déplacements entre une origine et une destination par la constitution de matrices origine-destination (O-D). Le second, quantitatif et moins descriptif, permet quant à lui de mesurer plus directement le trafic routier par la mise en œuvre de dispositifs de comptage permis par l'implantation de capteurs placés dans l'environnement à mesurer :

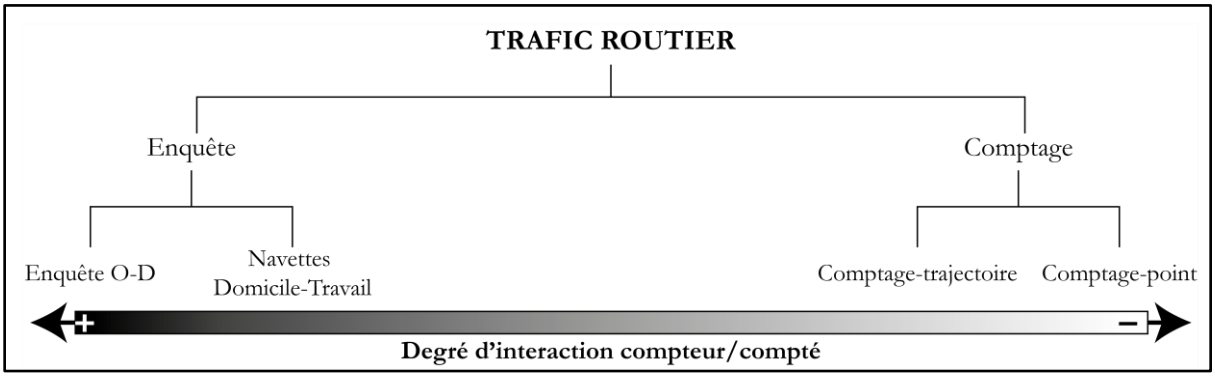



**Figure 1.12 : Typologie des dispositifs de quantification du trafic routier (Source : Commenges, 2013, p. 45)**

Les enquêtes présentent un degré d'interaction entre le compteur et le compté plus important que les comptages. C'est-à-dire que le lien entre le dispositif de comptage et l'objet compté permet une description précise des déplacements des individus, par exemple, en fonction du motif de déplacement et de la catégorie socio-professionnelle des individus. Au sein des enquêtes, Hadrien Commenges (2013) distingue des navettes domicile-travail les enquêtes O-D. Sous le terme de O-D, il regroupe « *toute enquête qui produit des couples origine-destination, couples de lieux orientés et temporellement estampillés (time-stamped)* » (p. 45). Le critère de distinction entre enquête O-D et navette domicile-travail est donc principalement temporel. Au sein des enquêtes, pour la mesure du trafic, la principale donnée disponible est le déplacement repéré par son origine et sa destination, permettant ainsi de construire la matrice O-D utilisée dans les modèles de trafic. Ceci explique en partie l'usage dominant des enquêtes O-D pour les études de trafic routier actuel autant sur le long que le court terme. Si l'enquête O-D présente de nombreux avantages pour la planification du transport routier en milieu urbain (Bonnel, 2002), elle montre à notre sens une de ses limites pour des applications sur court terme, par exemple, en ne fournissant qu'une mesure discrète du trafic routier entre les heures creuses et les heures de pointe. Tout d'abord, elles ne permettent pas d'appréhender les trafics de transit dans leur globalité. De plus, sur le court terme, la planification des transports n'est pas nécessaire ; il convient plutôt de rendre compte du trafic existant que de devoir l'anticiper. De par ses caractéristiques, il est courant de faire un aller-retour entre les estimations de la matrice O-D et les comptages pour réaliser des vérifications et des ajustements sur les enquêtes O-D (Dupuy, 1975 ; Bonnel, 2002). L'usage des données de comptage ne permet que l'évaluation des trafics routiers existants, alors que les enquêtes montrent leur utilité pour planifier des trafics futurs. Par conséquent, les données produites par les comptages ne sont pas antagonistes mais plutôt complémentaires des enquêtes.

C'est pourquoi, il est d'usage de distinguer les comptages des enquêtes. En effet, le comptage est au cœur de la mesure, mais il ne permet pas de faire un lien direct entre les caractéristiques de l'individu et la donnée recueillie. En revanche, les comptages présentent l'atout d'être plus directement lié à l'objet de la mesure. Le comptage est une mesure ancienne et, comme le relève

Gabriel Dupuy (1975, p. 14) : « *le comptage des individus, du bétail ou des marchandises en déplacement a été largement utilisé dans l'histoire comme mode de contrôle politique ou administratif* […]. *Depuis qu'il existe des véhicules et des routes, on compte donc les véhicules qui circulent sur ces routes* ». En milieu urbain, les comptages routiers ont plus particulièrement vocation à mieux réguler les cycles de feux tricolores. Les comptages produisent une information qualitativement pauvre mais quantitativement plus importante. La principale distinction entre les deux types de comptages est spatiale : des comptages faits en un point donné, on distingue les comptages-trajectoires permettant d'observer le déplacement à l'aide de plusieurs points comme par exemple les données GPS. Dès lors, les données de comptage présentent des limites intrinsèques aux caractéristiques du système de comptage mobilisé ne permettant d'évaluer que des flux existants (Dupuy, 1975). C'est pourquoi lors des études de trafic routier ce système de mesure a vite présenté ses limites avec la croissance des flux automobiles et des externalités qui leur sont liés, comme la congestion ou la pollution atmosphérique. La question n'est plus seulement d'adapter et de gérer les infrastructures routières, mais aussi de les anticiper et d'en évaluer leurs impacts socio-économiques sur l'ensemble d'un réseau urbain (Gallez et Kaufmann, 2009).

Dans le cadre de cette thèse, nous nous appuierons plus spécifiquement sur les données issues des capteurs urbains qui permettent de mesurer le trafic routier en milieu urbain à différentes tranches de temps. Ce système de mesure se caractérise par la localisation de capteurs dans l'espace permettant de transmettre les données vers un système de collecte à des périodes régulières. Les données produites par ces capteurs présentent différentes caractéristiques et se définissent par les critères suivants (Servigne *et al.*, 2009) :

- Le réseau de capteurs localisés constitue le système d'acquisition de la mesure
- Les capteurs mesurent en temps réel ou à fréquence régulière un phénomène
- Les données sont transmises vers un système de collecte centralisé
- Les capteurs peuvent être fixes ou mobiles.

La connaissance des déplacements en milieu urbain par ces sources présente une opportunité pour la recherche et permet un nouvel accès à une connaissance fine du territoire (Aguiton *et al.*, 2009). Si l'intégration des données d'enquête pose un problème de fond pour désagréger les données sur l'espace urbain (Banos, 2006 ; Banos et Thévenin, 2008), les données de capteurs inversent la tendance et il est souvent nécessaire de les agréger pour produire une information pertinente en vue d'analyser ces mobilités (Fen-Chong, 2012). Avec ces données issues des capteurs, il apparaît plus faisable d'appréhender le trafic routier sous sa composante dynamique, en plaçant le mouvement au centre du phénomène à étudier. Cette tendance est à présent possible grâce aux données de comptage issues de capteurs qui permettent un accès à une source de données de plus en plus massive grâce à l'enregistrement en temps réel. En outre, ces données s'insèrent actuellement dans une véritable demande auprès des décideurs locaux, qui à travers les concepts de *smart city* (ville intelligente) et de *big data* cherchent à disposer de données exhaustives temporellement et sur des thèmes multiples. L'application des données issues des capteurs disséminés dans l'espace urbain au profit des collectivités territoriales vise à améliorer la compréhension, le suivi, l'analyse et la planification de la ville en temps réel (Batty, 2013). Les données relatives aux comptages routiers figurent en avant de ces concepts et constituent comme le relève Michael Batty (2013), les premières données automatisées de la ville mises en place pour observer les mouvements du territoire.

### 2.1.3. Les notions de *big data* et *les smart cities* pour replacer les capteurs au service de l'information urbaine

Le bref panorama que nous avons dressé précédemment induit différents questionnements sur la nature et la production des données issues des systèmes de comptage. Il entraîne, tout d'abord, une nouvelle relation ; l'usager ne consomme plus seulement la donnée, il en est aussi le producteur et le diffuseur. Chaque individu est un « *prosumeur* »[28], c'est-à-dire qu'il est à la fois le producteur et le consommateur de l'information. Avec le développement des nouvelles technologies et des logiciels open-source, l'accès à une multitude de données collectées en temps réel ouvre la voie à de nouvelles problématiques pour la géographie quantitative et théorique (Batty, 2013 ; Audard *et al.*, 2014 ; Kitchin, 2014 ; Pumain, 2014a). Par exemple, les concepts de *big data* et de *smart cities* (villes intelligentes) offrent actuellement un cadre de réflexion intéressant pour répondre non seulement aux attentes sociales et économiques, mais aussi environnementales, de la politique urbaine. Michael Batty (2013, p. 276) soulève ce point de synergie entre les deux notions « *il y a une coïncidence entre qui est maintenant appelé les villes intelligentes (smart cities) et les big data, avec l'intelligence (smartness) dans les villes, portant principalement sur la manière dont les capteurs peuvent générer de nouveaux flux de données spatialisées en temps réel, et comment les bases de données qui sont ensuite générées peuvent être intégrées de sorte qu'elles puissent apporter une plus-value à la politique urbaine* ». En d'autres termes, l'idée centrale des *smart cities* vise à centraliser et à utiliser l'ensemble des données produites par un vaste ensemble de capteurs autant au niveau des usagers que des collectivités en vue d'améliorer l'information non seulement à destination des usagers mais aussi pour les collectivités territoriales.

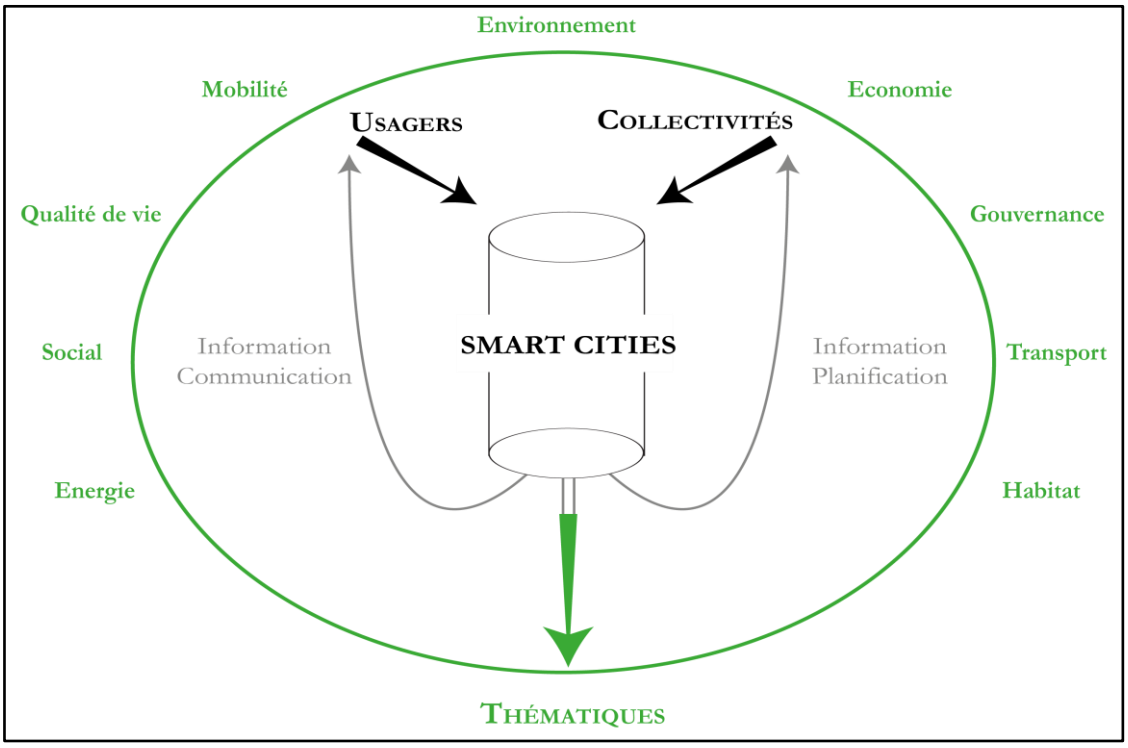



**Figure 1.13 : Le concept de *smart city* au service de l'information publique pour et par les acteurs de l'espace urbain (les usagers et les collectivités)**

[28] Ce terme anglo-saxon est la contraction des termes de « ***pro****ducer* » et « *con****sumer*** » et a été introduit en 1980 par Alvin Toffler dans son livre : *The Third Wave*, ed. Bantam Books, New York, 560p.

En fournissant des flux de données autant sur les mouvements des personnes que des marchandises, les capteurs s'insèrent dans cette problématique d'instrumentation de la ville intelligente envers les politiques publiques. Les données issues des capteurs peuvent, comme illustré en figure 1.13, être de natures différentes portant autant sur l'environnement ou l'énergie, que la mobilité, le transport ou l'habitat. Toutefois, la ville ne devient intelligente que si les données sont intégrées, synthétisées dans l'optique d'améliorer l'efficacité, l'équité, la durabilité et la qualité de vie dans les villes (Batty, 2012). C'est aussi la réappropriation des données produites et créées par l'espace public qui constitue une opportunité pour une nouvelle information territoriale. En effet, les *smart cities*, à travers les capteurs, visent à dépasser les problèmes de l'accessibilité des données, qui sont majoritairement produites par des entreprises privées, pour les replacer d'une part au service des usagers par l'information et la communication, et d'autre part au service de la collectivité par l'information et la planification (*cf.* figure 1.13). Les *smart cities* s'inscrivent donc dans une double boucle, où les *prosumeurs* se situent sur deux niveaux, celui de l'individu, et celui de la collectivité territoriale. L'inscription des *smart cities* dans les problématiques actuelles se fait en lien avec le développement croissant des données massives comme le *big data* (Batty, 2012, 2013).

Le terme de *big data*, quant à lui, est utilisé pour désigner un grand ensemble de données, dont la masse collectée ne permet pas d'être traitée par les outils informatiques classiques (Manovich, 2011). Michael Batty (2013) apporte une définition plus précise, en décrivant les *big data* comme étant toutes les bases de données qui ne tiennent pas dans une feuille de calcul Excel [29]. Leur quantité n'est pas seulement la seule caractéristique intéressante de ces données pour la géographie. C'est particulièrement leur caractère quasi-exhaustif, spatial et dynamique qui en fait un domaine d'étude de plus en plus séduisant pour les géographes (Audard *et al.*, 2014). Comme le relève Denise Pumain (2014a), *« la plupart des systèmes qui les enregistrent comportent une géolocalisation, du GPS aux téléphones mobiles , des caméras de surveillance aux capteurs embarqués dans les véhicules ou disséminés dans l'environnement, jusqu'aux localisations virtuelles des textes et des dialogues affichés sur la toile, dans des espaces de proximité connective ou sémantique ; la géographie est partie prenante pour l'analyse et le traitement de ces informations nouvelles* ». Au-delà de cet aspect spatial, il faut relever que les *big data* se caractérisent aussi dans le temps par un estampillage des enregistrements géolocalisées à plusieurs moments. Si la majorité de ces données sont issues de capteurs, il faut être conscient que les *big data* n'y sont pas uniquement associées (Batty, 2013).

Généralement lorsque l'on parle des *big data*, on pense plus aisément au nombre considérable de données qui sont enregistrées par notre utilisation des navigateurs de recherche en ligne comme Google ou l'utilisation des réseaux sociaux comme Facebook ou Twitter. Cependant, ici, nous restreignons notre définition à tous les systèmes de capteurs (fixes ou mobiles) permettant la collecte d'informations spatio-temporelles. Ainsi, on entend prendre en compte toutes les données qui permettent de se rattacher à un lieu et donc d'en faire une analyse géographique. Les sources de ces grands ensembles de données se déclinent en trois catégories (Kitchin, 2013) :

[29] La taille limite d'une feuille de calcul Excel est définie 1 048 576 lignes et 16 384 colonnes, soit un peu plus de 17 milliards de cellules : https://support.office.com/fr-fr/article/Sp%C3%A9cifications-et-limites-relatives-%C3%A0-Excel-1672b34d-7043-467e-8e27-269d656771c3?ui=fr-FR&rs=fr-CA&ad=CA.

- Les **sources de données directes** sont issues des dispositifs de surveillance individuels ou collectifs, c'est-à-dire dont le recueil de données est focalisé sur l'individu ou le lieu d'implantation du capteur (GPS, GSM).
- Les **sources de données automatisées** sont générées par un enregistrement automatique induit par l'usage d'un service (application ou service en ligne). Cette fonction est inhérente au capteur qui enregistre l'historique des usages que nous en faisons. À ce niveau, la collecte des données ne nécessite pas d'action de l'usager ou de l'opérateur. Les enregistrements de données s'opèrent directement par l'usage d'une application ou d'un site web.
- Les **sources de données volontaires** (*volunteered data*) sont produites par l'utilisateur. Elles se caractérisent par l'action de l'utilisateur à travers un réseau social et les sites de *crowdsourcing*. Les utilisateurs génèrent et contribuent à l'enrichissement des données par une action volontaire comme lors de l'utilisation de l'application Waze (*cf.* p. 67).

Ces données apportent de nouvelles informations sur les pratiques sociales et spatiales que les enquêtes ne peuvent révéler. La collecte régulière de ces données spatio-temporelles est aussi un enjeu sociétal et au service de la mémoire collective (Thévenin, 2011), mais qui est aussi public et politique à travers les *smart cities*. De plus, les *big data* n'ont, à notre sens, pas vocation à faire disparaitre les enquêtes. Elles s'insèrent plutôt comme un dispositif complémentaire permettant par leur nature, massive, d'enrichir la faiblesse des enquêtes. Lev Manovich (2011) relève la capacité de ces *big data* à être articulées à d'autres jeux de données. Outre cette tendance, la mobilisation des *big data* permet la mise en place d'approches plus spécifiques en se laissant guider par les données (*data-driven*) (Miller et Goodchild, 2014). Cette approche dévie de l'approche traditionnelle en générant des hypothèses et des idées nées à partir des données plutôt qu'issues de la théorie (Kitchin, 2014). Les *big data* offrent ainsi des perspectives nouvelles pour la recherche, afin de passer de données relativement rares à des données plus riches et disponibles, mais aussi de passer plus aisément de situations statiques à des situations dynamiques.

Il demeure une seconde question autour de ces données qui porte sur leur accessibilité. Comment accéder et analyser ces grandes bases de données ? Si l'individu en est le producteur, il est encore loin d'en être le détenteur. De même que les collectivités perdent la mainmise sur les données qu'elles produisent et plus particulièrement lorsque celles-ci délèguent certaines de leurs compétences à des organismes privés chargés de la gestion des grandes infrastructures comme par exemple le transport urbain. Les données géographiques et sociales sont depuis longtemps centralisées au niveau de l'État (Desrosières, 2000), mais aujourd'hui on assiste plutôt à une désinstitutionalisation des données individuelles vers le domaine privé. Les usagers des diverses applications donnent des informations les concernant. Celles-ci sont collectées et stockées par les propriétaires des applications. Pour donner un ordre de grandeur, on estime que les plus grandes entreprises détiennent près de 85 % de ces données individuelles[30]. Et seules ces entreprises ont accès à ces grandes masses de données. Lev Manovich (2011, p. 5) illustre cette tendance en montrant que les « *anthropologues travaillant pour Facebook ou un sociologue travaillant pour Google auront un accès aux données sociales que le reste de la communauté scientifique ne peut obtenir* ». Si l'accès aux

[30] Big Data et données personnelles : http://www.cil.cnrs.fr/CIL/spip.php?article1804.

données privées reste plus problématique, les données publiques s'ouvrent progressivement[31].

Néanmoins, nous n'en sommes qu'au début et il est encore difficile d'accéder à un grand nombre d'informations géographiques portant autant sur l'espace urbain que sur le trafic routier (CETE, 2013). C'est aussi en réponse à cette désinstitutionalisation que les concepts de *smart cities* se construisent, permettant aux politiques publiques de (re)centraliser l'ensemble des données produites sur leurs territoires en s'affranchissant des grandes entreprises productrices de données. À ce décalage entre les données détenues par les secteurs privé et public, on peut ajouter un dernier type de données issu des processus collaboratifs. Ces dernières constituent une alternative aux données privées (Audard *et al.*, 2014). Elles imposent toutefois, une double limite, tant sur l'harmonisation et la validation des données collectives, que sur leur caractère hétérogène (Audard *et al.*, 2014). Si la collecte et le recueil de ces données se fondent sur le principe participatif, le plus intéressant avec ces données est leur caractère volontaire et leur potentiel de développement. Michael Goodchild (2007) y voit par exemple un réseau potentiel de 6 milliards de capteurs humains, où chacun interprète et apporte une information locale.

Pour terminer, les *big data* couvrent aujourd'hui des domaines multiples allant de notre rapport à l'espace à nos relations sociales, en passant par nos usages du web. L'ensemble des arguments que nous avons mis en avant permet alors d'inscrire les données issues des capteurs dans le champ des *big data*. Ces systèmes de mesure répondent aux critères mis en avant : géolocalisés, les capteurs mesurent directement et automatiquement le phénomène à observer. Ils se fondent sur un recueil de données estampillées autant dans le temps que dans l'espace. À travers cette définition, nous verrons que le dispositif de quantification mobilisé dans cette thèse, les capteurs fixes, s'inscrit autant dans le champ des *smart cities* que des *big data*.

## 2.2. Les systèmes de capteurs mobiles pour capter la trajectoire des véhicules

L'intérêt actuel qui est porté autour des données de capteurs, qu'ils soient mobiles ou fixes, connaît de multiples applications pour les études géographiques. Les données issues de ces systèmes de comptages sont de plus en plus nombreuses dans notre environnement. Déployés en réseau, les capteurs transmettent des données au plus près du phénomène à étudier. Avec la *smart city*, tout comme l'homme avec les nouvelles technologies, c'est aussi la ville qui s'enrichit dans le même temps de ces différents capteurs. Ainsi envisagés, les capteurs fournissent différentes informations sur nos pratiques de l'espace urbain et nos habitudes.

La typologie présentée, précédemment, en figure 1.12 distingue deux types de dispositifs de comptage auquel on peut rattacher un système de mesure : les comptages-trajectoire plutôt rattachés à des capteurs dits *mobiles* et les comptages-points à des capteurs dits *fixes*. En complément, la typologie présentée en figure 1.14 recense les différents capteurs rencontrés en milieu urbain sur deux axes visant, d'une part, à distinguer les observations produites activement par l'usager de celles qui sont produites passivement par l'usager ; et d'autre part, à distinguer les capteurs mobiles des capteurs fixes. Ainsi, en fonction de la nature des capteurs mobilisés et des

[31] Par exemple, la mise en place de la directive Européenne de 2007 INSPIRE vise à faciliter l'accès aux données géographiques détenues par les autorités publiques (de l'État aux collectivités territoriales) en imposant la diffusion en *open access* de leurs données : Directive INSPIRE (2007/2/CE) relative à l'établissement d'une infrastructure d'information géographique dans la communauté européenne : http://inspire.ec.europa.eu/.

données qui en sont produites, quatre grandes typologies de capteurs se distinguent. À chacune de ces zones est rattaché un exemple de sources de données au sein de la figure ci-dessous :

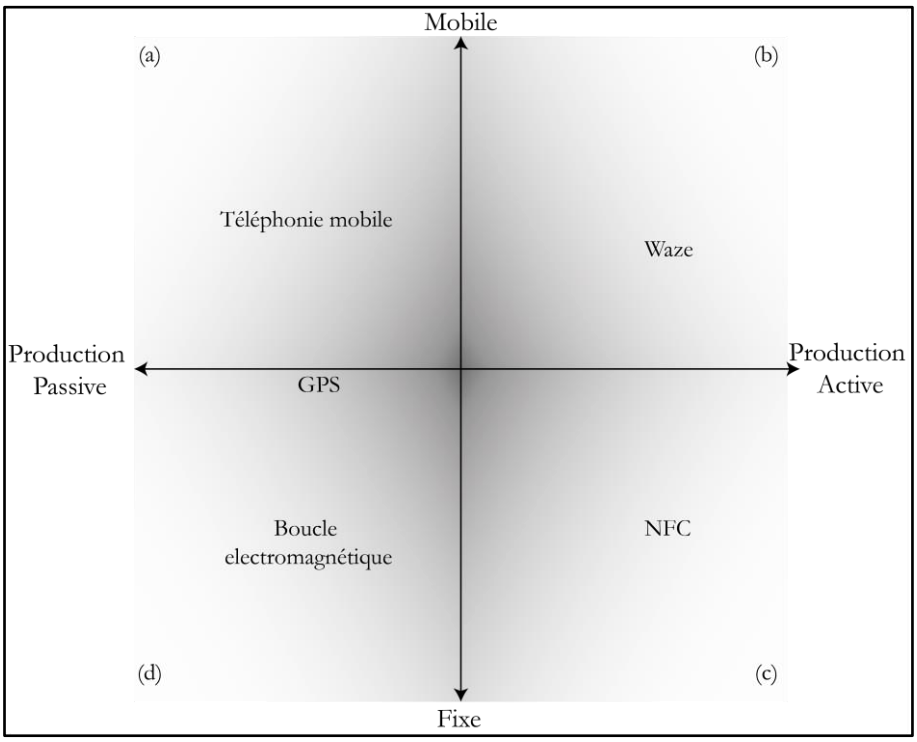



**Figure 1.14 : Les différents types de capteurs en milieu urbain pour le trafic routier (d'après Aguiton *et al.*, 2009, p. 2 et Fen-Chong, 2012, p. 31)**

### 2.2.1. **Les caractéristiques des capteurs mobiles**

En suivant la typologie proposée en figure 1.14, les capteurs mobiles permettent une mesure embarquée en appréhendant la trajectoire d'un objet par la succession de coordonnées localisées dans le temps et dans l'espace. Ces systèmes de mesures se fondent sur l'utilisation de capteurs sans fil équipés de technologies de géolocalisation (GPS, Bluetooth, téléphones mobiles…) permettant de transmettre et de stocker les données récoltées vers une ou plusieurs stations de base (Servigne *et al.*, 2009). Formées en un réseau véhiculaire, ces technologies mobiles rendent compte d'un ensemble de déplacements individuels (Faye *et al.*, 2012). En étant individualisé, la mobilisation de ces types de capteurs apparait mieux répondre aux problématiques de la complexité des mobilités urbaines en « *permettant d'appréhender les mutations en cours, de disposer d'outils de diagnostic nécessaires à la gestion et la planification de systèmes de transport intermodaux et d'imaginer des systèmes de transport durables* » (Gwiazdzinski et Drevon, 2014, p. 8). D'un point de vue thématique, l'utilisation des capteurs mobiles pour l'étude du trafic routier se définit sous le terme de Floating Car Data (FCD), qui permet à partir des GPS ou des téléphones mobiles de connaitre la position, la direction et la vitesse d'un véhicule sur les infrastructures de transport (Pfoser, 2008 ; Servigne *et al.*, 2009). Les FCD se réfèrent à l'utilisation des nouvelles technologies légères comme les téléphones mobiles ou les GPS en vue d'acquérir des informations sur l'état de la circulation routière. Il existe plusieurs technologies permettant d'extraire ces informations pour l'étude du trafic routier à une échelle individuelle, dont nous proposerons une synthèse plus bas (p. 70). Comme illustré en figure 1.14, deux types de technologies mobiles peuvent être distingués : les premiers sont le fruit d'une production passive

des individus, alors que les seconds sont produits activement par les individus via des communautés d'usagers, informant de leur plein gré de leur localisation, mais également des conditions de trafic routier auxquelles ils sont confrontés.

Tout d'abord, les capteurs mobiles à production passive constituent la première catégorie identifiée au sein de la figure 1.14. Ils regroupent les dispositifs de capteur mobiles dont la production de données se caractérise par l'utilisation passive d'un service pour la mesure du trafic routier. C'est, par exemple, le cas lorsque l'on utilise les outils de navigation GPS, ou lors de communications mobiles durant des déplacements routiers. L'utilisation des données de téléphonie mobile constitue une bonne illustration des capteurs mobiles, afin de montrer la structuration des déplacements en lien avec les axes de transports (Fen-Chong 2012). La mesure issue de la téléphonie mobile informe sur la localisation des individus au cours de la journée en fonction de leurs communications. Celle-ci peut être appréhendée de deux manières (*cf.* tableau 1.3, p. 70) : les communications FMD reliées à des cellules, et les communications GSM reliées aux antennes du réseau de téléphonie mobile. Si ces données apportent une information sur les trajectoires des individus pour l'analyse des flux de mobilité, leur exploitation n'est pas directe et nécessite la mise en place de traitements spécifiques principalement issus des outils de géovisualisation et de géoexploration (Andrienko *et al.*, 2009). Il faut souligner que pour des raisons de protection de la vie privée, les données de téléphonie ne permettent pas de localiser un individu. Elles sont collectées au niveau des antennes du réseau de téléphonie mobile et présentent donc un écart entre les trajectoires individuelles et les données collectées (Fen-Chong, 2012). Cette incertitude empêche une analyse de la mobilité et donc du trafic routier à l'échelle de l'individu et de ses localisations exactes. Par exemple, il est d'usage d'en agréger les données pour identifier les flux de mobilité à un niveau plus haut. Au cours de sa thèse, Julie Fen-Chong (2012) soulève aussi les défis techniques et méthodologiques à mettre en place pour rendre compte des déplacements à l'échelle d'un territoire. En effet, les données exploitées à l'échelle individuelle sont trop nombreuses pour être visualisées dans leur totalité ; il est alors primordial de structurer ces données pour en extraire une information via la mobilisation des outils rattachés à la *géocomputation* (*cf.* section 1.3, p. 45) : « *L'analyse exploratoire prend ici tout son sens : la grande quantité de données nécessite des agrégations spatiales et temporelles pour produire une information pertinente. Plusieurs types d'objets permettent d'appréhender ces mobilités : des cartographies des flux de mobilité, mais aussi des vecteurs moyens donnant les grandes tendances observées à l'échelle du territoire et non plus des seules stations du réseau de téléphonie mobile. L'utilisation de ces représentations est particulièrement efficace pour montrer à la fois les régularités existantes dans l'organisation des mobilités […] mais aussi la forte structuration de ces dernières par les axes de transports* » (Fen-Chong, 2012, p. 148). L'ensemble de ces critères montre que les capteurs mobiles constituent non seulement un système de comptage intéressant, mais aussi original pour l'étude du trafic routier. Néanmoins, face aux critères mis en avant lors de l'introduction, les informations collectées à partir des capteurs mobiles passifs ne sont pas explicitement dédiées à la mesure du trafic routier. De plus, les leviers éthiques, techniques et méthodologiques amènent à nous intéresser à la seconde typologie des capteurs mobiles illustrée en figure 1.14.

En ce qui concerne les capteurs mobiles à production active se rattachant à la typologie (b) présentée en figure 1.14, on regroupe tous les systèmes qui sont explicitement dédiés à la collecte de données de trafic et où l'usager prend une place active dans le processus de suivi. Ils peuvent être de deux ordres : soit l'individu est directement équipé d'une technologie GPS et participe

activement à partager ses déplacements routiers pour les besoins d'une étude ou d'une expérience (Gwiazdzinski et Klein, 2014) ; soit c'est directement l'individu qui participe par l'utilisation d'une plate-forme collaborative de *crowdsourcing*. Dans le premier cas, cette méthode de comptage se rapproche d'une enquête classique, où chaque individu volontaire est équipé d'une technologie de localisation lui permettant de fournir des informations, sur une période plus ou moins longue, de ses différents parcours. Toutefois, la mise en place de ces enquêtes nécessite une phase de formation et d'explication sur le fonctionnement du GPS, par exemple avec l'usager porteur de la technologie (Drevon *et al.*, 2014). Si cette méthode d'acquisition revêt de multiples intérêts, elle ne constitue pas un système de comptage valide pour notre travail. En effet, la mise en place d'une enquête par suivi GPS ne permet pas de répondre au jeu de critères dont nous avons identifié les besoins pour le travail présenté en introduction de ce premier chapitre (p. 55). Dans le second cas, les données produites activement se fondent sur le principe d'une collecte participative où chaque usager apporte une information géolocalisée sur son déplacement. Pour le trafic routier, nous pouvons citer l'application collaborative Waze[32] qui permet une interaction entre utilisateurs afin d'apporter des informations en temps réel sur l'état du trafic routier sous la forme d'une carte illustrée en figure 1.15 :

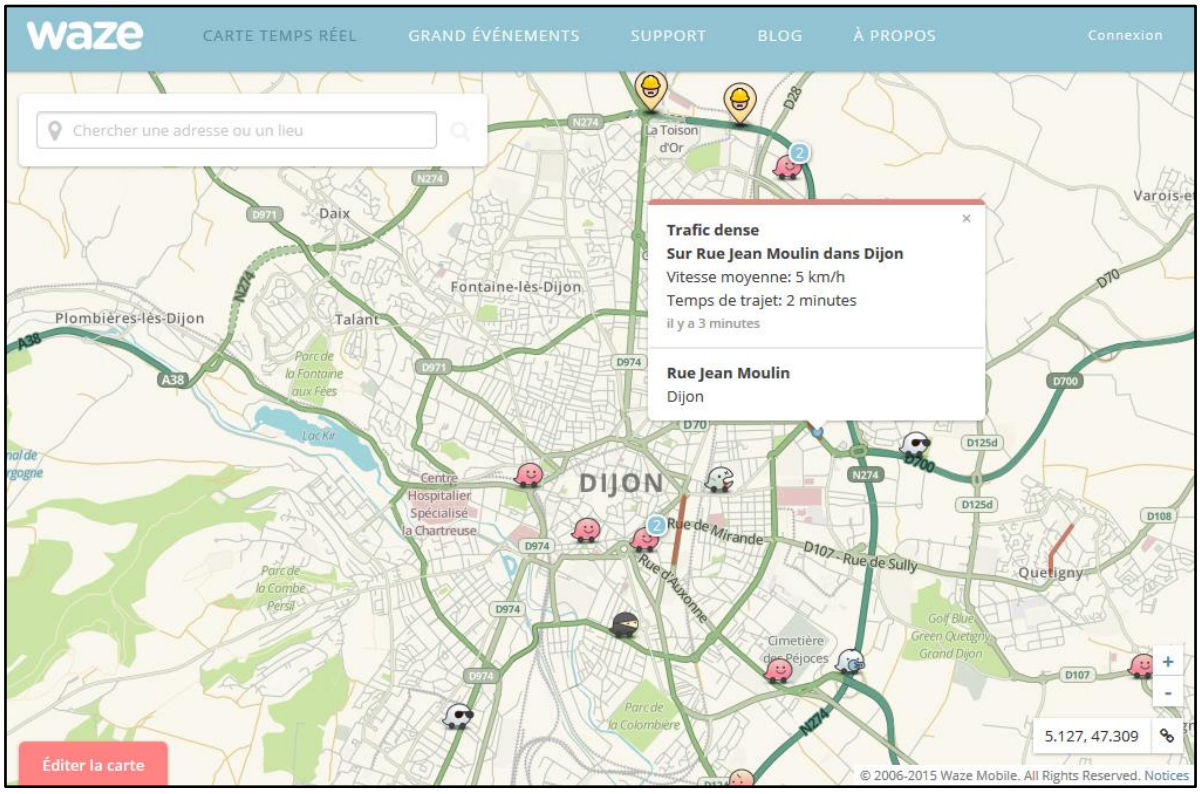



**Figure 1.15 : Extrait d'une carte du trafic routier sur la ville de Dijon depuis le site Waze. La localisation des utilisateurs est individualisée ; les événements de trafic routier sont décrits par le motif de l'évènement, la vitesse du trajet, le temps de trajet impacté (source : https://www.waze.com/fr/livemap)**

Les données de localisation sont automatiquement collectées par les téléphones mobiles qui sont équipés de GPS. À l'inverse des données de téléphonie collectées sans l'accord de l'usager, les données issues de ces systèmes permettent une représentation à l'échelle individuelle. Les

[32] Site de l'application Waze : https://www.waze.com/fr/.

données issues de ces systèmes de capteurs se fondent sur le principe du *crowdsourcing* via l'utilisation de *smartphones* (Chatzimilioudis *et al.*, 2012). Il implique une mobilisation active et participative des utilisateurs permettant d'enrichir en temps réel une information spatiale au même titre que les données issues d'OpenStreetMap (http://openstreetmap.fr/). La mise en place de ces systèmes de capteurs amène à un changement pour l'étude géographique. En effet, les données géographiques ne sont plus produites et détenues par l'État (Desrosières, 2000) ; les usagers s'approprient la production de l'information par eux-mêmes (Bakis et Valentin, 2010). Cette tendance induit une évolution dans l'utilisation des SIG, avec l'émergence de nouveaux utilisateurs créant une information géographiquement volontaire (Goodchild, 2009, 2010). Le terme VGI (*Volunteered geographic information*) introduit par Michael Goodchild permet de « *décrire l'action des utilisateurs contribuant à enrichir l'information géographique* » (Goodchild, 2009, p. 94).

### 2.2.2. **Les comptages-trajectoires : une mesure individuelle à appréhender**

Que la collecte des données issues des technologies mobiles soit active ou passive, l'information produite à partir de ces données nécessite d'expliciter les variables accessibles à partir de ces technologies. Les variables individuelles se caractérisent par la trajectoire des véhicules relevés à plusieurs moments dans l'espace. Le comptage est mobile et nécessite d'équiper les véhicules de capteurs individuels relevant le positionnement des véhicules à plusieurs moments donnés, comme par exemple les GPS (Hadachi, 2013) ou les téléphones mobiles équipés de GPS (White *et al.*, 2004). Les variables recueillies permettent de schématiser la trajectoire du véhicule au fil du temps comme illustré dans la figure 1.16 (ci-dessous) pour les données GPS :

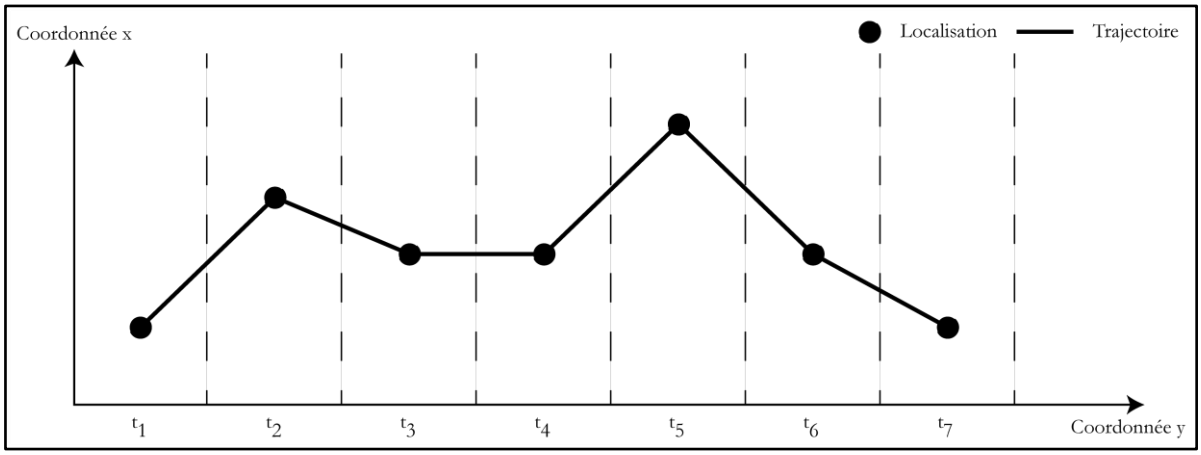



**Figure 1.16 : L'exemple de comptage-trajectoire du trafic automobile par relevés GPS (illustration d'après Hadachi, 2013)**

Chaque trace est utilisée pour récupérer les données de localisation issues des capteurs. La position des véhicules se caractérise par sa localisation $(x,y)$ à un instant $t$. Sa vitesse $v$ de déplacement est estimée en comparant les différents positionnements des véhicules sur le réseau routier par la relation suivante $t_1(x,y)- t_2(x,y)$. La trajectoire du véhicule est ainsi mesurée par l'association des différentes localisations relevées au cours du temps sur le réseau routier (Hadachi, 2013). Les principales variables individuelles du trafic routier sont les suivantes (Buisson et Lesort, 2010) :

- La **localisation,** qui s'exprime en coordonnées spatiales et définit le positionnement du véhicule à un instant donné
- La **vitesse de déplacement** exprime la distance parcourue par le véhicule dans le temps et s'exprime le plus habituellement en kilomètres par heure (km/h)
- La **trajectoire** est l'ensemble des déplacements relevés par les capteurs dans le temps.

Ces données fournissent une description fine du mouvement des véhicules sur les infrastructures routières. Il apparait, toutefois, évident que la mise en place de ce dispositif nécessite d'équiper un grand nombre de véhicules pour rendre compte de la circulation routière dans sa totalité. On peut, néanmoins, relever que la démocratisation des technologies mobiles équipées pour la grande majorité des systèmes de géolocalisation pallie ce problème. Elle ouvre plutôt la voie à de nouvelles problématiques d'avantage liées à la protection de la vie privée qu'à la capacité à équiper les flottes de véhicules de ces dispositifs. En outre, ses systèmes offrent l'opportunité d'un transfert de données bidirectionnel (Borsu, 2004). Ils permettent de récolter des données de trafic et d'apporter au conducteur des informations sur les conditions de circulation.

Les comptages-trajectoires soulèvent différents écueils dont le plus caractéristique porte sur leur granularité spatiale. En effet, dans le cas des données personnelles géolocalisées, la localisation est établie à partir de la localisation du capteur, que ce soit pour le téléphone mobile, les GPS ou le capteur environnemental. Elle permet d'appréhender les mouvements à un niveau fin. Mais la recherche d'une granularité spatiale plus fine rentre souvent en conflit avec la législation visant la protection des données à caractère personnel et aux traitements des informations à caractère nominatif[33]. Le passage à des données individuelles soulève régulièrement cette question de protection de la vie privée. Par exemple, dans le cas des données de téléphonie mobile, l'utilisation de ces dernières est très réglementée et il n'est pas permis pour un opérateur de faire le lien entre les données de localisation et les caractéristiques socio-économiques des individus (Dupuy et Benguigui, 2015).

Les enjeux autour des données géolocalisées se situent aussi là : comment assurer une déconnexion entre la localisation des individus et leur identité, tout en préservant leur anonymat. Holmes Miller (2013) propose, par exemple, d'agréger les données pour assurer un compromis entre la préservation de la granularité fine des données individuelles et celle de la vie privée. En effet, mobiliser les données individuelles à un niveau agrégé est une réponse aux contraintes éthiques et juridiques liées à l'utilisation des données personnelles. De surcroit, pour l'étude des mobilités et du trafic routier, il n'est pas nécessaire de se rattacher à l'individu et le niveau individuel n'est pas toujours l'échelle la plus pertinente : « *l'agrégation permet alors de changer de perspective, en s'intéressant non plus aux individus mais aux lieux qui sont fréquentés et mis en relation par ces individus* » (Fen-Chong, 2012, p. 92).

### 2.2.3. **Les limites des capteurs mobiles pour la mesure du trafic routier**

Le tableau 1.3 permet de dresser une liste des différentes technologies rattachées aux capteurs mobiles. Pour chaque technologie est spécifiée leurs mesures, leurs principes de fonctionnement ainsi que leurs limites pour la mesure du trafic routier en milieu urbain. En effet, si les problèmes

[33] Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés : https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000886460.

de protection de la vie privée constituent la limite la plus souvent mise en avant, force est de constater que d'autres contraintes liées autant au taux de couverture ou d'équipement des capteurs mobiles constituent aussi des limites non-négligeable pour la mesure du trafic routier :

| Technologie de capteurs mobiles | Mesures | Principes de fonctionnement | Principales limites |
|---|---|---|---|
| **GNSS (Global Navigation Satellite System)** | Vitesse moyenne, sens de déplacement, temps de parcours, congestion, matrices O/D | Utilisation d'un récepteur GPS traceurs pour calculer la position du véhicule | Pas de distinction des voies, Coût d'équipement important (transfert et stockage des données) |
| **Satellites** | Vitesse moyenne, sens de déplacement, temps de parcours, congestion | Communication par satellite directe | Taux d'équipement des véhicules, Disponibilité des satellites |
| **FMD (Floating Mobile Data)** | Vitesses instantanées et moyennes, temps de parcours, congestion | Localisation issues des réseaux de téléphonie mobile au sein de cellule | Précision faible, Disponibiité des mobiles en communication, pas de distinction des voies |
| **GSM (Global System for Mobile)** | Localisation précise (150 mètres avec un nombre d'antennes suffisant), temps de parcours, suivi des déplacements urbains et interurbains. | Localisation issues des réseaux de téléphonie mobile en lien aux antennes relais | Densité des antennes relais |
| **Bluetooth** | Vitesses moyennes, temps de parcours, sens de déplacement, matrice O/D | Interaction dispositif Bluetooth et station en bord de route | Taux d'équipeement à bord des véhicules et le long des voies |
| **WiFi** | Localisation (précision < 50m), vitesses moyennes, temps de parcours, suivi des déplacements urbains | Interaction dispositif WiFi et borne d'accessibilité WiFi | Couverture du territoire |
| **Communautés d'usagers** | Vitesses moyennes, temps de parcours | Interaction des usagers entre eux via une plateforme collaborative | Fiabilité des données, Taux d'équipement |

**Tableau 1.3 : Les différentes technologies de capteurs mobiles pour la mesure du trafic routier en milieu urbain (source : Borsu, 2004 ; SETRA, 2012)**

En fonction de la technologie employée, les mesures disponibles peuvent présenter une finesse pour l'étude du trafic qui n'est pas négligeable. C'est par exemple, le cas des technologies GNSS (Global Navigation Satellite System)[34] dont la précision est plus importante comparée à la mesure permise par la simple localisation mobile comme les GSM (Global System for Mobile). Toutefois, pour l'ensemble des technologies, on peut relever une limite commune qui se situe au niveau du coût d'implantation ou du degré d'utilisation (ou d'équipement) de ces technologies. Le taux de couverture et la disponibilité des dispositifs de recueil ou de détection sont souvent limités en l'état. En effet, si la démocratisation des téléphones mobiles ou des GPS de navigation permet aujourd'hui de couvrir un grand nombre d'utilisateurs, tous les individus ne sont pas équipés de ces technologies. De surcroît, la grande hétérogénéité des produits commercialisés ou des réseaux de communication mobilisés complexifie l'accès aux données[35]. En effet, dans la majorité des études utilisant ces technologies, un partenariat est mis en place entre les chercheurs et les

[34] Le terme de GNSS est un terme générique qui englobe l'ensemble des systèmes de navigation par satellites. Sous ce terme, on regroupe autant le système de positionnement américain GPS que celui européen GALILEO.
[35] Par exemple, une étude exhaustive portant sur l'ensemble des utilisateurs de téléphone mobiles nécessiterait de se rapprocher des quatre opérateurs mobiles français : Bouygues, Free, Orange, et SFR.

professionnels des télécommunications en vue de récolter ces données. Ce partenariat induit un coût supplémentaire en vue d'exploiter les données issues des appareils mobiles individualisés. Ces propos nous amenant alors à porter notre regard sur une seconde typologie de comptage : les comptages fixes.

## 2.3. Les données issues des capteurs fixes : une mesure collective

Les données issues des capteurs fixes se situent, elles-aussi, au cœur du phénomène que l'on souhaite étudier, le trafic routier, mais l'objet de leurs mesures diffère des capteurs mobiles. En effet, la mesure n'est pas effectuée à l'intérieur du phénomène (l'individu ou le véhicule), elle se situe plutôt à l'extérieur en étant placée au niveau de l'infrastructure routière (Champion, 2003), c'est pourquoi on parle plutôt d'une mesure collective. Dans ce travail, nous avons fait le choix de travailler avec les capteurs fixes pour plusieurs raisons que nous développerons au cours de cette section. Mais c'est aussi parce que pour la mesure du trafic routier, ces capteurs constituent une réponse aux limites intrinsèques des capteurs mobiles. Plusieurs points justifient ces propos :

- Une des réponses à la **protection des données individuelles** se situe au niveau des capteurs fixes qui fournissent une mesure dite collective
- La **propriété et la gestion des technologies** : les capteurs fixes sont le plus généralement installés et mobilisés par les collectivités territoriales
- Le **recueil de données** collectées au niveau des services publics facilite autant la collecte, que le traitement et la communication des données
- Un dispositif **exclusivement dédié au trafic routier** : les caractéristiques de ces systèmes sont le plus souvent explicitement destinées à la détection des véhicules comme les caméras de détections ou les boucles électromagnétiques.

Par ailleurs, les capteurs fixes sont des dispositifs anciens, ce qui leur donne une force pour la mesure du trafic routier en les positionnant comme un dispositif de comptage fiable, et de surcroît cela permet aussi de disposer d'une profondeur temporelle sur plusieurs années que les capteurs mobiles ne permettent que rarement d'apporter. Pourtant, les capteurs fixes ne sont actuellement que peu fréquemment mis en avant par rapport aux systèmes mobiles. Ces données ne sont que peu valorisées en tant que telles pour l'étude de la dynamique du trafic routier car, en partie, elles ne proposent que peu de descriptions sémantiques du fait de la nature de leurs mesures. Pourtant, elles se révèlent intéressantes pour l'étude de la circulation routière en milieu urbain. C'est pourquoi nous proposons de valoriser cette source de données en exploitant ces potentiels pour observer le trafic routier en milieu urbain. L'idée que nous souhaitons défendre ici est que le dispositif de mesure du trafic routier demeure un instrument fiable pour l'observation de la circulation routière, mais dont l'exploitation est sous-utilisée à des fins de contrôle et/ou de validation. Pourtant les capteurs fixes fournissent un ensemble d'informations structurées sous la forme d'un réseau et présentant un potentiel non-négligeable pour reproduire une carte de la dynamique de la circulation routière en milieu urbain.

### 2.3.1. **Les caractéristiques des différents capteurs fixes :**

Les capteurs fixes constituent la deuxième catégorie identifiée au sein de la figure 1.14 (p. 65). À l'inverse des capteurs mobiles, ils regroupent l'ensemble des technologies mesurant un phénomène en un point donné. Les capteurs fixes au contraire des mobiles se caractérisent par

un emplacement figé dans l'espace mais qui demeure estampillé dans le temps.

Tout d'abord, les capteurs fixes à production active regroupés sous la catégorie (c) en figure 1.14 constituent un système de comptage original et atypique par rapport aux autres catégories illustrées. En effet, cette typologie ne permet pas actuellement de placer un dispositif de capteurs fixes où les données produites permettent de saisir le déplacement des individus et des véhicules routiers. Toutefois, on peut voir émerger de nouveaux dispositifs proches de cette typologie dans l'espace urbain. L'exemple de capteurs capables de scanner les *smartphones* des passants voit son émergence avec le développement des communications NFC (Near Field Communication). Cette technologie, dont l'usage le plus connu est le paiement sans contact se développe aussi dans une optique publicitaire (http://www.clearchannel.fr/la-balise-connect). Les capteurs permettant cette mesure sont des balises passives placées sur le mobilier urbain et nécessitent l'action de l'utilisateur d'un *smartphone* pour activer le service. La vocation de ces systèmes est plus commerciale, en vue de proposer aux usagers des services spécifiques et des publicités ciblées. Néanmoins, on peut facilement envisager une collecte des données de déplacement individuel issues de ces systèmes placés sur le mobilier qui jalonne l'espace urbain.




**Figure 1.17 : Un système NFC placé sur le mobilier urbain (source : http://www.orange-business.com/fr/clear-channel)**

C'est pourquoi, pour les capteurs fixes nous nous focaliserons sur la dernière catégorie présentée sous la typologie (d) en figure 1.14 (p. 65). Pour la mesure du trafic routier, ces capteurs fixes à production passive sont depuis longtemps mis en place dans l'espace urbain en vue de gérer et d'optimiser la gestion des carrefours à feux (*cf.* section 3.1, p. 83). Cependant, ce système de comptage n'est que très rarement mis en avant lors des études sur le trafic routier. Pourtant, les informations collectées à partir de ce système de mesure fonctionnent continuellement au cours de la journée, des semaines et des années. Malgré toutes ces qualités, ce système de mesure est souvent utilisé à des fins de contrôle et de validation des études relatives à la circulation routière. C'est particulièrement le cas des capteurs fixes qui englobent les systèmes de mesure les plus classiques pour la mesure du trafic routier : les capteurs à boucle électromagnétique (*cf.* figure 1.18). Ces capteurs reposent sur un système de comptage placé à des lieux stratégiques et structurés en un réseau de capteurs fixes (Akyildiz *et al.*, 2002).

Ce capteur ponctuel permet de détecter le passage d'un véhicule lorsqu'il passe au-dessus de la boucle électromagnétique (Borsu, 2004). Il s'agit d'une utilisation passive du capteur, qui permet de mesurer une flotte de véhicules passant en un point. Cela permet ainsi d'obtenir directement le nombre de véhicules au cours d'une période donnée sur une ou plusieurs voies. Ce système de

comptage est le plus répandu pour la mesure du trafic routier, autant en milieu urbain qu'interurbain, sous le terme de SIREDO (Système Informatisé de REcueil de DOnnées de trafic). Déployées pour la mesure du trafic routier depuis les années 80, plus de 2 000 stations SIREDO sont placées sur les routes nationales françaises (Borsu, 2004). Gérées par les conseils départementaux[36] pour le réseau national, il apparait plus difficile d'estimer le nombre de comptages déployés en milieu urbain car les compétences de gestion de réseau sont du domaine des communes et des intercommunalités. En revanche, par leur ancienneté, ils constituent l'un des dispositifs de mesure les plus fiables et les plus pérennes pour évaluer le trafic routier passant sur une infrastructure de transport (Buisson et Lesort, 2010). Cette méthode de mesure est la plus répandue du fait de sa longue durée de vie et de son faible coût (Buisson et Lesort, 2010). Elle se constitue de 3 ou 4 bobines de fil de cuivre enfouies sous la chaussée. Cette boucle constitue la partie inductive du système de mesure qui, lorsqu'un véhicule passe sur la boucle, produit une oscillation (ou inductance) qui est associée au passage d'un véhicule. Le signal est ensuite comptabilisé par un bloc de comptage situé sur le trottoir qui permet de stocker ces informations, permettant alors de disposer d'une profondeur temporelle relativement importante. Les données sont ensuite centralisées par un système informatisé de recueil de données. Du fait de leur implantation, insérées au sein de la chaussée, ces capteurs présentent l'inconvénient de devoir être remplacés lors de la mise en place de travaux sur la chaussée.

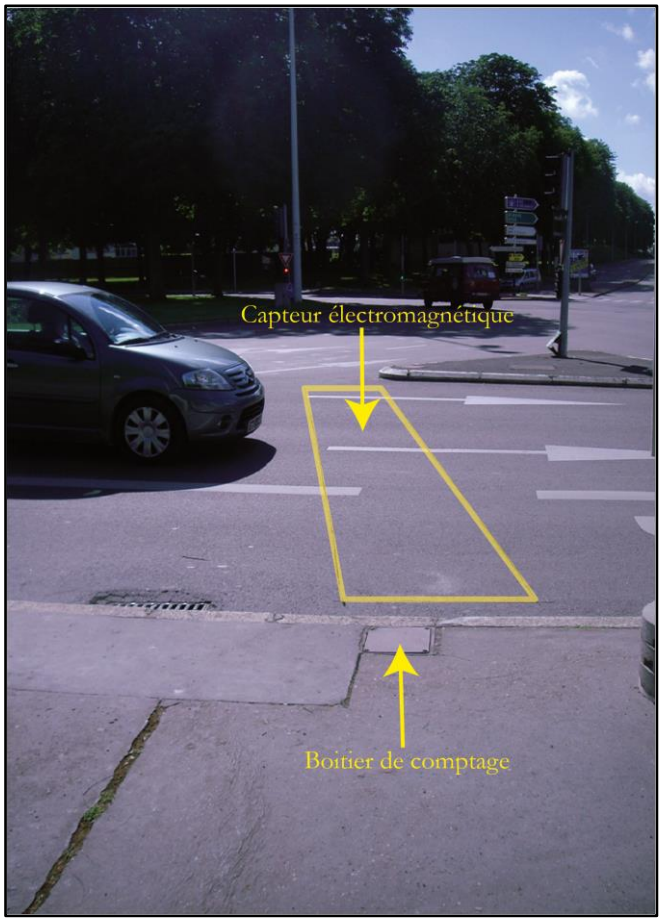



**Figure 1.18 : Illustration d'une boucle électromagnétique située sur le Boulevard Paul Doumer, à l'est de la commune de Dijon**

[36] À noter que depuis mars 2015, les « conseils généraux » sont redevenus des « conseils départementaux » : https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000027414225.

### 2.3.2. **Les comptages ponctuels : une mesure collective pour le trafic routier**

Les comptages-points ou comptages ponctuels constituent le cœur du corpus de données utilisé dans cette thèse (*cf.* Chapitre 3, p. 81). Ce type de données est le plus répandu pour le suivi du trafic en milieu urbain et permet de mesurer le passage des véhicules sur une infrastructure de transport. À l'inverse des comptages-trajectoires, le support de la mesure n'est pas le véhicule mais bien la route. En conséquence, les capteurs fixes fournissent une mesure située dans l'environnement proche du capteur. Les comptages-ponctuels permettent alors de caractériser la flotte de véhicules passant sur la route en fournissant des variables globales de mesure du trafic routier. À travers cet instrument de mesure fixe, deux grands indicateurs sont mobilisés (Buisson et Lesort, 2010) :

- Le **débit** est le nombre de véhicules passant pendant une période au point de comptage. Il s'exprime en nombre de véhicules pour un temps donné (par heure, jour, année) et décrit la répartition des véhicules dans le temps. Le débit est la variable la plus élémentaire des comptages routiers.
- La **concentration** est la densité de véhicules présents sur une infrastructure. Elle s'exprime en nombre de véhicules par unité de longueur (par kilomètre ou mètre) et décrit la répartition des véhicules dans l'espace.

Les données de comptages sont issues de systèmes de mesure transmettant des données au plus près du phénomène à étudier. Dès lors, les variables mesurées, individuellement ou collectivement, par ces systèmes permettent, non seulement, d'observer le trafic routier au cours du temps, mais aussi d'observer les conditions de circulation automobile sur un ensemble de section (*cf.* encart 1, p. 75). Elles se distinguent d'autant plus des données issues des enquêtes dont la résolution temporelle n'est pas aussi fine. Elles se distinguent aussi par leur profondeur temporelle (historicité), alors que les enquêtes ne sont que ponctuelles dans le temps. Organisés et déployés en réseau et au plus près du phénomène à étudier, les comptages routiers se rattachent à des systèmes de capteurs positionnés dans l'espace. Formés en réseau, ils constituent un système de mesure d'autant plus intéressant pour la représentation spatio-temporelle du trafic routier en milieu urbain.

***Encart 1 La relation entre Débit, Vitesse et Concentration pour caractériser les conditions générales de la circulation routière par section***

Ces variables caractérisent un flux de véhicules sur la route et ont amené à une vision statistique et mathématique du trafic (Champion, 2003). Ces méthodes s'inspirent de la dynamique des fluides et ont donné naissance à une notion essentielle pour étudier les conditions de circulation routière en mettant en relation le débit (*D*), la concentration (*Q*) et la vitesse (*V*) des véhicules comptés (Quinet, 1998). La figure 1.19 illustre cette relation sous trois formes, et permet de mettre en exergue les phénomènes de congestion routière. La formulation que nous avons choisie illustre les conditions de circulation dans des conditions optimales. Ainsi, la forme des courbes peut présenter d'autres formes en intégrant les paramètres des voies et le temps de trajet (Quinet, 1998 ; Buisson et Lesort, 2010). Le lien entre vitesse et débit montre que dans une situation fluide, la vitesse de déplacement décroît en fonction du débit (a), mais également en fonction de la concentration (c). Enfin, la dernière relation entre débit et concentration (b) montre que le débit de véhicules décroît selon la capacité des axes et inversement qu'il augmente en fonction de la concentration dans le cas d'un trafic fluide. Pour ces trois relations, deux conditions de circulation routière s'illustrent : lorsque le trafic est fluide et lorsque celui est congestionné. Par exemple, pour la relation (b) : au départ, le débit croît sans contrainte jusqu'à un optimum où celui-ci décroît. À ce niveau, le débit est supérieur ou égal à la capacité de l'infrastructure et les conditions de circulation se dégradent pour atteindre une situation de congestion. On observe les mêmes tendances pour les deux autres illustrations qui représentent les conditions de trafic sous des formes différentes.

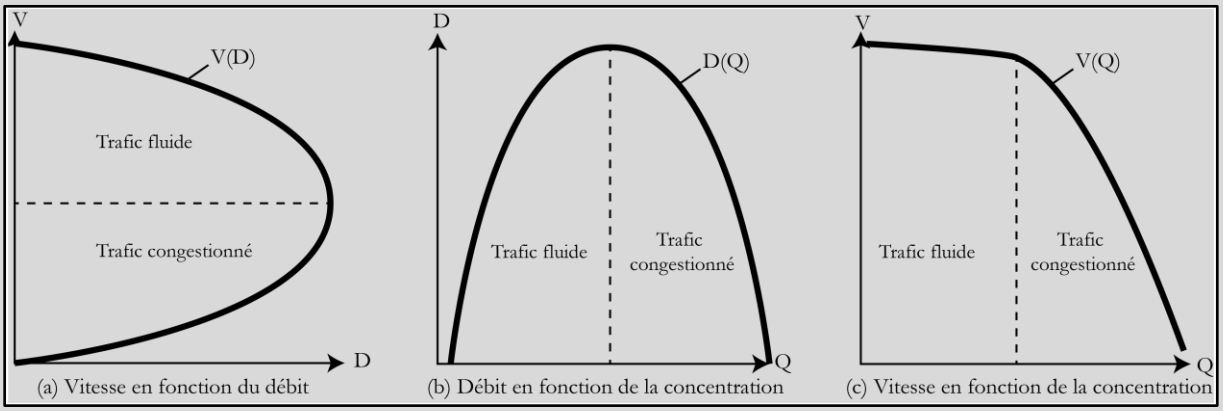



**Figure 1.19 : Les relations entre débit Q, vitesse V et concentration D (d'après Quinet, 1998 ; Buisson et Lesort, 2010)**

### 2.3.3. **Les avantages des capteurs fixes pour la mesure du trafic**

Les technologies de capteurs permettent de recueillir un ensemble de données géolocalisées et se situent au plus près de l'événement à observer (Gwiazdzinski et Drevon, 2014, p. 8) : « *elles permettent également d'imaginer de nouveaux croisements, de nouvelles hybridations, de nouveaux projets entre chercheurs, professionnels, acteurs publics et privés dans une logique de plate-forme d'innovation ouverte et de poser la question du nouveau rôle de l'usager co-concepteur des systèmes de mobilité* ». Par cette citation, l'utilisation de plus en plus intensive des outils de géolocalisation répond aussi à des besoins pour les acteurs des transports (décideurs, gestionnaires, exploitants, …) en vue de mieux saisir la complexité de l'espace intraurbain. Par exemple, ces systèmes peuvent se positionner comme un outil complémentaire aux enquêtes classiques de déplacements (Drevon *et al.*, 2014). Si les capteurs

mobiles s'avancent comme un nouvel outil en vue de capter la complexité des mouvements intraurbains (Gwiazdzinski et Klein, 2014), leur avènement peut-il remplacer, pour autant, les données produites via les enquêtes classiques ? À travers cette dernière section, c'est cette dernière question que nous souhaitons explorer en portant le regard sur une analyse duale des deux grandes typologies de capteurs (fixes et mobiles) présentées *a posteriori*.

Nous avons vu que les données de capteurs ne peuvent, par leurs natures apporter une information sémantique aussi importante que les données issues des enquêtes (*cf.* figure 1.12, p. 59). Pourtant les capteurs présentent différents atouts pour appréhender et quantifier les dynamiques urbaines avec un degré spatio-temporel relativement fin. Certaines études cherchent même à mobiliser exclusivement les données issues des capteurs, ce qui permet de s'affranchir de la lourdeur de la passation des enquêtes de mobilité (Drevon *et al.*, 2014 ; Gwiazdzinski et Klein, 2014). C'est typiquement le cas des enquêtes dites « *GPS only* » initiées à Cincinatti (Giamo *et al.*, 2010) et à Paris (Nguyen-Luong *et al.*, 2012). Ces études montrent la volonté de disposer d'un corpus de données issues des dispositifs de capteurs mobiles en vue de s'affranchir de la lourdeur des dispositifs d'enquête sur la mobilité. Toutefois, dans le même temps, très peu d'études mettent en avant les données issues des capteurs fixes alors que celles-ci sont très souvent utilisées lors des études relatives au trafic routier. Par exemple, elles servent de référence lors de la mise en place de modèles de trafic routier ou de déplacements urbains (Bonnel, 2002 ; Ortúzar et Willumsen, 2011). Ces données ne sont que très peu valorisées en tant que telles au profit des données issues des capteurs mobiles qui s'inscrivent probablement dans un effet de « mode » que nous pouvons relier à la démocratisation des différentes capteurs individuels (*smartphone*, GPS d'aide à la navigation…). C'est pourquoi les capteurs fixes s'avancent tout de même comme une source d'information fiable et intéressante pour l'étude du trafic routier.

En effet, d'un côté leurs caractéristiques se rapprochent des capteurs mobiles, et plus particulièrement lorsque l'on se focalise sur les incertitudes relatives à la validité du jeu de données à disposition (erreurs de mesure, d'enregistrement, de temporalité…). Par exemple, les réseaux de capteurs fixes peuvent présenter des ruptures de mesure rattachées à un dysfonctionnement des boucles électromagnétiques ou à des modifications structurelles du réseau de comptage. Mais, d'un autre côté, comparativement aux capteurs mobiles, ils présentent d'autres atouts. Au premier plan, les capteurs fixes présentent un avantage considérable par leur mesure qui est unique, évitant ainsi toute ambiguïté sur la mesure à laquelle ils se rattachent et qui fait leur spécificité. Par exemple, le capteur électromagnétique est un capteur fixe dont les composants technologiques visent exclusivement à comptabiliser le nombre de véhicules ; il ne peut avoir vocation à d'autres types de mesure que celle du trafic routier. Ce qui est différent lorsque l'on utilise les capteurs mobiles qui peuvent présenter un caractère vague, voire approximatif, si on n'apporte pas de précision. Julie Fen-Chong (2012, p.86) illustre ces propos lors du traitement des données de téléphonie mobile, qui par l'apport de connaissances extérieures au système de mesure permet « *d'inférer d'avantage de sens autour des données* ».

En outre, les capteurs fixes apportent une information plus exhaustive autour de leur environnement (leur positionnement spatial). Dans ce cadre, les utilisateurs n'ont pas besoin d'être équipés et la couverture de l'événement mesuré peut être considérée comme totale. Cette couverture s'explique, en outre, par la propriété de ces systèmes de mesures. À l'inverse des

capteurs mobiles qui relèvent plutôt du domaine privé[37], les systèmes de capteurs fixes sont majoritairement du domaine public en charge de la gestion des voiries[38] (Communes, Intercommunalités, Départements et Directions Interdépartementales des Routes : DIR). Ces derniers propos permettent de faciliter l'accès aux données de trafic routier moins délicates à obtenir que les données issues des capteurs mobiles[39]. Cette tendance devrait d'autant plus s'accroître à travers les *smart cities* (Batty, 2013) qui visent à centraliser l'ensemble des informations territoriales issues de la diversité des capteurs présents dans l'espace urbain (*cf.* section 2.1.3, p. 61), mais également en s'insérant dans un contexte plus général, comme par exemple avec la Directive INSPIRE relative à l'établissement d'une infrastructure d'information géographique. Néanmoins nous pouvons avec un certain regret souligner les difficultés à obtenir des données lorsque l'on se tourne vers des services privés. Les données ne sont que rarement gratuites et nécessitent la mise en place de conventions et de financements pour pouvoir les exploiter. Pourtant, l'acquisition et la manipulation des données est une véritable opportunité pour les organismes les possédant. Ouvrir l'accès aux données permet d'entrevoir des perspectives de développement non-négligeables ; à l'inverse les garder pour soi ne laisse que peu de chances de les transformer en un service. Par exemple, Google propose de multiples applications issues de l'usage des données individuelles dans des domaines relativement surprenants comme *Google Flu* pour le suivi de l'évolution de la grippe ou de la dengue dans le monde[40].

[37] Par exemple, pour les données issues des GPS, le service de géolocalisation TomToM propose un accès à ces données. Pour les données de téléphonie mobile, la propriété des données est la possession des opérateurs de téléphonie mobile (Orange, Bouygues, SFR et Free).

[38] On peut toutefois relever le cas plus spécifique des autoroutes concédées qui relève du domaine semi-public.

[39] L'indépendance du secteur public par rapport aux entreprises privées reste un argument important pour assurer une bonne gestion des infrastructures. Par exemple, la communication des données issues de la SNCF sur ses données commerciales, comme l'opacité totale des données sur les consommations d'eau et d'énergie montre bien les limites du tout privé face au secteur public.

[40] Le service n'est actuellement plus visualisable depuis Septembre 2015. Les données recueillies restent cependant disponibles : https://www.google.org/flutrends/about/. À noter que les données géolocalisées sont le résultat des requêtes de santé faites à partir du moteur de recherche Google, dont la fiabilité a été souvent remise en cause.

## Conclusion :

L'observation des mouvements urbains par les capteurs urbains permet d'avoir accès à une autre connaissance du territoire (Aguiton *et al.*, 2009). Elle revêt un intérêt fort pour l'étude d'un objet inscrit dans l'espace et se déroulant dans le temps. Se situant au cœur de la mesure du phénomène en étant localisée et estampillée dans le temps, c'est aussi son caractère exhaustif et continu qui en fait un système de mesure d'autant plus intéressant (Servigne *et al.*, 2009). L'usage des différentes technologies de capteurs (*cf.* figure 1.12, p. 59) constitue des nouvelles sources de données fécondes pour la représentation des territoires et des populations en mouvement.

Le bilan présenté en tableau 1.4 (ci-dessous) sur les systèmes de mesure par comptage tend à montrer que les données issues des systèmes de mesures fixes fondent un dispositif de quantification adéquat et encore d'actualité pour une étude dynamique du trafic routier. Ils sont explicitement tournés vers le comptage des véhicules routiers et se fondent sur des caractéristiques spatio-temporelles originales. Structurés en réseau, ces capteurs offrent l'opportunité d'être au cœur du phénomène qu'ils mesurent permettant de fournir une image du trafic routier tout au long de la journée. De plus, ce chapitre a permis d'introduire les principaux éléments des dispositifs de quantification du trafic routier. Tout d'abord, nous avons défini le trafic routier permettant de définir l'objet d'étude de notre travail. La démarche que nous proposons d'entreprendre se structure autour d'un réseau de capteurs fixes permettant de relever le débit de véhicules à intervalles réguliers. Ces données issues de capteurs électromagnétiques s'inscrivent au même titre que les capteurs mobiles et individuels dans le champ des *big data* et *des smart cities*. Elles présentent un caractère en temps réel et en temps différé, permettant le recueil historiques des débits, cette historicité facilitant en outre sa collecte. La typologie proposée a permis de les distinguer des données individuelles sur plusieurs critères résumés dans le tableau ci-dessous :

| | Capteurs mobiles | Capteurs fixes |
|---|---|---|
| **Collecte** | Automatique | Automatique |
| **Cible mesurée** | Individus équipés et utilisant le service | Tous les utilisateurs du service |
| **Echelle d'observation** | Individuelle (Trace) | Collective (Infrastructure) |
| **Temporalité** | Temps réel | Temps quasi-réel et Historicité |
| **Type de mesure** | Multiple | Unique |
| **Centralisation** | Privée | Public et semi-privée |

**Tableau 1.4 : Les caractéristiques des capteurs fixes et mobiles pour la mesure du trafic routier en milieu urbain**

Le tableau 1.4 permet d'observer que les capteurs fixes constituent une autre réponse à l'utilisation des capteurs mobiles pour notre travail. Si l'usage des données individuelles pose encore de multiples questions autant sur le plan méthodologique, qu'éthique ou juridique, le développement de méthodes de mesures « mixtes » comme les Virtual Trip Line (Hoh *et al.*, 2008) montrent l'intérêt de se rattacher à des réseaux de capteurs fixes plutôt que mobiles. Cette méthode utilise les données de capteurs individuels alors que la collecte des données est rattachée à des mesures fixes. Elle se base sur l'implantation de comptages virtuels (*Virtual Trip Line : VTL*) placés en des points stratégiques sur le réseau routier (*cf.* section 9.1.3, p. 308). Ces méthodes visent à apporter une autre réponse que les méthodes d'agrégation pour contourner les

contraintes éthiques reliées aux données individuelles (Fen-Chong, 2012). Le concept des VTL a fait l'objet d'un développement opérationnel dans le cadre du projet de recherche « *Mobile Millenium* [41]», afin de générer un dispositif de mesure du trafic routier basé sur la géolocalisation GPS des téléphones mobiles. Avec cette méthodologie, la mesure du trafic se fait au niveau de l'infrastructure et se rapproche des méthodes de comptage classiques du trafic routier utilisées dans ce travail. Elle permet aussi d'illustrer le positionnement de notre travail. En effet, l'utilisation des capteurs fixes apparait être une autre réponse aux enjeux liés à la manipulation des données individuelles. De surcroît, elle permet aussi de remettre au goût du jour l'utilisation des comptages électromagnétiques pour l'étude du trafic routier dont nous aborderons plus spécifiquement les enjeux et les traitements au cours du Chapitre 3, p. 81.

[41] Site du projet Mobile Millenium : http://traffic.berkeley.edu/.

# Chapitre 3. DU COMPTAGE ROUTIER AU DISPOSITIF DE QUANTIFICATION DU TRAFIC ROUTIER : LE CAS DE LA VILLE DE DIJON

La mobilisation des données de capteurs, fixes, dans le domaine des transports (urbains ou non) s'inscrit dans les Systèmes de Transports Intelligents (STI). Les STI se définissent comme « *l'application des technologies de l'information et de la communication dans le domaine des transports* » (Faye *et al.*, 2012, p. 2). Sous cette définition relativement vague, on rassemble l'ensemble des moyens de type TIC (Technologies de l'Information et de la Communication)[42] mis en place pour la régulation du trafic routier et les études liées à la circulation, à la sécurité ou à la pollution atmosphérique. L'utilisation des STI permet au gestionnaire de trafic routier de bénéficier d'outils de mesure dont les objectifs sont multiples (Faye *et al.*, 2012). Par exemple, ils sont utilisés pour surveiller et fluidifier le trafic routier en milieu urbain, lors de la mise en place des plans de feux tricolores, ou de politiques de stationnement… En milieu urbain, les réseaux routiers sont gérés par des Centres d'Ingénierie et de Gestion du Trafic (CIGT) chargés d'élaborer et de mettre en œuvre la stratégie d'exploitation de la route. Le rôle de ces structures est de trois natures [43] :

- **Gérer le trafic** de la manière la plus fluide possible par une bonne gestion des carrefours importants de la ville. Cette gestion passe principalement par l'application de cycles de feux tricolores sur les intersections.
- **Informer les usagers** de l'état du trafic et des conditions de circulation au cours de la journée. Cette compétence couvre aussi la gestion des parkings publics, et permet par exemple, aux usagers de connaitre le nombre de stationnements encore disponibles.
- **Maintenir les routes dans un bon état de viabilité** en limitant toutes les perturbations comme lors de la viabilité hivernale avec des actions de dégagement des chaussées ou lors de la mise en place de déviations pendant des périodes de travaux de voirie.

Sébastien Faye *et al.* (2012) identifient deux grandes catégories d'application des STI : ceux contribuant pleinement et directement au domaine de la régulation du trafic routier (usage opérationnel) et ceux contribuant plus indirectement et sans y faire référence à ce domaine, comme lors de la modélisation ou de la simulation du trafic routier (usage de recherche). Le travail présenté s'inscrit dans cette seconde catégorie. Pour les deux catégories, la connaissance des volumes de trafic est un élément essentiel pour alimenter, d'une part les systèmes d'aide à la gestion du trafic, et d'autre part, pour apporter des données de calibrage ou de validation pour la modélisation du trafic routier (Dupuy, 1975 ; Bonnel, 2002). Les comptages routiers se situent donc à l'interface de ces deux grandes catégories.

Dans leur vocation première, les capteurs fixes placés au cœur du réseau routier dijonnais ont pour objectif de fournir des valeurs de trafic pour la mise en place de plans de circulation. Dans un second lieu, ces capteurs peuvent être des sources de données précieuses pour la

[42] Pour le terme de TIC, nous utiliserons les définitions faites par Galit Cohen *et al.* (2002, p. 34) « *l'ensemble des technologies et des applications qui permettent le traitement électronique, le stockage et le transfert de l'information à des tiers* » et Georges A. Giannopoulos ( 2004, p. 302) : « *les TIC sont utilisées pour délimiter les différentes technologies de l'information qui ont été utilisées dans le domaine des transports depuis les années 80. Elles comprennent un grand nombre de technologies et de systèmes développés dans le domaine des transports* ».

[43] CEREMA « *Exploitation routière : gestion du trafic, information routière, maintien en état de viabilité de la route* » : http://www.ouest.cerema.fr/exploitation-routiere-r17.html.

représentation et la visualisation des conditions de circulation sur le réseau urbain. Néanmoins, en l'état ces données ne peuvent apporter qu'une information ponctuelle des conditions de trafic. Ces systèmes de mesure ont pour principale vocation de surveiller le trafic routier pour optimiser la gestion globale des infrastructures de transport routier. Comme nous l'avons vu précédemment (*cf.* section 1.2.1, p. 36), il convient donc de les réagencer et de les structurer pour en extraire une information viable pour d'autres applications en les détournant de leur vocation première. Ce glissement, progressif, d'un usage opérationnel vers un usage de stratégique ou de planification du trafic routier constitue l'objet de ce chapitre, en vue d'introduire le travail présenté à la suite. Les données issues des capteurs fixes de trafic routier constituent le principal jeu de données mobilisé au cours de cette thèse.

Nous avons vu dans le chapitre précédent que l'*objet* de la mesure de ces capteurs est centré sur la route et non pas sur le véhicule. Ainsi constitués, ils permettent de connaître en un point du réseau la fréquentation du tronçon observé sous la forme d'un débit : nombre de véhicules pour une tranche de temps donné. Formés en un réseau plus ou moins dense, les capteurs fixes permettent d'obtenir une photographie relativement fidèle de la circulation routière en milieu urbain. L'état des lieux que nous avons fait *a posteriori* a permis de montrer les avantages de ces capteurs fixes par rapport aux capteurs mobiles. Il convient alors d'en explorer le potentiel d'exploitation comme *dispositif de quantification* : dans quels champs d'application sont-ils mis en place ? Que nous révèlent les données issues de ces capteurs ? Que mesurent-ils ? Comment les envisager pour reconstituer les dynamiques urbaines ?

## 3.1. Un rôle pivot des données de comptages routier pour la régulation de la circulation routière en milieu urbain

Les STI sont principalement mobilisés lorsque des mesures de gestion dynamique du trafic routier sont mises en place. Fournissant des informations au niveau des CIGT, ils permettent de transmettre en temps réel des informations sur les conditions de trafic routier, et de stocker et archiver les historiques de trafic mesurés. Le recueil de données et la mesure du trafic routier jouent ici un rôle essentiel pour l'exploitation des routes. Il permet au gestionnaire d'avoir une connaissance fine des conditions de circulation sur son réseau routier à un moment donné. Pour cela, le gestionnaire a besoin d'obtenir des données dynamiques de comptage (Bréheret *et al.*, 2000). Ces données sont essentiellement fournies par les postes de mesures (PM) implantés sur le réseau routier. Placés en des lieux stratégiques, ces capteurs fournissent des informations sur les conditions de circulation sous la forme de comptages permettant ainsi aux acteurs compétents de mettre en place des actions sur le court et moyen termes pour optimiser le passage des flux de trafic routier. Avant de s'intéresser plus spécifiquement aux données de comptage mises en place sur le réseau routier dijonnais, il est important d'éclairer les principes de fonctionnement généraux des boucles électromagnétiques pour la quantification et la gestion du trafic routier en ville.

### 3.1.1. Les capteurs électromagnétiques : fonctionnement et nature des mesures pour le suivi du trafic routier

Lors de la partie 2.3, nous avons positionné les capteurs électromagnétiques en tant que système de comptage à capteur fixe, en les distinguant des autres capteurs que l'on peut rencontrer dans l'espace urbain. Ils représentent une grande part des technologies employées pour le suivi du trafic routier sur les carrefours à feux ainsi que sur les routes les plus fréquentées (Borsu, 2004 ; Buisson et Lesort, 2010). Les boucles électromagnétiques sont constituées de boucles inductives (bobines de cuivre) placées sous la chaussée à une profondeur de 10 cm. En étant alimenté par un courant électrique, le capteur fonctionne sur un principe d'auto-induction au sein d'un circuit fermé (Borsu, 2004). En d'autres termes, ce sont les perturbations et la variation du courant électrique qui vont permettre de détecter la présence des véhicules sur la chaussée par l'observation d'une oscillation. La liaison de la boucle électromagnétique à un boitier de détection permet non seulement la conversion du signal analogique en un signal électrique mais aussi la transmission des données centralisées vers un Système de Régulation du Trafic (SRT).

Le fonctionnement des boucles électromagnétiques suit donc deux phases illustrées en figure ci-dessous. La première phase dite « analogique » concerne la détection du véhicule ; elle se caractérise par la détection d'un signal analogique des essieux passant au-dessus de la boucle électromagnétique pendant une tranche de temps. La seconde phase dite « digitale » permet le traitement du signal analogique vers le comptage de véhicules en un temps de présence. Enfin le signal digital est transmis et centralisé à des intervalles réguliers (de la minute à l'heure) vers un poste central au sein du SRT sous la forme d'un nombre de véhicules/tranche de temps : le débit. Il existe plusieurs types de capteurs électromagnétiques, illustrés en entête du tableau 1.5. Les plus simples sont constitués d'une seule boucle inductive, alors que d'autres, plus complexes, sont équipés de deux boucles successives, voire d'un capteur de charge.

En fonction du nombre de boucles inductives placées sous la chaussée (*cf.* tableau 1.5, ci-dessous), différentes mesures peuvent être faites en complément du débit de véhicules (Buisson et Lesort, 2010) :

- **Le taux d'occupation** de la chaussée correspond au temps pendant lequel un véhicule est détecté. Cette mesure permet de calculer la concentration de véhicules présents sur le réseau routier.
- **L'intervalle inter-véhiculaire** correspond au temps de détection entre deux véhicules comptés.
- La **vitesse des véhicules** est le temps qu'un véhicule va mettre à traverser deux boucles de comptages successives.
- Les **types de véhicules** et leurs **longueurs** sont le résultat du temps de passage (*cf.* figure 1.20) du véhicule entre deux boucles permettant de distinguer la détection des poids-lourds (temps de passage long), des véhicules légers (temps de passage court).
- La **charge des véhicules** nécessite un équipement supplémentaire mesurant la charge du véhicule passant sur la route.

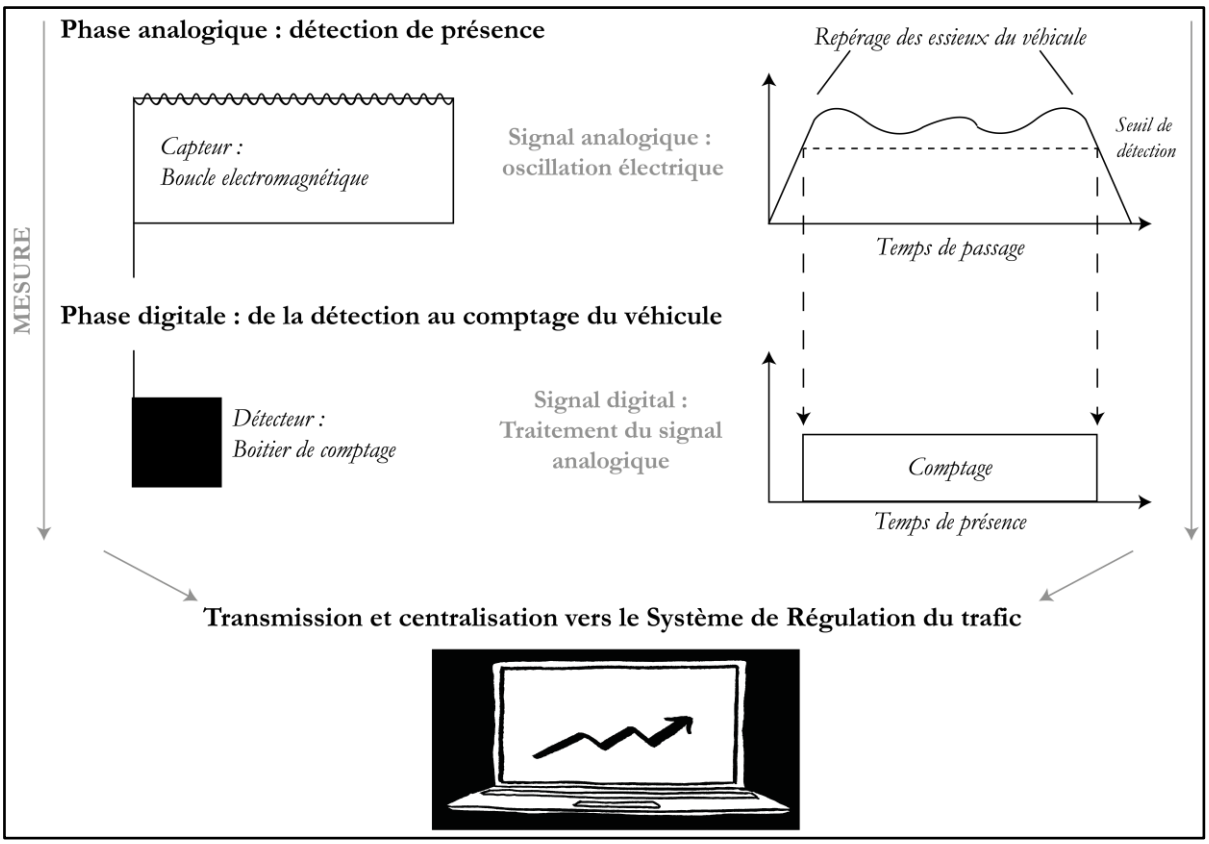
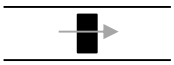
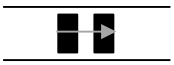
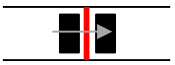



**Figure 1.20 : Principe de fonctionnement d'une boucle électromagnétique pour le comptage du trafic routier (d'après Borsu, 2004)**

Généralement, les capteurs sont constitués d'une seule boucle inductive et permettent de mesurer le débit de véhicules, le taux d'occupation et la distance inter-véhiculaire. Ces capteurs relativement simples permettent de relever les informations primaires pour le trafic routier. Ils sont le plus souvent rencontrés en milieu urbain pour la gestion des carrefours à feux ou pour des besoins ponctuels de mesure du trafic routier. Les autres capteurs, constitués de plusieurs boucles successives, permettent de mesurer les mêmes variables, mais en apportant des informations

complémentaires pour caractériser le flux routier présent comme la dimension des véhicules, leur type ou leur vitesse de déplacement. Ces capteurs sont le plus habituellement présents sur les axes à fort trafic, autoroutes et routes nationales. À l'inverse des capteurs simples, ceux-ci sont plutôt utilisés pour le suivi et la mesure du trafic routier. L'ensemble des mesures de trafic par capteurs électromagnétiques est synthétisé au sein du tableau ci-dessous :

| Nature des mesures |  Boucle inductive simple |  Boucle inductive double |  Boucle inductive double + capteur de charge |
|---|---|---|---|
| **Débit tous véhicules** | X | X | X |
| **Taux d'occupation** | X | X | X |
| **Distance inter-véhicule** | X | X | X |
| **Débit par type de véhicule (PL et VL*)** | | X | X |
| **Longueur** | | X | X |
| **Vitesse** | | X | X |
| **Charge** | | | X |

* PL : Poids lourds / VL : Véhicules légers

**Tableau 1.5 : La nature des mesures par type de capteurs électromagnétiques (d'après Buisson et Lesort, 2010)**

À partir des données issues des systèmes de capteurs électromagnétiques, différents indicateurs sont établis pour caractériser le trafic routier en milieu urbain, notamment le Trafic Moyen Journalier Annuel (TMJA) [44] correspondant à l'unité de mesure la plus souvent rencontrée lors des études de trafic. Plus généralement, les capteurs fournissent des informations essentielles au régulateur de trafic pour établir des stratégies de gestion de la circulation routière urbaine, qui passe principalement par la gestion des carrefours à feux tricolores (CEREMA, 2014). Formés en réseau, ils permettent aussi de suivre l'évolution de la circulation routière autant sur le court que le long terme. D'une manière générale, le capteur électromagnétique joue un rôle central pour estimer avec précision le débit de trafic observé en vue de gérer et réguler la circulation routière à une échelle locale.

### 3.1.2. La gestion et la régulation de la circulation routière en milieu urbain : le rôle central des comptages routiers

En milieu urbain, la régulation du trafic routier passe principalement par la gestion des carrefours équipés de signalisation tricolore. La gestion des carrefours à feux repose sur une gestion temporelle des flux pour permettre l'écoulement du trafic routier dans des conditions optimales (CEREMA, 2014). En effet, ces « portes d'entrée » du réseau routier jouent un rôle essentiel dans la régulation des flux de véhicules passant sur les différents axes d'une agglomération. Lorsque les flux de trafic routier sont soumis à d'importantes variations dues au rôle des carrefours tricolores, on parle le plus habituellement de « *trafic pulsé* ». En reprenant la métaphore du métabolisme

[44] Le TMJA correspond à l'unité de mesure synthétique la plus utilisée pour caractériser le trafic routier depuis les débits comptés. Cette unité de mesure correspond à la somme du débit journalier observé sur une année divisé par 365 et s'exprime en UVP (Unité de véhicule particulier) par jour.

urbain (Barles, 2008), on pourrait considérer le trafic routier comme un flux sanguin permettant d'alimenter nos organes en ressources vers différents lieux d'activités (domicile, travail, loisirs…). Les feux tricolores, quant à eux, jouent un rôle comparable à des valves cardiaques permettant de rythmer et d'orienter les flux de trafic routier dans le bon sens. Ainsi envisagés, les carrefours jouent un rôle essentiel en régulant les flux de trafic entrants et sortants au niveau des carrefours.

La régulation du trafic est approchée par une fluidification du trafic avec une alternance de cycle de feux (du vert au rouge) pour minimiser au maximum les attentes. Toutefois, la régulation passe aussi par la mise en place de cycles plus contraignants, par exemple, en augmentant volontairement la durée des feux rouge afin de ralentir les accès aux axes les plus fréquentés pour éviter des congestions trop importantes. Le fonctionnement d'un carrefour repose alors sur un système cyclique qui se définit par un temps donné où les différentes phases d'un feu tricolore (du vert au rouge) vont s'enchaîner sur les différents axes. L'ensemble de ces éléments se formalise sous la forme d'un plan de feux illustré en figure 1.21 correspondant « *à la description d'un cycle donné, et défini par les différentes phases que le cycle va dérouler ainsi que leur durée* » (Faye *et al.*, 2012, p. 9) :

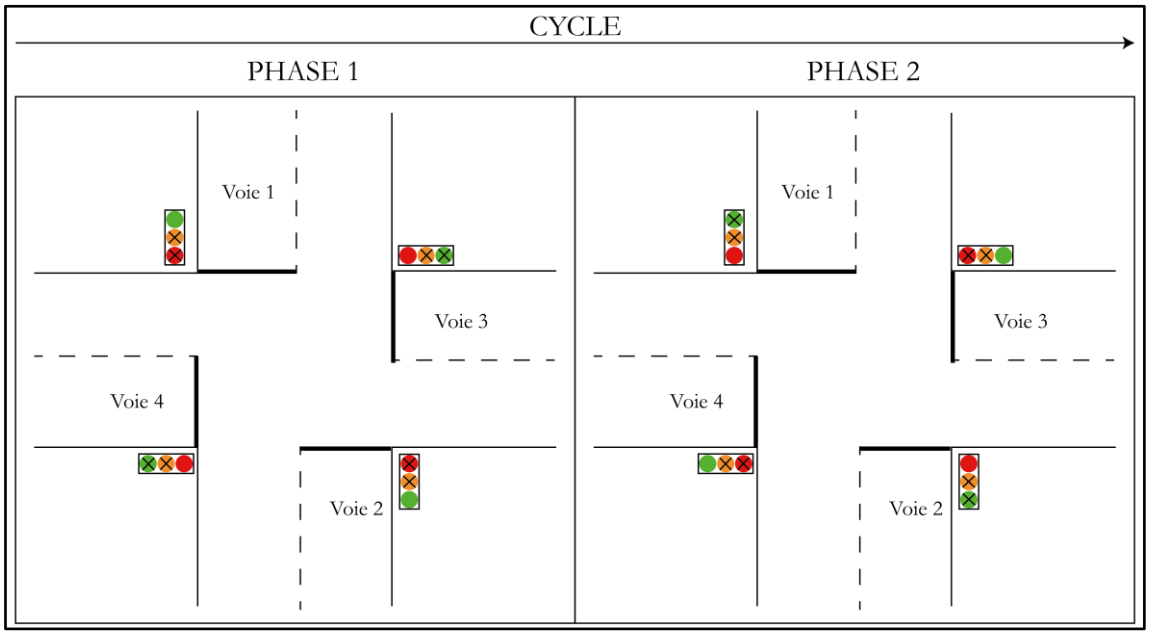



**Figure 1.21 : Exemple type d'un plan de feux en un cycle de deux phases sur un carrefour en milieu urbain (d'après Faye *et al.*, 2012)**

La stratégie de régulation est définie individuellement par l'exploitant en charge du réseau routier en fonction des contextes locaux, d'où le besoin de comptage routier. Ces séquences sont toutefois soumises à une réglementation définissant les temps minimal de passage entre les différentes phases (*Article LA11-6 Instruction interministérielle sur la signalisation routière*, 1991) : la durée minimale du vert et du rouge est de 6 secondes ; la durée maximale d'attente au feu rouge est de 120 secondes. En respectant ces principes, un recueil de l'ensemble des pratiques de gestion du trafic routier urbain en France produit par le CEREMA (2014) montre que l'ensemble des agglomérations utilisent un logiciel d'aide à la gestion de feux. L'exploitation de ces systèmes nécessite de s'adapter rapidement aux conditions de circulation observées et s'opère sur des temporalités très courtes, d'où le besoin d'automatiser la gestion de ces flux à l'aide d'outils informatiques et des comptages routiers.

Pour la gestion automatique, deux types de régulation peuvent être distingués (Dupuy, 1993) :

- **Une régulation locale** : dans cette situation les feux tricolores sont commandés par des capteurs détectant le trafic sur les axes arrivant au carrefour. La gestion est considérée comme autonome lorsqu'un véhicule arrivant est détecté, le passage au vert se fait automatiquement.
- **Une régulation globale** : l'ensemble des feux tricolores est géré par un système central (PC de circulation) collectant l'information sur les différents carrefours afin d'optimiser la répartition des trafics par la mise en place de cycle de feux dont l'un des plus connus est l'onde verte (*cf.* encart 2, p. 88). Le fonctionnement de ces systèmes est relié à des boucles de comptage permettant une gestion par carrefour ou alors sur un ensemble de carrefours. Le pilotage de l'ensemble de ces informations (cycle de feux et comptage automatique) est centralisé au sein d'un SRT.

Il faut noter que la régulation du trafic routier peut être combinée : dans le cas de la ville de Dijon, les cycles de feu sont gérés localement lors des périodes de faibles fréquentations, principalement la nuit. À l'inverse, lors des périodes les plus fréquentées, la régulation est gérée à un niveau global permettant d'optimiser le trafic routier s'insérant sur les différents carrefours urbains. La gestion des carrefours à feux repose ainsi sur une approche capacitaire des flux routiers (CEREMA, 2014). En mobilisant les données de comptages routiers, la régulation du trafic (locale ou globale) permet de s'adapter aux conditions locales de manière dynamique. Au sein des SRT, le capteur électromagnétique joue un rôle central pour estimer avec précision le débit de trafic en vue d'établir des cycles de feux adaptés au contexte observé. Pour cela, le régulateur de trafic utilise des données de comptages routiers afin de calibrer ceux-ci. Afin d'adapter les cycles de feux au contexte local, la majorité des carrefours à feux sont équipés d'un capteur fixe afin d'ajuster les cycles de feux au débit maximal observé sur les axes routiers (CEREMA, 2014).

La particularité des SRT se situe dans leur besoin de réactivité et de dynamisme, afin de prendre des décisions en conséquence. Néanmoins, ce degré de réactivité est à relier aux systèmes de contrôle utilisé par le régulateur. Trois catégories de systèmes de contrôle se distinguent en fonction du degré d'intégration entre le système de contrôle et les boucles électromagnétiques (Faye *et al.*, 2012) :

- **Les systèmes de contrôle à temps fixe** : Un plan de feux est prédéfini en fonction des horaires de la journée identifiés par l'agent de régulation. Ces systèmes constituent la première génération des systèmes de gestion. Statiques, ils ne permettent de s'adapter au contexte local que par une action humaine.
- **Les systèmes de contrôle à temps dynamique** : De seconde génération, ces systèmes sont couplés à des capteurs de trafic routier recueillant les données à des intervalles réguliers. Les données sont ensuite employées pour optimiser, ou mettre en place, des plans de feux préétablis.
- **Les systèmes de contrôle à temps réel** : Correspond au même système que le précédent avec une gestion et une adaptation en temps réel des plans de feux. L'ensemble des plans de feux est calculé et modifié par le système central.

Le premier système fonctionne sur un principe de découplage à l'inverse des deux derniers

systèmes de contrôle qui eux sont directement couplés aux comptages routiers. Les systèmes de contrôle à temps dynamique et à temps réel permettent d'introduire un caractère dynamique dans la gestion du trafic routier en milieu urbain. Ils peuvent être décomposés en deux types (Faye *et al.*, 2012) : soit le contrôle est réactif aux informations fournies par les comptages routiers ; soit il est adaptatif en temps réel par rapport aux données reçues au niveau du SRT.

Ce contrôle des carrefours à feux nécessite l'usage de données de comptage localisées aux niveaux des intersections avec un taux de recueil relativement fin pour pouvoir réagir ou s'adapter aux flux de circulation routière relevés. Formée en réseau ces données issues en très grande partie des technologies de capteurs électromagnétiques permettent de répondre aux besoins du gestionnaire de trafic sous un angle spatio-temporel pour optimiser le cycle de feux tricolores (CEREMA, 2014).

***Encart 2 L'exemple des ondes vertes : une régulation globale de la circulation urbaine***

Les ondes vertes permettent de synchroniser un ensemble de feux sur une même section afin de fluidifier le trafic routier. La synchronisation de la signalisation permet de propager progressivement un feu vert de l'amont vers l'aval d'un axe routier. Principalement appliquée aux voies à sens unique (Héran, 2002), cette technique de régulation du trafic permet de réduire les vitesses de circulation sur un axe (de 50 km/h à 30 km/h par exemple).

En minimisant le nombre d'arrêts sur une même section, l'application des ondes vertes permet de parcourir un même tronçon de route à une vitesse modérée. Le principe des ondes vertes est illustré en figure ci-dessous pour un temps de parcours de 1 minute :

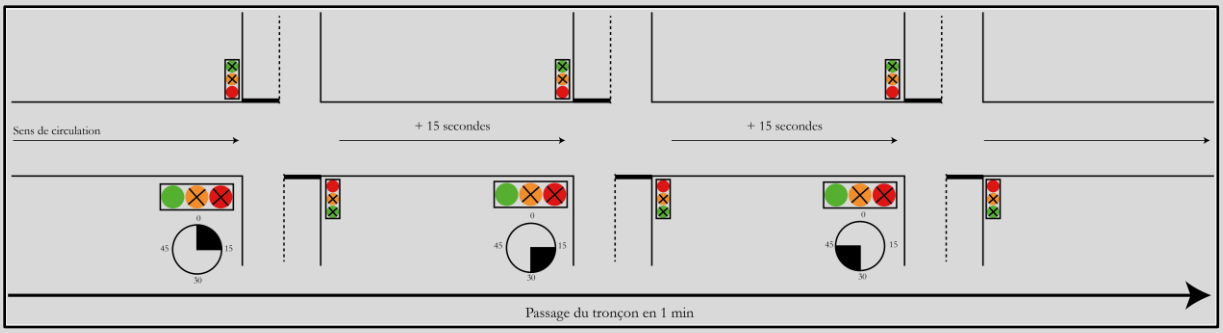



**Figure 1.22 : Principe de fonctionnement et de synchronisation d'une onde verte sur un tronçon routier à sens unique**

Le pilotage des ondes vertes se fait grâce à des logiciels d'aide à la régulation du trafic comme celui proposé par le CEREMA : le logiciel ONDE V[45]. La mise en application des ondes vertes permet aussi de répondre à des objectifs environnementaux comme la pollution atmosphérique ou le bruit. En effet, en diminuant les vitesses de circulation, les décélérations et les accélérations, on réduit sensiblement la consommation de carburant et donc les émissions de polluants atmosphériques.

[45] Guide descriptif du logiciel ONDE V : http://www.certu-catalogue.fr/ondev-logiciel-version-1-0-1999.html.

## 3.2. Présentation du réseau de mesure de la ville de Dijon

Le dispositif de quantification que nous mobilisons pour l'étude du trafic routier doit répondre aux différents critères que nous avons mis en avant au cours des sections précédentes. Tout d'abord, il doit permettre d'observer le trafic routier autant dans l'espace que dans le temps. La granularité temporelle est une des conditions nécessaires pour approcher les PAA de manière adéquate (*cf.* section 1.1.3, p. 31). Dès lors, il est essentiel de disposer d'un réseau de capteurs automatiques relevant le trafic routier de manière continue ou quasi-continue. Dans ce sens, le réseau de mesures mis en place par la cellule régulation de la circulation apparait intéressant à exploiter, en particulier les données issues des stations permanentes. En outre, il repose sur un système de capteurs fixes constitué de capteurs électromagnétiques dont nous avons énoncé les avantages au cours du précédent chapitre. L'objet de cette section est de présenter le dispositif de quantification utilisé par la commune de Dijon en précisant les différentes étapes de traitement réalisées lors de la constitution du dispositif de quantification employé au cours de ce travail.

### 3.2.1. **La gestion du trafic routier de la ville de Dijon**

#### a. ***La cellule régulation de la circulation de la ville de Dijon***

Le trafic routier dijonnais est géré depuis 1978 au niveau de la commune par l'intermédiaire de la cellule « régulation de la circulation » de la ville de Dijon[46]. Ce CIGT est constitué de 4 agents dédiés totalement à la cellule de régulation et dont les champs d'intervention sont de trois ordres : la régulation, le contrôle des accès des voies réservées (zones piétonnes, accès aux parkings,…) et la vidéosurveillance. Le service est plus particulièrement en charge de la surveillance et de la régulation de la circulation routière sur plus de 200 carrefours à feux.

Pour réaliser ses missions, la cellule dispose d'un Poste de Commande (PC) géré par l'intermédiaire d'un Système de Régulation du Trafic (SRT) centralisant les différents outils nécessaires à la gestion du trafic routier :

- Le **SRT** en lui-même et ses fonctionnalités automatisées se structurent autour d'un système de *contrôle à temps réel* sous la forme d'un logiciel centralisé de gestion de trafic et d'information des voyageurs (SERELEC[47])
- Les **différents équipements de communication** avec le SRT : les caméras de surveillance, les postes de mesures (PM) du trafic routier, les carrefours tricolores, la signalétique à destination des usagers, …
- Un **système d'alerte** en cas de dysfonctionnement ou de problème de l'un des équipements de communication, reporté sur un des écrans de contrôle central comme illustré en figure 1.23.

[46] Depuis 2015, la compétence de la cellule régulation de la circulation routière est rattachée à la communauté urbaine du Grand Dijon. Cette restructuration étant récente, seul le responsable de la cellule de régulation de la circulation est pour le moment rattaché à la communauté urbaine. La restructuration, à terme, vise à réorganiser et à mutualiser l'ensemble des dispositifs de surveillance, vidéosurveillance et gestion du trafic routier à l'échelle du Grand Dijon sous une même cellule mutualisée S3C (Sécurité Civile, de Circulation et de Coordination).

[47] Il existe deux grands logiciels pour la gestion automatique des feux qui se partagent l'ensemble du marché français (CEREMA, 2014). Le plus utilisé est le logiciel GERTRUDE (http://www.gertrude.fr/) mobilisé, par exemple, dans les villes de Bordeaux, Reims ou Montpellier ; celui utilisé par la ville de Dijon est SERELEC (http://www.serfim.com/activites/tic/).




**Figure 1.23 : Illustration du PC de la circulation routière de la ville de Dijon (au centre, l'écran de contrôle à diodes des carrefours à feux ; de chaque côté de l'écran de contrôle des écrans numériques permettent de visualiser l'état du trafic routier en direct par l'intermédiaire de caméras)**

L'ensemble de ces outils permet au CIGT dijonnais de gérer les flux de circulation routière en temps réel et en fonction du contexte local. La stratégie de régulation du trafic routier dijonnais vise à mettre en place « *des cycles de feux les plus courts possibles, mais aussi les plus simples* » (CEREMA, 2014, p. 56), tout en évitant au maximum les trafics de transit passant par le centre-ville de Dijon. Cette stratégie générale se décline en deux stratégies particulières. La journée, l'ensemble des carrefours est géré globalement au niveau du SRT, afin d'assurer une coordination optimale des flux de trafic routier. À l'inverse, la nuit, les carrefours sont gérés localement, c'est-à-dire que les feux sont autonomes et se déclenchent seulement lorsqu'un véhicule est détecté par un capteur électromagnétique, ou lorsqu'il n'y a pas de capteur par un radar de détection situé au-dessus du feu tricolore. La figure ci-dessous illustre ces deux dispositifs de détection autonomes des véhicules pour les feux tricolores sur un même carrefour :




**Figure 1.24 : Illustration des dispositifs de détection autonomes pour les feux tricolores sur un même carrefour entre le Boulevard Gabriel et l'Avenue Alain Savary (à gauche : un radar de détecteur de présence ; à droite : un capteur électromagnétique)**

### b. *Le réseau de comptage routier de la ville de Dijon*

Outre la détection de véhicules pour le déclenchement des feux tricolores et la gestion des cycles de feux, les capteurs électromagnétiques permettent aussi le comptage des véhicules (*cf.* figure 1.20, p. 84). L'ensemble des comptages routiers réalisés par la commune de Dijon est illustré en figure 1.25, il se décline en deux types :

- les **comptages permanents** fournissant des relevés de trafic régulier par ¼ heure ;
- les **comptages ponctuels** fournissant des données de trafic sous la forme de TMJA.

Tout d'abord, pour les comptages permanents, le PC de circulation dispose d'un réseau de 215 capteurs électromagnétiques permanents disposés sur l'ensemble de la commune depuis 2001[48]. Le système de comptage est constitué de boucles inductives simples (*cf.* tableau 1.5, p. 85) permettant d'obtenir les valeurs de débit de véhicules récoltées à un intervalle régulier de 15 minutes. Du fait de son implantation (sous la chaussée), le réseau de capteurs est considéré comme relativement stable, permettant de le qualifier de permanent. On peut toutefois noter des modifications, voire des suppressions de systèmes de comptage lors de travaux de voirie, comme par exemple lors de la mise en place du tramway dont nous verrons l'impact sur le réseau de mesure dans la section 3.3.3 (p. 114). L'ensemble des mesures permanentes se situe à proximité des grands carrefours à feux de la commune auxquels elles sont initialement dédiées pour la gestion des plans de feux.

En s'appuyant sur la carte présentée en figure 1.25, on peut observer que l'implantation du réseau de mesure permanent (en noir) est principalement orientée sur les axes routiers les plus structurants du réseau routier (liaison départementales et interurbaines). Son implantation suit d'ailleurs la structure du réseau routier dijonnais qui est organisé sous la forme d'un réseau radioconcentrique qui est « *une forme relativement répandue des réseaux urbains. À partir d'une certaine distance les 6 ou n grandes voies radiales sont reliées par des rocades ou par des boulevards de ceinture périphérique* […] ; *ce dessin est rationnel, mais est souvent critiqué car il conserve, accroit et exagère la fréquentation de ces axes amenant à un encombrement de ceux-ci* ». (Brunet *et al.*, 2005, p. 374). En termes d'infrastructures routières, la ville de Dijon présente un réseau relativement dense permettant une bonne desserte de l'espace urbain, qui se fait principalement autour de deux ceintures de boulevards périphériques (proche du centre-ville et en périphérie de la ville) desservant des axes routiers moins importants. L'accès à ces boulevards est assuré par différentes artères urbaines permettant une liaison Nord/Sud et une liaison Est/Ouest. Plus de 70 % des mesures de trafic sont réalisées sur les liaisons départementales et interurbaines[49] (*cf.* tableau 1.6, p. 94). Enfin, il est possible de

[48] Année de la numérisation totale de la cellule régulation de la circulation routière, les données antérieures à 2001 sont stockées sous la forme de disquettes dont l'accessibilité est aujourd'hui obsolète.

[49] La hiérarchie du réseau routier suit la nomenclature utilisée par l'IGN (2014) pour la BD TOPO :

- Les *liaisons nationales* sont principalement composées des autoroutes et des quasi-autoroutes et assurent les liaisons au niveau européen
- Les *liaisons départementales* sont les routes nationales qui ont pour fonction d'assurer les liaisons à fort trafic entre les agglomérations et d'assurer la continuité des liaisons interurbaines en agglomération
- Les *liaisons interurbaines* sont majoritairement composées des routes départementales, structurent la circulation en agglomération et assurent les liaisons à l'intérieur d'un même département
- Les *liaisons urbaines* structurent elles-aussi la circulation en agglomération en permettant de se déplacer rapidement à l'intérieur d'une même commune
- Les *liaisons résidentielles* permettent de desservir l'intérieur d'une commune et font le lien entre les différents quartiers résidentiels d'une même commune.

voir sur la figure 1.25 quelques mesures permanentes, situées au Sud, qui sont réalisées en dehors de la commune dijonnaise :

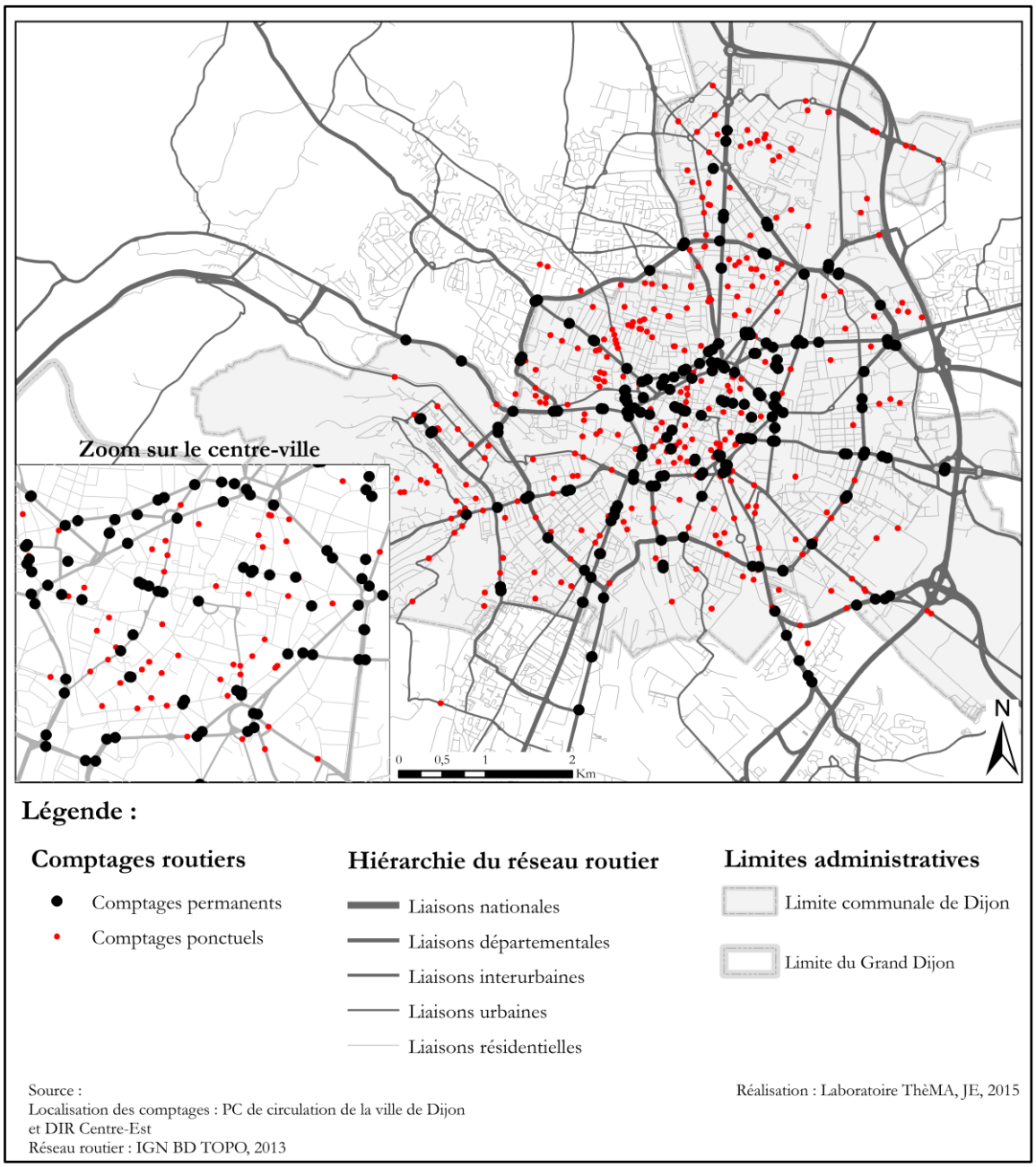



**Figure 1.25 : Le réseau de comptage routier permanent et ponctuel géré par la cellule de régulation de la circulation de la ville de Dijon**

Il est important de resituer le réseau de mesure dijonnais comparativement à plusieurs communes disposant du même système de comptage (c'est-à-dire dédié à la régulation de la circulation routière)[50] : en d'autres termes, le réseau de comptage permanent dijonnais présente-t-il une couverture spatiale significative et conforme à ce que d'autres villes proposent ?

Notons, tout d'abord, que la localisation des comptages routiers pour la régulation du trafic

[50] Les villes sélectionnées sont celles de *Belfort, Besançon, Bordeaux, Dijon, Montpellier, Nantes, Reims, Rouen, Strasbourg, Toulouse.*

routier n'est que rarement disponible (CETE, 2013). C'est pourquoi nous avons fait le choix d'établir un indicateur synthétique, illustré en figure 1.26, tenant compte de la densité de capteurs rapporté à la superficie totale de l'espace communal (nombre de capteur/km²) au détriment d'un indicateur tenant plutôt compte du réseau routier (nombre de capteur/km de réseau). Ce dernier nous aurait permis de confronter plus finement la structure des réseaux de comptage recensés (comme illustré en tableau 1.6). Cependant, si l'indicateur proposé est effectivement discutable, il permet néanmoins d'obtenir un ordre de grandeur afin de comparer des villes de taille très différente de celle de Dijon comme par exemple Bordeaux, Strasbourg ou Toulouse. Pour l'ensemble des villes pour lesquelles nous avons pu obtenir des informations relatives aux comptages routiers locaux[51], la densité moyenne des réseaux de capteur dédiés à la régulation du trafic routier est de 4,4 capteurs/km². Comme illustré ci-contre, le réseau géré par la ville de Dijon présente une couverture de mesure plus importante (5,32 capteurs/km²) que d'autres villes. Il se place d'ailleurs en seconde position de notre recensement, derrière la ville de Belfort (6,02 capteurs/km²). Le réseau de comptage dijonnais est comparable à des villes de même rang comme Rouen ou Nantes, qui présentent une couverture spatiale relativement similaire (respectivement 5,52 et 5,28 capteurs/km²). À travers la synthèse graphique illustrée en figure 1.26, la ville de Dijon, en se situant près de la frontière haute, présente un réseau de comptage permanent relativement mieux doté que des communes plus peuplées comme Toulouse (3,38 capteurs/km²), Montpellier (3,53 capteurs/km²) ou Bordeaux (4,05 capteurs/km²). L'ensemble de ces informations généralistes montre que le terrain d'étude utilisé au cours de cette thèse présente des caractéristiques conformes à ce que nous retrouverions sur d'autres terrains d'études, notamment français, bénéficiant du même dispositif de mesure permanent pour la régulation des feux tricolores. De surcroît, il conforte aussi le caractère généraliste du travail qui sera présenté tout au long de cet écrit.

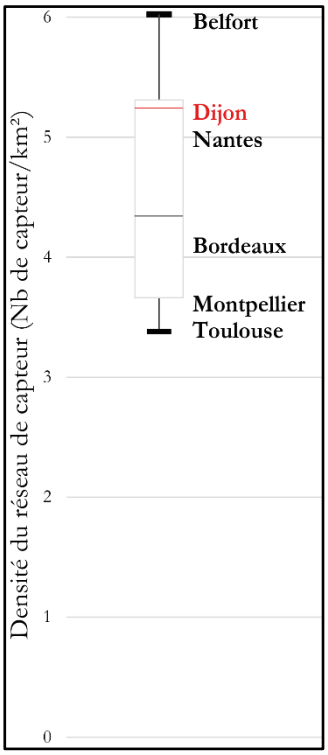



**Figure 1.26 : Répartition des villes en fonction de la densité des réseaux de comptage**

Parallèlement et en complémentarité des mesures permanentes relevées à des périodes régulières, la CIGT dijonnais procède à des comptages ponctuels qui sont représentés en rouge au sein de la figure 1.25. D'une durée plus courte (n'excédant pas les trois semaines), les comptages sont réalisés dans différents cas, par exemple, à la demande de riverains exprimant une gêne occasionnée par la circulation routière. Ces comptages peuvent encore être mis en place pour examiner l'impact d'un nouveau plan de circulation. C'est particulièrement le cas d'actions visant, par exemple, à réduire la circulation routière par la mise en place d'une circulation à sens unique,

[51] Comme il est mentionné dans le rapport du CETE de Lyon (CETE, 2013, p. 12) seules les villes de Bordeaux, Strasbourg, Nantes et Montpellier mettent en accès libre des données relatives à la circulation routière. Pour d'autres villes, nous avons mobilisé d'autres canaux d'information, par exemple pour Rouen à travers le projet de simulation du trafic routier MOSAIIC (Czura *et al.*, 2015), ou pour Toulouse via le projet CADDY (Bréheret *et al.*, 2000 ; Bauzer-Medeiros *et al.*, 2006). Enfin pour d'autres villes nous nous sommes tournés vers les services en charge de la gestion de la circulation routière urbaine : Besançon, Belfort et Reims. L'ensemble des données recensées est présenté en annexe 1 (p. ii).

ou alors, lors d'actions sur la réduction des vitesses de circulation, voire lors de l'installation de ralentisseurs de trafic pour l'application de zone à 30km/h à proximité des établissements scolaires. Depuis 2003, la cellule a procédé à plus de 650 comptages ponctuels sur l'ensemble de la commune de Dijon. Comme on peut l'observer sur la carte ci-dessus, les comptages ponctuels sont plus particulièrement réalisés sur les routes de moindre importance, et principalement axés sur les liaisons résidentielles pour près de 70 % des comptages effectués. Une synthèse des comptages réalisés par type de réseau routier est proposée en tableau 1.6. Les liaisons nationales, absentes du maillage dijonnais, mises à part, les mesures réalisées par les gestionnaires de trafic de la ville de Dijon permettent une couverture spatiale représentative de l'ensemble des types de routes présents sur la commune de Dijon :

| | **Comptages permanents (215 PM)** | **Comptages ponctuels (655 PM)** |
|---|---|---|
| **Liaisons nationales** | 0% | 0% |
| **Liaisons départementales** | 22% | 4% |
| **Liaisons interurbaines** | **50%** | 14% |
| **Liaisons urbaines** | 10% | 14% |
| **Liaisons résidentielles** | 18% | **68%** |

**Tableau 1.6 : Pourcentage des comptages routiers réalisés par la ville de Dijon en fonction de la nature du réseau routier**

### 3.2.2. Les données de comptages recueillies : un réseau de comptage hiérarchisé à l'image du réseau routier

L'ensemble des comptages mobilisés pour ce travail repose exclusivement sur les capteurs électromagnétiques permanents (figure 1.25). Les mesures réalisées se concentrent sur la commune de Dijon. Néanmoins, comme illustré en figure 1.27, la ville ne procède pas à des mesures sur deux axes importants de l'agglomération dijonnaise situés à l'Ouest (A38, illustré en rouge) et à l'Est de la commune (Rocade Georges Pompidou, illustré en orange). Ceci s'explique par la compétence de la commune qui ne gère pas ces types de route, qui sont du domaine de compétence des Directions Interdépartementales des Routes (DIR). Dans notre cas, c'est la DIR Centre-Est[52] qui se charge du suivi du trafic routier sur ces deux axes majeurs de l'agglomération. Ces deux axes sont primordiaux pour la visualisation du trafic routier car ils permettent de relier la commune de Dijon aux principaux réseaux autoroutiers locaux en direction de l'A6 (Paris-Lyon) par l'A38 et en direction de l'A31 (vers le Luxembourg) et de l'A39 (vers la Suisse) pour la Rocade Georges Pompidou. Pour acquérir les données de comptage sur ces deux axes (A38 et Rocade Georges Pompidou), nous nous sommes alors tournés vers la DIR Centre-Est. Il faut toutefois relever que les relevés effectués par la DIR Centre-Est diffèrent par nature de ceux effectués par la ville de Dijon. Ces comptages ont pour vocation de suivre l'évolution du trafic routier et reposent sur des capteurs électromagnétiques à double boucle inductive permettant de distinguer les types de trafic entre les véhicules légers et poids lourds (*cf.* tableau 1.5, p. 85). En outre, la fréquence de recueil des données est à une résolution horaire alors que ceux de la ville de

[52] Site internet de la DIR Centre-Est : http://www.enroute.centre-est.developpement-durable.gouv.fr/.

Dijon s'opèrent à une granularité plus fine (le quart d'heure). L'ensemble du réseau de comptage automatique collecté est présenté en figure ci-dessus :

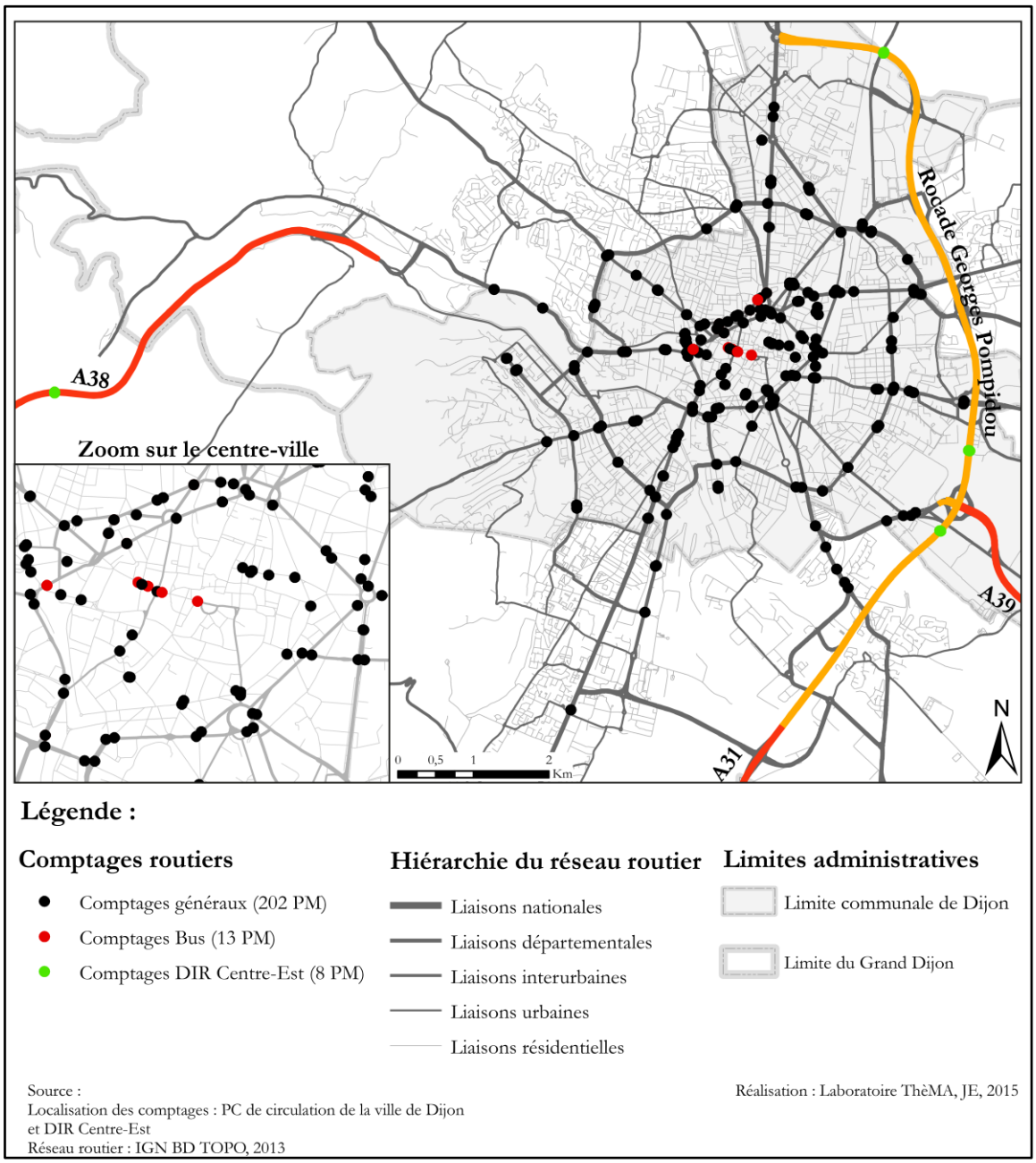



**Figure 1.27 : Localisation des comptages automatiques effectués sur le territoire de l'agglomération dijonnaise**

S'agissant du réseau de mesures effectuées par la ville de Dijon, il se compose d'une part des comptages généraux (202 PM) permettant de relever l'ensemble des flux de trafic passant, et d'autre part il est complété par des comptages dédiés aux bus (13 PM) sur les tronçons routiers exclusivement empruntés par les bus. Ce réseau de mesures, dédié aux comptages des bus, se localise principalement à proximité du centre-ville dijonnais avec 8 PM présents sur ce secteur. En ce qui concerne le réseau géré par la DIR Centre-Est, le réseau se compose de 8 PM toujours réalisés dans les deux sens de circulation : la rocade Georges Pompidou (à l'Est) présente un réseau de mesures plus dense (6 PM) que l'A38 (2 PM) dont le trafic routier est relevé à plus de

10 km de la ville de Dijon. Notre problématique portant sur la représentation du trafic routier, les données relevant exclusivement les passages de bus ne seront pas prises en compte. Cela nous permettra d'évaluer le trafic routier sans distinction des catégories de véhicules. De plus, les bus ne circulant pas exclusivement sur des voiries dédiées, ils peuvent être pris en compte au sein des comptages généraux. De même les stations de comptage sur l'A38 étant situées trop loin de la ville, celles-ci ne seront pas prises en compte lors des prochaines analyses. Au total, nous avons sélectionné 208 stations de mesures permanentes pour les analyses qui suivront (le nombre exact de stations mobilisées sera toujours précisé dans la description des figures). L'ensemble des données de trafic collecté se formalise sous la forme de valeurs de débit de véhicules par quart d'heure.

L'ensemble du réseau de mesures présente une faible emprise sur l'ensemble des tronçons routiers de l'agglomération et de la ville de Dijon : respectivement 5,3 % des tronçons routiers de l'agglomération sont dotés de stations de comptage, et 12,2 % des tronçons de la ville de Dijon sont équipés de capteurs. Si cette emprise du réseau de mesure apparait très faible au premier abord, il faut nuancer ce résultat en s'intéressant à la fréquentation des axes routiers observée sur l'ensemble des stations de comptages permanentes et ponctuelles illustrées en figure 1.25 et en figure 1.27. Afin d'obtenir une information synthétique de la répartition des axes routiers mesurés en fonction de leur fréquentation, ceux-ci ont été regroupés par la méthode des K-means présentée en figure 1.28 (ci-dessous). Cette méthode présente l'avantage de regrouper les observations par clusters (classes d'observations) dans lesquels chaque observation appartient au cluster avec la moyenne de TMJA la plus proche. Le principe des K-means se fait sur la base d'un critère des plus proches voisins en choisissant, au préalable, le nombre de classes à établir se faisant en fonction d'une Classification Ascendante Hiérarchique (CAH). Le graphique présenté ci-dessous offre ainsi l'opportunité d'observer dans une même classe les axes routiers présentant des caractéristiques analogues en termes de TMJA observé sur l'ensemble des axes routiers (828 axes) ayant eu au moins une mesure au cours de la période de comptage allant de 2003 à 2013 :

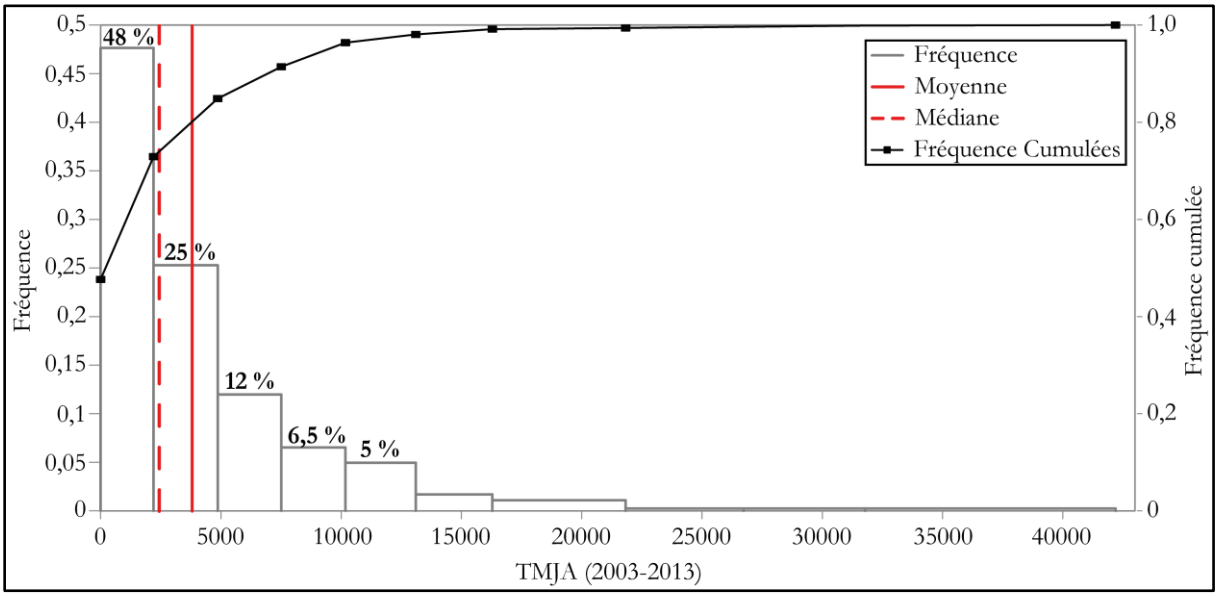



**Figure 1.28 : Classification en 9 classes selon la méthode des K-means des axes routiers en fonction du TMJA observé pour 828 axes routiers mesurés sur la période 2003-2013 (deux sens confondus)**

L'analyse a été portée sur neuf classes[53] réparties en fonction des TMJA relevés sur l'ensemble des axes routiers possédant au moins une station de comptage routier. Globalement, la fréquentation des axes routiers de la ville de Dijon est relativement hétérogène avec un TMJA moyen de 3 800 véhicules/jour et un TMJA médian de 2 433 véhicules/jour pour un écart-type de près de 4 400 véhicules/jour. Ceci montre que la majorité des axes du terrain d'étude concentre une faible part des trafics routiers observés : plus de 70% des axes routiers ont une fréquentation inférieure à 5 000 véhicules/jour. À l'inverse, un faible nombre d'axes (moins de 10 %) concentre les flux routiers les plus importants, allant de 40 000 à 10 000 véhicules/jour. Ainsi, il est possible de distinguer trois grandes tendances d'occupation des routes (Appert et Chapelon, 2003) :

- Les **axes de faible fréquentation**, inférieure à 5 000 véhicules/jour, sont les plus présents dans le réseau urbain, mais ils concentrent une faible part du trafic routier de la ville de Dijon. Ces axes sont caractéristiques des liaisons résidentielles assurant les dessertes à l'intérieur d'un même quartier.
- Les **axes de moyenne fréquentation**, de 5 000 à 10 000 véhicules/jour sont caractéristiques des routes urbaines assurant le lien entre les axes de faible fréquentation et ceux de forte fréquentation ; ils concentrent près de 20% des axes routiers de la commune.
- Les **axes de forte fréquentation** sont les moins nombreux mais ils captent les flux de trafic les plus importants de la ville, supérieurs à 10 000 véhicules/jour. Ces réseaux se dispersent aux entrées Sud et Nord de la ville et s'étendent sur les périphéries du centre-ville sous la forme de boulevards. Leurs caractéristiques se rapprochent des autoroutes urbaines et des voies express urbaines.

Le trafic routier dijonnais se structure autour d'un réseau routier radioconcentrique qui est « *une forme relativement répandue des réseaux urbains. À partir d'une certaine distance les 6 ou n grandes voies radiales sont reliées par des rocades ou par des boulevards de ceinture périphérique […] ; ce dessin est rationnel, mais est souvent critiqué car il conserve, accroit et exagère la fréquentation de ces axes amenant à un encombrement de ceux-ci* » (Brunet *et al.*, 2005, p. 374). Suivant cette définition, le trafic dijonnais aura donc naturellement tendance à se reporter d'abord vers les axes de faible fréquentation, puis à s'orienter vers ceux de moyenne fréquentation pour se retrouver ensuite sur les axes les plus structurants de la ville. De plus, en lien avec le tableau 1.6 (ci-dessus), on peut observer que la répartition des comptages routiers automatiques effectués par la ville de Dijon reflète cette organisation : les stations de comptage sont positionnées sur les réseaux les plus importants ; plus on se dirige vers les routes les moins importantes et plus la couverture du réseau de mesures est restreinte.

### 3.2.3. **Du comptage au dispositif de quantification du trafic routier**

Les données issues des comptages automatiques présentées en figure 1.27 constituent le dispositif de quantification mobilisé pour notre travail. Il est composé, d'une part, des comptages généraux réalisés par la CIGT de la ville de Dijon, et d'autre part, des mesures réalisées par la DIR Centre-

[53] Le nombre de classe a été définit selon l'indicateur de discrétisation de Sturges : *log(N+1, base=2)*. Soit dans notre cas pour 828 valeurs, 9 classes ont été définies. Cependant, nous pouvons noter que les indicateurs de Brook-Carruthers et Huntsberger présentent des indicateurs de classes plus importantes, respectivement 11 classes et 15 classes.

Est. Comme nous l'avons vu précédemment 208 capteurs ont été pris en compte. L'ensemble des données issues des capteurs permanents permet de couvrir une période relativement vaste allant de 2001 à 2012 sur une granularité de 15 minutes. Cette période court de l'année de lancement du système informatisé du CIGT de Dijon au lancement opérationnel du tramway de Dijon (septembre 2012). Sa mise en place a profondément modifié la structure du réseau routier dijonnais, ainsi que le réseau de comptage permanent implanté sur la commune de Dijon comme nous le verrons lors de la dernière section. Avant cela, une première étape du travail a visé à établir une chaîne de traitement des données de comptage afin de les insérer dans un Système de Gestion de Base de Données (SGBD), ainsi qu'au sein d'un SIG, en vue de faciliter la visualisation et la gestion des données disponibles. L'objet de cette section vise à présenter la chaîne de travail qui s'est opérée en deux temps (*cf.* figure 1.29): une première étape de prétraitement visant à mettre en forme et à homogénéiser les bases de données brutes recueillies sous la forme de fichiers Excel par station de comptage, mais aussi pour localiser le réseau de mesures au sein d'un SIG ; la seconde étape spatio-temporelle constitue la mise en place de la base de données « trafic » au sein d'un SGBD permettant de gérer d'une part la dynamique temporelle des données de comptage, et d'autre part la localisation du réseau de mesures lui étant liée.

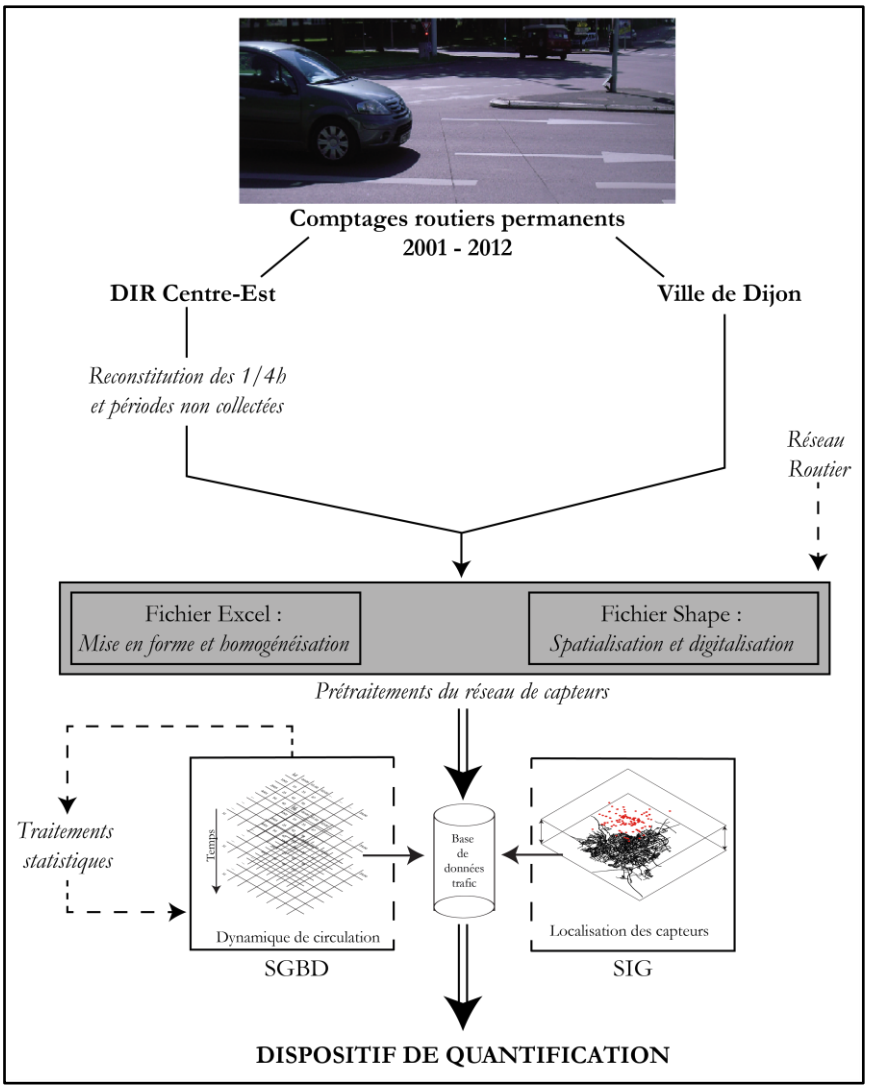



**Figure 1.29 : Du dispositif de comptage au dispositif de quantification : traitement des comptages routiers permanents collectés au sein des CIGT**

### a. *Homogénéisation et stockage des données de comptages routier*

L'homogénéisation et le stockage des données de comptage constituent l'étape de traitement préalable à l'analyse du réseau de comptage. Elle vise d'une part à formaliser les données issues de la DIR Centre-Est et de la ville de Dijon sous un même format (*cf.* encart 3), puis, d'autre part à organiser les données sur un critère temporel pour en extraire les données dynamiques plus aisément. En effet, les données extraites depuis les SRT se formalisent sous la forme de tableurs organisés dans des dossiers annuels[54]. Au sein de chaque dossier « *année* », on retrouve un fichier par station de comptage contenant un classeur de données stockées sous un tableur et organisé par mois, au sein duquel les données de comptage sont structurées pour les jours en colonne et en ligne pour les quarts d'heure de comptage.

Cette formalisation ne permet pas en l'état d'extraire l'ensemble des données de comptage aisément et sans avoir à développer l'ensemble des arborescences des fichiers. Il apparait essentiel de restructurer les données en vue de faciliter leur manipulation par l'intermédiaire d'un SGBD au sein duquel les données sont organisées sur la base d'un critère temporel hiérarchisé illustré en figure 1.30 : *Année, Mois, Jour, Quart d'heure, Station de comptage*. Il est à noter qu'un premier traitement automatique des fichiers a été nécessaire pour les mettre en forme par l'intermédiaire d'une macro dont l'ensemble des traitements réalisés est présenté en annexe 2 (p. iii). L'ensemble des architectures des données de comptage décrites précédemment est synthétisé en figure ci-dessous :

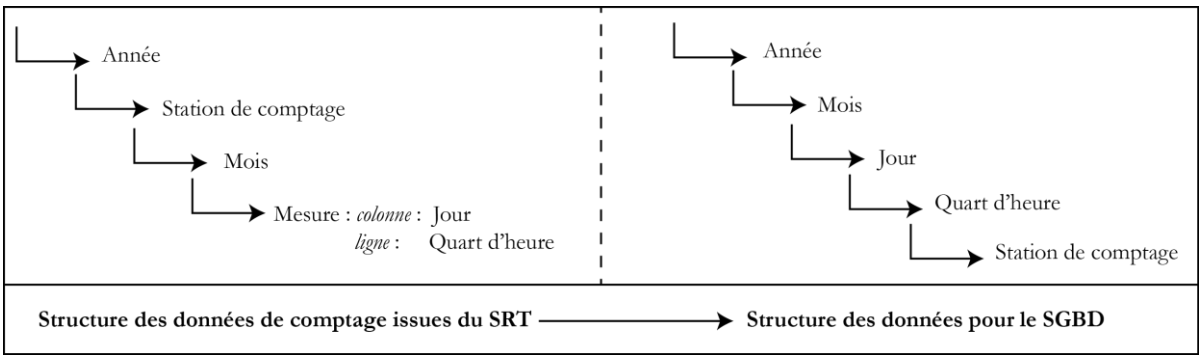



**Figure 1.30 : Architecture des données de comptages routiers au sein du SRT (à gauche) et pour le SGBD (à droite)**

### b. *Spatialisation et caractéristiques des stations de mesures au sein du SIG*

L'ensemble des stations n'étant pas spatialisé par les services auprès desquels les données ont été recueillies, un second prétraitement a permis de les localiser sur le réseau routier de la ville de Dijon au sein d'un SIG. En collaboration avec le SRT, des adresses ont été définies pour chacun des capteurs permettant par l'intermédiaire d'un outil de géocodage[55] d'obtenir leur localisation en coordonnées spatiales (*x ; y*). Par ailleurs, l'ensemble des éléments permettant de caractériser les comptages d'une station a été formalisé sous la forme de *booléen*. Par exemple, pour des comptages généraux (CG), 0 signifie l'absence de ce type de comptage, et inversement pour le 1.

[54] Au total, pour la période allant de 2001 à 2014, plus de 2 500 fichiers Excel ont été collectés depuis le SRT de la ville de Dijon.
[55] L'outil de géocodage utilisé est disponible en ligne : http://dehaese.free.fr/Gmaps/testGeocoder.htm.

La formalisation du réseau de mesure s'illustre à travers une structure orientée *objet* entre deux éléments : les capteurs et le réseau routier. En outre, en considérant les capteurs comme un *puits* stockant les véhicules comptés, cette formalisation reflète la définition du trafic routier mise en avant précédemment (*cf.* section 2.1.1, p. 56) : le trafic routier se constitue de deux composants à l'origine du mouvement, **le véhicule et l'infrastructure**. La construction des données spatiales du réseau routier sous le SIG suit cette définition, dont la formalisation UML (Unified Modeling Language) est présentée sur la partie droite de la figure 1.31. Les couches géographiques « Réseau routier » et « Capteur » sont interfacées en fonction de leur association spatiale, définie à partir du SIG :

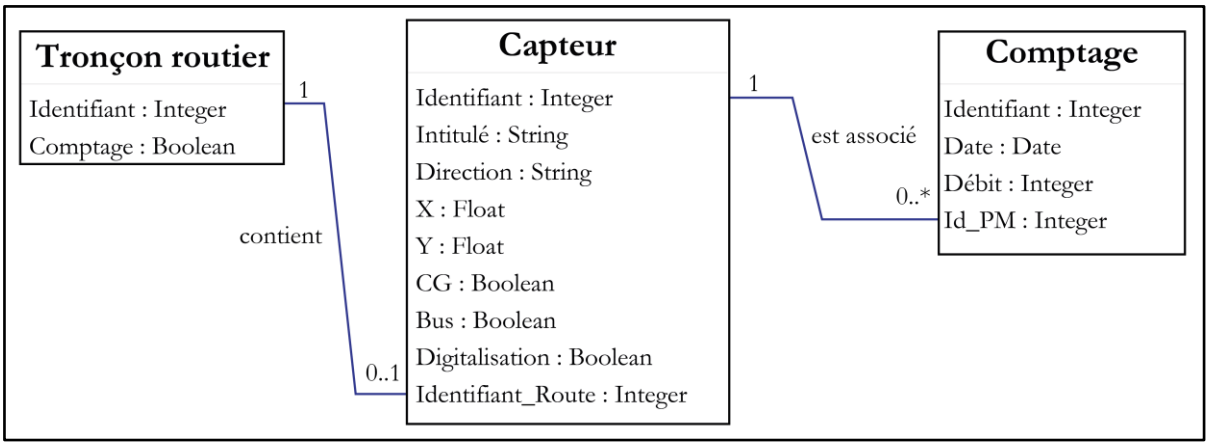



**Figure 1.31 : Diagramme de classe UML formalisant le réseau de comptage routier sous PostGIS : à gauche la structuration dite « temporelle » (la relation entre les capteurs et les données de comptages) ; à droite, la structuration dite « spatiale » (la localisation des capteurs sur le réseau routier)**

***Encart 3 Un traitement spécifique pour les données de la DIR Centre-Est***

Les données collectées auprès de la DIR Centre-Est présentent un format différent des données issues du SRT de la ville de Dijon. D'une part, les données sont collectées toutes les heures au lieu d'une granularité au ¼ heure pour les données de la ville de Dijon ; d'autre part, l'échantillon recueilli couvre seulement les années allant de 2010 à 2013. Les données antérieures à 2010 ont donc été recalculées pour couvrir la même période que les données de la ville de Dijon. Pour cela, les bilans de trafic annuels publiés par la DIR Centre-Est (de 2005 à 2010)[56] ont été utilisés afin d'obtenir les évolutions de trafic des années précédentes.

Un coefficient de variation sur les données de comptages de 2010 a été défini en fonction des informations disponibles de la manière suivante, en sachant que $V_{t-1}$ est la valeur à reconstruire, $\frac{p}{100}$ est le coefficient de variation, et $V_t$ la valeur mesurée :

$$V_{t-1} = \left(1 + \frac{p}{100}\right) V_t$$

Une fois, les valeurs reconstruites, les données horaires ont été passées en ¼ heure en divisant les comptages horaires de manière uniforme par 4, par exemple, pour 100 véhicules/heure, le nombre de véhicules ¼ heure est de 25.

[56] Descriptif du réseau de mesure exploité par la DIR Centre-Est : http://www.enroute.centre-est.developpement-durable.gouv.fr/donnees-de-trafic-r87.html.

Une fois le réseau de capteurs décrit et localisé, une information complémentaire a été mise en place manuellement afin de déterminer le sens de comptage d'un capteur en fonction du sens de digitalisation (construction numérique des vecteurs) du réseau routier par un *booléen* en vue de mettre en relation les capteurs et le réseau routier comme illustré en figure 1.32 ci-dessous :

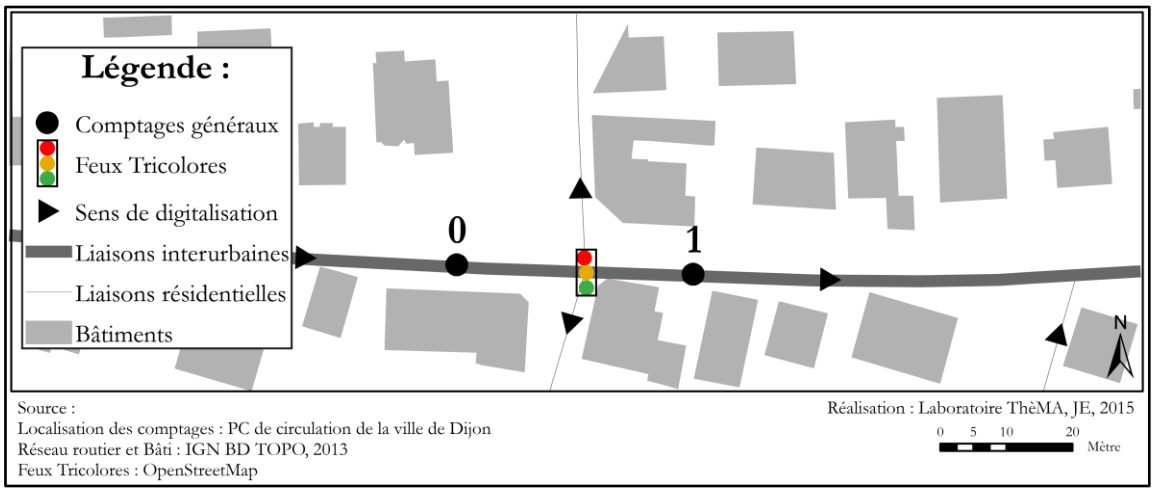



**Figure 1.32 : Codage du sens de comptage des stations de mesures sur le réseau routier (1 : correspond au sens de comptage identique au sens de digitalisation du réseau routier et inversement pour un sens de comptage inverse au sens de digitalisation)**

### c. *Une gestion spatio-temporelle des données de comptages par l'intermédiaire d'un SGBD et d'un SIG intégrés*

La dernière étape de traitement du réseau de comptage vise à insérer les informations collectées au sein d'une base de données unique permettant, d'une part, de gérer l'ensemble des données dynamiques de circulations collectées, et d'autre part, de permettre une représentation des stations de mesures sur le réseau routier sous un SIG. En d'autres termes, la base de données permet de faire l'interface entre les données de comptages routiers et les capteurs physiques. Le choix du SGBD s'est naturellement tourné vers le logiciel PostGIS qui est une extension du SGBD PostgreSQL permettant d'inclure des données spatiales à côté des données habituellement stockées dans un SGBD classique (Corti *et al.*, 2014). Dans notre cas, comme illustré en figure 1.29 et formalisé en UML au sein de la figure 1.31, la base de données se structure en deux parties. D'une part, on retrouve l'ensemble des données de comptage constituant la partie temporelle de notre dispositif de quantification. Et d'autre part, l'ensemble des stations référencées est localisé sous la forme d'une couche géométrique permettant la représentation spatiale des données. La structure de la base de données sous PostGIS suit la restructuration temporelle des données de comptages décrite en figure 1.30. Les données de circulation routière sont ensuite, comme illustré en figure 1.31, stockées au sein d'une table « Comptage » à laquelle est associée la couche géographique « Capteur », permettant de récupérer au sein d'un même fichier, les données de débit relatives au capteur ainsi que ses caractéristiques physiques.

Comme illustré dans la seconde partie de la figure 1.31 (à gauche), la gestion des comptages est gérée individuellement sur la base d'un critère temporel (*timestamp*) contenu dans un champ « Date ». L'extraction et la sélection des données s'opèrent par des procédures SQL imbriquées (*cf.* annexe 3, p. v). Cette structuration permet une gestion souple des données de comptages en

permettant de faire des allers et retours entre ces données et leur localisation. De plus, cela permet de répondre plus rapidement à des besoins ponctuels, comme lors de la demande de données agrégées à des pas de temps agrégés, par exemple, les débits horaires ou les TMJA. Une illustration de ces traitements est proposée en annexe 3 (p. v) lors d'un échange avec les services du pôle transport du Grand Dijon dans le cadre d'une analyse des effets du tramway sur la circulation routière à Dijon en mai 2014.

## 3.3. Que nous révèlent les capteurs de la ville de Dijon pour la visualisation du trafic routier en milieu urbain ?

L'insertion des données issues des dispositifs de comptage de la ville, centralisée par l'intermédiaire d'une base de données, permet d'établir un état des lieux du réseau de mesure sur des critères spatio-temporels. En entrant plus directement dans la donnée, l'objet de cette section vise à explorer les informations qu'il est possible d'extraire de la base de données mise en place pour la représentation du trafic routier en milieu urbain : que mesurent les capteurs électromagnétiques en milieu urbain ? Quelles en sont les forces et les faiblesses ? Comment envisager ce dispositif de quantification pour la simulation du trafic routier ?

### 3.3.1. État des lieux des capteurs électromagnétiques : un réseau de mesures tributaire de l'évolution du réseau urbain

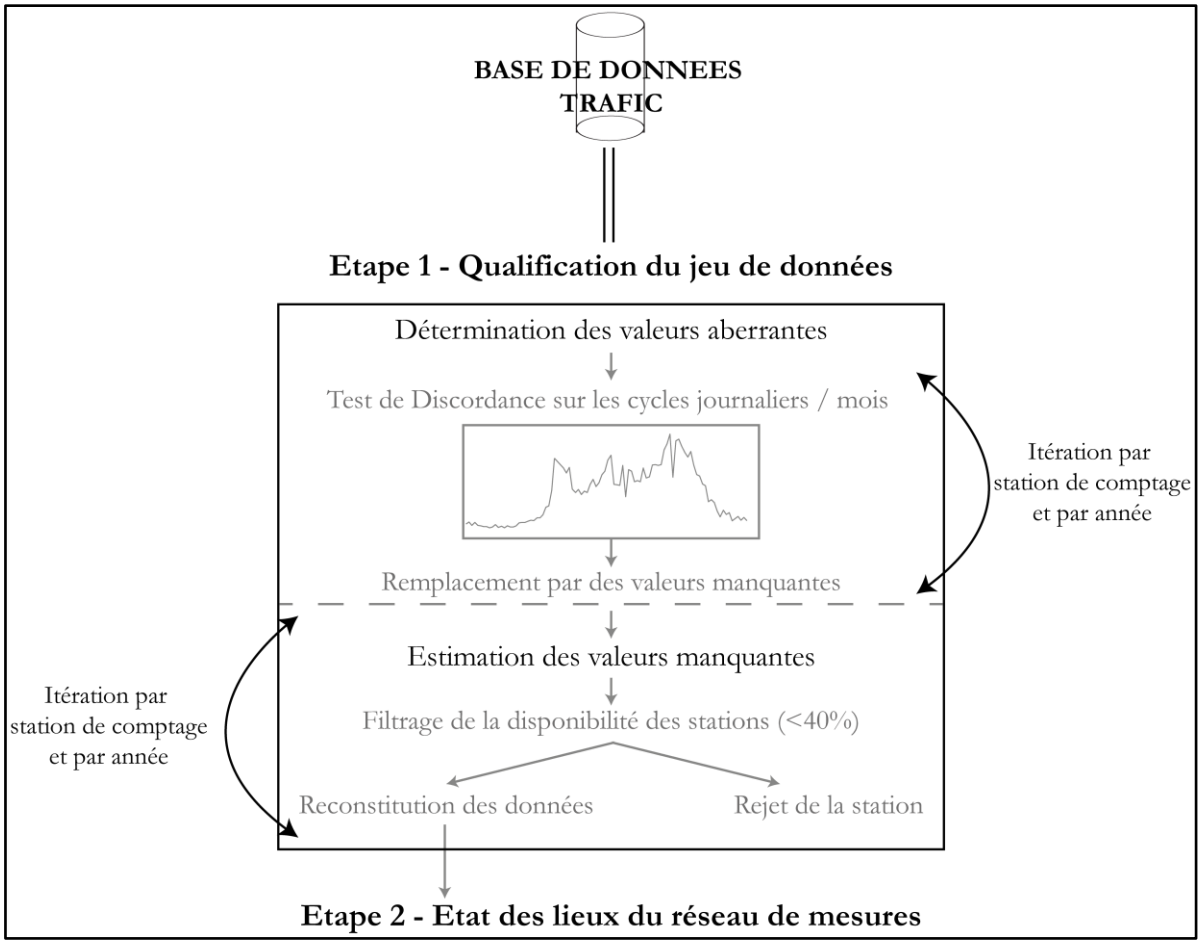



**Figure 1.33 : Processus pour le traitement statistique des données de comptages routiers**

L'état des lieux global du réseau de comptage que nous avons à disposition nous amène à envisager une analyse sur deux niveaux : le premier se situe au niveau des stations de comptage afin d'apprécier la qualité du jeu de données collecté, tandis que le second niveau porte sur l'évolution du réseau de capteurs autant dans le temps que dans l'espace. L'analyse sur ces deux niveaux permettra d'apporter un état des lieux global du réseau de capteurs collectés au niveau des stations, puis au niveau du réseau de mesures. En outre, l'objectif de cette analyse est de proposer, par l'intermédiaire d'un traitement statistique, une vue générale du dispositif de quantification mis en place afin de proposer une analyse spatio-temporelle du réseau de mesures de la ville de Dijon. La procédure mise en place pour l'état des lieux du réseau de mesures est synthétisée au sein de la figure 1.33 (ci-dessus).

#### a. *Qualification des données de comptage : une analyse temporelle de la fiabilité du réseau de capteurs*

Au préalable de l'état des lieux du réseau de capteurs, il est nécessaire de vérifier la qualité des données collectées au sein de la base de données constituée. Ce premier travail se situe à l'intersection du stockage des données au sein du SGBD présenté précédemment (*cf.* figure 1.29, p. 98). Le processus statistique mis en place se base sur une analyse des valeurs manquantes (VM) et aberrantes (VA) par station de comptage. Elle permet d'apporter des informations statistiques en vue de qualifier les données de débit recueillies en prêtant une attention particulière aux spécificités de chacune des stations. En effet, comme pour toutes les bases de données, il est fréquent de trouver des erreurs de mesures au niveau des stations. Ces erreurs peuvent provenir, d'une part, de pannes du réseau de capteurs, et d'autre part, de mesures aberrantes dues à un dysfonctionnement du système de mesure.

Pour les données de comptage, la qualification s'établit en deux temps (Bréheret *et al.*, 2000) : une première phase d'identification des valeurs dites aberrantes (VA) et une seconde analyse des valeurs dites manquantes (VM). La chaîne de traitement mise en place suit la procédure présentée en figure 1.33. Elle se structure en un processus incrémental réalisé sous un logiciel de calcul matriciel[57] permettant de gérer plus aisément l'ensemble de la base de données de trafic routier.

Tout d'abord, il convient de définir et d'illustrer ces deux types de valeurs. Les VA se définissent comme étant « *une observation (ou un ensemble d'observations) qui semble être inconsistante avec le reste des données ou d'une manière, il y a VA lorsque l'une ou l'autre observation d'un ensemble de données, détonne ou n'est pas en harmonie avec les autres observations* » (Planchon, 2005, p. 20). Dans le cas des données de trafic routier, les VA revêtent quelques spécificités qu'il convient de définir. En effet, des phénomènes exceptionnels de trafic routier peuvent être observés lors, par exemple, de congestion spécifique ou de déviation routière. Les données ainsi recueillies peuvent présenter des incohérences avec l'ensemble des données habituellement observées, que l'on qualifiera plutôt de valeurs extrêmes ou alors de VA (Bréheret *et al.*, 2000). Il est donc nécessaire d'établir un processus de traitement et de repérage des VA en fonction des spécificités de chacune des stations auxquelles les données sont liées. Ce processus vise à caractériser si ces incohérences sont à relier à une erreur de mesure ou si celles-ci sont inhérentes à un événement de trafic

[57] Pour l'ensemble des traitements statistiques, nous avons utilisé le logiciel Matlab, néanmoins d'autres logiciels libres comme R ou Scilab, ou payants comme SAS permettent de réaliser ces traitements sur des grands ensembles de données.

routier que l'on peut qualifier d'extrême (Planchon, 2005). Dans ces conditions, une VA sera définie si celle-ci apparait lors d'une seule période de mesure et en un seul point du réseau de mesures.

Comme illustré en figure 1.33, les VA repérées sont ensuite considérées comme des VM au sein de la base de données en vue de les corriger. Pour réaliser ce traitement, nous avons fait appel à un test de discordance en se basant sur la méthode des quartiles (définis à 99 %) afin d'identifier et de rejeter les valeurs tirant fortement la distribution des données de débit observées comme illustré en figure 1.34[58]. Le traitement des données a permis de détecter 7 % de VA sur l'ensemble de la période d'étude (2001-2012) ce qui représente un taux convenable pour notre base de données. De surcroît, ce résultat montre que le dispositif de mesures à disposition présente une bonne fiabilité pour la mesure du trafic routier, par rapport à d'autres dispositifs de comptage, ce qui apparait être conforme à d'autres études portant sur les capteurs électromagnétiques (Bréheret *et al.*, 2000). La fiabilité du système de mesures explique aussi sa pérennité et son implantation sur un grand nombre de réseaux urbains et interurbains.

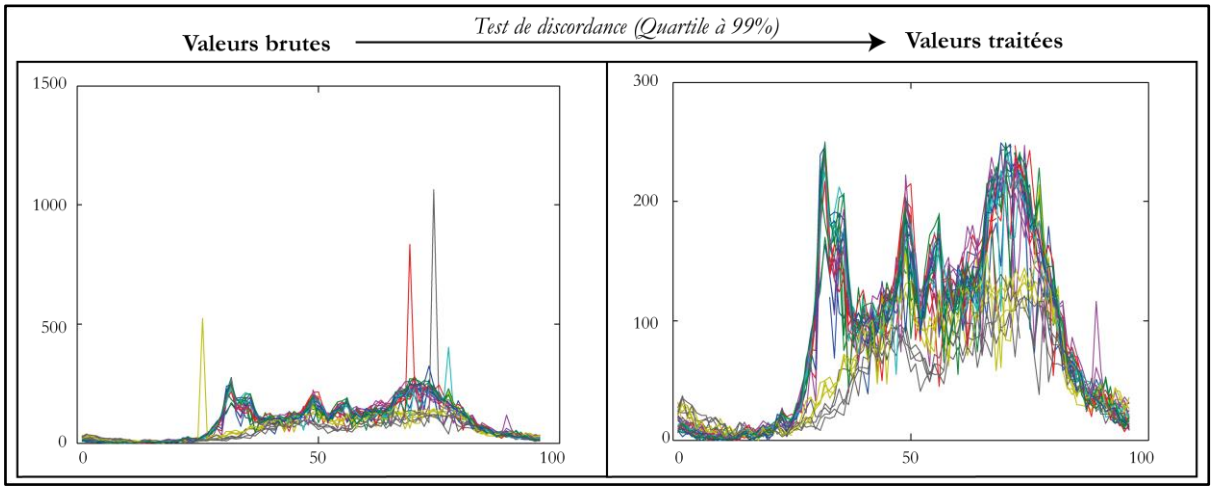



**Figure 1.34 : Illustration du traitement des VA d'une station de comptage pour le mois de Septembre 2008 (en abscisse les valeurs de débit ; en ordonnée une journée de comptage correspondant à 96 comptages)**

Les VM, quant à elles, sont des données totalement absentes de la base de données. Elles correspondent essentiellement à des pannes ou à des dysfonctionnements temporaires, voire permanents, des capteurs électromagnétiques. Il est aussi possible d'identifier l'absence totale du capteur lorsque le nombre de VM est trop important ; dans ce cas cela signifie que la station de mesure est hors service ou a été supprimée du réseau de comptage routier pendant une période donnée. Afin de considérer l'absence ou non d'une station pour une année donnée, un seuil de 40 % de VM a été défini (le seuil correspond à près de 5 mois sans mesure). Au-delà de ce critère, une station est considérée absente pendant la période considérée. À l'inverse, lorsqu'une station présente un taux de VM inférieur à 40%, un processus de reconstruction de ces données manquantes a été appliqué dont la formule est explicitée en fin de section (p. 106).

[58] À noter que les données ont été traitées via un processus itératif en fonction des stations par jour de comptage pour un même mois et pour une année considérée afin de garder la spécificité de comptage de chacune des stations. Le test de discordance permet alors de détecter si les VA font partie, ou pas, de la population à considérer.

La figure 1.35 illustre ce processus de filtrage des VM établi par station de mesure. En rouge sont illustrées les valeurs de débit manquantes (VM), alors que les valeurs de débit présentes sont en bleu. Cette représentation s'inspire des représentations proposées par Harvey J. Miller (2015, p. 103, fig. 6.2) pour la représentation des données spatio-temporelles. De plus, cette représentation présente l'avantage de visualiser très rapidement l'état de fonctionnement d'une station de comptage. Les données sont organisées, pour les abscisses, par année, et pour les ordonnées, par mois. Sous cette forme, il est possible de visualiser l'état d'une station de mesures au cours de la période de recueil. Pour illustrer cela, deux stations de comptages « types » et situées sur des boulevards périphériques ont été sélectionnées (leur localisation est reportée en vert sur la figure 1.38, p. 108) : : l'une située sur le boulevard Kennedy (PM 331) et dont le capteur n'a été soumis qu'à très peu de perturbation depuis son lancement ; l'autre située sur Boulevard Trimolet (PM 342), dont le lancement a été opéré en août 2002, et qui est actuellement démontée, car elle se situe sur le tracé du tramway (*cf.* figure 1.38, p. 108) :

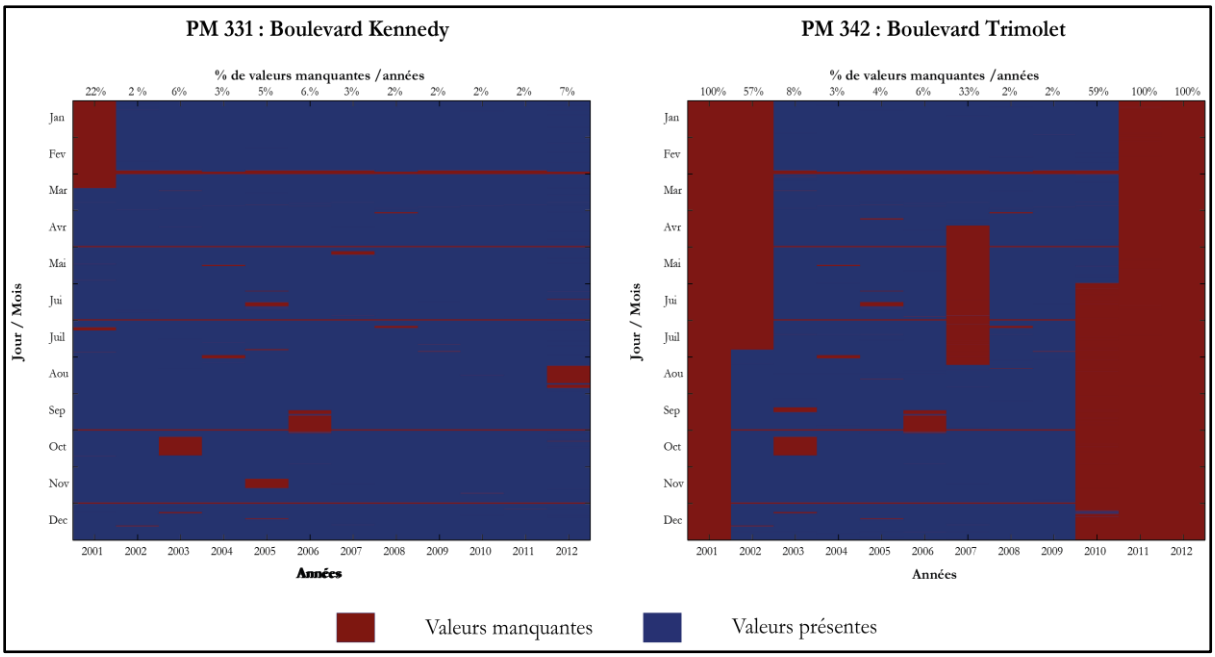



**Figure 1.35 : Traitement des VM par station de comptage (à gauche, une station disponible pour l'ensemble de la période 2001 à 2012 ; à droite, une station disponible pour 7 années : 2003 à 2009)**

La distribution des VM au sein de la base de données est synthétisée dans la figure ci-dessous. Leur distribution est de deux natures : soit celles-ci sont isolées (du quart d'heure à la journée) ; soit elles sont additives, allant de la semaine à plusieurs mois de dysfonctionnement. Respectivement, le taux de VM isolées est de 91 % et celui de VM additives est de 9 %. Ce premier résultat signifie, en outre, que les capteurs de mesure présentent principalement des dysfonctionnements et pannes sur le court-terme, plus proches des dysfonctionnements illustrés pour la station de mesure du boulevard Kennedy à gauche de la figure 1.35. Alors que dans le cas de VM additives, on se retrouve confronté à des stations présentant des caractéristiques plus proches de la station du boulevard Trimolet.

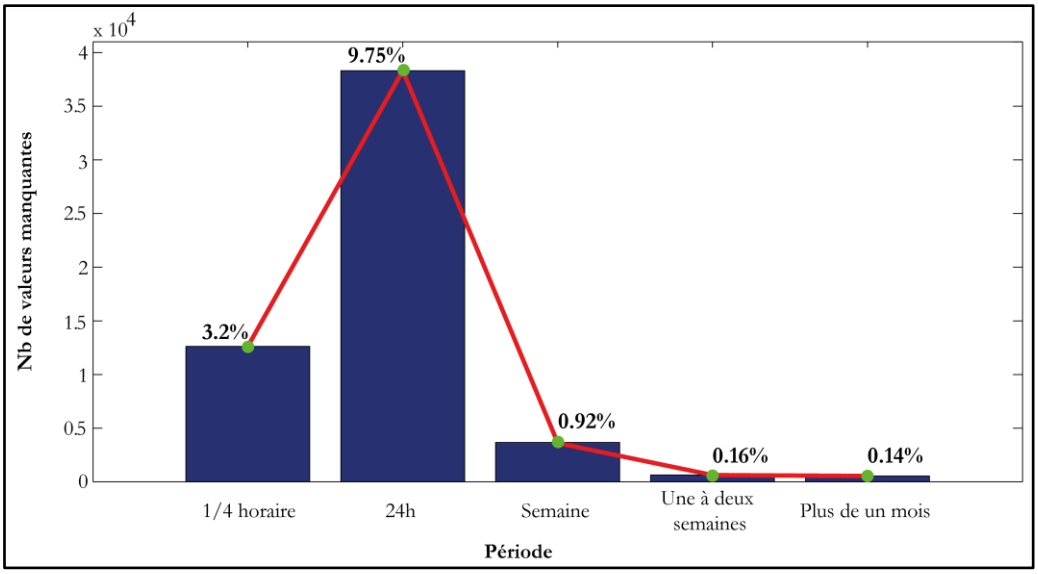



**Figure 1.36 : Périodicité des valeurs manquantes sur l'ensemble de la base de données trafic (les pourcentages sont définis sur l'ensemble des valeurs de débit au quart d'heure de la période 2001-2012)**

Les VM étant peu nombreuses sur l'ensemble de la base de données, celles-ci ont été réestimées par un traitement de moyenne mobile. Cette méthode permet de calculer une moyenne qui au lieu d'être calculée sur l'ensemble des $N$ valeurs d'un échantillonnage, est calculée tour à tour sur chaque sous-ensemble des $n$ valeurs choisies. Le sous-ensemble utilisé pour calculer chaque moyenne se déplace sur l'ensemble des données. Cette méthode nous permet de garder la typicité d'une journée, les données corrigées étant au ¼ horaire. Le remplacement des valeurs manquantes s'est fait de cette manière : si une donnée $x_i$ est manquante, elle est remplacée par la moyenne des valeurs $x_{n-1}$ et $x_{n+1}$, et si $k$ données sont manquantes, celles-ci sont remplacées par la moyenne des $k$ valeurs entre $x_{n-k}$ et $x_{n+k}$ afin de s'adapter à la spécificité de chacune des stations :

$$x_n = \sum_{i=n-k}^{i=n+k} \frac{x_i}{2k}$$

### b. *État des lieux du réseau de mesures sur la période d'étude : disponibilité et évolution du réseau de capteurs dans le temps*

Sur l'ensemble des données traitées, la base de données se compose de 16 % de VM, ce qui apparait relativement élevé. Néanmoins, ce premier résultat est à attribuer au fonctionnement du réseau de capteurs dans le temps, dont le nombre de capteurs évolue au cours de la période. L'ensemble des 208 stations ne fonctionne pas sur l'ensemble de la période d'étude, mais année par année ; il est alors important d'avoir une approche temporelle pour apprécier la qualité des données recueillies. Pour nuancer cette première estimation, il faut apporter une information plus précise de la répartition des VM en mettant en place un filtrage sur le nombre de stations disponibles, comme représenté à travers la figure 1.37. En effet, le réseau de mesure connait différentes évolutions en termes d'implantation et de modification. Une approche par station de mesure est donc nécessaire et permet de voir que le réseau de mesures présente un taux de VM moyen plus convenable de 6,5 % sur l'ensemble de la période d'étude. La synthèse de ces traitements est présentée en figure ci-dessous :

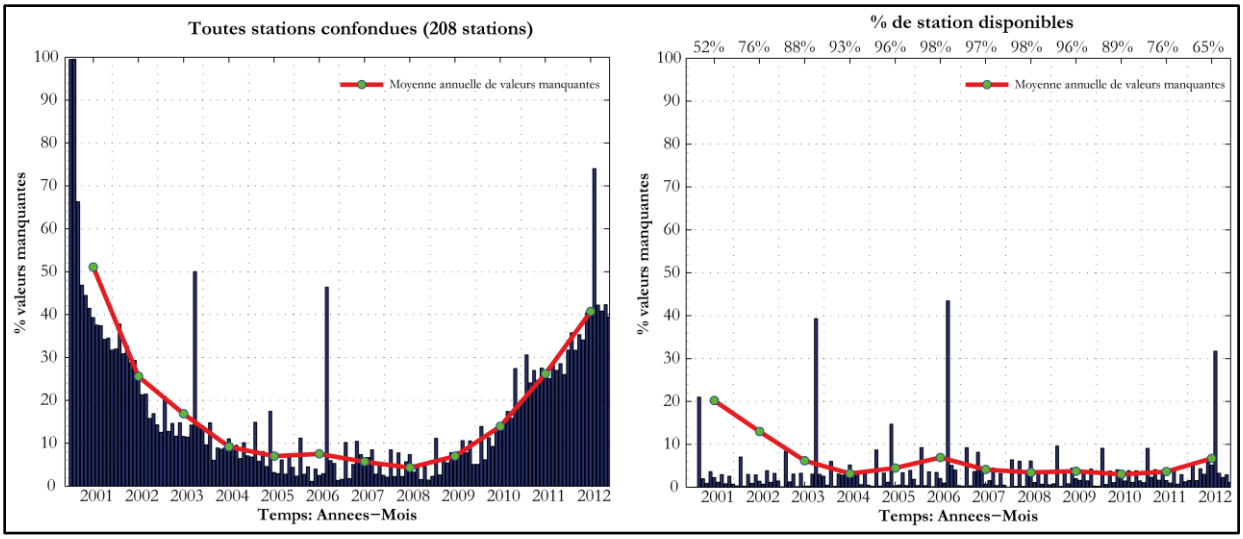



**Figure 1.37 : Représentation des valeurs manquantes sur la période 2001-2012 (à droite, traitement sans distinction de la disponibilité des stations ; à gauche, traitement avec filtrage des stations disponibles)**

Le réseau de capteurs connait différents changements au cours de la période 2001-2012. Ces modifications sont de deux ordres : soit elles sont liées à une densification du réseau de capteurs, soit elles sont liées à des modifications structurelles du réseau de voirie urbaine. Les illustrations de la figure 1.37 permettent de commenter ces évolutions au cours de la période recueillie. On peut y observer trois phases distinctes de mesure :

- Une **phase de lancement** allant de 2001 à 2002 ; elle correspond à la mise en place du réseau de mesures qui démarre à partir du mois de mars de l'année 2001 avec le lancement de 52 % des stations du réseau de mesures. À partir, de l'année 2002, le nombre de stations augmente, faisant ainsi diminuer le nombre de VM sur ces deux périodes.
- Une **phase de routine** allant de 2003 à 2009 ; elle correspond à la période de mesure où le nombre de stations est le plus élevé : entre 88 % et 98 % de stations en fonctionnement. Logiquement, le nombre de valeurs manquantes associé à cette deuxième période diminue fortement à des taux inférieurs à 10 % de VM.
- Une **phase de perturbation** à partir de 2010 ; elle présente une dégradation de la qualité des données, avec une perte de plus de 20% des stations de mesure. Cette dernière phase correspond aux modifications de la voirie routière de la ville de Dijon lors de la mise en place du tramway, qui a débuté à partir du mois d'octobre 2010 (*cf.* figure 1.38, p. 108). En effet, plus de 40 postes de mesures sont localisés sur le tracé du tramway, modifiant ainsi le réseau de mesures que nous avons à disposition, avec une perte de plus de 20 % des capteurs.

La carte illustrée en figure 1.38 permet de visualiser l'état du global du réseau de comptage sur la période d'étude, par poste de mesure. Globalement, les stations présentant une bonne disponibilité se situent le plus généralement sur les plus grands axes de l'agglomération. À l'inverse, les moins disponibles sont plutôt situées en marge du réseau de la ville de Dijon. La représentation graphique en figure 1.39 est complémentaire de la carte proposée et permet d'illustrer le nombre de stations disponibles à une échelle de temps plus fine (mois par mois). En

s'appuyant sur cette représentation, on peut observer que sur l'ensemble de la période d'étude, la période allant de 2003 à 2009 présente la densité de capteurs la plus importante avec une disponibilité variant de 75 à 98 % de stations de mesure à disposition :

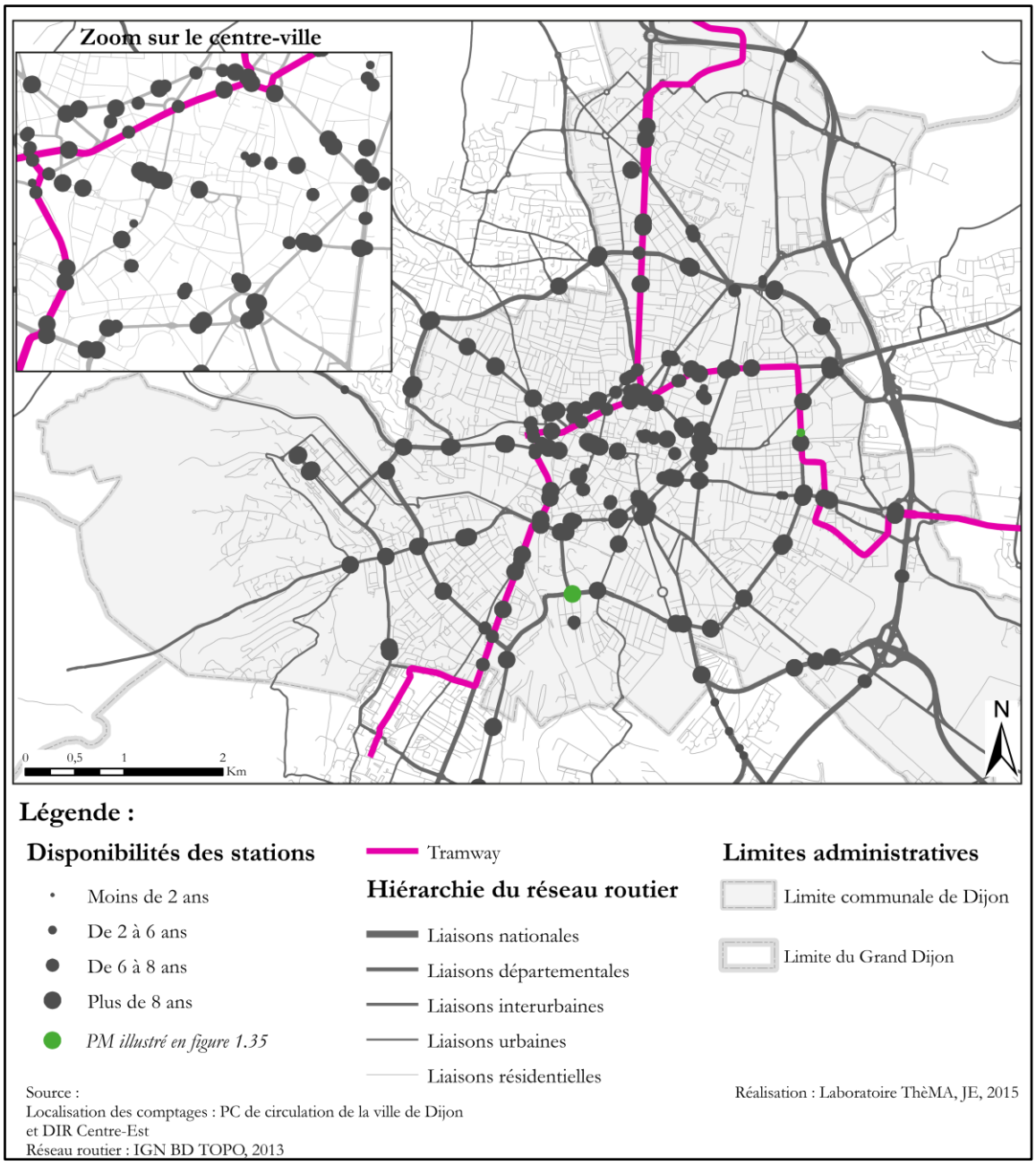



**Figure 1.38 : Disponibilité des 208 stations de comptage sur la ville de Dijon pour la période 2001-2012 : une illustration des impacts de l'implantation du tramway de Dijon sur le réseau routier**

L'ensemble de cette analyse sur la disponibilité des données permet de commenter la période que nous avons sélectionnée en proposant un état des lieux global du réseau de mesure. Si le début de la collecte correspond au lancement du système informatisé (2001), ce sont les modifications et perturbations rattachées à l'implantation du tramway qui nous ont permis de définir une période de fin de recueil. Cette période qui débute à partir de 2010 (*cf.* figure 1.39) montre une baisse progressive et significative du réseau de mesure. De plus, elle a eu un impact sur la localisation de

capteurs qui ont été déplacés, supprimés ou réimplantés en d'autres lieux. La figure ci-dessus permet de conclure nos propos en permettant de visualiser le tracé du tramway en juxtaposition de la localisation et de la disponibilité des capteurs sur la période d'étude prise en compte :

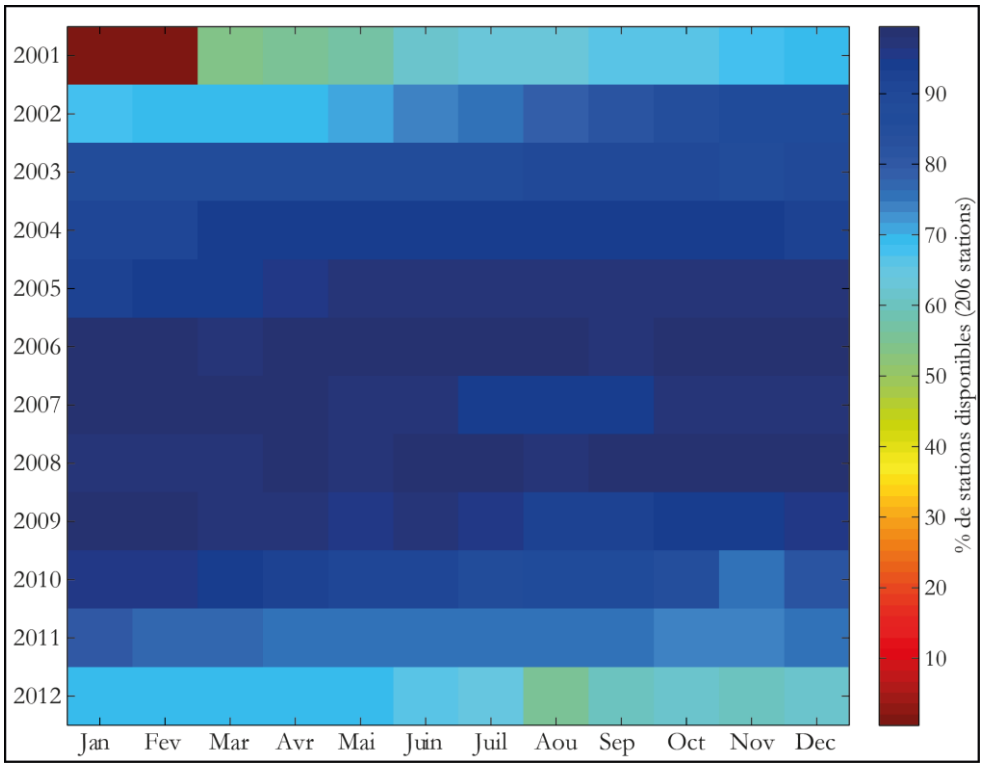



**Figure 1.39 : Pourcentage de disponibilité des stations par mois et année sur l'ensemble de la période de recueil (en abscisse les mois ; en ordonnée les années ; le gradient de couleur illustre le pourcentage de disponibilité des stations pour 208 stations)**

### 3.3.2. Se focaliser sur la temporalité des données pour révéler les rythmes de l'espace urbain

L'intérêt et la force du réseau de capteurs permanents se situent sur sa granularité temporelle permettant une observation à un pas de temps fin (le quart d'heure), ainsi que sur sa profondeur temporelle supérieure à 10 ans. En revanche, le signal mesuré par les capteurs ne permet pour le moment pas d'en apporter une définition : que mesurent les capteurs ? Comment interpréter les comptages qui leur sont associés ? L'objet de cette section est de comprendre dans leur globalité les données issues du réseau de mesures par une approche chronologique.

Ce traitement s'insère dans les propos de la partie 1.2 (p. 35) afin de construire une information à partir des données brutes par la décomposition de notre série de données. En effet, analyser une série chronologique est loin de permettre d'accéder à toutes les informations qu'elle contient, surtout à travers des séries longues. Il est donc souvent nécessaire de transformer la série de données, c'est-à-dire d'en donner une autre représentation qui fasse apparaitre plus clairement telle ou telle de ses caractéristiques dans le temps. Dans notre cas, nous avons souhaité différencier le bruit et le signal de notre série par l'extraction de cycles temporels caractéristiques des mesures de trafic réalisées sur la ville de Dijon.

Présentée en figure 1.40 (en haut), la méthode employée ici est celle de l'analyse en ondelettes

basée sur la méthode des transformées de Fourier (ou FFT pour *Fast Fourier Transform*) qui est un outil fondamental dans l'étude des fonctions périodiques (Torrence et Compo, 1998). La FFT, proche de l'analyse harmonique, permet de décomposer, dans le spectre de Fourier, les fonctions périodiques d'une série en une somme de sinus et cosinus affectés de coefficients. La méthode des ondelettes présente l'avantage de décomposer le signal à la fois dans l'espace réel et dans l'espace de Fourier et permet de prendre en compte et de regrouper des variations de même type qui sont homogènes dans les deux espaces (Torrence et Compo, 1998). Pour cette analyse, nous avons appliqué la méthode des ondelettes sur l'ensemble des données de comptage brutes (au quart d'heure) pour la période 2002-2012, comme présenté en haut de la figure 1.40. L'analyse graphique se constitue de deux éléments, le *spectre de puissance locale* (en haut à gauche) et le graphique de *puissance globale* (en haut à droite).

Les différents éléments associés à la projection en ondelette se lisent en deux temps (*cf.* figure 1.40). En s'appuyant, dans un premier temps, sur le *spectre de puissance locale,* on observe en abscisse la série chronologique où apparaissent les années considérées pour l'analyse (2002-2012[59]) ; en ordonnée apparaissent les périodes exprimées et décomposées sous la forme de fréquence par quart d'heure ; tandis que la courbe en pointillés placée sur le graphique indique la limite de significativité des résultats (ici p=0,05). Ainsi, toutes les données situées en dessous de ces pointillés ne sont pas prises en compte pour l'interprétation des résultats. Les cycles temporels mis en évidence par les ondelettes sont représentés par des blocs d'une même couleur relativement homogènes. Le gradient de couleurs allant du bleu au rouge représente l'intensité du spectre (à relier au graphique de puissance globale) ; par conséquent plus les valeurs sont proches du rouge plus l'intensité du cycle est importante et caractéristique de la série. Sous cette forme le temps est décomposé de la même manière que l'on déroulerait un *volet roulant* où chacune des *lamelles* correspondrait à une fréquence sous la forme d'un quart d'heure et l'ensemble du *tablier* à l'ensemble de ces fréquences cumulées par quart d'heure. Dans un second temps, l'interprétation de ces gradients se fait aussi en fonction du graphique de *puissance globale*. Illustrée par la courbe continue en bleu, la puissance globale correspond à la moyenne du spectre local avec son seuil de significativité défini à 95 % (en pointillés bleus). C'est pourquoi, pour interpréter un cycle comme significatif, il est nécessaire d'avoir une double lecture de la figure 1.40 (de gauche à droite) en regardant, tout d'abord, sa ligne de significativité au sein du spectre de puissance locale et en se focalisant, ensuite, sur sa significativité au sein du graphique de puissance globale.

Cette double lecture entre l'analyse des spectres de puissance locale et globale permet de mettre en avant trois périodicités du trafic routier sur la série chronologique (encadrés en A, B et C en bas de la figure 1.40) :

- En rouge (A), un cycle journalier (valeur de 96 quarts d'heures pour un jour en ordonnée),
- En vert (B), un cycle hebdomadaire (valeur de 672 quarts d'heures pour une semaine en ordonnée)
- En violet (C), un cycle annuel (valeur de 35 040 quarts d'heures pour une année en ordonnée).

[59] Les mesures de trafic commençant à partir de mars 2001, l'année 2001 n'a pas été incluse au sein de l'analyse spectrale pour ne pas modifier les résultats au sein du spectre de puissance présenté en figure 1.40.

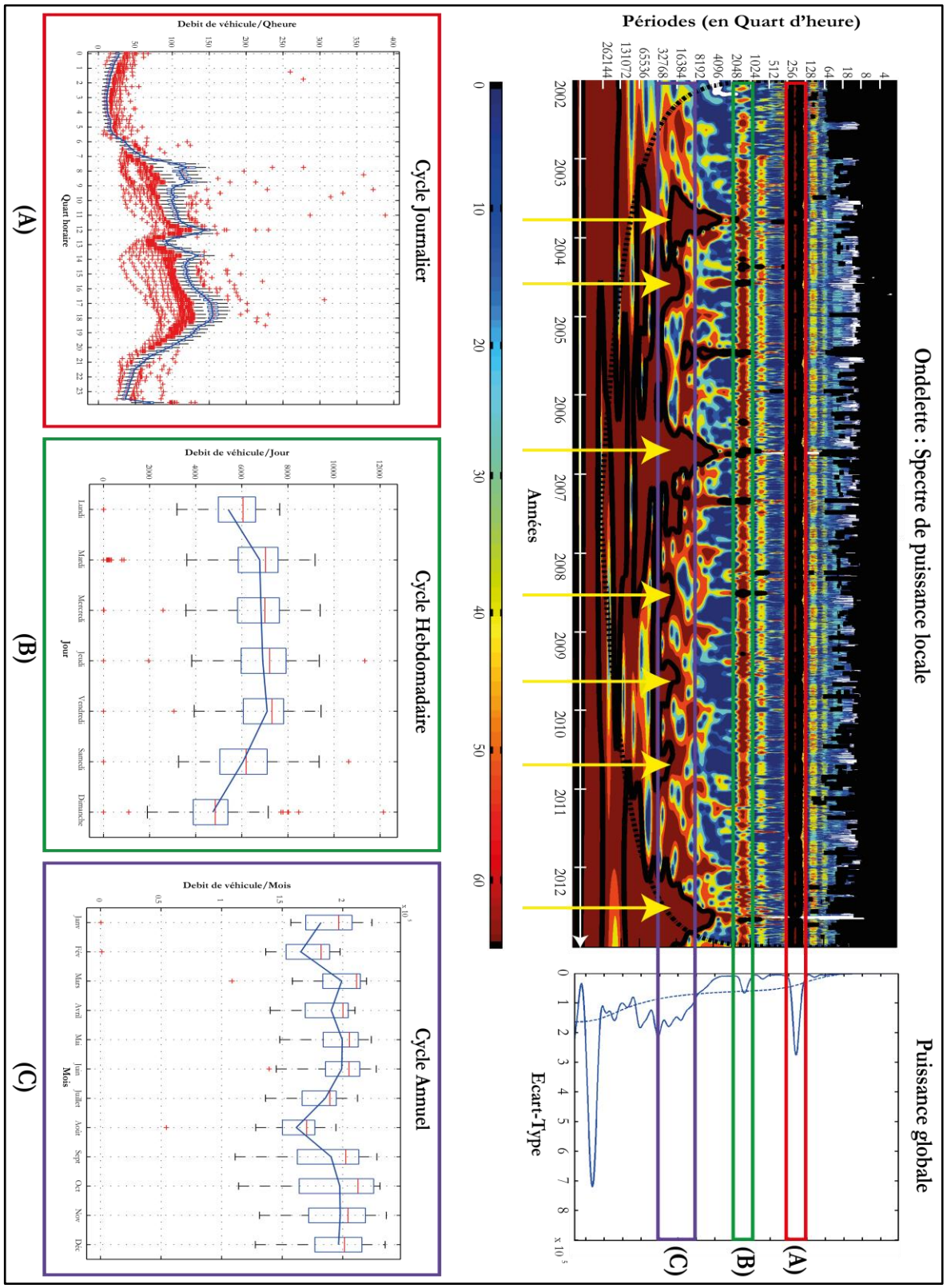



**Figure 1.40 : Caractéristiques de la fréquentation routière de la ville de Dijon sous la forme de cycles temporels pour la période 2002 -2012 en quarts horaires pour les 208 PM permanents:**

**<u>En haut</u> : Projection en ondelettes sur l'ensemble du réseau de comptage routier de la ville de Dijon (moyenne des 208 stations permanentes, significativité : 95%) ; <u>En bas</u> : Extraction des cycles mis en évidence par les ondelettes (A, B, C) sous la forme de boxplot (la moyenne des débits pour les 208 stations de mesures est illustrée par la courbe en bleu)**

L'extraction des périodes (A, B et C) sous cette seconde forme permet de commenter la structure du trafic routier au cours du temps. De plus, cela permet aussi d'apporter une information supplémentaire aux données de comptage collectées en y apportant une définition plus sémantique :

- Le **cycle journalier** (A) s'illustre par un rythme diurne avec une fréquentation du trafic routier qui augmente lors de la journée et inversement la nuit. Par ailleurs, au cours de la journée, le trafic est rythmé par trois « pics » de fréquentation : le premier correspond au début de la journée (de 6 h à 9 h du matin) ; le second correspond à la pause méridienne (11 h 30 à 14 h) ; enfin le troisième correspond à la fin de journée, s'étalant de 16h à 19h.
- Le **cycle hebdomadaire (B)** s'illustre par une alternance de la fréquentation des axes en fonction de la semaine et des week-ends (WE), les WE correspondant aux périodes les moins fréquentées, alors que les jours de la semaine, le trafic routier tend à augmenter progressivement du lundi au vendredi.
- Le **cycle annuel ou saisonnier (C)** se caractérise principalement par une alternance entre les mois de juillet et d'août qui présentent les débits de trafic les plus faibles et les autres mois de l'année. Toutefois, il est à noter que les mois de janvier, de février[60] et d'avril présentent un trafic légèrement plus faible que les autres mois (mars, mai, juin, septembre, octobre, novembre et décembre) ; cette baisse peut être en partie reliée aux vacances scolaires.

Dans le cas C, il est possible d'observer un changement de fréquence des cycles dont les intensités les plus importantes s'observent lors de la seconde moitié d'une année, reflétant donc l'apparition d'une alternance entre les hivers et les étés (comme annoté par les flèches jaunes sur la figure 1.40) ; c'est particulièrement le cas des années 2003, 2006 et 2012. Ceci illustre le fait que l'analyse en ondelettes apporte aussi une information à un niveau interannuel. En effet, si les années 2003, 2006 et 2012 se distinguent des autres, les valeurs du spectre de puissance local, sous le cadre violet (C) en figure 1.40, apparaissent être significativement en phase pour la période s'étalant de 2003 à 2008 (en abscisse). Ceci permet de montrer qu'en termes de fréquence le trafic routier sur cette période est stable en suivant toujours la même configuration cyclique entre les jours, les semaines et les saisons. Ce dernier constat tend à montrer que la configuration du trafic routier observée sur la ville de Dijon ne connait pas de modifications majeures sur la période d'étude. En d'autres termes, nous pouvons considérer que les rythmes mis en évidence par l'analyse en ondelettes (journalier, hebdomadaire et saisonnier) se répètent au cours de chaque année sous une même configuration.

Globalement, les comptages routiers apparaissent refléter une rythmique urbaine qui se répète plusieurs fois et à différentes échelles temporelles, que ce soit sur les temps les plus courts (le quart d'heure, la journée) ou les plus longs (la semaine, les saisons, les années). Intuitivement, la description de ces trois cycles de fréquentation routière apparait comme une traduction, sous une forme temporelle, des rythmes urbains : néanmoins, quelle(s) information(s) pouvons-nous y apporter ? À quelle(s) typologie(s) de déplacement et de trafic routier correspondent ces signaux ?

[60] À relever que le mois de février présente des spécificités qui sont aussi inhérentes à sa structuration concentrant moins de jours de comptage (de 28 à 29 selon l'année considérée).

Les caractères normatif et répétitif de ces cycles illustrent des rythmes contraints sous leurs formes temporelles : jour/nuit, semaine/WE, été/hiver ou travail/vacances… Plus particulièrement, leur caractère récursif sur plusieurs années laisse penser que nous sommes en présence d'un cadre *routinier*. Dans ce sens, la routine telle qu'elle est appréhendée rejoint la définition proposée par Thomas Buhler (2012, p. 137) lorsque celui-ci distingue les termes d'« *habitude* » et de « *routine* » : « *Les routines peuvent être définies comme des actions répétées et exécutées sous le contrôle des habitudes, et non l'inverse. La routine renverrait donc plutôt à une répétition d'actions identiques ou comparables, et surtout aux seules dimensions observables, corporelles et matérielles de l'habitude* ». Appliquées aux cadres de la mobilité et des transports, les routines entrainent une rigidification des emplois du temps (Juan, 2015). De plus, pour Frédéric De Coninck (2015, p. 5), la mobilité quotidienne est « *un lieu de construction central des routines* », dont les temporalités sont rythmées par nos emplois du temps quotidien et hebdomadaire.

Ici, ce sont bien les activités régulières qui ont été mises en avant par les trois cycles présentés au sein des deux figures précédentes. Plus précisément, ces trois périodes mettent en avant les activités « *routinières liées à la structuration institutionnelle* » (Juan, 2015, p. 9). Sous ce terme, on regroupe l'ensemble « *des activités régulières, pour soi ou la famille, qui est concerné par cette routinisation de la mobilité quotidienne. Ces dernières couvrent aussi bien les déplacements relatifs au soin des personnes, que d'approvisionnement, d'accompagnement ou pour le travail, les activités culturelles, de plaisir ou sportives ».* Les données de comptage illustrent bien ces différents motifs de déplacement sans pouvoir y apporter de distinction sémantique, et à un niveau que l'on peut qualifier de macro, c'est-à-dire que les événements plus ponctuels (donc non-routiniers) ne peuvent être appréhendés aussi aisément par les comptages routiers. Ainsi, si les données de capteurs ne peuvent apporter d'information sémantique au même titre que les enquêtes dont nous avons discuté les écueils au cours de la partie 2.1.2 (p. 58), le corpus de données que nous avons à disposition permet de révéler des informations, auxquelles le système de collecte n'est pas destiné. Pour terminer, les comptages routiers sont donc bien le reflet de différents rythmes urbains, dont les signatures moyennes (en seconde partie de la figure 1.40) montrent clairement les effets de plusieurs rythmes de déplacements imbriqués, au premier rang desquels on peut associer les rythmes domicile-travail et domicile-école. Par exemple, pour le cycle annuel, les baisses de trafic pour les mois de janvier, février, avril, juillet et août apparaissent être soumises aux rythmes des vacances scolaires. De même que les alternances hebdomadaires entre semaine et WE sont une traduction « combinée[61]» des jours de travail et des jours d'ouverture des écoles.

A ces trafics, il faut ajouter l'ensemble des déplacements dits de « transit » que les comptages routiers permettent d'appréhender. L'ensemble des informations issues des stations de comptages de la ville de Dijon offrent une photographie chronologique des rythmes urbains sans cadre qualitatif à proprement dit. Sous cette acception, la base de données mise en place apparait intéressante à mobiliser comme un dispositif de quantification adéquat pour la représentation du trafic routier sur les axes les plus structurants de l'agglomération. Au final, les capteurs fixes révèlent, au même titre que tous les capteurs urbains, comme les mobiles (Fen-Chong, 2012), des informations permettent autant de mesurer que d'appréhender les rythmes diurnes en s'inscrivant

[61] Nous préférons utiliser le terme de « combinaison », car pour les déplacements scolaires, on pourrait s'attendre à une baisse du trafic pour la journée du mercredi, qui peut être atténuée, d'une part par le trafic domicile-travail et, d'autre part, par le trafic de loisirs du mercredi après-midi.

dans une logique spatio-temporelles des temporalités urbaines quotidiennes (Mille, 2000). Toutefois, il faut ici préciser les défis à relever pour une représentation complète du trafic routier sur l'ensemble du réseau, car en l'état ces données permettent seulement de visualiser le trafic routier en différents points du réseau routier.

### 3.3.3. **Les défis à relever pour l'étude spatio-temporelle du trafic routier**

Le premier défi mis en avant précédemment est de détourner l'usage des données de régulation des feux tricolores pour les agencer dans une perspective de simulation du trafic routier. Le travail présenté vise à passer d'un usage dit « opérateur » vers un usage de « modélisation » via l'observation du trafic en temps quasi-réel. La construction d'une chaîne de travail orientée *objet* et décrite précédemment tend à nous montrer un potentiel d'exploitation des données de comptage routiers pour la représentation des rythmes de la circulation en milieu urbain. Le développement précédent nous a permis d'envisager le réseau de capteurs comme un dispositif de quantification (Commenges, 2013, p. 18) en convenant d'une part des différents *objets* associés aux stations de comptage et formalisés sous la forme d'une base de données de trafic (le réseau routier, les stations de mesures, les véhicules). Puis, en entrant directement dans la donnée, pour en expliciter la *mesure numérique* exprimée.

Le second défi est de proposer une approche géographique pour la simulation du trafic routier. Ici nous soulevons un problème éminemment géographique afin de passer d'une perception ponctuelle des conditions de circulation à une vision dynamique de celles-ci sur l'ensemble d'un réseau routier. La volonté est de passer du point à la ligne, plus précisément de la station de comptage au réseau routier, en se formalisant sous la forme d'une problématique eulérienne. Issue du domaine de la science physique pour l'étude du mouvement des fluides et des particules[62], l'approche eulérienne est centrée sur le milieu dans lequel évolue une particule, un *objet* ou des individus. Cette approche est généralement opposée à celle de Lagrange qui se focalise plutôt sur le suivi du mouvement en étant centrée sur la particule, l'individu. De nature, ces deux approches reviennent à opposer les capteurs mobiles qui appartiennent aux approches lagrangiennes (Fen-Chong, 2012 ; Buard, 2013) aux capteurs fixes focalisés en un point donné et qui se rattachent quant à eux aux approches eulériennes.

Les enjeux liés aux problématiques eulériennes se rapprochent de celles liées à la représentation du trafic routier depuis les stations de comptage routier. En effet, les stations de mesure permettent de considérer pendant une tranche de temps donnée les flux de circulation passant sur une infrastructure routière. Dans le cadre d'une démarche de modélisation, par nature, la méthode des comptages est spatialement incomplète, imprécise et incertaine (De Runz, 2008) : incomplète car elle ne permet pas de relever le trafic routier sur l'ensemble du réseau de mesure ; imprécise et incertaine car les comptages ne permettent pas d'apporter des informations sémantiques sur l'origine ou la destination des flux de trafic routier. Pour pallier à ces différentes limites, l'approche que nous envisagerons d'explorer au cours de la partie suivante vise à se rapprocher du domaine de la *géocomputation* et des méthodes associées à l'Intelligence Artificielle, et plus particulièrement des outils de géosimulation. En nous reposant d'une part sur la richesse des différents comptages (ponctuels et permanents) présentée au cours des sections précédentes,

[62] Voir le « *Chapitre 3 : Cinématique des fluides* » pp. 52 à 90 de l'ouvrage de Guyon, E., Hulin, J.-P., Petit, L., Hinch, J., 2012. *Hydrodynamique physique*, 3e édition. ed. Coédition CNRS Éditions, Les Ulis, 694p.

et en combinant, d'autre part ces informations à une représentation fine du réseau routier de la ville de Dijon, nous souhaitons explorer les potentiels d'utilisation des capteurs fixes pour la simulation du trafic routier. L'approche que nous proposons vise à replacer les données de capteurs fixes au premier plan de la démarche de simulation, en assumant non seulement, un positionnement ancré dans la géographie, mais aussi dans une perspective empirique dont nous avons discuté des enjeux au cours de la partie 1.2 (p. 35).

Enfin, le troisième défi se situe au niveau de la représentation du trafic routier à une échelle fine de temps et du cycle le plus caractéristique des PAA (*cf.* section p. 1.1.3, p. 31), le cycle journalier. Cette dernière étape constitue l'enjeu opérationnel du travail présenté. Dans quelle mesure sommes-nous capable de mobiliser les données de capteurs fixes ? En effet, les stations de comptage montrent d'ores et déjà leurs avantages pour une représentation des conditions locales de circulation à une granularité temporelle fine (*cf.* figure 1.40, p. 111). D'autant plus que les données de comptages sont largement exploitées pour les études d'impact de la PAA par les AASQA[63] (PCIT, 2012), mais à un niveau d'agrégation (la journée) ne permettant pas une représentation foncièrement dynamique du système de transport urbain.

[63] Les AASQA (Associations Agréées de la Surveillance de la Qualité de l'Air) assurent la surveillance de la qualité de l'air en France. Elles sont en charge de la surveillance, l'information, la prévision, ainsi que l'information et la compréhension des pollutions atmosphériques à l'échelle locale (région et ville).

## CONCLUSION DE LA PREMIÈRE PARTIE

Cette première partie, assume un positionnement empirique en entrant directement dans les questionnements et enjeux liés à la quantification du trafic routier intraurbain. Les propos de cette partie se sont structurés en trois blocs afin d'explorer les enjeux conceptuels/thématiques du sujet dans un premier temps, empiriques dans un second temps et enfin heuristiques en vue de sélectionner et hiérarchiser les données issues des capteurs fixes. Rappelons ici que l'objectif exploratoire de notre travail se structure non seulement autour de la notion de pollution de proximité, mais aussi sur un travail de terrain sur le trafic routier de la ville de Dijon pour l'estimation des polluants issus des véhicules routiers.

Le premier objectif de cette partie a été de proposer un cadre de réflexion conceptuel en abordant les différents enjeux thématiques de ce travail, liés autant aux dynamiques intraurbaines, qu'à la géographie des transports et à l'étude de la pollution atmosphérique automobile (PAA). En outre, les questionnements épistémologiques sur les approches guidées par les données ont permis de positionner le travail présenté dans une perspective empirique et abductive. En se situant dans ce champ, nous avons tenté de replacer au premier plan le rôle des données pour la représentation conjointe de l'espace et du temps. Nous avons alors positionné les enjeux de notre travail au sein de la *géocomputation* qui se situe à la frontière de trois disciplines (la statistique, l'informatique et la géographie). Ce positionnement n'est pas anodin car il permet de mettre en avant les enjeux contemporains des approches guidées par les données en analyse spatiale (Anselin, 1989). Le principal constat de ce premier développement tend à montrer le rôle central que jouent l'ordinateur et la donnée pour l'étude des dynamiques spatiales et plus généralement en géographie.

Le second chapitre a, quant à lui, visé à mettre en exergue les enjeux empiriques des dispositifs de mesures du trafic routier (les comptages et l'enquête). À travers ces deux cadres de mesures, nous avons pu explorer la définition du trafic routier qui s'est construite entre, d'une part, une vision sociologique (Lannoy, 2003), et d'autre part, une vision plus techniciste issue de la socio-économie des transports (Commenges, 2013). Cette insertion n'est pas neutre car elle permet de mieux comprendre les usages des données d'enquête et de comptage pour l'étude du trafic routier. Longtemps mises à part, les données de comptage connaissent aujourd'hui un nouvel essor, en particulier, avec le développement des *big data* et des systèmes de mesures y étant associés. Mobiles ou fixes, les systèmes de capteurs appartiennent au champ des *big data* et impliquent aujourd'hui de mieux les exploiter pour en extraire une information géographique. À travers la typologie proposée en figure 1.12 (p. 59), nous avons introduit l'hétérogénéité des données de capteurs rencontrées en milieu urbain. Cette grille de lecture comparative nous a amené à resituer les données issues des capteurs fixes, dont la mise en application depuis les années 60, est ancienne, alors que pourtant leur usage n'est que très peu mis en avant dans la littérature et dans les problématiques actuelles de mesure du trafic routier. Si les données issues des capteurs mobiles connaissent plus particulièrement un renouveau avec la démocratisation des GPS et des *smartphones*, l'intérêt porté par les aménageurs et les décideurs locaux aux données issues de capteurs mobiles montre des potentiels opérationnels de plus en plus importants pour compléter les informations issues des enquêtes classiques (Drevon *et al.*, 2014 ; Gwiazdzinski et Drevon, 2014). Pour autant, les différents leviers éthiques et méthodologiques tendent à nous

montrer que les données issues des capteurs fixes ne sont pas à mettre de côté (Herrera *et al.*, 2009). Elles présentent un autre type de réponse aux problèmes inhérents aux capteurs mobiles, sur plusieurs points tant thématiques et économiques, qu'éthiques. De surcroît, comme nous l'avons vu au sein du tableau 1.4 (p. 78), les données issues des capteurs fixes présentent des atouts qui portent autant sur leur collecte, la cible qu'ils mesurent, l'échelle d'observation appréhendée, leur temporalité, que sur la mesure qui leur est associée. Enfin, on peut insister sur la tendance émergente des *smart cities* qui porte sur l'intégration et la diffusion de plus en plus forte de la multitude des données issues de capteurs à destination du secteur public et des usagers, ce qui permet de concevoir l'information autrement que par la sphère économique. Sous cet angle, les réseaux de capteurs fixes ne sont pas morts et ils peuvent servir de parangon pour reprendre la main sur la diffusion et la production de nouvelles informations spatio-temporelles tournées vers de multiples thématiques, portant par exemple autant sur le transport, que l'énergie ou la consommation.

À partir de l'analyse présentée au cours du troisième chapitre, l'exploration des données de comptage de la ville a permis, sous un angle heuristique, d'explorer les potentiels d'exploitation et de mobilisation des données issues du réseau de comptage routier de la ville de Dijon. Ce système de mesure dédié initialement à la gestion de la circulation et des feux de signalisation ouvre la porte à une approche plus générale pour l'étude de la circulation routière. Proche de l'*objet* de la mesure (le trafic routier), il présente une exploitation limitée pour la compréhension qualitative des déplacements routiers. Toutefois, en reposant sur les caractéristiques quantitatives et temporelles du réseau de mesures, il ouvre la voie à une approche spatio-temporelle du trafic routier plus générale, en vue de reproduire les « palpitations » de l'espace urbain. C'est au sein de ce dernier chapitre que le dispositif de quantification mobilisé prend tout son sens pour la mise en place d'un processus d'entitisation du trafic routier. En effet, la mesure d'un phénomène dans le temps nécessite de construire les *objets* à partir de ce qui est observable et observé depuis les jeux de données, afin de leur donner un sens aussi bien dans le temps que dans l'espace (Mathian et Sanders, 2014).

À travers l'exemple de la ville de Dijon, c'est cet usage dit descriptif par les données de comptage que nous avons souhaité questionner tout au long de cette partie. Il apparait intéressant d'exploiter ces données en vue de s'affranchir de l'usage des modèles habituellement utilisés dans les études d'impact du trafic routier (modélisation à quatre étapes : encart 4, p.136). La deuxième partie de la thèse se poursuit en se focalisant plus spécifiquement sur le second dispositif de la matrice technique (Commenges, 2013) : le dispositif de modélisation. Cette seconde partie s'insère dans la continuité des réflexions menées autour du dispositif de quantification qui alimente, façonne et conditionne la construction du dispositif de modélisation. En effet, les enjeux liés au rapprochement entre les données et l'observation des dynamiques urbaines ne constituent qu'une première clé d'entrée du travail de modélisation et de simulation du trafic routier. L'idée est d'explorer les aspects techniques de la modélisation du trafic routier initiée au cours des années 70 par cette interdépendance entre données et modélisation, permettant d'une part d'expliquer la stabilité des modèles de déplacements urbains issus de la socio-économie des transports et dont « l*es principales limites sont maintenant humaine et technique »* (Ortúzar et Willumsen, 2011, p. 1), et d'autre part, de situer les enjeux de la démarche de simulation entreprise en rapprochant plus fidèlement dispositif de quantification et dispositif de modélisation.

# PARTIE 2 -

# DE LA MODÉLISATION À LA SIMULATION MULTI-AGENTS DU TRAFIC ROUTIER POUR LA PAA : UNE APPROCHE ORIENTÉE SUR LES DONNÉES


INTRODUCTION DE LA DEUXIÈME PARTIE ..... 121

CHAPITRE 4. MODÉLISER ET SIMULER LE TRAFIC ROUTIER EN MILIEU URBAIN POUR LA QUALITÉ DE L'AIR : DÉFIS ET ENJEUX ..... 125

*4.1. Un retour sur les concepts de modèle(s) et de simulation en sciences humaines ..... 127*

*4.2. Quel modèle de trafic retenir pour l'étude de la PAA ? ..... 133*

*4.3. Les enjeux liés à une approche de simulation du trafic routier à partir des comptages routiers ..... 141*

*Conclusion : ..... 150*

CHAPITRE 5. LES APPORTS DE LA SIMULATION MULTI-AGENTS POUR L'ÉTUDE DU TRAFIC ROUTIER ..... 151

*5.1. Le renouveau de la modélisation du trafic routier par la simulation-multi-agents ..... 152*

*5.2. Les approches de géosimulation pour la représentation du trafic routier : quelle méthode privilégier ? ..... 159*

*5.3. La SMA pour la modélisation du trafic routier à partir des capteurs de trafic routier : le modèle SCAUP ..... 165*

*Conclusion : ..... 182*

CHAPITRE 6. EXPÉRIMENTATION ET MISE EN MOUVEMENT DU TRAFIC ROUTIER PAR LA SIMULATION MULTI-AGENTS ..... 183

*6.1. Les plates-formes de simulation multi-agents : ..... 184*

*6.2. Formalisation du modèle de simulation multi-agents : du projet à l'expérimentation de SCAUP ..... 190*

*6.3. Expérimentation et plan d'expérience du modèle SCAUP : ..... 205*

CONCLUSION DE LA DEUXIÈME PARTIE ..... 217

# INTRODUCTION DE LA DEUXIÈME PARTIE

Au cours de la précédente partie, nous avons entrevu que les données de comptage constituent un outil adéquat pour la représentation du trafic routier. En revanche, la mobilisation de ces données en l'état ne permet pas de pouvoir rendre compte des flux de circulation urbains sur l'ensemble des axes routiers de la ville de Dijon. Il faut donc généraliser l'information issue des capteurs sur les infrastructures routières. C'est pourquoi, dans la plupart des applications nécessitant l'usage de données produites par des capteurs, la place prise par les modèles dans leur ensemble[64] n'est pas négligeable en vue de « *décrire, comprendre, et anticiper les phénomènes qui se déroulent dans le monde réel* » (CGDD, 2010, p. 3).

Les études relatives au trafic routier constituent un des premiers domaines où l'usage des modèles a été largement employé dans une logique de planification et de quantification des volumes de déplacements. Les investissements pour le développement d'une infrastructure de transport routier étant lourds, il a été nécessaire de s'entourer d'outils pour anticiper et justifier le développement de nouvelles routes (Dupuy, 1975). À titre d'exemple, un audit du conseil général des ponts et chaussées de 2006 (Cazala *et al.*, 2006) estime que le prix moyen au kilomètre de la construction d'une route 2x2 voies est de près de 5 millions d'euros/km[65]. On comprend aisément que le besoin d'anticiper et de quantifier les investissements routiers apparait primordial. C'est dans cette optique que la modélisation du trafic routier à quatre étapes s'est imposée comme un outil de référence. Néanmoins, en privilégiant une approche sur le long terme (*cf.* figure 2.3, p. 134), ces modèles fournissent une approche globale des déplacements qui est ensuite décomposée par mode de transport, pour en construire des volumes de trafic routier. À cet égard, avant de pouvoir rendre compte du trafic routier, la démarche de modélisation s'attèle d'abord à reconstruire les déplacements des individus dans l'espace urbain à partir de données socio-économiques (Bonnel, 2002). C'est aussi pour cela que dans la littérature, on parle d'abord d'un modèle de transport (Meunier, 2013), voire de mobilité quotidienne (Commenges, 2013), avant de parler d'un modèle de trafic à proprement dit.

Les objectifs visés par ces modèles couvrent aussi bien ce pourquoi ils ont été construits, le dimensionnement et le développement des infrastructures de transport, que d'autres champs d'application, comme par exemple la pollution atmosphérique ou sonore. Pour la pollution atmosphérique, ces modèles de trafic dits « *classiques* » se définissent aujourd'hui comme un référentiel pour la construction des cadastres d'émissions relatives à la circulation routière (PCIT, 2012). Toutefois ces modèles sont une bonne illustration de l'éloignement entre ce qui est observé à partir d'enquêtes de déplacement et ce qui en est modélisé (*cf.* section 2.1.2, p. 58). De plus, de manière plus large, l'ensemble de ces modèles repose sur une formalisation qui n'a que très peu évolué conceptuellement depuis la création des premiers modèles de trafic (Bonnel, 2010). En effet, c'est à travers un cadre conceptuel en quatre étapes que le trafic routier est construit puis restitué par modélisation. Ces derniers propos constituent l'une des critiques des modèles de trafic *classiques*, considérés comme un outil « *fourre-tout* » pour des applications très

[64] Ici nous englobons sous le terme de modèle aussi bien les modèles conceptuels, que de données ou de simulation ; néanmoins, nous explorerons plus en détails les notions de *modèle* et de *démarche de modélisation* au cours de la section 4.1.1 (p. 119).

[65] L'estimation moyenne des investissements routiers est établie sur 35 projets d'infrastructures 2x2 voies réalisés entre 1997 et 2002 présentée en annexe 2 du rapport de l'audit (Cazala *et al.*, 2006, pp. 32–34).

diversifiées. C'est particulièrement le cas lors de leur utilisation à des fins d'évaluation de la pollution atmosphérique locale (Fouillé *et al.*, 2012). C'est pourquoi rapprocher la donnée et le modèle peut constituer une bonne réponse à ces critiques pour une meilleure estimation des émissions de PAA en milieu urbain.

Pour Arnaud Banos (2013), les modèles en géographie doivent mieux appréhender la donnée dans le sens où leur vocation vise à reproduire des processus et structures observés. Les bases de données nécessaires pour la simulation d'un système dynamique requièrent la mobilisation de données inscrites autant dans le temps que dans l'espace. Le développement des données issues des capteurs ainsi que l'avènement des approches de simulation en sciences humaines (Amblard et Phan, 2006) constituent des potentiels d'utilisation de plus en plus multiples et variés pour la modélisation des dynamiques intraurbaines. Avec les méthodes de géovisualisation ou de géoexploration, les méthodes de géosimulation constituent l'un des moyens en vue d'explorer et de représenter une information spatio-temporelle à partir du dispositif de quantification (Thévenin, 2011). De surcroît, ces outils se rattachent aux méthodes guidées par les données que nous avons rattachées plus haut à la *géocomputation* (*cf.* section 1.3, p. 45).

Pour notre problématique, ces différents outils de modélisation à disposition du chercheur offrent, en outre, l'opportunité de passer de l'observation à une information spatio-temporelle (Mathian et Sanders, 2014). Ainsi, après nous être intéressé au dispositif de quantification pour la mesure du trafic routier, nous nous focaliserons plus spécifiquement dans cette partie sur les dispositifs de modélisation du trafic routier. Le travail présenté repose sur une démarche de géosimulation empirique du trafic routier en milieu urbain. Par nature générative (Epstein, 1999 ; Benenson et Torrens, 2004), la démarche de géosimulation « *combine les apports de l'intelligence artificielle et des systèmes d'information géographique pour simuler la dynamique de système* » [66]. Cette approche géographique repose sur quatre éléments (Daudé, 2004 ; Thévenin, 2011) :

- 1. Une démarche générative en faisant le lien entre interactions locales et globales ;
- 2. Une simulation microscopique à l'échelle du véhicule ;
- 3. Des interactions d'ordre hiérarchique ;
- 4. Une appréhension du temps synchrone (diffus) et asynchrone (par cycle).

Ces quatre éléments restreignent d'autant plus le choix des données nécessaires à la modélisation du trafic routier. Dans une logique de simulation multi-agents, l'approche guidée par les données (*data-driven*) doit contribuer à obtenir des résultats de simulation qui sont plus proches de l'observation (Hassan *et al.*, 2010b). Améliorer l'adéquation entre le modèle de trafic et le trafic observé par des dispositifs de mesures adéquats présente donc une véritable opportunité pour les géographes en vue de reproduire les déplacements des véhicules sur le réseau routier. Le concept de *data-driven* en géographie a été initié en réponse aux flux massifs de données issus des capteurs et des individus inscrits dans l'espace (Miller et Goodchild, 2014). Cette interdépendance entre données et simulation montre l'intérêt d'une approche empirique de simulation du trafic routier. La donnée n'est plus seulement utilisée pour la validation, mais elle permet aussi de concevoir et d'initialiser le modèle de simulation dans la perspective d'évaluer les polluants atmosphériques automobiles (PAA).

[66] Lena Sanders : définition de **« géosimulation » :** http://www.hypergeo.eu/spip. php?article311.

Dans ces conditions, en vue d'obtenir un instrument de mesure plus précis pour l'évaluation de la PAA, quelle(s) approche(s) mobiliser et envisager pour représenter le trafic routier à partir des données de comptages routiers ? Il apparait important de relever certains questionnements qui orientent les choix méthodologiques pour la démarche de modélisation qui sera mise en place: quelles sont les approches existantes ainsi que leurs évolutions récentes pour la modélisation du trafic routier ? Quelle place prennent les données de capteurs dans le processus de construction et de modélisation du trafic routier ? Enfin quelles sont les pistes méthodologiques et les moyens permettant de prendre en compte les données issues des capteurs urbains ? L'ensemble de ces questionnements nous permettra ensuite d'expliciter l'ensemble du processus de simulation multi-agents du trafic routier entrepris au cours de cette thèse pour l'évaluation de la PAA à travers l'approche de modélisation proposée ici *via* le modèle SCAUP (*Simulation multi-agents à partir des CApteurs Urbain pour la Pollution atmosphérique automobile*).

# Chapitre 4. MODÉLISER ET SIMULER LE TRAFIC ROUTIER EN MILIEU URBAIN POUR LA QUALITÉ DE L'AIR : DÉFIS ET ENJEUX

L'évaluation des impacts environnementaux du trafic routier en milieu urbain s'inscrit dans les problématiques de la planification urbaine (CGDD, 2012). En effet, depuis les lois du Grenelle I et du Grenelle II[67], le cadre réglementaire oblige les collectivités territoriales à se doter d'outils pour mesurer et évaluer les impacts de la circulation routière sur l'environnement urbain. L'établissement des Plans de Déplacements Urbains (PDU) constitue une bonne illustration du lien entre les problématiques de la qualité de l'air, de l'urbanisme et du transport routier, car ils visent à « *assurer un équilibre durable entre les besoins en matière de mobilité et de facilité d'accès, d'une part, et la protection de l'environnement et de la santé, d'autre part* »[68]. Pour cela, il est nécessaire d'apporter une quantification des niveaux de la pollution atmosphérique automobile sur l'ensemble de l'espace urbain. Préalablement, il est indispensable de disposer des débits de véhicules présents sur les infrastructures routières pour ensuite alimenter le calcul des émissions de polluants atmosphériques. Cette tâche relève de la compétence des associations de qualité de l'air locales (AASQA) qui privilégient « *les données issues d'un modèle de trafic si elles existent, en combinaison des données d'observation* » et plus particulièrement les modèles à quatre étapes (PCIT, 2012, p. 168).

Pourtant, force est de constater que la modélisation à quatre étapes, en l'état, n'apparait pas la plus adaptée pour fournir une image fidèle de la circulation routière urbaine (Fouillé *et al.*, 2012). Il est d'usage de retravailler les sorties issues du modèle de trafic. Le guide méthodologique[69] pour l'élaboration des inventaires territoriaux des émissions atmosphériques (PCIT, 2012) illustre bien ces écueils en incitant le modélisateur à mobiliser les données de comptages routiers pour s'assurer de la validité des données modélisées autant sur les plans temporel que spatial. À ce niveau, les données issues des capteurs électromagnétiques s'avancent comme un dispositif de référence en vue de valider et de caler les données modélisées. Valeurs de référence pour la quantification du trafic routier (PCIT, 2012), ces données ne sont que rarement valorisées pour la modélisation du trafic routier. Dans l'approche que nous proposons, les données sont placées au cœur du dispositif de modélisation, les rendant indispensables pour la simulation du trafic routier.

Initialement développée pour une seule application, la modélisation du trafic a progressivement évolué vers des usages très diversifiés s'éloignant « *de la mission de dimensionnement des infrastructures à laquelle ils répondaient initialement* » (Fouillé *et al.*, 2012, p. 199). Malgré toutes ses limites, la modélisation à quatre étapes dite « *classique* » reste dominante pour l'évaluation des politiques environnementales (Masson, 2000 ; Fouillé *et al.*, 2012). En effet, au contraire des modèles de simulation du trafic, comme SUMO (Behrisch *et al.*, 2011) ou VISSIM (Fellendorf et Vortisch, 2010), qui ne traitent qu'une portion relativement réduite du territoire (un carrefour à feux tricolores, une intersection ou un quartier), l'approche à quatre étapes présente un avantage indéniable pour couvrir un espace d'étude plus large que ce qui est permis à travers la simulation

[67] http://www.developpement-durable.gouv.fr/-Le-Grenelle-de-l-environnement-de-.html.
[68] Loi n° 96-1236 du 30 décembre 1996 sur l'air et l'utilisation rationnelle de l'énergie (LAURE) http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000005622536&dateTexte=vig.
[69] Ce document rédigé par le Pôle national de Coordination des Inventaires Territoriaux (PCIT) est le guide méthodologique de référence employé par les AASQA pour l'élaboration des inventaires d'émissions atmosphériques ainsi que des cadastres d'émissions de polluants atmosphériques à l'échelle locale (PCIT, 2012).

du trafic routier, qui est limitée à des zones plus restreintes. Si le modèle à quatre étapes présente un formalisme avantageux, le trafic routier issu de cette approche n'est pas donné, il est plutôt construit en « *utilisant des hypothèses sur la demande des usagers potentiels du système de transport et des hypothèses sur l'offre de transport, pour simuler le comportement des usagers et obtenir des estimations de trafic résultant de l'interaction entre offre et demande* » (Meunier, 2013, p. 5). Actuellement, cette conception demeure encore majoritaire. Toutefois, la diversité des approches montre une évolution des outils et des méthodologies employés pour la modélisation du trafic routier, au premier rang desquelles les approches microscopiques et dynamiques constituent sans doute les avancées les plus caractéristiques.

En portant un regard sur cette différenciation entre l'approche classique et la simulation du trafic routier, ce chapitre permettra d'illustrer la diversité des applications qui s'inscrivent actuellement dans le champ de la modélisation du trafic routier, en distinguant d'une part les modèles macroscopiques des modèles microscopiques, et d'autre part les modèles statiques des modèles dynamiques. Ce chapitre illustrera également l'éloignement entre l'*objet* de mesure et l'*objet* modélisé, éloignement qui constitue un des constats de départ de cette thèse. Dans quelle mesure les modèles de trafic actuels peuvent-ils répondre aux problématiques de la pollution atmosphérique automobile ? Dans quel cadre méthodologique (re)valoriser les données de comptages routiers au sein de ces approches ? Quelle approche de modélisation spatio-temporelle du trafic routier envisager ?

## 4.1. Un retour sur les concepts de modèle(s) et de simulation en sciences humaines

En Géographie, le modèle occupe une place prépondérante dans l'analyse spatiale, non seulement avec ses outils, mais aussi dans son ancrage thématique (Anselin, 1989 ; Daudé, 2005 ; Mathian et Sanders, 2014). La géographie étudie ce qui est observable en se rattachant à l'espace ainsi qu'au temps (Daudé et Langlois, 2006). On peut aussi relever ici que le premier domaine où les modèles se sont développés porte généralement sur des problématiques liées aux transports et aux déplacements (Dupuy, 1975). La polysémie du terme de « *modèle* » implique de s'y arrêter pour en délimiter les contours, tant la notion de modèle peut être mobilisée dans des cadres et des contextes très différents. Ici, sous le terme de modèle, nous nous attacherons plus à nous insérer au sein de la « *démarche de modélisation* » plutôt qu'à la signification du terme de « *modèle* » en lui-même. En effet, l'étude d'un système quelconque implique la mise en place d'un modèle au sein duquel on mobilise différents sous-modèles (Sanders, 2001a), ce qui est le fruit d'une démarche de modélisation (Chorley, 1964) où à chaque étape on mobilise des modèles de différente nature : le modèle conceptuel, le modèle de données, le modèles de simulation…

### 4.1.1. Un « *modèle de modèles* » pour construire une démarche de modélisation

Dans son sens le plus général, le modèle, se définit comme étant « *ce qui sert d'objet d'imitation pour faire ou reproduire quelque chose* » : or le sens scientifique qui lui est inféré est plus précis, voire contraire, et se définit comme « *une représentation simplifiée d'un processus, d'un système* » [70]. D'ores et déjà, la définition distingue deux types de modèles : les modèles statiques « *qui sont les représentations d'un existant quelconque* », comme le système urbain ; et les modèles dynamiques « *dont le fonctionnement permet de simuler des processus* »[71]. Qu'il soit statique ou dynamique, le modèle est à associer à une démarche de construction de la connaissance (*cf.* section 1.2.1, p. 36, et à droite de la figure 2.1) au sein de laquelle on manipule des modèles aux différentes étapes de la démarche scientifique (Ramat, 2006).

À travers l'ensemble de la littérature portant sur les modèles, la première définition qui semble faire consensus pour la Géographie vient de Peter Haggett (1965, p. 19) : « *in model we created an idealized reresentation of reality in order to demonstrate certain of its properties* », dont la définition qui en est la plus couramment retenue est « *a simplified version of reality, built in order to demonstrate certain of the properties of reality* ». De ce point de vue, le terme renvoie à un *objet* ou plusieurs *objets réels* auquel on tente de donner une représentation plus ou moins simplifiée en fonction de ce qui parait faire sens. La simplification telle qu'elle est envisagée n'est pas à entendre dans le sens d'une réduction du phénomène à étudier. Par exemple, en s'appuyant sur la théorie des systèmes complexes pour l'étude des dynamiques intraurbaines, le but est de fournir un modèle aussi formalisé que possible pour en faciliter sa reproduction (*cf.* section 1.1.1.b, p. 26). Plus qu'une simplification, le terme renvoie plutôt à un processus d'abstraction et de mise en hypothèses des éléments à mettre en place pour l'étude d'un phénomène (Batty, 2007 ; Varenne, 2008).

Le modèle permet ainsi de donner une représentation structurée, et souvent hiérarchisée, du système à étudier. Il s'avance donc comme un outil d'aide à la réflexion pour étudier un

[70] Définition de « **modèle** » issues du Petit Robert 2006.
[71] Henri Chamussy : définition de « **modèle »** : http://www.hypergeo.eu/spip. php?article9#.

phénomène dont l'objectif fondamental est de proposer une démarche face à une problématique donnée où s'agencent des fragments du monde *réel* arbitrairement isolés (Haggett, 1965). C'est pourquoi, le modèle s'organise avant tout comme un instrument de construction de la connaissance permettant progressivement de sélectionner et d'identifier les composants élémentaires de la construction d'une *démarche de modélisation* au sein de laquelle on manipule différents modèles (Sanders, 2001a). En s'inspirant de la grille de lecture proposée par Richard J. Chorley (1964, p. 129), il est possible de formaliser la *démarche de modélisation* en trois grandes étapes illustrée au sein de la figure 2.1 : **l'abstraction, la formalisation et l'interprétation**.

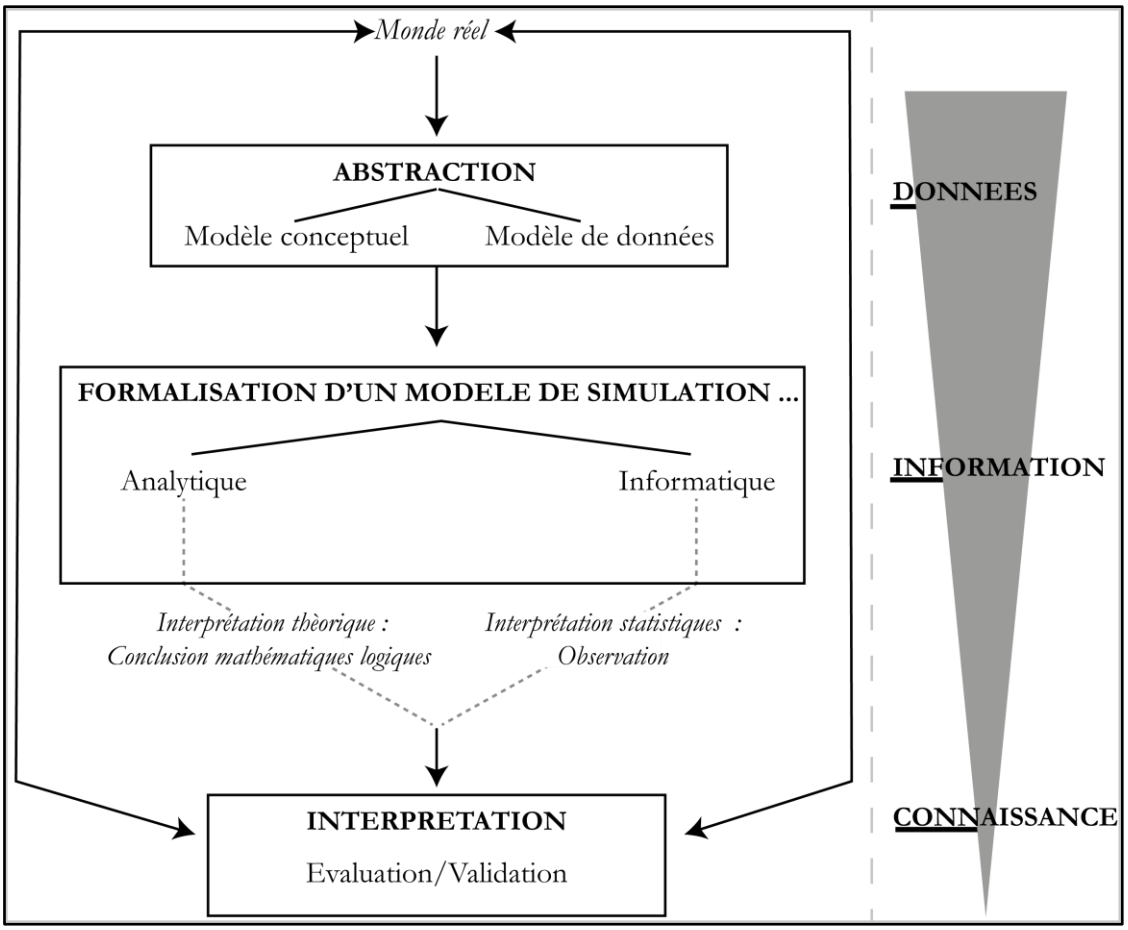



**Figure 2.1 : La démarche de modélisation : un processus de formalisation en trois étapes (D'après Chorley (1964, p. 129) : « *Un modèle pour les modèles* »)**

L'**abstraction** correspond à une première phase de conceptualisation et de simplification du monde réel ou d'une partie du monde réel à étudier. Cette première étape permet d'identifier les *objets* élémentaires à conceptualiser et à recueillir sous la forme d'un modèle conceptuel et d'un modèle de données. Cette étape constitue un prérequis au traitement de l'information spatiale sous une forme thématique constituée d'une part d'un modèle conceptuel et d'autre part d'un modèle de données (Sanders, 2001a). Cette première partie d'abstraction d'un modèle conceptuel et d'un modèle de données a été abordée tout au long de la première partie, qui nous a permis de convenir d'une définition du trafic routier afin d'en identifier ses *objets* élémentaires : les *véhicules* et l'*infrastructure* de transport ; ainsi que d'identifier les enjeux liés à la constitution empirique d'un dispositif de quantification issu des données de comptage (*cf.* Chapitre 3, p. 81).

La **formalisation** constitue la seconde phase de la démarche de modélisation. Ici, nous nous restreignons à la formalisation des *modèles de simulation,* qui permettent la « *mise en action d'un modèle qui s'accompagne de l'observation du comportement de ce dernier* » (Ramat, 2006, p. 51). En d'autres termes, la simulation permet la « *computation* » d'un modèle, et ce « *quels que soient le principe et le type de modèle* » (Varenne, 2013, p. 39). Deux grandes approches de simulation se distinguent[72] (*cf.* figure 2.1) : les approches *analytiques,* reposant traditionnellement sur un ensemble de fonctions et d'équations finies décrivant le phénomène à modéliser ; et les approches *informatiques* permettant de décrire le modèle à étudier en tentant de reproduire physiquement ce qui fait sens. Les modèles sont le plus souvent construits à travers un langage informatique permettant la reproduction d'un « monde réel ». L'exemple le plus caractéristique de cette typologie est celui de la simulation multi-agents construite à travers les plates-formes qui permettent de définir des modèles à l'aide d'un langage de modélisation dédié[73] (Ferber, 1995). La formalisation permet la mise en action du modèle (Chorley, 1964). Dans le cas d'une approche mathématique celle-ci est appréhendée par une résolution à base d'équations définissant le comportement du modèle, et ne faisant pas nécessairement appel à l'utilisation d'un ordinateur ; alors que dans le cas d'un modèle informatique, il s'agit plutôt d'une résolution algorithmique, et qui se fonde sur l'utilisation systématique d'une infrastructure informatique.

La dernière phase est celle de l'**interprétation**, permettant par l'observation de tirer des conclusions adaptées au monde réel, et qui peut s'opérer soit par une interprétation théorique dans le cas d'un modèle mathématique, soit par une interprétation statistique dans le cas d'un modèle expérimental (*cf.* figure 2.1). L'interprétation constitue une étape primordiale afin d'évaluer la capacité des modèles à se confronter au monde réel. L'ensemble de ces étapes s'inscrit dans un système fermé, où les issues du processus de modélisation sont ensuite confrontées au monde réel et inversement. Comme nous le verrons plus bas, cette dernière phase n'est pas cantonnée à une simple validation des sorties du modèle mais bien à un processus d'évaluation du modèle visant à apporter une vision moins restrictive de son interprétation (*cf.* Chapitre 7, p. 223).

Comme évoqué précédemment, pour nous le terme de *modèle* fera référence à un cadre de réflexion facilitant l'étude d'un phénomène réel, et qui, comme nous l'avons vu pour la donnée, s'épure au fur et à mesure de la démarche (*cf.* figure 2.1). Cette épuration fait référence au processus de simplification en plusieurs phases qui a été mis en avant précédemment. Ensuite, il faut être en mesure de le formaliser en vue de lui donner un sens intelligible pour tous (Marilleau, 2006). C'est bien dans ces deux dernières affirmations que pour nous le modèle se définit avant tout comme un cadre de réflexion qui n'est pas cantonné à un seul modèle, mais bien étendu à des modèles agencés et juxtaposés ensemble pour constituer un cadre de construction de la connaissance. En d'autre termes, la définition que nous retiendrons est que le modèle constitue un support conceptuel fournissant une image hypothétique de la réalité simplifiée et rationnelle fournissant au chercheur « *une source d'hypothèses de travail à mettre à l'épreuve de la réalité* » (Haggett, 1965, p. 19).

[72] À noter que Richard J. Chorley (1964, p. 129) y ajoute les modèles dits « naturels » ; ici nous avons fait le choix de seulement distinguer les formalisations dites « analytiques » de celles dites « informatiques ».

[73] Les plates-formes orientées agents comme NetLogo (Tisue et Wilensky, 2004) ou GAMA (Taillandier *et al.*, 2014) offrent des langages de modélisation dédiés permettant de faciliter la construction de modèle à une plus grande diversité d'utilisateurs.

Au final, il en ressort que le « *modèle »* est plutôt le résultat d'une démarche intellectuelle et normative, d'un raisonnement, où chaque stade de la démarche de modélisation correspond à une opération traitant de certains aspects du monde réel. Le « *modèle* » permet alors de bénéficier d'un cadre de réflexion simplifié, construit et formalisé à partir d'un « *système réel* » à étudier (Haggett, 1965). Appréhender la notion de modèle sous cet angle, nous permettra d'introduire ensuite la notion de « *simulation* » et de « *géosimulation* » qui constitue une des formes d'exploration et d'animation des modèles en géographie et *a fortiori* en sciences humaines. Cela permettra aussi de situer les objectifs et enjeux du travail présentés ensuite en vue d'affecter les données de comptages routiers au sein d'un modèle de simulation multi-agents dans une optique d'évaluation de la PAA. En effet, la démarche de modélisation SCAUP (*Simulation à partir de Capteurs Urbains pour la Pollution atmosphérique automobile*) sera par la suite présentée sous la forme d'un modèle. Cependant, l'approche que nous défendons dans cette thèse constitue avant tout une démarche de modélisation du trafic routier (*information*) à partir de comptages routiers (*données*) destinée à la mesure de la PAA (*connaissance*).

#### 4.1.2. **La simulation pour animer et explorer les modèles : le cas des *modèles de simulation* et de *géosimulation***

Pour les modèles de trafic, nous avons distingué les modèles statiques des modèles dynamiques, ces derniers visant à mettre en mouvement les véhicules sur une ou plusieurs infrastructures de transport. Si les modèles dynamiques constituent une bonne piste d'investigation pour notre travail d'évaluation de la PAA, il existe plusieurs méthodes pour animer le modèle, au premier rang desquels la *simulation* constitue aujourd'hui une approche largement utilisée. Toutefois, en fonction de l'objectif visé, la délimitation du terme de simulation ne couvre pas les mêmes acceptions (Briassoulis, 2000). Ici, nous porterons un regard plus particulier sur les *modèles de simulation* et la *géosimulation*, qui se distinguent d'autres modèles comme les modèles statistiques, les modèles d'interaction spatiale, ou les modèles d'optimisation (Briassoulis, 2000 ; Bouquet *et al.*, 2015).

Le modèle de simulation tel que nous l'entendons se définit avant tout comme étant « *l'animation d'un modèle* » (Bouquet *et al.*, 2015, p. 23). Son objectif fondamental vise à assurer la mise en œuvre d'un modèle afin d'en observer le comportement et les processus (Ramat, 2006). Il se définit, ensuite, à travers le concept de *géosimulation* comme étant l'animation d'*objets* inscrits dans l'espace géographique et permettant l'étude des dynamiques urbaines (*géo*). Sous cet angle, la (*géo*)simulation telle que nous l'appréhenderons constitue une piste d'investigation non seulement conceptuelle et opérationnelle, mais aussi méthodologique pour l'étude et la représentation spatio-temporelles des dynamiques en géographie (Thévenin, 2011). Dans notre cas, le *modèle de simulation* vise à intégrer des données observées pour ensuite les transformer en une représentation structurée à travers différentes techniques permettant de mettre en mouvement les véhicules et les individus dans un espace construit (Marilleau, 2006).

Historiquement, la simulation s'est rattachée à un cas particulier de certains modèles, les modèles physiques et phénoménologiques (Varenne, 2008). Actuellement, le modèle et la simulation sont deux approches bien différenciées. Jacques Lévy et Michel Lussault (2003, p. 926) soulignent ce point et distinguent les deux notions en positionnant la simulation « *à mi-chemin entre la reproduction et le modèle. Dans la reproduction, seul compte la correspondance terme à terme avec la réalité, quel que soit le*

*moyen par lequel elle a été réalisée. Dans le modèle, seule compte le dispositif technique ou mathématique qui a permis d'obtenir une approximation de la réalité. La simulation utilise des relations moins profondes et moins structurelles que le modèle, mais elle se rapproche plus de la réalité que lui. Symétriquement, elle est moins précise qu'une reproduction, mais conserve une cohérence interne assez proche d'un modèle* ». Sous cette seconde définition, on aura plutôt tendance à parler de « *modèle de simulation* » en opposition aux « *modèles de compréhension* » (Varenne, 2008 ; Banos, 2013). Le « *modèle de simulation* » vise à reproduire un système au sein duquel les processus ne sont pas complétement compris (Banos, 2013). Il permet alors l'animation des processus à observer, qui ne sont que partiellement ou peu compris : « *ce qui est simulé, ce n'est plus le modèle à proprement parler, mais directement le réel visé par la simulation* » (Varenne, 2008, p. 23), et plus spécifiquement pour la géographie (Daudé et Langlois, 2006).

Aujourd'hui, d'une manière générale, la simulation repose plus particulièrement sur les outils informatiques. Pourtant, il existe différentes approches de simulation dont certaines ont vu le jour dès le XVIII[e] siècle avec la résolution d'équations différentielles pour l'études des systèmes dynamiques (Gilbert et Troitzsch, 2005) . En sciences humaines, Nigel Gilbert et Klaus Troitzsch (2005) distinguent deux grandes approches de simulation, illustrées ci-dessous (*cf.* figure 2.2) : celle reposant sur les mathématiques et celle reposant sur l'informatique. La figure ci-dessous permet de commenter l'évolution des approches de simulation qui historiquement reposaient majoritairement sur les formalismes issus des mathématiques, à l'aide par exemple des équations différentielles ou de la théorie des jeux. Ces simulations reposent sur un calcul numérique, ou non, d'un ensemble d'équations finies permettant de formaliser le modèle étudié (*formalisme analytique, cf.* figure 2.1). Parallèlement, à partir de la seconde moitié du XX[e] siècle, d'autres approches de simulation se sont développées en s'appuyant plutôt sur une résolution *totalement* numérique du système cible étudié (*formalisme informatique, cf.* figure 2.1) ; c'est particulièrement le cas des modèles de queues, de la microsimulation, ou de la géosimulation :

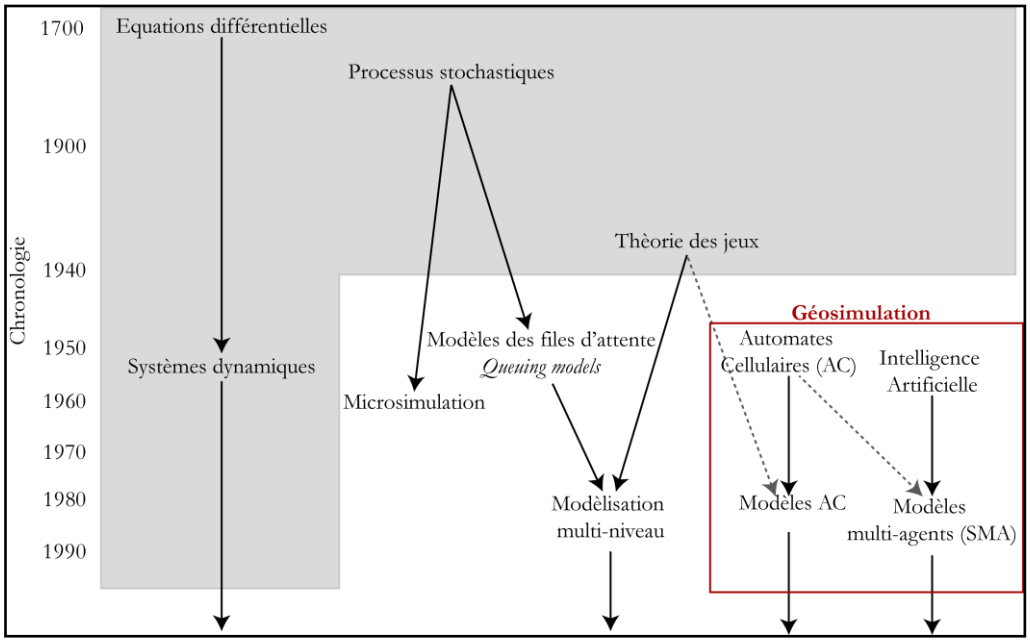



**Figure 2.2 : Chronologie des approches de simulation en sciences humaines : les apports des outils informatiques (d'après Gilbert et Troitzsch, 2005, p. 7)[74]**

[74] La zone illustrée en gris représente les méthodes issues des mathématiques ; tandis que la zone blanche représente les méthodes issues de l'informatique.

Ces deux grandes approches de simulation (*mathématique et informatique*[75]) permettent d'évaluer le comportement d'un modèle et tiennent plutôt compte des choix faits par le chercheur (Ramat, 2006). L'objectif de notre démarche n'est pas de décrire l'ensemble des approches de simulation[76], mais plutôt de positionner la géosimulation par rapport aux autres approches en s'appuyant sur la figure 2.2. En effet, les approches de géosimulation ont vu le jour plus tardivement, à partir des années 60, grâce aux apports de l'intelligence artificielle et des automates cellulaires issus de l'informatique.

En outre, les approches de géosimulation permettent aussi de proposer un formalisme construit en dehors du cadre de la modélisation *classique*, par exemple, en reposant sur les formalismes UML plutôt que sur un formalisme reposant sur un système d'équation (*analytique*) du modèle de simulation. Cette formalisation permet d'autant plus de faciliter le dialogue entre informaticien et modélisateur en offrant un langage commun et compréhensible à chacun (Ferber, 2006), ce qui apparait plus difficile avec la formalisation mathématique (Banos, 2013). Comme le souligne Franck Varenne (2013, p. 42) lorsqu'il aborde la notion des *modèles de simulation* auquel la géosimulation se rattache : « *ce type de modèle de simulation est un modèle pour la simulation du système cible. Significativement, le rôle de l'infrastructure informatique, des langages dédiés et des métamodèles y est de plus en plus grand* ». Dans cette acception, la géosimulation offre l'avantage, comme le soulève Arnaud Banos (2013, p. 82), de se libérer des contraintes de la formalisation mathématique qui n'est pas « *le langage universel des modèles* ». Pour la géographie, c'est plus particulièrement le cas des outils issus de la géosimulation, sur lesquels nous reviendrons ensuite. Pour Itzhak Benenson et Paul Torrens (2004, p. 2), la géosimulation se distingue des autres méthodes de simulation car elle prend en compte de « *manière explicite l'espace et la géographie. Elle s'illustre particulièrement bien lors de la représentation des objets dans les simulations, la spécification de leurs comportements, et la délimitation des objets dynamiques* ». De plus, la géosimulation se rattache aux démarches liées à la *géocomputation*, dans lesquelles ce travail s'inscrit (*cf.* section 1.3, p. 45), en combinant à la fois « *les apports de l'intelligence artificielle et des systèmes d'information géographique pour simuler la dynamique de système* »[77]. Plusieurs points justifient ce rattachement à la *géocomputation* (Benenson et Torrens, 2004) :

- La géosimulation repose sur la théorie des systèmes complexes, fournissant un cadre d'expérimentation et d'étude de phénomènes émergents (*cf.* section 1.1.1, p. 24)
- La géosimulation se construit en complémentarité des SIG, qui eux permettent non seulement l'encodage et l'intégration des différents objets spatiaux et de leurs attributs, mais qui en fournissent également un ensemble de méthodes pour relier les objets au sein d'un modèle de simulation.
- La géosimulation se construit sur une approche de programmation orientée objet [78] (POO) fournissant un cadre « intuitif pour représenter le "monde réel" en offrant de nouveaux langages pour le développement de modèles de simulation » (Benenson et Torrens, 2004, p. 4).

[75] Notons ici que la simple délimitation entre simulation mathématique et simulation informatique est simpliste tant les avancées informatiques ont été permise par les disciplines mathématiques et inversement (Varenne, 2008, 2013). Néanmoins, pour ce travail, nous nous arrêtons sur cette distinction afin de mettre en avant les grandes évolutions des méthodes et outils de simulation en sciences humaines.
[76] Pour plus de détails, nous renvoyons le lecteur vers l'ouvrage de Gilbert et Troitzsch (2005), qui traitent en détail chacune des approches au sein des chapitres 3 à 9 (pp. 28 à 215).
[77] Lena Sanders : définition de **« géosimulation »** : http://www.hypergeo.eu/spip. php?article311.
[78] En Anglais : *Object-Oriented Programmation (OOP).*

La simulation informatique permet de proposer un cadre de modélisation qui s'est construit en parallèle et en complément des approches existantes pour la représentation spatio-temporelle du trafic. Elle présente, à notre sens, un renouvellement des approches de modélisation des dynamiques urbaines, et dans notre cas de simulation du trafic routier afin d'en évaluer la pollution atmosphérique automobile (PAA).

## 4.2. Quel modèle de trafic retenir pour l'étude de la PAA ?

L'étude de la pollution atmosphérique de proximité nécessite de s'intéresser aux impacts locaux du trafic routier se manifestant à des échelles fines (bords de route) et sur des temporalités allant de la minute à la journée. Le modèle de trafic doit donc répondre à ces deux critères pour quantifier les volumes de trafic présents sur les routes au cours de la journée. Il est alors nécessaire de travailler avec une précision importante sur les dimensions spatiale et temporelle (Roussel, 2010 ; Coria *et al.*, 2015). Avant tout, il est nécessaire de disposer de données relatives à la circulation routière pour ensuite en estimer les niveaux de PAA (*cf.* figure 1.4, p. 35). Si les données relatives aux capteurs électromagnétiques permettent d'obtenir ces informations en différents points d'une zone d'étude, ces informations ne permettent que rarement de relever l'ensemble des flux de circulation routière ; c'est pourquoi le plus généralement on se tourne vers les modèles de trafic pour généraliser ces informations sur un espace plus vaste.

Il existe deux grandes catégories de modèles de trafic routier (Leclercq, 2002 ; Bourrel, 2003 ; Champion, 2003 ; Buguellou, 2012) : les premiers dits *macroscopiques* fournissent un niveau de détail moins important sur les conditions de circulation, mais permettent une approche plus globale des problèmes de déplacement (Masson, 2000). Les seconds dits *microscopiques* permettent de fournir une description plus précise des conditions de circulation réelles : les caractéristiques cinématiques (accélération, vitesse…) des véhicules. Toutefois, les données et les calculs nécessaires à son application limitent le plus souvent ce type d'approche à un ensemble restreint de tronçons routiers et de carrefours (Maciejewski, 2010). C'est pourquoi le plus généralement on se tourne vers les modèles macroscopiques qui sont les outils les plus répandus au sein des agences en charge de la qualité de l'air (AASQA) pour estimer les volumes de trafic routier (PCIT, 2012). Comme le relève Ludovic Leclercq (2002, p. 73) : « *choisir un modèle de trafic, pour une application donnée, consiste à déterminer le bon équilibre entre l'échelle souhaitée de représentation des phénomènes et la complexité du modèle retenu* ». Pour cette section, c'est l'ensemble de ces critères que nous souhaitons explorer afin de choisir l'approche de modélisation du trafic routier la plus adéquate au travail présenté.

### 4.2.1. **Quels modèles de trafic, pour quels usages ?**

L'étude et la gestion de la circulation routière en milieu urbain couvrent des applications multiples dans le domaine du trafic routier allant autant sur le court que sur le long terme. La diversité des modèles permet un ensemble d'applications, destinées autant aux usagers (informations sur les conditions de circulation) qu'aux décideurs locaux (régulation et planification urbaine). Que ce soit pour l'une ou l'autre des applications, la question de la quantification du nombre de véhicules est transversale : comment quantifier le trafic routier ? Par exemple, l'information à destination des usagers ne peut se faire sans une connaissance fine des conditions de circulation en différents points du réseau. De même la mise en place d'outils de planification nécessite l'utilisation d'outils

pour prévoir et évaluer l'évolution des déplacements automobiles.

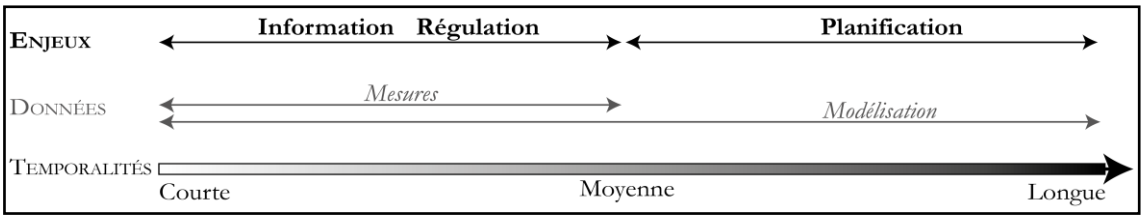



**Figure 2.3 : Enjeux, données et temporalités des études sur la circulation routière : une place prépondérante de la modélisation**

Notons ici que les enjeux liés à la pollution de proximité automobile, dont les caractéristiques spatio-temporelles sont fortement volatiles, s'avancent plutôt dans des applications allant du court au moyen terme. Néanmoins, face à ces enjeux, on se retrouve souvent confronté à un manque de données relatives à la circulation routière. Le choix du modèle dépend donc aussi des données à disposition du modélisateur. Deux stratégies liées aux données peuvent être différenciées : la plus simple est de s'entourer des mesures de terrain issues de différents capteurs. Lorsque les données ne peuvent recouvrir l'ensemble du terrain d'étude, la seconde stratégie est plutôt de mobiliser des modèles de trafic qui permettent de couvrir un espace plus vaste. Malgré tout, en s'appuyant sur la figure 2.3 (ci-dessus), quelle que soit la temporalité visée, et que des données de comptage soient disponibles ou pas, on constate que la modélisation occupe une place prépondérante lors des études relatives à la circulation routière (Chatzis, 2013). Le guide méthodologique pour l'élaboration des inventaires territoriaux des émissions atmosphériques montre bien cette tendance en relevant que « *les données issues d'un modèle de trafic sont à privilégier si elles existent, en combinaison avec des données d'observation* » (PCIT, 2012, p. 168).

Généralement, au sein de la modélisation globale des transports, on distingue deux types de modèles (Bonnel, 2002 ; Leclercq, 2002) : les *modèles macroscopiques* qui cherchent à répartir les flux de déplacement estimés sur un ensemble d'infrastructures de transport et les *modèles microscopiques* qui cherchent à décrire et à simuler l'écoulement des véhicules sur une infrastructure de transport (*cf.* section 4.2.2, p. 137). À cette typologie, on rattache respectivement aux modèles *macroscopiques* des enjeux portant sur le long terme, ce qui permet non seulement de prévoir mais aussi d'anticiper la mise en place d'aménagement routiers (Bonnel, 2002) ; et aux modèles *microscopiques* des applications portant plutôt sur des temporalités allant du moyen au court terme permettant, par exemple, de proposer des stratégies de régulation des feux tricolores (Leclercq, 2002). Dans la pratique, lorsque les mesures ne permettent pas de couvrir l'ensemble du terrain d'étude, le passage d'un modèle *macroscopique* à un modèle *microscopique* s'opère le plus souvent par l'intermédiaire des premiers (*macroscopiques*). C'est pourquoi, avant de porter un regard sur la classification des modèles de trafic routier, il est important d'apporter une distinction entre ce qui relève des modèles de trafic dits de *planification* et ceux dits de *simulation*. Le plus généralement, les modèles de planification du trafic routier se rattachent aux modèles *macroscopiques* (Bonnel, 2002). S'ils permettent d'obtenir une image de la répartition des flux de véhicules, ces modèles ne cherchent pas à représenter avec finesse le déplacement des véhicules le long d'une infrastructure. Au sein de ces modèles, le trafic routier est construit à partir d'un processus en quatre étapes (*cf.* encart 4, p. 136), dont les trois premières (Génération, Distribution, Répartition) permettent, à partir des données d'enquêtes, d'estimer les flux de déplacements ; l'étape d'affectation quant à elle se rattache plus réellement à une représentation du trafic routier (Henn, 2001). Cette

approche permet de passer du *flux de déplacement* entre deux espaces (origine-destination) au *trafic routier* sur l'ensemble du réseau routier (*cf.* section 2.1.1, p. 56). À cet égard, avant de pouvoir rendre compte du trafic routier, la démarche de modélisation construit les déplacements des individus dans l'espace urbain à partir de données socio-économiques (Bonnel, 2002).

Cette conception du trafic routier est aujourd'hui largement acceptée dans la littérature comme étant l'approche dite de *modélisation classique*, dont la formulation en quatre étapes en est la forme la plus connue (Bonnel, 2002 ; Ortúzar et Willumsen, 2011 ; Commenges, 2013). De ce point de vue, dans sa vision la plus classique, un modèle de trafic sert d'abord à localiser et à modéliser des déplacements en « *utilisant des hypothèses sur la demande des usagers potentiels du système de transport et des hypothèses sur l'offre de transport, pour simuler le comportement des usagers et obtenir des estimations de trafic résultant de l'interaction entre offre et demande* » (Meunier, 2013, p. 5). Toutefois, cette première définition ne permet pas de relever la multiplicité des modèles de trafic routier. En effet, aux modèles de planification se juxtaposent les modèles de simulation du trafic routier. Entre ces deux catégories, c'est bien la stratégie d'approche qui les distingue (Costeseque, 2011) : la première approche, de planification, est dite *statique* en se contentant de répartir l'ensemble des flux origine-destination sur le réseau routier à partir des quatre étapes ; alors que la seconde approche, de simulation, va plus loin et est plus complète en décrivant le trafic routier plus finement par l'intermédiaire d'un modèle d'écoulement (Henn, 2001), en se centrant sur une vision plutôt *dynamique* du trafic routier (Leclercq, 2002 ; Bourrel, 2003). Cette distinction rejoint de cette façon la définition générale apportée sur les modèles, selon laquelle la première distinction s'établit entre deux catégories de modèles, qui sont les modèles *statiques* et les modèles *dynamiques* (*cf.* section 4.1.1, p. 127). Ainsi, aux enjeux liés à la résolution des modèles de trafic, se juxtaposent des enjeux liés à leur résolution temporelle.

En s'appuyant sur la figure 2.3, avec les données issues de comptages routiers, nous disposons d'un dispositif de quantification adéquat pour l'étude du trafic sur les court et moyen termes. Contrairement aux EMD alimentant les modèles de planification, les comptages ne permettent pas d'envisager des applications sur le long terme. Par conséquent, les comptages routiers ne peuvent être envisagés que pour des applications allant du court aux moyens terme et dans l'optique de dresser un état des lieux de la circulation routière actuellement comptabilisée. La démarche de représentation du trafic routier envisagée doit se faire en concordance avec le dispositif de quantification que nous avons à disposition. La simulation spatio-temporelle du trafic routier, dont les problématiques s'inscrivent dans le champs des modèles de simulation du trafic comme PARAMICS (Sykes, 2010), VISSIM (Fellendorf et Vortisch, 2010) ou SUMO (Behrisch *et al.*, 2011), est le plus généralement formalisée à un niveau microscopique (Barceló, 2010). Ces propos tendent à nous montrer que les gains en termes de résolution temporelle ne peuvent se faire sans une descente des échelles de représentation des véhicules sur les infrastructures routières.

***Encart 4 Le formalisme de la modélisation à quatre étapes***

Tout d'abord, afin d'éviter toute ambiguïté, il faut noter que la modélisation à quatre étapes n'est pas un modèle à proprement parler, mais qu'elle constitue plutôt une approche, un cadre et un formalisme de la modélisation (Ortúzar et Willumsen, 2011). Le formalisme du modèle à quatre étapes est issu de l'emboitement séquentiel de quatre modules (Bonnel, 2002 ; Ortúzar et Willumsen, 2011) :

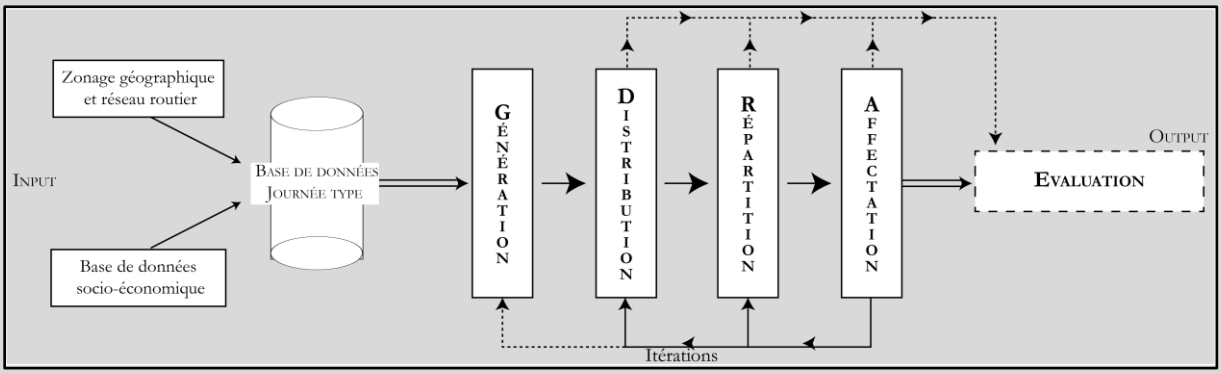



**Figure 2.4 : La structure séquentielle du modèle à quatre étapes (d'après Ortúzar et Willumsen 2011, p. 21)**

Un **module de génération**, qui vise à prédire le nombre total de déplacement générés par des lieux d'origine (O) et des lieux de destination (D). Cette étape permet la génération des populations et des zones d'activités de l'espace à étudier sous la forme d'un maillage territorial. Elle consiste à déterminer le nombre de déplacements émis et reçus par chaque zone sous la forme de flux (potentiels) de déplacement.

Un **module de distribution**, qui vise à répartir les flux de déplacement générés entre les zones O et D. Lors de cette étape, on considère que les masses de population induisent, statistiquement, des flux de circulation sous la forme d'une matrice O-D correspondant au nombre de déplacements sortant ou entrant dans chacune des zones de l'espace urbain. L'estimation des matrices O-D est issue le plus généralement d'un modèle gravitaire supposant que l'attraction de deux zones est fonction de la masse de population présente dans chaque zone. Avec le modèle gravitaire, les flux ne sont plus des populations mais des attributs d'émission et d'attraction définis lors de la phase de génération (Ortúzar et Willumsen, 2011).

Un **module de répartition**, qui vise à passer du flux de déplacement à une série finie de modes de transport donnés. La phase de répartition constitue la première phase de création du trafic routier. Cette étape correspond à la phase de choix d'un mode de transport, pour Patrick Bonnel (2002, p. 19) : « *mathématiquement, elle permet d'éclater la(les) matrice(s) O-D en autant de matrices O-D par mode qu'il y a de modes pris en compte* ».

Un **module d'affectation** qui vise à répartir les trafics sur les infrastructures de transport. Cette dernière étape permet d'assigner les itinéraires de déplacement en calculant un ou plusieurs plus courts chemins à emprunter pour relier physiquement O à D pour un mode et un motif de déplacement issu de la matrice O-D calculée préalablement. L'une des méthodes d'affectation des itinéraires les plus connues est l'algorithme de Dijkstra permettant de déterminer un seul plus court chemin pour se rendre d'un point à un autre (Dijkstra, 1959).

Les sous-modèles ne sont pas à considérer comme des isolats : chacune des étapes peut être reliée par des processus itératifs et de rétroaction (*feedback*). Certains modèles vont présenter des caractéristiques itératives tandis que d'autres vont présenter une structure procédurale plus simpliste. Par exemple, la prise en compte des phénomènes de congestion induit des itérations et rétroactions entre les étapes de distribution et d'affectation, l'encombrement d'une voie entrainant une modification de la matrice O-D et nécessitant de recalculer les chemins les plus courts entre les différentes zones du terrain d'étude. C'est pourquoi, il est important de séparer ce qui relève du modèle à quatre étapes et de ce qui relève du logiciel de modélisation du trafic routier, le logiciel servant plutôt à mettre en œuvre « *la construction de modèles* […] *en choisissant certains des procédés ou règles disponibles dans le logiciel* » (Bonnel, 2002, p. 36). Certains logiciels comme DAVIS ou VISUM vont intégrer différentes fonctions foncièrement différentes, permettant d'appliquer différents types de sous-modèles, en intégrant éventuellement une forme plus ou moins sophistiquée du modèle gravitaire, un nombre de modes de transport plus ou moins important, des méthodes d'affectation plus ou moins complexes…

### 4.2.2. **Du macroscopique au microscopique : quelle classification retenir ?**

Si les objectifs des deux typologies se distinguent, les modèles de trafic présentent tous des caractéristiques similaires au niveau des données nécessaires à leur mise en place (Bonnel, 2002 ; Champion, 2003 ; Ortúzar et Willumsen, 2011) : intégrant, d'une part, une description de l'offre de transport (le réseau routier et ses caractéristiques), et d'autre part, une description de la demande de transport sous la forme d'une matrice O-D permettant de décrire le nombre de véhicules empruntant un itinéraire.

La distinction entre chaque modèle s'oriente ensuite dans la manière avec laquelle la matrice O-D est appréhendée pour alimenter le modèle : dans le cas des modèles de planification celle-ci est stabilisée et on parle alors de modèles *statiques* ; alors que pour les modèles de simulation, les flux sont redéfinis à chaque pas de temps du processus de simulation, et on parlera alors de modèle *dynamique*. Cependant, le cloisonnement entre modèle *dynamique* et modèle *statique* n'est pas suffisant, et il apparait relativement limité pour rendre compte de la diversité des modèles de trafic routier existants. Le choix du modèle tient aussi compte de l'échelle de représentation avec laquelle le temps et l'espace sont appréhendés, dont la synthèse[79] est illustrée en figure 2.5 (ci-dessous).

Usuellement, la classification des modèles de trafic s'opère en fonction de la manière avec laquelle les véhicules sont représentés (Leclercq, 2002, p. 73) : « *cette approche permet de distinguer les grands types de modèles : macroscopiques, mésoscopiques ou microscopiques* ». Il faut noter que cette classification n'est pas toujours la plus adaptée, et plus particulièrement pour les modèles de simulation du trafic pour lesquels différents critères de classification peuvent être cités (Bourrel, 2003 ; Champion, 2003 ; Ksontini, 2013) : le temps (continu ou discret), le niveau de détail (plus ou moins fin), les processus utilisés (déterministes ou stochastiques), et la conception/formalisation utilisée (mathématique ou comportementale). Emmanuel Bourrel (2003), quant à lui, propose de

[79] À noter qu'il existe des modèles macroscopiques qui intègrent la dynamique ; c'est particulièrement le cas des modèles intégrant la congestion (Bonnel, 2010). Toutefois, l'état de l'art tend à nous montrer que les gains en termes de résolution temporelle se font le plus généralement en lien avec une finesse de la représentation des véhicules sur les infrastructures de transport routier.

classifier en retenant plutôt des critères liés à la représentation du trafic (sous la forme de véhicules ou de flux) et à la loi de comportement utilisée (flux ou comportements individuels). L'ensemble de ces critères montre toute la difficulté à convenir d'une classification figée des différents modèles de trafic routier, tant leur diversité et leurs applications sont multiples. Par simplification et en s'appuyant sur la figure 2.5, nous retiendrons le premier critère de classification qui s'opère selon la finesse avec laquelle les véhicules sont représentés (Leclercq, 2002 ; Marilleau, 2006) :

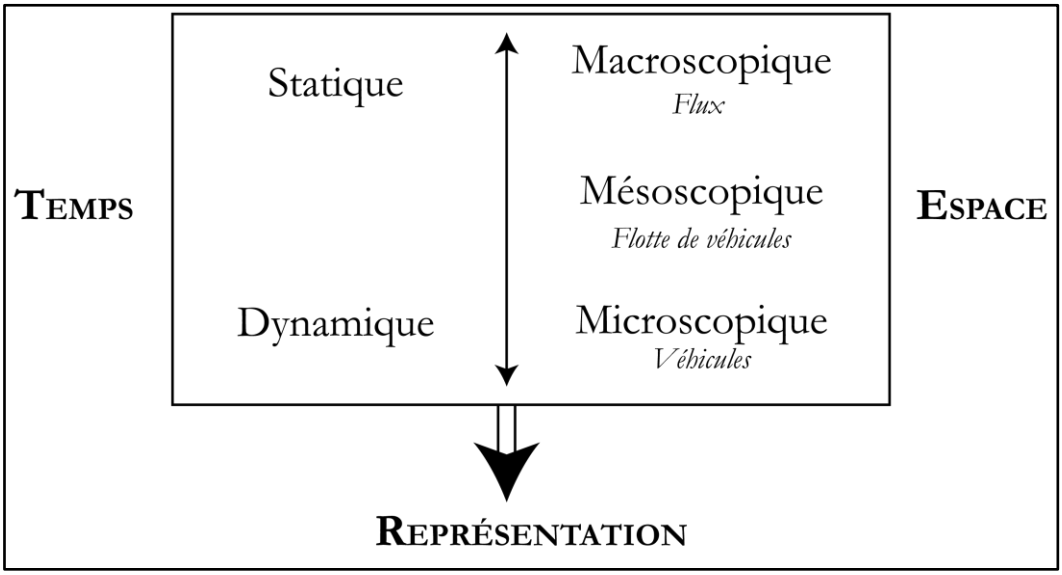



**Figure 2.5 : Le lien entre échelle de temps et représentation des véhicules : une descente des échelles spatio-temporelles à prendre en compte.**

- Les **modèles macroscopiques** présentent le degré de finesse le moins important. Au sein de ces modèles comme DAVIS[80], VISUM ou EMME/2 (CERTU, 1999), la représentation des véhicules se fait sous la forme d'un flux agrégé de circulation ne permettant pas de distinguer les véhicules individuellement. Les modèles macroscopiques sont couramment employés pour les études globales et s'inscrivent dans la continuité des modèles de trafic à quatre étapes classiques. Les véhicules ne sont pas décrits individuellement et présentent un degré d'abstraction relativement important.
- Les **modèles microscopiques** se fondent sur une représentation fine des véhicules pris individuellement. Cette représentation permet de mieux appréhender les dynamiques d'un espace d'étude (Marilleau, 2006). Les véhicules sont dotés d'un comportement, plus ou moins complexe, permettant de leurs affecter une capacité d'interaction et de perception de l'environnement dans lequel ils se déplacent. Pour le trafic routier, les modèles microscopiques les plus caractéristiques de cette catégorie sont PARAMICS, VISSIM, MATSim ou SUMO (Barceló, 2010). À noter que ces modèles sont mis en place pour comprendre les dynamiques qui s'opèrent sur un espace restreint comme une intersection ou un tronçon autoroutier.

[80] Acronyme de Distribution, Affectation et VISualisation développé depuis les années 1970 en France, il est le modèle de trafic qui a été le plus mobilisé pour les études de transport jusqu'aux années 1990. Exclusivement dédié au trafic, son évolution vers la multimodalité a donné naissance au modèle DAVISUM (contraction de deux modèles : DAVIS et VISUM). Une description détaillée de l'évolution de DAVIS est présenté dans Chatzis, 2009 (p. 41).

- Les **modèles mésoscopiques,** à l'image de METROPOLIS, sont moins répandus et se fondent sur une formalisation mixte du modèle de trafic (De Palma, 2013). À mi-chemin entre les modèles macroscopiques et microscopiques, les déplacements sont appréhendés à un niveau microscopique alors que les flux de trafic sont représentés à une échelle macroscopique. Ils se fondent ensuite sur une description des véhicules sous la forme d'un ensemble de flottes ou paquets de véhicules.

À ces trois représentations, on peut ajouter les modèles nanoscopiques qui « *représentent de manière fine les mobiles par leurs comportements, leurs interactions mais aussi par leurs actions internes* » (Marilleau, 2006, p. 17). Plus proche du simulateur de conduite que du modèle de trafic, nous n'aborderons pas cette dernière catégorie, de même que la modélisation mésoscopique, moins répandue et plus marginale, ne sera pas retenue ici. Ici, nous nous intéresserons plus particulièrement aux modèles macroscopiques et microscopiques, en portant un regard particulier sur leurs avantages et inconvénients pour notre problématique de travail : l'estimation des émissions de polluants issues des véhicules routiers.

#### 4.2.3. **Quelle approche de modélisation du trafic routier pour l'étude de la pollution atmosphérique automobile ?**

La modélisation macroscopique permet d'obtenir une quantification du trafic sous la forme de flux, les modèles de trafic se distinguent ensuite dans la représentation qui sera faite des véhicules (*cf.* figure 2.5, p. 138). Dans une optique de planification, les sorties de ces modèles peuvent suffire à obtenir une bonne photographie de la circulation routière urbaine sous la forme de flux de véhicules (Bonnel, 2002). Dans d'autres cas, cette échelle spatio-temporelle n'est pas suffisante et il faut convenir d'un niveau de détail plus fin en engageant un raffinement de la représentation des véhicules sur le réseau routier. C'est particulièrement le cas lorsque l'on s'intéresse aux problématiques de la pollution sonore liée au trafic routier (Leclercq, 2002). Pourtant cette approche macroscopique reste dominante et largement privilégiée pour les études d'impacts environnementaux locaux comme la pollution sonore ou la pollution atmosphérique (Leclercq, 2002 ; Fouillé *et al.*, 2012 ; PCIT, 2012). Il est alors important d'expliciter les différentes limites de cette approche pour notre problématique de travail :

- **Des limites sur la lourdeur du dispositif de modélisation** : La mise en application du modèle à quatre étapes nécessite la collecte d'un grand nombre de données pour alimenter le modèle, ces données portant autant sur les caractéristiques de déplacements des individus que sur la localisation des lieux d'activité, de loisir ou de domicile (Ortúzar et Willumsen, 2011). Notons qu'à l'inverse, la mise en place d'un modèle exclusivement dédié à la simulation du trafic réduit considérablement le nombre de données d'entrée nécessaire à la modélisation en se centrant exclusivement sur le trafic routier (Champion, 2003).
- **Des limites sur le plan temporel** : Ces modèles statiques ne permettent pas d'appréhender la variabilité temporelle du trafic routier au cours de la journée. L'état de la circulation est restitué pour une situation de référence sur la base d'une journée type (le jour ouvrable) pour une tranche de temps type (l'heure de pointe). L'heure de pointe se présente comme un standard de la modélisation macroscopique (Commenges, 2013). La variabilité temporelle du trafic routier ne peut donc être approchée que sur une situation

de référence ou à une résolution horaire (Meunier, 2013).

- **Des limites sur le plan spatial** : Le territoire pris en compte au sein des enquêtes est constitué sur la base spatiale des zonages OD définis lors de l'étape de génération. Les données ne sont donc pas rattachées à un lieu mais sont comprises dans des zonages plus ou moins fins (Bonnel, 2001, Commenges, 2013). En outre, l'accent est mis sur le déplacement de l'individu sur les infrastructures de transport les plus caractéristiques. Par exemple, les réseaux routiers de catégorie moindre ne sont pas pris en compte. L'ensemble du réseau routier n'est alors pas couvert par ces types de modèles. Ce défaut peut être rattaché au formalisme de l'approche macroscopique, dans le sens où cette échelle d'observation empêche de voir l'ensemble des déplacements, en particulier les plus courts : « *plus les zones sont vastes, et moins les flux entre zones sont importants, puisque des déplacements interzones deviennent alors intrazones et ne sont plus comptabilisés dans les échanges* » (Pumain et Saint-Julien, 2010b, p. 17).
- **Des limites sur le trafic routier restitué** : Ces limites sont en partie à relier aux limites spatiales évoquées ci-dessus pour deux raisons (Pumain et Saint-Julien, 2010b, p. 17) : « *d'une part, la mobilité estimée à partir des flux entre deux zones est toujours au plus égale et souvent inférieure à la mobilité totale ; d'autre part, la mesure de la mobilité dépend de la forme et de la finesse du découpage en zones géographiques* ». En effet, les sorties générées par les matrices OD ne rendent pas compte de la totalité du trafic routier. Il demeure limité à une logique d'échange entre chaque couple origine et destination construit à partir des données et des zonages disponibles. Les déplacements restitués sont donc limités au zonage d'enquête, en excluant par nature les trafics d'échange et de transit qui traversent une agglomération. Il faut alors ajuster et caler les estimations avec des données de comptages routiers très peu mises en valeur dans la chaîne de modélisation.

L'ensemble des limites liées aux modèles macroscopiques tient compte des dispositifs de quantification et de modélisation mobilisés : l'EMD et la modélisation à quatre étapes. Celles-ci illustrent bien l'éloignement entre l'*objet* de la mesure et l'*objet* restitué par modélisation pour l'estimation des volumes de trafic routier. En effet, le formalisme des EMD présente de nombreux écueils autant au niveau spatial (les zonages) qu'en termes de résolution temporelle (une journée type). Ceci nous amène à envisager une autre réponse que la modélisation macroscopique pour l'étude de la pollution de proximité automobile. En effet, il manque au modèle macroscopique une représentation de l'écoulement des véhicules à l'intérieur des tronçons routier en vue d'identifier les *hot-spots* (*cf.* figure 1.2, p. 32).

Méthodologiquement, la principale distinction entre l'approche macroscopique et l'approche microscopique s'établit sur leur capacité à intégrer un modèle d'écoulement ou non au sein du modèle de trafic (Buisson, 1996 ; Leclercq, 2002). Le modèle d'écoulement vise à décrire la manière avec laquelle les véhicules vont se déplacer sur le réseau routier (Buisson, 1996). En d'autres termes, il permet de représenter l'avancement des véhicules sur la route. Cette dernière étape est spécifique aux modèles de simulation du trafic routier et peut s'ajouter en complémentarité du modèle à quatre étapes (Henn, 2001). L'écoulement du trafic routier varie ainsi au cours du temps en fonction des conditions de circulation routière et d'utilisation des infrastructures de transport. Les flux de trafic connaissent des variations de vitesse, et donc de déplacement, tout au long du processus de simulation, et ils se rapprochent des conditions réelles de la circulation routière en milieu urbain à un instant *t*. Ceci nous amène à nous positionner en

faveur des modèles microscopiques, voire vers les modèles mésoscopiques, qui présentent des avantages autant sur le plan spatial que sur le plan temporel (Leclercq, 2002 ; Bourrel, 2003 ; Champion, 2003):

- **Sur le plan spatial**, l'apport des approches microscopiques permet d'envisager les déplacements non plus sur la base des zonages origine-destination, mais à un niveau désagrégé en apportant un regard sur les entités élémentaires du trafic routier : les véhicules (Leclercq, 2002). Au sein des modèles de simulation, l'attention est portée à « *l'intérieur du véhicule* » (Champion, 2003, p. 5) dans une optique généralement individu-centrée se caractérisant par l'interaction des véhicules avec les différents éléments du réseau routier.
- **Sur le plan temporel**, les outils de simulation se fondent sur un formalisme intégrant le temps en tant que variable de base du modèle de simulation (Champion, 2003). En effet, le formalisme de la simulation microscopique s'intéresse à la dynamique des véhicules au niveau individuel (Bourrel, 2003). Ainsi, à chaque pas de temps de la simulation, il est nécessaire de recalculer l'ensemble des états des véhicules sur le réseau routier. Ludovic (Leclercq, 2002, p. 66) note à ce niveau une des limites de la simulation : « *dès que le nombre de véhicule est important, les temps de calcul deviennent conséquents* ».

Ainsi, les modèles dédiés à la simulation du trafic se distinguent des modèles de planification sur plusieurs points (Bonnel, 2002 ; Leclercq, 2002 ; Costeseque, 2011) :

- ils sont explicitement dédiés au trafic routier ;
- ils s'appuient sur l'ingénierie du trafic plutôt que sur la socio-économie des transports ;
- ils s'imposent comme un outil de gestion des trafics routiers en milieu urbain ;
- ils couvrent des approches plutôt tournées sur les court et moyen termes ;
- ils reposent sur une représentation spatio-temporelle du trafic routier.

Ces derniers propos nous amènent à nous intéresser à la modélisation dynamique et plus particulièrement aux modèles de simulation du trafic routier qui se fondent sur une représentation dynamique du trafic routier. En se focalisant sur le déplacement et le mouvement des véhicules, cette catégorie de modèle apparait apporter une autre réponse à la représentation spatio-temporelle du trafic routier dans l'optique de mieux en estimer la PAA, à laquelle nous allons plus spécifiquement nous intéresser dans la section suivante.

## 4.3. Les enjeux liés à une approche de simulation du trafic routier à partir des comptages routiers

La simulation du trafic s'est plus particulièrement développée depuis les années 2000 grâce aux apports des outils informatiques et des simulations multi-agents qui, comme le souligne Nicolas Marilleau (2006), constituent l'un des outils le plus approprié pour la simulation des mobilités géographiques, et donc a *fortiori* du trafic routier. Ces modèles de simulation présentent un raffinement plus profond de la représentation du trafic routier en faisant appel d'une part à un modèle d'affectation et d'autre part à un modèle d'écoulement en vue de simuler plus fidèlement le comportement de conduite d'un conducteur à l'échelle microscopique (Buisson, 1996 ; Leclercq, 2002).

### 4.3.1. Les enjeux liés à la matrice origine-destination : un élément clé pour la simulation du trafic routier ?

Dans les usages, la simulation du trafic routier entretient un lien étroit avec la modélisation *classique* (*cf.* figure 2.6), car il est nécessaire de disposer de données permettant d'appréhender les volumes de trafic routier sous la forme de matrices origine-destination (O-D). La matrice O-D joue ici un rôle pivot pour la quantification des flux de trafic routier en permettant d'alimenter le modèle de trafic en données d'entrée.

Comme illustré en rouge au sein de la figure 2.6, la matrice O-D se situe en aval de la collecte de données pour estimer les volumes de trafic et se positionne en amont de la modélisation du trafic routier pour alimenter celui-ci. En s'appuyant sur cette figure, on peut ainsi identifier deux itinéraires en vue d'estimer les volumes de trafic pour alimenter le modèle de trafic : lorsque les données de comptage couvrent le domaine d'étude, on privilégiera leur usage, ce qui est particulièrement le cas lorsque l'on s'intéresse à une portion réduite du territoire comme un carrefour ou un quartier. Toutefois, lorsque l'on s'intéresse à un espace plus vaste, les données de comptage ne sont pas suffisantes. On fait alors appel aux estimations issues des trois premières étapes de la modélisation à quatre étapes (*cf.* encart 4 et figure 2.6) pour construire la matrice O-D sur un espace plus vaste. Cette relation est révélatrice du fonctionnement de nombreux modèles de simulation du trafic routier[81]. À titre d'exemple, le modèle de simulation VISSIM est construit de manière à pouvoir intégrer la matrice O-D issue de son modèle *parent* VISUM (Fellendorf et Vortisch, 2010), de même que le logiciel de simulation AIMSUN peut intégrer les matrices O-D des modèles classiques comme VISUM ou EMME/2 (Casas *et al.*, 2010).

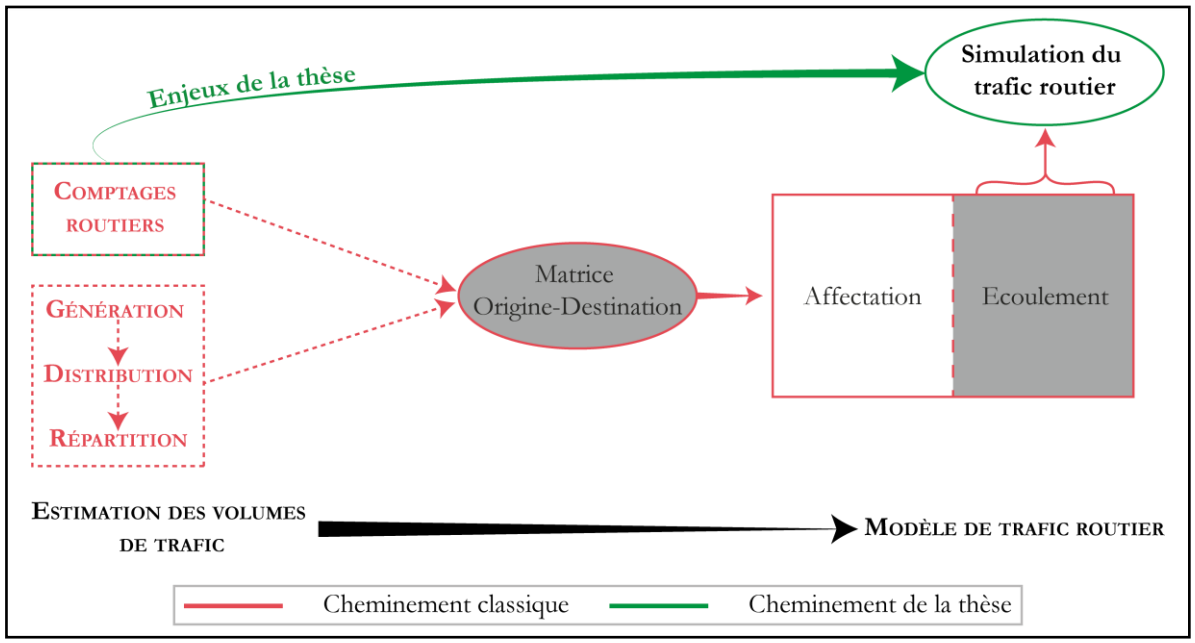



**Figure 2.6 : Le rôle pivot de la matrice origine-destination pour la simulation du trafic routier (en rouge, le cheminement classique ; en vert, le cheminement de la thèse)**

[81] L'ensemble de ces propos illustre parfaitement les difficultés que nous avons rencontrées pour établir une lecture et une classification exhaustive des modèles de trafic routier, tant les usages et leur mise en applications sont étroitement liés.

Au final, que ce soit à partir des comptages routiers ou à partir du modèle à quatre étapes, il n'existe à notre connaissance pas de modèle permettant de simuler le trafic routier sans faire appel à une matrice O-D. C'est sur cette dernière acception que nous situons les enjeux de notre travail de simulation du trafic routier, illustrés en vert sur la figure 2.6. L'approche envisagée vise à apporter une réponse à ce niveau pour passer de volumes de trafic routier comptabilisés à la simulation du trafic routier plus directement, sans faire appel à une matrice O-D, mais plutôt en mobilisant sans intermédiaire les comptages routiers dans le processus de simulation.

Toutefois, la résolution temporelle des données d'observation ne permet que rarement la mise en place d'un modèle de simulation, ce qui explique la relation entre les modèles de simulation et les modèles traditionnels à quatre étapes. En effet, les EMD sont des enquêtes établies sur une journée type de la semaine (Commenges, 2013), ce qui constitue une des limites de la résolution temporelle de ces modèles de trafic, qui ne permettent d'appréhender la variabilité du trafic routier que sur deux périodes caractéristiques (les heures de pointe et les heures creuses). C'est pourquoi l'utilisation des données de capteurs, dont la force réside dans leur granularité temporelle (*cf.* section 3.3, p. 102), apparait être une alternative pour alimenter différemment les modèles de simulation de trafic routier. De plus, différents arguments nous ont amené à ce positionnement en faveur des comptages :

- Le processus de constitution des matrices O-D peut s'avérer très vite chronophage, c'est pourquoi proposer une approche de simulation plus directe peut permettre d'envisager un cadre de modélisation du trafic routier plus simplifié pour l'évaluation de la PAA.
- Toutes les villes ne disposent pas de données d'enquête permettant la constitution de matrices O-D. En effet, si en France de nombreuses villes peuvent, par l'intermédiaire de l'INSEE[82] ou du CEREMA[83] obtenir ces informations, ce n'est pas le cas de toutes les villes dans le monde. Par contre, les données de comptage, dont les caractéristiques de mise en place sont non seulement plus simples, mais aussi moins onéreuses, sont des informations plus généralement accessibles pour un grand nombre de villes.

Des modèles opérationnels de simulation du trafic comme SUMO (Behrisch et al, 2010), ou VISSIM (Maciejewski, 2010) permettent d'intégrer directement des comptages routiers. L'intégration des données s'applique d'ailleurs aisément sur les réseaux autoroutiers dont les entrées et sorties sont couvertes par des données de comptage. Néanmoins, à l'image de SUMO, son application est limitée pour des réseaux de centre-ville et de boulevards qui présentent une couverture de mesure plus faible (Maciejewski, 2010). Avoir recours à ces modèles demande un temps de recueil de données qui peut s'avérer long : en effet, il est souvent nécessaire de construire des matrices OD afin de générer les flux de trafic routier (*cf.* figure 2.6, p. 142).

C'est pourquoi proposer une approche de simulation du trafic routier à partir des capteurs de trafic nécessite de revoir le positionnement de cette donnée dans le processus de construction d'un modèle de trafic. L'idée que nous souhaitons défendre vise à replacer les données issues des

[82] Description des données détaillées MOBPRO qui permet de recueillir des données sur le thème des déplacements entre le domicile et le lieu de travail : http://www.insee.fr/fr/bases-de-donnees/default.asp?page=recensement/resultats/doc/presentation-fichiers-detail-mob-pro.htm.

[83] Note de présentation des enquêtes déplacements « standard CERTU » : Enquête Ménage-Déplacement, Enquêtes Déplacements Villes Moyennes, Enquête Déplacements Grand Territoire, Janvier 2013 : http://www.territoires-ville.cerema.fr/IMG/pdf/Les_enquetes_Deplacements_site_web_cle5964c6.pdf.

capteurs électromagnétiques au cœur du dispositif de modélisation du trafic routier. Les enjeux empiriques et inductifs du travail se situent à ce niveau : comment intégrer les données d'observation au centre du dispositif de modélisation pour la simulation du trafic routier ?

Au contraire du cadre traditionnel de modélisation qui vise d'abord à modifier les données d'observation via un processus *top-down* ou *bottom-up* pour ensuite alimenter le modèle de trafic, nous souhaitons positionner l'observation au centre du processus en nous reposant sur les avantages des comptages routiers, sans en modifier ni leurs caractéristiques ni leur nature. Si les données de comptage n'alimentent que peu les modèles de trafic – du moins directement – l'évolution des approches de modélisation du trafic routier tend à montrer que les données de comptages routiers jouent un rôle secondaire pour la représentation du trafic routier, au profit des enquêtes de mobilité (*cf.* figure 2.7). On peut constater que ces données de référence, les comptages, ne sont que rarement mises au premier plan, et sont seulement utilisées pour valider et ajuster les estimations issues du processus de modélisation. Or, comme nous l'avons vu au cours du chapitre 3 (p. 81), les comptages revêtent un intérêt particulier pour rendre compte de la dynamique du trafic routier :

- ils permettent d'obtenir une information sur les flux de circulation routière en différents points de l'espace urbain.
- ils permettent de s'affranchir des trois premières étapes des modèles classiques, dans le sens où la demande (le trafic) et l'offre de transport (la route) sont mesurées et données en temps quasi-réel.
- ils permettent de répondre à des enjeux liés autant à la régulation des feux tricolores qu'à des applications environnementales sur le court et moyen terme.

Dans une approche de simulation classique, les informations issues des comptages routiers sont positionnées comme des valeurs de référence et sont utilisées à différents niveaux du processus pour calibrer les matrices O-D, ajuster les valeurs d'affectation et ensuite valider, ou pas, les sorties du modèle de trafic (*cf.* figure 2.7, à gauche). En d'autres termes, elles s'insèrent à toutes les étapes sans pour autant être valorisées en tant que telles.

Ainsi, comme illustré en figure 2.7 (ci-dessous), en mobilisant les comptages routiers, il est nécessaire de sortir du cadre traditionnel de la modélisation du trafic routier. Il demeure toutefois des écueils à soulever en vue de simuler et de reconstruire une image du trafic routier sur l'ensemble d'un réseau routier. Les données sont ponctuelles et il est nécessaire de passer à une information généralisée sur les différents axes du réseau routier urbain. On se retrouve ici confronté à une problématique proche de celle des modèles de simulation du trafic routier, avec les modèles d'affectation et d'écoulement, en vue d'une part de définir la direction des flux de véhicules recensés par les stations de comptage ; et d'autre part, d'en appréhender leurs déplacements dans le temps sur les infrastructures routières (Buisson, 1996). L'approche proposée ici est de repositionner cette valeur de référence au centre du dispositif de modélisation dans une optique empirique, les informations issues des comptages routiers servant à alimenter le modèle de simulation. Nous proposons de simplifier l'approche en mettant les données de comptage au premier plan, comme illustré à droite de la figure 2.7 :

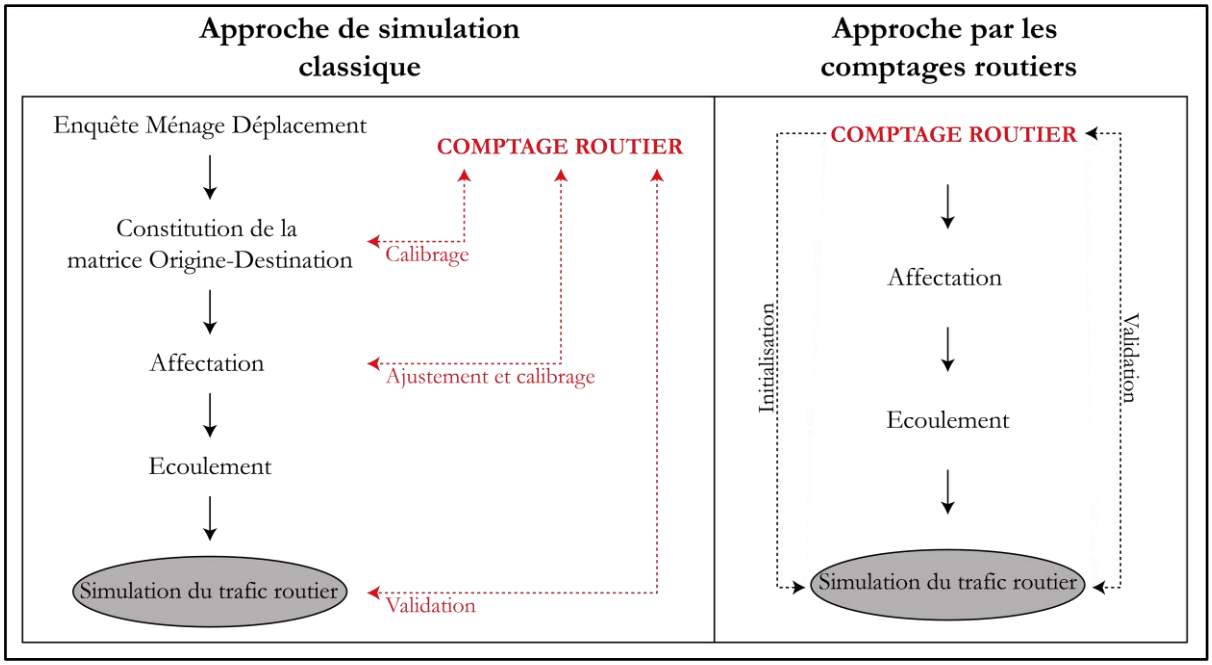



**Figure 2.7 : Revaloriser les comptages routiers pour la simulation du trafic routier : quel positionnement de la donnée dans le processus de modélisation ?**

Ce positionnement permet de se dédouaner du cadre commun (rationalisme, *cf.* figure 1.6, p. 38) qui vise au préalable à reconstituer une matrice origine-destination à partir des enquêtes ménages déplacement. En partant d'un positionnement empirique, celui-ci doit permettre de repositionner l'information au centre du processus de construction envisagé. Ainsi, les données d'observation sont directement intégrées dans la constitution du modèle, en partant de l'observation sans faire appel à un processus *top-down* ou *bottom-up*. L'adéquation entre le modèle de données et le modèle de représentation doit s'établir en cohérence avec l'information collectée par le dispositif de quantification, afin de se rapprocher, d'une part, de l'*objet* de la mesure, et d'autre part, de l'*objet* de modélisation et de simulation. C'est pourquoi il est nécessaire de s'arrêter sur ces différents éléments.

### 4.3.2. Les enjeux liés à la matrice technique : de la quantification à l'évaluation de la pollution atmosphérique automobile

Pour identifier les leviers expliquant la stabilité de la modélisation *classique*, Hadrien Commenges (2013) propose de positionner celle-ci au sein d'un triptyque qui se constitue : 1. d'un **dispositif de quantification** à travers l'enquête O-D ; 2. d'un **dispositif de modélisation** à travers le modèle à quatre étapes ; 3. d'un **dispositif d'évaluation** économique. De là, il replace, sous le terme de « *matrice technique* », ces trois dispositifs qui correspondent à « *un ensemble de dispositifs techniques qui se tient, c'est-à-dire qui assure la stabilité de l'ensemble, […] plus qu'une juxtaposition des trois dispositifs techniques : elle forme un tout cohérent par le tissu de relations qui les unissent* » (Commenges, 2013, pp. 75–76). L'ensemble s'illustre au sein d'un système fermé, où le dispositif de quantification alimente celui de modélisation, qui permet ensuite d'alimenter le dispositif d'évaluation (*cf.* figure 2.8). Cette relation est réciproque dans le sens où l'évaluation façonne la modélisation à quatre étapes, qui elle aussi façonne le dispositif de quantification sous la forme d'enquête O-D. Le dernier lien permet quant à lui de confronter la mesure aux estimations pour valider les estimations et inversement.

La grille de lecture, proposée par Hadrien Commenges (2013), révèle à notre sens les limites de l'approche dite *classique* pour notre travail. En effet, l'abandon des enquêtes O-D au profit des données de capteurs fixes modifie sensiblement les caractéristiques de la matrice technique, autant au niveau du dispositif de modélisation que de celui de l'évaluation, de même que la modification des objectifs visés par le dispositif d'évaluation (de l'évaluation économique à une évaluation environnementale) modifie lui-aussi la forme du dispositif de modélisation du trafic routier :

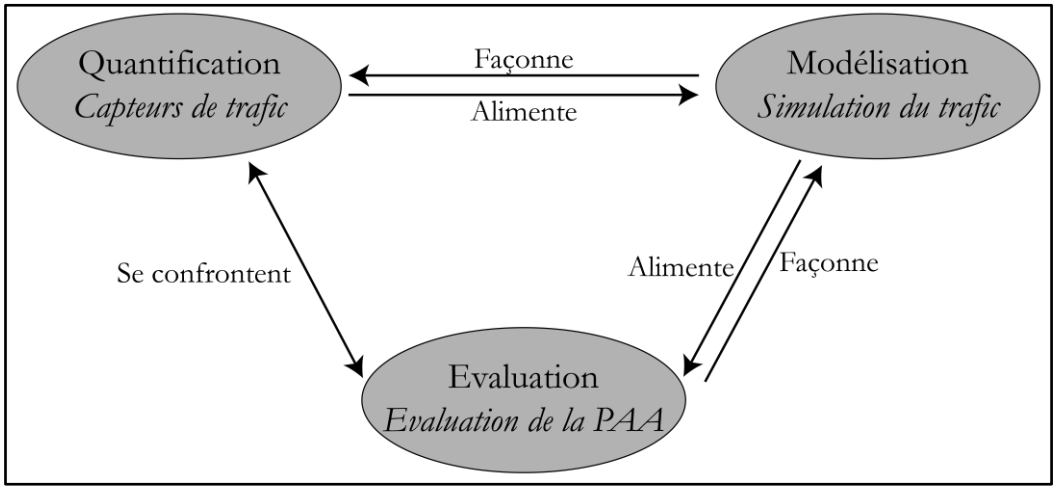



**Figure 2.8 : Une grille de lecture de la modélisation classique du trafic routier à travers le concept de la matrice technique (source : Commenges, 2013, p. 76)**

Il faut donc être en mesure de réagencer le dispositif de modélisation non seulement en adéquation avec les données que nous avons à disposition, mais aussi avec l'objectif final de la démarche qui vise à proposer une évaluation des niveaux de PAA sur l'ensemble d'un espace urbain. Ici, en reprenant le concept de la *matrice technique*, nous nous focaliserons sur les relations qu'entretiennent les trois dispositifs techniques dont la synthèse est illustrée en figure 2.9 (ci-dessous). En outre, sous cette forme, il est possible d'observer les liens de dépendance (illustrés en gris) entre chaque dispositif appliqué à notre problématique de travail. D'une manière générale, on observe que chaque dispositif en amont alimente celui qui est en aval. Les données subissent des modifications tout au long de la démarche, en partant de la donnée d'observation en allant vers la donnée de simulation, pour ensuite s'orienter vers des données d'évaluation à des fins environnementales. Ainsi, à chaque étape du processus les données relatives au trafic routier s'affinent et s'épurent pour aboutir à une connaissance des niveaux de polluants atmosphériques émis par les véhicules. On se retrouve ici au sein du processus de construction de la connaissance (*cf.* section 1.2.1, p. 36) qui vise à passer de la donnée d'observation (les comptages) à une information sur les flux de trafic routier puis à une connaissance sur les niveaux d'émissions des polluants atmosphériques automobiles.

Chaque dispositif est connecté par une relation qu'il est possible de qualifier de *substantielle*, dans le sens où chacun alimente en substance le dispositif suivant. C'est pourquoi le dispositif de quantification permet d'alimenter le dispositif de modélisation du trafic routier qui, lui, sert ensuite à alimenter le dispositif d'évaluation. Cette interconnexion s'illustre par plusieurs relations. La première relation, allant du dispositif de quantification au dispositif de modélisation, permet de spatialiser les informations relatives au trafic routier à partir des capteurs

électromagnétiques. Ce n'est que dans un second temps, par l'intermédiaire du dispositif de modélisation, que les informations spatio-temporelles du trafic routier seront appréhendées. En effet, cette seconde relation permet à travers le dispositif de modélisation de généraliser un plus grand nombre d'informations relatives au trafic routier qui s'observe, autant au niveau des infrastructures de transport que dans le temps.

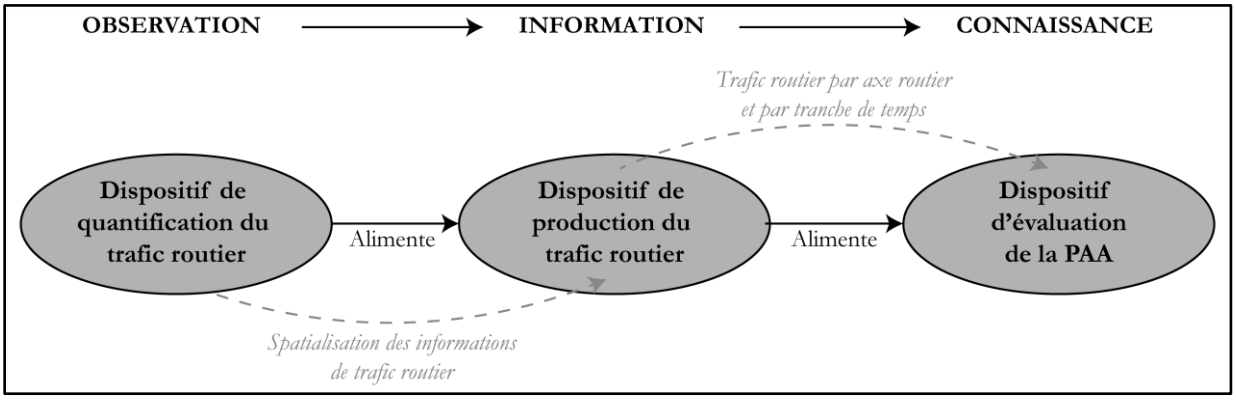



**Figure 2.9 : L'imbrication des dispositifs de la matrice technique pour l'évaluation de la PAA : de l'observation à la connaissance**

Ces propos nous distinguent du cheminement de la modélisation classique qui, comme nous l'avons introduit, vise d'abord à estimer les flux de déplacements entre des zones d'origine et de destination pour estimer les flux de trafic routier. Dans notre cas, à travers les comptages routiers, nous disposons déjà des informations relatives aux flux de trafic routier ; le modèle de trafic doit alors nous permettre de mettre en mouvement ces informations sur le réseau routier. Dans ce sens, le trafic routier n'est plus à construire, mais il faut plutôt être en mesure de le reproduire, voire de le produire. De surcroît, ainsi alimenté par les données de comptage, le dispositif de modélisation remplit plus un rôle de *production* du trafic routier et des données associées que de construction (*cf.* figure 2.9). Le dispositif de modélisation joue ici un rôle central permettant de faire le pivot entre les données de capteurs et les données de polluants atmosphériques. Toutefois, le passage des données de trafic aux émissions de PAA s'établit de manière indépendante ; en effet, comme nous le verrons plus bas, les données sont ensuite mobilisées pour alimenter le logiciel d'émissions CIRCUL'AIR utilisé par les AASQA (*cf.* section 5.3.3, p. 179).

Porter le regard sur la matrice technique nous permet d'aborder les objectifs de notre travail relatif au dispositif de modélisation pour la simulation du trafic routier. Les données issues des capteurs fixes ne permettent pas d'appréhender les flux sous la forme de couple origine-destination, de même qu'elles ne peuvent pas être rattachées aux enquêtes de déplacement qui servent à alimenter le dispositif de modélisation (*cf.* Chapitre 2, p. 55). Les données issues des capteurs fixes restreignent donc nos choix relatifs au dispositif de modélisation du trafic routier. En conséquence, il faut être en mesure de mettre en mouvement les informations issues des comptages en tentant de produire le trafic routier, plutôt que de le construire comme c'est le cas à partir des enquêtes de déplacement.

### 4.3.3. Les enjeux liés à la formalisation du modèle : entre approche mathématique à approche comportementale du trafic routier

Pour la simulation du trafic routier, on distingue deux approches pour simuler le mouvement des véhicules sur la route : d'une part, les modèles de simulation *mathématique* qui simulent le mouvement des véhicules par la résolution d'équations mathématiques ; et d'autre part, les modèles de simulation *comportementale* qui simulent le trafic routier à l'aide d'une approche informatique (Champion, 2003 ; Ksontini, 2013).

La simulation mathématique du trafic routier utilise des lois d'écoulement qui simulent le passage des véhicules sur une route en s'inspirant des théories issues de la mécanique des fluides (Quinet, 1998). C'est particulièrement le cas du modèle de simulation dit LWR dont la dynamique est appréhendée à travers la variation des courbes débits-vitesse[84] (*cf.* encart 1, p. 75). D'autres modèles incluent des dynamiques sous la forme de files d'attentes (*queuing model*) ou de lois de poursuite permettant ainsi de régir le déplacement des véhicules pris à un niveau individuel (Champion, 2003). Du côté opérationnel, il existe différents modèles de simulation employant un formalisme mathématique comme, par exemple, AIMSUN (Casas *et al.*, 2010), ou PARAMICS (Sykes, 2010). Conceptuellement, les modèles de simulation mathématique se construisent dans la continuité de la modélisation dite *classique*, en prenant en compte le trafic routier sous la forme de flux et en simulant le trafic routier par les interactions entre offre et demande. Dès lors, l'objet de la simulation est le flux et non pas le véhicule, ce qui constitue une des lacunes de cette approche, qui même au niveau microscopique n'est pas toujours satisfaisante pour rendre compte de la complexité du trafic routier (Champion, 2003, p. 11) : « *les phénomènes de trafic ne peuvent être décrits fidèlement par des équations mathématiques* […] *De ce fait et de par leur nature même, les modèles mathématiques interdisent toute interactivité avec un élément extérieur* ».

C'est pourquoi, une seconde approche de simulation du trafic routier a vu le jour à l'image des logiciels comme MATSim (Balmer, 2007), SUMO (Behrisch *et al.*, 2011) ou VISSIM (Fellendorf et Vortisch, 2010). Ceux-ci se fondent plutôt sur une approche comportementale, plus proche de la simulation informatique, permettant de prendre en compte le trafic routier à son niveau le plus élémentaire : les véhicules. À l'inverse des modèles mathématiques qui visent à reproduire le trafic, ces approches visent plutôt à « *produire le trafic* » (Champion, 2003, p. 12) : « *les phénomènes de trafic émergent car ils sont le résultat, d'une part, des pratiques individuelles et des interactions et, d'autre part, de la variété des comportements. La complexité des situations que l'on observe n'est donc pas issue de la complexité de l'algorithme mis en œuvre mais reflète les effets induits des multiples interactions qui s'opèrent entre l'entité et son environnement* ». Ici, se sont bien les apports de la théorie des systèmes complexes (*cf.* section 1.1.1, p. 24) qui permettent, par l'émergence des interactions locales, de rendre compte d'un trafic routier simulé. La simulation du trafic routier à partir d'une simulation informatique procure l'opportunité d'établir un modèle de simulation plus lâche, dans le sens où le modèle permet, non pas de construire, mais bien de produire un trafic routier. Comme nous l'avons vu, s'inscrire dans cette démarche de simulation (*informatique*), permet aussi d'envisager une approche expérimentale

[84] Les travaux sur le modèle LWR se basent sur cette analogie entre dynamique des fluides et dynamique du trafic routier : Lighthill, Whitham (1955) et Richards (1956) :

- Lighthill, M. Witham, G. On kinematic waves, a theory of traffic flow on long crowded roads, *Proceedings of the Royal society*, Series A, Vol. 229, London, 1955, pp. 317-345,
- Richards, P.J. Shock waves on the highway. *Operations Research* n°4, 1956, pp. 42–51.

(Banos, 2013). Ainsi alimenter le modèle de simulation directement par les données de comptage est une démarche peu commune dont les bases conceptuelles ne peuvent que très peu se rattacher aux approches de simulation existantes. Les contraintes liées à la manipulation de la matrice O-D au sein des modèles de trafic nous amènent à envisager la mise en place d'un modèle de simulation construit à travers une plate-forme de géosimulation, et plus spécifiquement la simulation multi-agents (SMA). Il apparait alors nécessaire de repenser les outils de simulation du trafic routier à travers ces outils, pour y intégrer plus directement les données d'observation, afin de s'adapter aux enjeux mis en évidence au cours de cette dernière section.

## Conclusion :

Le modèle est une représentation schématique de la réalité élaborée en vue de la comprendre et de la faire comprendre (Haggett, 1965 ; Mathian et Sanders, 2014). Le passage de la réalité au modèle n'est pas direct et s'insère dans un processus de construction de la connaissance en trois étapes, où à chaque étape on manipule différents *modèles* pour construire, formaliser et interpréter l'étude d'un phénomène (*cf.* figure 2.1, p. 128). C'est ici que se situe l'un des enjeux du travail proposé, qui vise à rapprocher *objet* de mesure et *objet* de modélisation en utilisant plus explicitement les données de comptages routiers pour la seule représentation du trafic routier en milieu urbain. Nous avons mis en avant les enjeux liés aux dispositifs de quantification pour l'étude des dynamiques urbaines. Nous avons vu que les données de comptage constituent un outil adéquat pour la représentation et à la modélisation du trafic routier. Toutefois, ce positionnement nécessite non seulement de conceptualiser, mais aussi de formaliser différemment le modèle de trafic routier, en s'éloignant de la conception *classique* à quatre étapes. De surcroît, nous avons vu que dans une optique d'évaluation de la PAA, il est nécessaire d'adopter de nouvelles approches de modélisation en vue de mieux répondre aux limites intrinsèques au formalisme à quatre étapes (Masson, 2000). La simulation apparait comme une réponse adéquate car elle permet de mettre en action, par la reproduction de processus le modèle.

Dans notre cas, elle présente l'opportunité de simuler le trafic routier à son niveau le plus élémentaire. Pour cela, le formalisme d'un modèle de simulation informatique apparait intéressant à mobiliser en vue de non seulement se centrer sur les composants élémentaires du trafic routier, mais aussi de dépasser les limites intrinsèques à la formalisation O-D. Ce dernier outil (la simulation informatique) constitue d'ailleurs une alternative intéressante à la modélisation à quatre étapes en mettant avant tout l'accent sur le véhicule en déplacement (microscopique) plutôt que le flux (macroscopique). Sur la base de ces constats, nous proposons de réorienter la matrice technique (Commenges, 2013) pour l'évaluation du trafic routier en milieu urbain. Ceci nécessite de concevoir deux dispositifs : celui de quantification en utilisant les relevés de comptages routiers plutôt qu'une matrice OD, et celui de la modélisation en mobilisant les SIG et les plates-formes de SMA, que nous traitons dans les parties suivantes. Se rapprocher de l'observation du trafic routier via les comptages routiers et les SMA apparait être une réponse aux critiques émises à l'encontre des modèles classiques. C'est dans ce but de simplification empirique, et guidés par la donnée, que nous proposons la mise en place d'un modèle de SMA. Par nature, cette modélisation permet une logique radicalement centrée sur l'individu, ou dans notre cas sur le véhicule. Elle implique un fonctionnement *bottom-up* permettant d'être plus directement en lien avec les données d'entrée du modèle.

Or, il demeure important de se demander dans quel cadre ces outils de simulation permettent d'apporter une autre réponse à ce qui était jusqu'à présent modélisé. De plus, dans ce cadre, l'utilisation des données de comptage parait adaptée afin d'alimenter un modèle de géosimulation pour l'évaluation de la PAA. Toutefois, elle soulève de nombreux questionnements afin de conceptualiser et de produire un trafic routier sur un espace plus vaste que celui qui est observé à partir des comptages. C'est un questionnent qui apparait foncièrement géographique, en vue de passer du point (*le comptage*) à la ligne (*le trafic routier*).

# Chapitre 5. LES APPORTS DE LA SIMULATION MULTI-AGENTS POUR L'ÉTUDE DU TRAFIC ROUTIER

La simulation du trafic routier permet une représentation spatio-temporelle du trafic bien plus avancée que pour les modèles *classiques*. En effet, ces modèles, que nous avons choisis au cours du chapitre précédent, permettent une représentation fine des conditions de circulation routière autant dans le temps que dans l'espace. Les modèles dynamiques s'inscrivent dans des formalismes *individu-centré* et tendent à montrer une évolution des usages pour la représentation spatio-temporelle du trafic routier. En effet, ces nouvelles approches se sont plus spécifiquement développées à travers des plates-formes de géosimulation offrant aujourd'hui de nouvelles opportunités pour l'étude et la reproduction des processus spatio-temporels (Thévenin, 2011). S'inscrire dans une optique de géosimulation s'agence alors en adéquation avec l'évolution des nouvelles approches de modélisation du trafic routier en vue de s'affranchir du cadre traditionnel. Arnaud Banos (2013) identifie d'ailleurs plusieurs avantages qui vont dans ce sens :

- Ce sont des méthodes s'inscrivant dans la complexité
- Elles reposent sur des approches de *bottom-up*
- Elles présentent un ancrage spatial et territorial plus important.

Dans cette acception, la simulation multi-agents (SMA) s'avance aussi comme outil expérimental ou la géographie apparait être partie prenante. C'est pourquoi, nous explorerons plus en détail les apports des outils de géosimulation, et plus spécifiquement de la simulation multi-agents, pour la simulation du trafic routier. Quels peuvent potentiellement en être les apports pour la modélisation du trafic routier et pour l'estimation de la PAA ? Comment repenser les outils de simulation du trafic à travers ce cadre de réflexion en vue de mieux quantifier les flux de trafic routier en milieu urbain ?

À travers la littérature sur l'évolution de la modélisation du trafic routier, on constate que l'insertion des approches de géosimulation constitue les derniers développements conceptuels et méthodologiques dans ce domaine. Par l'intermédiaire d'une première lecture chronologique, nous soutiendrons l'idée que les approches de géosimulation ont introduit un bouleversement méthodologique des approches de modélisation du trafic routier. Ces propos permettront de nous situer dans le prolongement des développements récents dans le domaine de la modélisation du trafic routier. Ensuite, nous explorerons plus en détail les apports des approches de géosimulation en portant un regard plus particulier sur leurs atouts pour la modélisation du trafic routier. L'ensemble de ces arguments permettra ensuite d'aborder les grandes lignes de la démarche de modélisation adoptée pour ce travail à travers le modèle de *Simulation multi-agents à partir de CApteurs Urbains pour la Pollution atmosphérique automobile* (SCAUP). Pour cela, nous nous focaliserons sur les éléments conceptuels, les hypothèses envisagées ainsi que les objectifs opérationnels pour l'application du modèle SCAUP.

## 5.1. Le renouveau de la modélisation du trafic routier par la simulation-multi-agents

L'implantation des approches de géosimulation dans le paysage de la modélisation du trafic routier est le fruit « *de la convergence entre une évolution théorique et une révolution technique. Le succès de ce type de modélisation apparait de ce fait comme le produit et le moteur d'un renversement dans la façon de concevoir le milieu urbain* » (Cambien, 2010, p. 42). Pour illustrer l'insertion et les développements méthodologiques qu'ont apporté les approches de géosimulation dans le champ de la modélisation du trafic routier, nous nous appuyons sur les évolutions de la modélisation classique, car ces dernières constituent une bonne illustration des apports récents de la géosimulation. Le paysage de la modélisation du trafic routier s'est construit autour de trois grandes périodes qui seront développées dans les sous-sections suivantes (Bonnel, 2002 ; Debizet, 2004, 2011 ; Chatzis, 2013) : la première période s'étend de 1950 à 1975, au cours de laquelle les premiers modèles de trafic routier ont été appliqués. La seconde période, allant de 1975 à 1990, se caractérise par une forte remise en cause des fondements de la modélisation classique : au cours de celle-ci ont été initiés les premiers modèles microscopiques grâce aux apports de l'approche désagrégée des choix discrets qui ont néanmoins connu un succès limité (McFadden et Reid, 1975 ; McFadden, 2000). Enfin la troisième période s'étalant des années 1990 à aujourd'hui, s'est traduite par un bouleversement conceptuel et méthodologique de la ville grâce à l'insertion des approches individu-centrées qui ont été initiées par l'intermédiaire des outils de géosimulation.

### 5.1.1. **Genèse de la modélisation *classique* du trafic routier**

Dans un contexte général de demande croissante des déplacements motorisés (Dupuy, 1975), les premières applications de la modélisation des transports ont permis d'adapter et de dimensionner les infrastructures routières existantes à ces évolutions (Debizet, 2011). Elles ont ainsi facilité dans un premier temps la circulation routière en permettant l'élargissement des infrastructures existantes et la mise en place des grandes déviations urbaines. Les modèles jusqu'alors mis en place se basaient sur les comptages routiers et les enquêtes cordons[85], permettant de comptabiliser les flux de trafic entrants et sortants d'une agglomération en vue de mieux quantifier les trafics de transit. Toutefois, l'utilisation de ces données a très vite présenté ses limites pour la création et l'implantation de nouvelles infrastructures où aucune donnée ne peut être disponible (Bonnel, 2010). Les comptages routiers sont alors vite mis de côté au profit des enquêtes de type Origine-Destination (O-D), comme les navettes domicile-travail (Commenges, 2013). Ce changement s'explique d'une part par l'adoption d'une prévision généralisée du trafic routier sur le long terme, et d'autre part, par l'application d'une nouvelle approche basée sur la méthode des facteurs de croissance et les enquêtes O-D. C'est dans ce contexte que les premiers modèles de trafic sont nés aux États-Unis et ont été importés à la suite en France à partir de 1960. La méthode des facteurs de croissance constitue un des prémices de la modélisation du trafic routier, permettant d'extrapoler le trafic futur à partir des trafics existants (Dupuy, 1975). Dès lors, comme le relève Gabriel Dupuy (1975, p. 19), « *le trafic routier n'est plus seulement considéré*

[85] L'enquête cordon est une enquête de circulation « origine-destination » qui permet d'obtenir des informations sur les déplacements routiers. Elle consiste ainsi à interviewer des automobilistes ou des chauffeurs de poids-lourds sur le déplacement qu'ils sont en train de réaliser. Elle fournit également des informations importantes sur le transport de marchandises (CERTU, 2003).

*comme un débit sur une voie mais comme des flux entre des origines et des destinations repérées dans l'espace, dont la combinaison donne un flux global sur l'axe en question […] et exprime en quelque sorte les "besoins" d'infrastructures* ». On peut noter dans cette modification les apports de la socio-économie des transports qui définissent le trafic routier non pas dans sa seule dimension physique, mais également dans ses manifestations socio-économiques (Lannoy, 2003). L'objet de recherche ne devient plus seulement le trafic routier en soi, mais bien les relations entre deux lieux, d'origine et de destination, qui produisent des flux de circulation au sein de l'espace urbain. Cet enrichissement constitue de surcroît les prémices d'une conception du trafic routier issue de la socio-économie des transports sous la forme de matrice O-D (Commenges, 2013). Néanmoins, au cours de cette période, les modèles se basant sur les facteurs de croissance ont, eux aussi, vite présenté des limites (Dupuy, 1975 ; Bonnel, 2002) :

- En se basant sur les trafics existants la modélisation est dépendante du contexte local, par exemple ; si le trafic relevé $T_{ij}$ est très faible voire nul, le trafic estimé $T'_{ij}$ restera lui aussi faible.
- Elle ne permet pas de simuler correctement le trafic d'échanges avec des zones où l'urbanisation est en forte croissance.
- Elle ne permet pas de prévoir des flux de circulation lorsqu'aucune donnée n'est disponible, par exemple, dans le cas de zones urbaines ouvertes à l'urbanisation.

Ces contraintes ont conduit à abandonner la méthode des facteurs de croissance au profit d'une méthode se basant sur une approche gravitaire (Dupuy, 1975). Cette modification a créé une véritable rupture dans les études de la circulation routière et se traduit par trois modifications interdépendantes (Dupuy, 1975, p. 30) : i. l'adoption d'un horizon long terme ; ii. l'abandon du comptage routier au profit des enquêtes O-D ; iii. la prise en considération de l'ensemble du réseau routier, qu'il soit existant ou en projet. C'est ainsi que les prémices de la modélisation à quatre étapes voient le jour. La mise en œuvre du modèle gravitaire induit aussi des modifications méthodologiques qui se traduisent par une mobilisation systématique des outils informatiques et une complexification de la méthode d'estimation des flux de circulation (Dupuy, 1975). À partir de ce moment, les deux premières étapes de génération-distribution voient le jour et permettent d'établir la demande de déplacement (*cf.* encart 4 p. 136). On comprend bien le contexte de structuration du modèle à quatre étapes, dont le socle de base se fonde sur la prévision du trafic routier, qui est ensuite complété par les deux autres étapes de répartition-affectation pour donner naissance aux modèles à quatre étapes dits de *première génération*, car ils s'intéressent exclusivement aux véhicules particuliers (Chatzis, 2013). Le contexte français suit aussi ce mouvement par l'intermédiaire du SERC[86] (Service d'Études et Recherches de la Circulation Routière) qui va

[86] Une brève histoire d'acronymes liés aux organisme en charge des études du transport routier (Chatzis, 2013) : Le SERC créé en 1954 est l'ancien nom donné au SETRA (*Service d'Études sur les Transports, les Routes et leurs Aménagements*), créé en 1968 lors de la fusion du SERC, du SCET (*Service Central d'Études Techniques*) et du SSA (*Service Spécial des Autoroutes*). Le SERC s'est dissocié entre une branche d'étude, le SETRA, et une branche de recherche, l'IRT (*Institut de Recherche sur les Transports*), devenu ensuite l'INRETS (*Institut National de Recherche sur les Transports et leur Sécurité),* puis à parti de 2011, l'IFSTTAR (*Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux*). En parallèle, à partir de 1976, la compétence « urbaine » des SETRA s'est structurée au sein du CETUR (*Centre d'Études des Transports URbains*), devenu à partir de 1994 le CERTU (*Centre d'Études sur les Réseaux, les Transports, et l'Urbanisme*) suite à la fusion du CETUR et des STU (*Services Techniques de l'Urbanisme*). Depuis 2014, le CERTU, le SETRA, les CETE (*Centres d'Études Techniques de l'Équipement*) et les CETMEF (*Centres d'Études Techniques, Maritimes et Fluviales*) sont regroupés au sein d'un même organisme sous le nom de CEREMA (*Centre d'Études et d'Expertise sur les Risques, l'Environnement, le Mobilité et l'Aménagement*).

jouer un rôle clé pour le développement d'une « *expertise française en matière de modélisation des déplacements urbains* » (Chatzis, 2013, p. 344). Dès lors, l'ingénierie du transport s'est dotée de nouveaux outils permettant de prendre en compte ces trois modifications tout en permettant « *non seulement de trancher dans les choix d'investissement mais aussi de justifier de façon rationnelle les choix effectués […] en ayant une efficacité difficilement contestable* » (Dupuy, 1975, p. 87). Les derniers propos de Gabriel Dupuy montrent aussi toute la légitimité économique que prend la modélisation des transports pour justifier les lourds investissements engagés dans le développement des infrastructures routières.

À partir de la fin des années 1970, on voit s'introduire les premiers logiciels opérationnels de modélisation du trafic routier. C'est aussi à partir de ce moment que le paysage de la modélisation française se distingue du contexte international. Il se traduit particulièrement par une forte institutionnalisation de la modélisation des transports en France, qui a cristallisé les innovations méthodologiques et technologiques qui ont pu voir le jour au même moment de l'autre côté de l'Atlantique (Chatzis, 2013). En France, avec le développement du modèle opérationnel DAVIS, on assiste à ce que Gilles Debizet (2011, p. 11) nomme la « *standardisation de la conception routière* », contribuant « *à ce que la France soit un moment à la pointe de la modélisation, elle a peut-être aussi contribué à son déclin. Privilégiant la circulation automobile par rapport à d'autres modes de déplacement et laissant peu de marge de manœuvre aux concepteurs pour adapter la voirie – et la rue – aux spécificités du contexte urbain* ». Ici l'auteur pointe particulièrement le retard que la modélisation des transports en France a pris pour intégrer une approche multimodale alors que les Nord-Américains engageaient déjà des réflexions sur la capacité du modèle à quatre étapes à prendre en compte un plus grand nombre de mode de transport, à s'intéresser à l'individu, ou à simuler le comportement des individus (Chatzis, 2013).

#### 5.1.2. **L'émergence des critiques de la modélisation *classique* du trafic routier : « *zone don't travel ; people travel !* »**

Au cours de la seconde période (1975-1990), on assiste à une légitimation des modèles de prévision du trafic autant sur le long terme (implantation de nouvelles infrastructures) que sur le court terme (plan de circulation urbain, étude de la pollution atmosphérique…). Gabriel Dupuy (1975, p. 107) identifie d'ailleurs cette période comme « *significative dans l'histoire du développement des modèles de trafic, […] les dossiers pilotes faisant tous référence à l'utilisation de modèles* ». On assiste donc à l'aboutissement méthodologique et institutionnel de la modélisation à quatre étapes positionnant ainsi les modèles de trafic comme un outil courant dans la planification des transports urbains français (Dupuy, 1975 ; Bonnel, 2010). Pourtant, au même moment, cette période se caractérise par une remise en cause de la modélisation à quatre étapes. Ses critiques sont de deux natures, portant autant sur le *fond*, les bases théoriques et les hypothèses sur lesquelles se construit le modèle à quatre étapes, que sur la *forme* des modèles, la structure et les techniques employées (Masson, 2000). Si ces critiques ont été émises autant du côté français que nord-américain, force est de constater que le paysage de la modélisation française s'est construit en dehors du contexte international, et qu'il a tardé à adopter les apports des nouvelles approches initiées depuis bien plus longtemps aux États-Unis (Chatzis, 2013).

Les critiques de fond portent plus spécifiquement sur les hypothèses sur lesquelles se fondent les modèles de trafic. Sophie Masson(2000, pp. 27–29) identifie trois critiques qui portent : 1. sur la vocation des modèles, établis au service de l'automobile ; 2. sur les problèmes de calibrage des

modèles ; 3. sur leur inadéquation à établir des prévisions sur le long terme. Ici, afin d'illustrer nos propos sur l'évolution méthodologique des modèles de trafic, nous nous focaliserons plus spécifiquement sur la première critique. En effet, après les années 1975, on arrive à une prise de conscience que le « *tout automobile* » qui avait accompagné une période économique florissante arrive à ses limites. Dès lors, outre les problématiques techniques de dimensionnement des infrastructures, les préoccupations en termes de qualité de vie émergent, tout autant que les problèmes de qualité de l'air ou de congestion routière. Le début de cette période se caractérise plus particulièrement par une « *triple crise des déplacements (financement, congestion, environnement)* » (Bonnel, 2002, p. 23). On assiste à une fin du développement des grandes infrastructures routières, faisant progressivement glisser les besoins en termes de gestion des infrastructures autant sur le long que sur le court terme. Ces tendances illustrent bien l'une des critiques les plus récurrentes émise à l'encontre des modèles à quatre étapes, considérés comme « *un édifice en faveur de l'investissement routier automobile* » (Bonnel, 2002, p. 17). En effet, l'application de la modélisation du trafic routier n'a permis d'entrevoir qu'un développement croissant et continu de la circulation automobile sans s'ouvrir à d'autres modes de transport, comme par exemple les transports en commun. De plus, les infrastructures étant maintenant développées, le regard se portera d'autant plus sur l'existant et sa gestion que sur le développement de nouvelles routes, faisant glisser les besoins vers le court et moyen terme. À partir de ce moment, on aura tendance à parler de modélisation des déplacements, alors qu'initialement le terme de modélisation du trafic était utilisé (Debizet, 2011 ; Meunier, 2013).

Parallèlement, on voit aussi émerger des critiques portant sur la pertinence de la modélisation classique. On retrouve ainsi des critiques de fond montrant le « *manque de pertinence* » du modèle à quatre étapes à être explicatif des comportements individuels (Bonnel, 2002, p. 21). Ces critiques principalement issues des sociologues de l'École de Chicago soulignent que la mobilité telle qu'elle est envisagée par la socio-économie des transports ne permet pas de rendre compte de la multiplicité des facteurs influençant la mobilité (Lannoy, 2003 ; Gallez et Kaufmann, 2009). Malgré tout, ces critiques montrent toutes les limites de la modélisation à quatre étapes, considérée comme trop simpliste. Patrick Bonnel (2010) relève qu'aucune solution opérationnelle n'a pour le moment été trouvée. Pour répondre à ces critiques, Ortúzar et Willumsen (2011, p. 22) conscients des limites du formalisme séquentiel, proposent de positionner le modèle à quatre étapes avant tout comme un outil de référence : « *l'évolution des emplois, le lieu de résidence, le choix des lieux d'achat et ainsi de suite ; sont autant d'éléments qui vont réagir, au moins partiellement, à des changements de l'accessibilité fournie par le système de transport. Malgré ces observations, le modèle séquentiel en quatre étapes fournit un point de référence afin de le comparer avec des méthodes alternatives* ».

Les critiques de forme portent, quant à elles, sur la structure et les techniques employées. Elles « *se situent à plusieurs niveaux et sont en partie liées les unes aux autres* » (Masson, 2000, p. 24). Elles portent autant sur les limites de la séquentialité en quatre étapes et la causalité linéaire du processus séquentiel, que sur la lourdeur des modèles ou leur capacité à intégrer une approche comportementale des individus… Les deux éléments les plus caractéristiques de ces critiques portent sur : i. le caractère agrégé des modèles qui ne permet pas d'intégrer l'hétérogénéité des comportements individuels ; ii. les zonages utilisés lors de la génération des populations : « *zone don't travel; people travel!* » (McFadden, 2000, p. 1). Pour répondre à ces deux limites, on assiste à la mise en œuvre d'une démarche de modélisation désagrégée initiée aux États-Unis (McFadden,

2000). L'approche désagrégée constitue l'avancée théorique et méthodologique la plus significative de cette période en permettant de centrer l'analyse non plus sur les flux mais plutôt sur l'individu en tenant compte autant de ses préférences que de ses choix[87] (McFadden et Reid, 1975). Le développement de cette approche se distingue de la démarche traditionnelle : « *d'une part, l'analyse est effectuée au niveau du décideur, tel qu'une personne ou un ménage, plutôt qu'à partir des flux zonaux. D'autre part, ces modèles se fondent sur la théorie micro-économique du comportement du consommateur* » (Masson, 2000, p. 36). Avec cette approche centrée sur les individus, on assiste au lancement des premiers modèles de microsimulation du trafic routier. Toutefois, leur pénétration au sein du paysage français se fera bien plus tardivement et avec une certaine frilosité au cours des années 1990, alors que le premier modèle opérationnel est lancé dès 1970 sur la ville de San Francisco (Chatzis, 2013). Ainsi, malgré les critiques, le modèle à quatre étapes s'est progressivement inséré dans le paysage de la modélisation comme une référence dans le champ des études de transport[88] (Ortúzar et Willumsen, 2011).

Cependant les modèles opérationnels employés en France évoluent peu sur cette période, et en dépit des critiques ils connaissent un franc succès dans le paysage de la modélisation des déplacements. Les évolutions se traduisent plutôt par un raffinement méthodologique des quatre étapes, en vue de mieux prendre en compte d'autres modes de transport (Chatzis, 2013), ainsi que d'élargir les champs d'application des modèles vers les problématiques plus générales des déplacements urbains. L'application du modèle DAVIS, largement privilégié et utilisé jusque dans les années 1990, illustre cette cristallisation des modèles de première génération du côté français. En effet, l'évolution de celui-ci vers l'intermodalité s'est faite tardivement par un croisement entre les modèles DAVIS et VISUM qui a donné naissance au modèle intermodal DAVISUM (Chatzis, 2009). Pourtant dans le contexte international, de nombreuses évolutions théoriques ont vu le jour à travers la modélisation désagrégée, qui vont s'introduire en France plus tardivement en s'entourant plutôt des apports de la géosimulation que de la microsimulation. Si la relation entre les apports de la microsimulation et l'insertion des approches de géosimulation n'est pas si simple, Aurore Cambien (2010, p. 40) relève quand-même que « *le développement de la microsimulation est parfois interprété comme la genèse des modèles actuellement basés sur les agents* ».

#### 5.1.3. **L'apport des outils de géosimulation dans les modèles de trafic : un bouleversement méthodologique**

Au cours de la troisième période (1990 à aujourd'hui), on assiste à une diversification des approches de modélisation du trafic. En effet, de nombreux modèles macroscopiques et aussi microscopiques s'introduisent dans le champ de la modélisation[89]. On retrouve deux innovations distinctes : les premières se situent dans la continuité des premiers modèles de trafic tout en intégrant les apports de la théorie des choix discrets à une échelle désagrégée (McFadden, 2000) ; alors que les secondes s'attachent plutôt à simuler le déplacement des individus à un niveau de détail spatial et temporel plus fin, en cherchant à simuler les déplacements des individus dans

[87] La théorie du consommateur étant, par exemple, l'une des approches privilégiée pour intégrer un comportement de l'individu dans le processus de modélisation (Quinet, 1998).
[88] Ces propos expliquant sans doute le fait que la notion de modèle « classique » à progressivement remplacé celui de modèle à quatre étapes.
[89] Notons ici que ces évolutions sont plus particulièrement introduites par l'intermédiaire de bureaux d'études privés ou de laboratoires de recherches (Chatzis, 2009 ; Debizet, 2011).

l'espace (Masson, 2000). Plus précisément, on voit s'implanter, d'une part, des modèles de microsimulation, et d'autre part, des modèles de simulation multi-agents. Leur développement est initié en réponse aux critiques émises depuis les années 1970-1980 à l'encontre de la modélisation classique. Si les bases théoriques trouvent leurs origines dans les années 1970 (Cambien, 2010), ce sont surtout les progrès en informatique qui ont permis d'envisager l'opérationnalité de ces approches.

Grâce aux apports théoriques des approches désagrégées et à l'essor des techniques liées à l'informatique (calcul haute performance, diversification des systèmes d'exploitation et des langages…), on voit s'introduire dans le champ des transports urbains des modèles novateurs, se basant plutôt sur une approche microscopique des déplacements urbains, permettant d'une part la modélisation de la demande de déplacement des usagers, et d'autre part de simuler le comportement des véhicules sur une infrastructure de transport (Cambien, 2010). Les approches de modélisation du trafic s'élargissent et s'enrichissent avec l'émergence d'une plus grande diversité de méthodes : certaines se situent dans la continuité des modèles classiques, alors que d'autres utilisent les apports de la modélisation désagrégée (De Palma, 2013) ou de la géosimulation (Fellendorf et Vortisch, 2010 ; Behrisch *et al.*, 2011). Avec ces nouveaux modèles, comme par exemple TRANSIMS (Balmer *et al.*, 2004 ; Balmer, 2007), on assiste à un renversement méthodologique et théorique des déplacements, qui jusqu'alors étaient appréhendés par un mécanisme *top-down*. Les apports de la théorie de la complexité (*cf.* section 1.1.1, p.24) et le développement des approches orientées objets permettent d'entrevoir la ville selon un processus *bottom-up* plutôt centré sur l'individu (Cambien, 2010, p. 44) : « *les phénomènes urbains ne sont plus conçus comme le seul produit de l'action planificatrice […] Désormais, ils sont envisagés comme le résultat des multiples interactions entre les mécanismes économiques, démographiques, juridiques, sociaux…* ».

Les modèles mis en place gagnent alors en résolution spatiale, qui tient compte de la finesse avec laquelle le zonage du terrain d'étude a été mis en place lors de l'étape de génération (Bonnel, 2010). On distingue deux approches de génération des populations : une *méthode d'agrégation* (*bottom-up*) qui consiste à restituer au sein de chaque zone un individu moyen qui sera représentatif des caractéristiques de la zone ; une *méthode de désagrégation* (*top-down*) qui consiste non plus à reproduire un individu moyen à l'intérieur d'une zone mais plutôt à redistribuer les informations socio-économiques à l'échelle du ménage, voire de l'individu (*cf.* figure 2.10, ci-dessous). Les deux méthodes présentent des forces et limites. Le processus d'agrégation s'avère le plus simple à mettre en place et se rattache à la modélisation macroscopique. Il présente des avantages indéniables face à la disponibilité des données. Néanmoins, il peut aussi être perçu comme une perte d'information depuis les données d'observation, qui ont une résolution plus fine[90] (Bonnel, 2010). La désagrégation, quant à elle, vise à répondre à cette limite en se fondant sur une approche *top-down*, afin de reconstituer une population dite synthétique en vue de restituer plus fidèlement les caractéristiques individuelles de la population de l'espace d'étude (Beckman *et al.*, 1996 ; Banos et Thévenin, 2008). Ce choix entre agrégation ou désagrégation est nécessairement tributaire des données accessibles en vue de répartir un échantillon de données socio-économiques représentatif du comportement de mobilité des individus au sein de chaque

[90] Pour répondre à ces limites, Patrick Bonnel (2010) cite plusieurs méthodes d'agrégation pour l'étape de génération : les modèles par régression qui consistent à construire un individu moyen et les modèles catégoriels qui consistent à segmenter la population initiale de la zone en plusieurs groupes ayant des caractéristiques de déplacement hétérogènes.

zone préalablement définie (Robert, 2001). Toutefois, avec les approches individuelles se pose de plus en plus la question des données, rarement disponibles à une échelle individuelle. Si la génération des populations synthétiques offre une alternative au manque de données, il est souvent relevé que ces approches présentent une faiblesse quant aux données accessibles qui vont d'une part alimenter le modèle, et d'autre part, permettre de valider ces modèles (Balmer, 2007 ; Banos et Thévenin, 2008). Par ailleurs, on peut noter à ce sujet sur le long terme un potentiel de développement non-négligeable à travers les *big data* et les données de capteurs (*cf.* Chapitre 2, p. 55), lors de la mise en œuvre d'une approche désagrégée en vue d'obtenir des données à une échelle individuelle. Pour le moment, il n'est pas possible d'y associer de nouvelles approches opérationnelles, mais on peut tout de même y voir une quatrième phase de développement pour la modélisation du trafic. Enfin, en se fondant sur un mécanisme *top-down*, la modélisation désagrégée permet d'envisager la modélisation sous un angle orienté-objet, en se focalisant sur les composants élémentaires du système à étudier. Cela induit nécessairement une modification des outils permettant de modéliser, mais surtout de simuler ces objets dans un espace géographique :

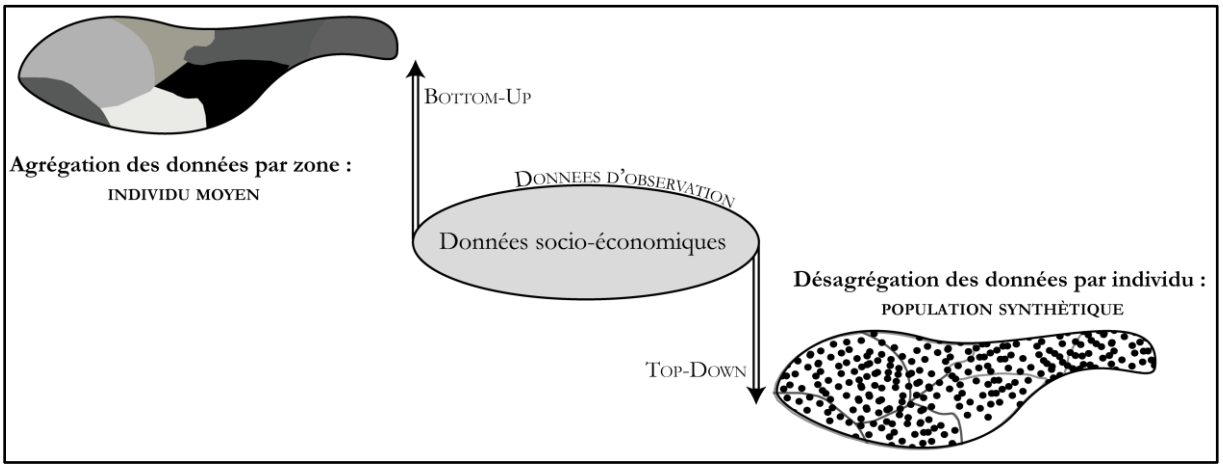



**Figure 2.10 : Un bouleversement méthodologique avec les approches désagrégées : l'approche *top-down***

Enfin, on peut noter que l'évolution des modèles de déplacement et de trafic se calque sur l'évolution des approches de simulation présentées précédemment (*cf.* section 4.1.2, p. 130). En effet, que ce soit pour les modèles de planification comme TRANSIMS ou MATSim (Balmer, 2007), ou pour les outils de simulation du trafic routier comme SUMO (Behrisch *et al.*, 2011) ou VISSIM (Fellendorf et Vortisch, 2010), ces nouvelles approches se fondent toutes sur un formalisme *top-down*, en intégrant en grande partie une approche de géosimulation où « *chaque individu est présenté comme une entité autonome non seulement au cours du processus de modélisation de la demande, mais aussi d'affectation* » (Balmer, 2007, p. 13). Ces derniers développements permettent de mieux appréhender les dynamiques intraurbaines autant au niveau spatial que temporel. Ils se fondent sur une double évolution : conceptuelle à travers l'approche *top-down* et méthodologique à travers la géosimulation. Ainsi, inscrire notre travail au sein de la géosimulation pour la simulation du trafic routier se fait en concordance avec les évolutions récentes, autant dans le domaine de la modélisation en sciences humaines et de la modélisation des déplacements urbains que dans le domaine de la simulation du trafic routier.

## 5.2. Les approches de géosimulation pour la représentation du trafic routier : quelle méthode privilégier ?

Notre objectif vise à reproduire un trafic routier émergent à partir d'un dispositif de quantification issu des capteurs fixes placés en milieu urbain. La démarche que nous envisageons est exploratoire et se fonde sur une approche abductive en positionnant les données au premier plan de la démarche, sans partir de la théorie (*cf.* figure 1.7, p. 39), mais en s'entourant plutôt de la richesse du dispositif de quantification qui a été construit. Dans un chapitre d'ouvrage, Pierre Livet, 2006 (pp. 194–195) liste huit conditions qui justifient l'usage de la simulation multi-agents. Nous retiendrons plus particulièrement la septième condition qui correspond à notre positionnement, non seulement abductif mais aussi guidé par les données : « *lorsque l'on n'a pas de théorie du domaine, que l'on connait certaines des formes des phénomènes réels, mais pas toutes, et que par ailleurs, on a des scénarios typiques de comportement des éléments qui interagissent, et que l'on dispose de processus qui ont des similarités avec ces types de comportement* ». Dans cette condition, l'usage des SMA pour notre travail devrait permettre de retrouver les formes des phénomènes réels et observés (le trafic routier) par un aller-retour avec les données d'observation (les comptages routiers) pour alimenter, puis valider notre démarche.

### 5.2.1. La géosimulation : des approches de simulation de l'espace géographique

Nous l'avons vu précédemment, la simulation se distingue du modèle dans le sens où elle utilise « *des relations moins profondes et moins structurelles que le modèle, mais elle se rapproche plus de la réalité que lui* » (Lévy et Lussault, 2003, p. 926). De plus, Itzhak Benenson et Paul Torrens (2004), qui sont à l'origine du terme de *géosimulation*, relèvent que celle-ci se distingue des autres approches de simulation en mettant l'accent sur les dimensions spatiale et géographique des objets. Sur son site internet, Paul Torrens[91] met en avant quatre aspects faisant la spécificité des approches de géosimulation vis-à-vis d'autres simulations :

- **La représentation des unités spatiales** à un niveau totalement désagrégé, mettant l'accent sur les entités élémentaires de l'espace comme par exemple les individus, les ménages, ou les véhicules. Dans ce sens, la géosimulation est une **approche objet-centrée**, en se concentrant sur la représentation et le comportement des objets en interaction avec les autres (Benenson et Torrens, 2004).
- **La représentation des relations spatiales**, en se concentrant sur les interactions entre les entités élémentaires du système étudié selon un principe *bottom-up* (partant du bas). Cette composante fait directement référence à la théorie de la complexité que nous avons abordée dans les grandes lignes précédemment (*cf.* section 1.1.1.b, p. 26).
- **Le traitement du temps** en géosimulation offre l'occasion de construire une simulation à des échelles de temps approchant un temps réel. À cet effet, elle ouvre la porte à une approche spatio-temporelle plus explicite que, par exemple, les outils de géovisualisation qui se fondent plutôt sur une représentation en *snapshot* ou que par l'intermédiaire d'une représentation composite (Klein, 2007 ; Thévenin, 2011).
- **Les objectifs de la géosimulation**, dont la portée n'est pas uniquement à vocation prédictive ; les approches de géosimulation ont tendance à être « *des outils pour penser avec* », plutôt conçus dans l'optique d'explorer différents scénarios des systèmes à étudier.

[91] Paul Torrens, « *An introduction of Geosimulation* » : http://www.geosimulation.org/geosim/.

La géosimulation se définit comme une démarche de simulation du monde réel désignant « *un ensemble de concepts, méthodes et techniques, qui combinent les apports de l'intelligence artificielle et des sciences de l'information géographique, pour simuler la dynamique des systèmes spatiaux* »[92]. Plus généralement, on pourrait alors situer la géosimulation à l'intersection de trois courants de pensée (Benenson et Torrens, 2004), que l'on peut relier aux éléments de la *géocomputation* présentés plus haut :

- **la science de l'information géographique,** par une prise en compte explicite de la dimension spatiale en vue d'initialiser, construire, calibrer et valider l'approche de géosimulation[93].
- **la science informatique,** par l'intégration de mécanismes de mémorisation, d'apprentissage, et d'évolution permettant de faire ainsi évoluer une entité au sein d'un système numérique[94].
- **la théorie des systèmes complexes** fournit un cadre de réflexion pour étudier les interactions qui se produisent à l'échelle locale. Les interactions sont d'ordre hiérarchique, organisant les entités les plus élémentaires en vue d'étudier et de comprendre les phénomènes émergeant à des échelles plus globales.

Enfin, les méthodes de géosimulation ont connu les développements les plus significatifs au sein de la recherche urbaine (Benenson et Torrens, 2004). Les auteurs défendent l'idée que la géosimulation constitue actuellement une nouvelle vague, voire un nouveau courant, pour l'étude des systèmes urbains. Il est vrai que les méthodes de géosimulation couvrent de multiples cadres d'études, que ce soit pour l'occupation du sol (Antoni, 2006), l'extension urbaine (Batty, 2007), les réseaux de villes (Pumain, 2014b), les mobilités quotidiennes (Banos *et al.*, 2010) ou les mobilités résidentielles (Badariotti et Weber, 2002 ; Hirtzel, 2015)… Pour l'ensemble de ces travaux, l'approche spatio-temporelle constitue un prérequis important pour en étudier les dynamiques urbaines. Néanmoins, la dynamique inhérente aux approches de géosimulation ne constitue pas le seul avantage de la géosimulation. Paul Torrens (2000) identifie cinq atouts de la géosimulation pour l'étude des systèmes urbains par rapport aux approches classiques de modélisation urbaine. Ces cinq points expliquent en grande partie le succès de ses nombreuses applications sur le milieu urbain :

- **Le niveau de détail** : la géosimulation est désagrégée, permettant de simuler les interactions entre les individus et avec les éléments de l'espace urbain.
- **La dynamique** : la géosimulation est intrinsèquement dynamique, permettant d'établir des simulations en « temps réel ».
- **La construction** : les algorithmes mobilisés peuvent être obtenus directement à partir des idées théoriques, permettant de comprendre le fonctionnement urbain.
- **La formalisation** : si elle est toujours complexe dans sa construction, la formalisation des modèles apparait plus intuitive que la plupart des techniques traditionnelles.
- **L'interface graphique** : les environnements sont pour la plupart du temps visuels et interactifs permettant de faciliter l'interprétation du modèle.

[92] Lena Sander : définition « **Géosimulation** » : http://www.hypergeo.eu/spip. php?article311.
[93] Les bases de données issues des logiciels de SIG fournissent des informations sur les entités et les *objets*, qui, une fois couplées à la géosimulation, offrent un cadre contextuel pour reproduire le monde virtuel.
[94] Ces mécanismes reposent en partie sur les fonctionnalités des automates et/ou objets, ainsi que sur les fonctionnalités issues de l'intelligence artificielle (AI).

Du côté opérationnel, comme illustré au sein de la figure ci-dessous, deux approches de géosimulation sont mises en avant (Benenson et Torrens, 2004 ; Batty, 2007) : la simulation multi-agents (SMA) et les automates cellulaires (AC). Que ce soit pour les SMA ou les AC, la démarche de géosimulation est considérée comme générative (Epstein, 1999). Il est alors important de relever que ces deux approches présentent de nombreuses similitudes (Batty, 2007), comme le montre, par exemple, les travaux de Dominique Badariotti et de Christiane Weber (2002) sur la mobilité résidentielle à Bogota

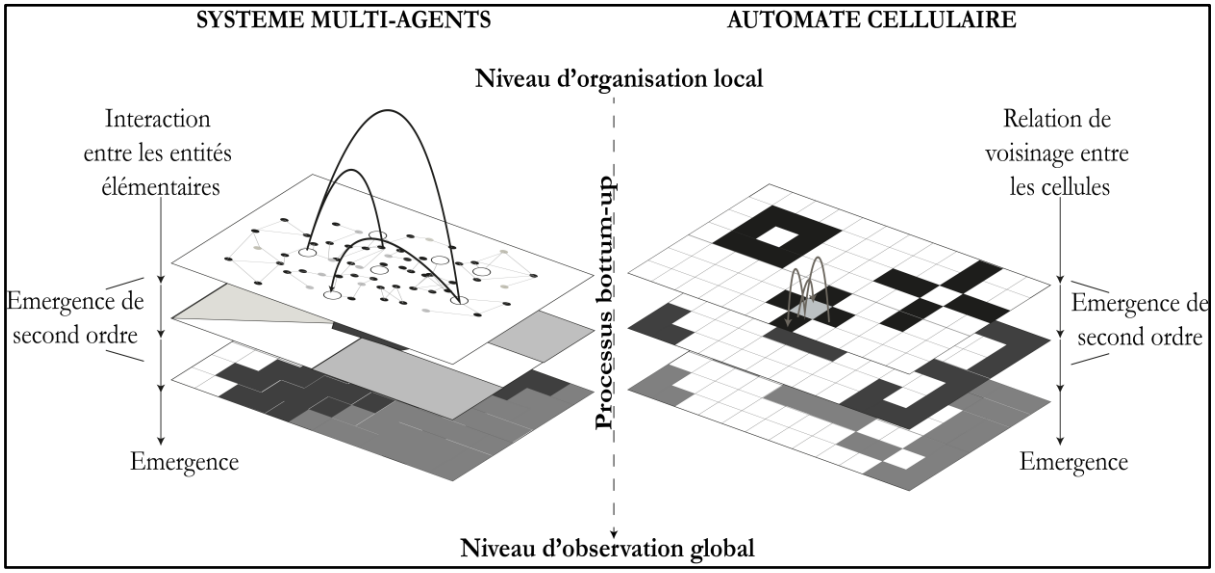



**Figure 2.11 : Processus d'organisation et d'observation de deux approches de géosimulation : le système multi-agents (à droite) et l'automate cellulaire (à gauche) (d'après Daudé, 2004 et Torrens, 2010)**

Comme illustré en figure 2.11, la géosimulation vise, à partir d'un processus ascendant (aux niveaux les plus fins), à identifier et analyser des situations à des échelles globales. À notre sens, l'hypothèse commune des deux approches se situe au niveau de cet aller-retour (Daudé, 2004, p. 5) : « *en se situant au niveau des entités élémentaires […] on postule ainsi que la répétition de processus plus ou moins élaborés au niveau individuel a une incidence sur le devenir de la diffusion à un niveau global, à la fois dans le taux de propagation, leur distribution spatiale et la vitesse de diffusion* ». Elles se définissent comme des modèles individu-centrés dont les particularités se fondent sur trois caractéristiques (Benenson et Torrens, 2004) : la **représentation** des objets ; le **comportement** des objets ; la **durée (temporalité)** des objets.

### 5.2.2. **Automate Cellulaire (AC) ou Simulation Multi-Agents (SMA) pour le trafic routier ?**

Les AC traduisent le monde virtuel sous la forme d'une grille régulière, où les objets constituent les différentes cellules de cette grille. La prise en compte du temps se fait par l'intermédiaire de processus de changement et d'évolution d'état des cellules en fonction de l'état des autres cellules qui lui sont adjacentes. La dynamique est alors approchée par la transition de variations locales de proche en proche à l'issue de chaque cycle de simulation (Batty, 2007). Elle constitue aussi une des limites des AC en ne pouvant intégrer des objets mobiles à proprement dit. Nous illustrons ces propos à travers les modèles de trafic routier. En effet, comme illustré en figure 2.12, en

fonction de l'approche de géosimulation choisie, l'espace va être formalisé soit sous la forme d'un espace discontinu, soit sous la forme d'un espace continu (Krajzewicz *et al.*, 2002) :

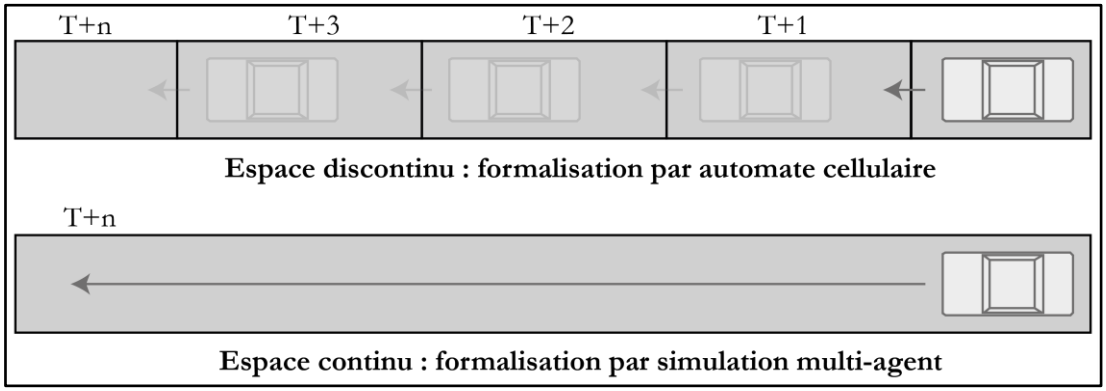



**Figure 2.12 : La représentation d'une infrastructure routière sous la forme d'une espace discontinu (en haut) et d'un espace continu (en bas) pour le déplacement d'un véhicule (Source : d'après Krajzewicz *et al.*, 2002, p. 1)**

Typiquement, l'approche par AC permet de formaliser l'infrastructure routière sous la forme d'un ensemble de cellules ; dans ce sens, les objets mobiles ne peuvent y être explicitement intégrés. En effet, au sein des AC, les véhicules se déplacent discrètement cellule après cellule, c'est pourquoi on parlera plutôt d'espace discontinu. Inversement, les SMA offrent l'avantage de se construire sous la forme d'un espace continu où le déplacement des véhicules et leur position évoluent tout au long de la route. Il faut noter que les deux formalismes présentent leurs avantages et inconvénients selon l'échelle de travail choisie. Ainsi, dans une perspective macroscopique ou mésoscopique, on privilégiera une représentation des véhicules dans un espace discontinu ; alors que dans une perspective microscopique, la représentation d'un espace continu apparait plus adéquate. Dans cette acception, la mise en place d'une approche par les SMA constitue, à notre sens, une bonne réponse pour reproduire la circulation routière en milieu urbain sous une forme continue. Les SMA présentent l'avantage de représenter le trafic routier sous sa forme la plus élémentaire : le véhicule en déplacement sur une infrastructure de transport.

C'est pourquoi les SMA contribuent à combler les lacunes des AC pour la simulation du trafic routier à une échelle microscopique (*cf.* figure 2.12). Les modèles multi-agents ou simulations multi-agents (SMA) constituent le dernier développement des approches de simulation identifiées en sciences humaines (*cf.* figure 2.2, p. 131), et connaissent un véritable succès dans notre discipline (Amblard et Phan, 2006). Certains auteurs voient dans les SMA une véritable alternative à la modélisation (Benenson et Torrens, 2004 ; Ferber, 2006 ; Marilleau, 2006 ; Torrens, 2010 ; Bouquet *et al.*, 2015 ; Torrens, 2015). Pour justifier ces propos, Jacques Ferber (2006) voit dans les SMA plus qu'une technique de modélisation (p. 27) : « *l'utilisateur d'un SMA réalise un modèle en deux temps : dans un premier temps il réalise un laboratoire virtuel qui reprend en miniature les caractéristiques du domaine empirique objet de la simulation, ou "système-cible" ; dans un deuxième temps, il pratique des expérimentations, déplaçant des individus, changeant leur comportement, modifiant les conditions environnementales. De ce fait, son modèle, sa vision du système cible, s'élabore au fur et à mesure. Elle n'est pas issue d'une abstraction d'un monde réel, mais d'une reconstruction, d'une recréation d'un monde artificiel. Il élabore des hypothèses et peut les tester in vitro en utilisant les mêmes types d'observation que dans le monde réel […] Ainsi, à la différence des approches classiques, une SMA n'est pas une boite noire, et la SMA ne se réduit pas à l'implémentation d'un modèle et à l'analyse de sa réponse en fonction de paramètres d'entrée, mais participe au*

*processus de recherche de modèles* ». Par conséquent, s'inscrire dans ce courant et considérer les SMA comme un outil de modélisation en géographie, c'est aussi s'inscrire dans un cadre de pensée, voire un « *paradigme agent* » qui est « *avant tout plus une manière de penser qu'une technique d'implémentation […]. Il organise simplement notre réflexion par analogie avec le monde qui nous entoure* » (Bouquet *et al.*, 2015, p. 18). Les SMA se fondent sur une structuration plus complexe que les AC en permettant de représenter différents types d'agents (ménages, individus, véhicules, piétons…) au sein d'un environnement réel ou virtuel. Michael Batty (2007, p. 209) définit l'*agent* comme étant « *une "cellule mobile" située par rapport aux cellules mais qui peut se déplacer entre les cellules* ». Toutefois, cette définition ne montre que la complémentarité des SMA avec les AC et apparait relativement restreinte. Nous retiendrons plutôt la définition apportée par Jacques Ferber (2006, p. 24) : « *on appelle "agent" une entité informatique (un processus) située dans un environnement, plongée dans une structure sociale, qui est capable de : 1. agir dans son environnement ; 2. percevoir et partiellement se représenter son environnement (et les autres) ; 3. communiquer directement ou indirectement avec d'autres agents ; 4. mue par des tendances internes (buts, recherche de satisfaction, drives, utilité) ; 6. se conserver et se reproduire ; et qui représente un comportement autonome qui est la conséquence de ses perceptions, de ses représentations et de ses communications* ».

### 5.2.3. **Un pas nécessaire vers la simulation multi-agents pour la mise en mouvement des dynamiques intraurbaines ?**

La dynamique des SMA diffère de celle des AC en se fondant autant sur une interaction entre les agents (agents-agents) qu'avec l'environnement (agent-environnement). L'*agent* est un mobile capable de transformer sa position et son orientation au cours du temps sans en modifier sa forme sur la base de ces deux types d'interactions. Ainsi, les SMA apparaissent mieux répondre aux thématiques liées à la mobilité géographique et aux dynamiques intraurbaines (Marilleau, 2006, p. 33) : « *Les systèmes multi-agents renferment toutes les qualités requises pour la simulation de la mobilité géographique. Ils offrent la possibilité de décrire l'individu dans son unicité et de le mettre en compétition avec d'autres agents sur un même environnement représentatif d'un territoire réel. Ainsi, l'observation du comportement de chaque mobile, comme celui de la population toute entière est possible. Sur ce point, les approches agents surpassent allègrement les méthodes analytiques traditionnelles qui se limitent à une étude macroscopique de la mobilité* ».

Bernard Moulin et Walid Ali (2008) présentent la SMA en géographie comme le couplage de deux technologies : 1. la **simulation à base d'agents**, permettant de représenter les entités au sein d'un environnement virtuel ; et 2. les **SIG**, permettant d'introduire les différentes caractéristiques des données au sein du simulateur multi-agents. Ce couplage met l'accent sur la notion d'*agent géographique*, avant tout perçu comme un *objet* doté de différentes caractéristiques (Daudé et Langlois, 2006) :

- une **emprise spatiale**, qui est l'identité de l'*objet* permettant d'en délimiter ses contours (extérieurs et intérieurs)
- une **description matérielle**, qui est soit une description élémentaire (taille, forme,…) si sa forme est finale ; soit une description « non-terminale » si l'*objet* est composé de sous-ensembles (groupes/individus)
- des l**iens avec son environnement**, qui indiquent que l'agent est dépendant des objets qui lui sont proches spatialement (contraintes), ou de façon descriptive (même groupe

d'*objets*)

- un **comportement**, qui permet de décrire la capacité de l'*objet* à faire un choix parmi d'autres choix en fonction de son emprise spatiale, de sa description et de ses liens avec l'environnement.

La constitution d'une approche *objet-centrée* nécessite non seulement d'identifier les *objets* à simuler, mais aussi de constituer un corpus de données spatio-temporelles permettant d'alimenter le modèle de simulation. Deux éléments apparaissent déterminants pour la géosimulation multi-agents : d'une part, l'*environnement géographique* permettant de construire le monde virtuel, et d'autre part, les *agents géographiques* qui vont s'insérer au sein de cet environnement en fonction des données disponibles. L'environnement géographique correspond à « *un support géométrique (donnant à la fois sa forme, et sa localisation dans l'espace), auquel on associe une description matérielle et un identifiant* » (Langlois, 2006, p. 365). L'utilisation des SIG pour construire l'environnement géographique correspond à une formalisation numérique du territoire réel ou virtuel (la ville, les bâtiments, les routes, les lieux d'activités…). En mobilisant les outils de la géomatique, la constitution de l'environnement n'apparait pas complexe, d'autant plus qu'il existe un grand nombre de bases de données permettant de formaliser l'espace géométriquement. On peut, par exemple, citer les bases de données produites par l'IGN, comme la BD TOPO[95] qui est une base de données de référence pour les problématiques liées à l'aménagement, l'environnement ou l'urbanisme. Cette base de données 3D permet d'obtenir une photographie relativement fidèle du territoire sous une forme vectorielle et à une résolution spatiale très fine (de 50 à 500 mètres de résolution). Cette formalisation de l'environnement s'opère par des couches possédant chacune des propriétés particulières décrivant sa topologie, ainsi que ses caractéristiques (Langlois, 2006).

Construite à partir des SIG, la constitution de l'environnement géographique ne comporte aucune évolution (Langlois, 2006 ; Marilleau, 2006) : c'est un invariant du système multi-agents ; *a contrario*, la dynamique est plutôt portée par l'*agent géographique*, qui est perçu comme une entité génératrice (Epstein, 1999) de la dynamique. Il est donc nécessaire de construire des *objets mobiles* qui se déplacent selon différentes contraintes (Langlois, 2006), soit contenues par l'espace (une maille, un pixel, un élément du réseau), ou associées à la forme des objets (impact et/ou rencontre) ou alors liées à la description attributaire de l'environnement (vitesse de déplacement, accessibilité). Mobilisant aussi les SIG, la constitution de cette information en amont doit être individuelle ; or il apparait évident que les données ne sont que rarement exhaustives, de même qu'elles ne sont ni continues dans l'espace, ni dans le temps. Dans ce sens, la solution la plus habituellement utilisée est de faire appel aux populations synthétiques, permettant de créer une population d'individus virtuels (agents) qui se rapproche au plus près des caractéristiques des populations initiales (Balmer *et al.*, 2004 ; Banos *et al.*, 2010).

Si la géographie s'intéresse à la SMA depuis quelques années, c'est aussi parce qu'elles partagent un langage commun. La notion d'*objet/agents* que nous avons citée à plusieurs reprises dans les sections précédentes apparait être un concept central de l'approche de modélisation géographique (Daudé et Langlois, 2006). Par essence, l'intérêt porté aux approches spatiales, dont la géographie est partie prenante, s'allie de plus en plus avec le besoin de comprendre les mécanismes générateurs de l'espace géographique. Ainsi, le temps est aussi un domaine d'exploration, dont le

[95] Descriptif de la BD TOPO : http://professionnels.ign.fr/bdtopo.

géographe peut aujourd'hui se saisir plus aisément pour l'étude des dynamiques intraurbaines par l'intermédiaire des différents outils de la *géocomputation*.

Les approches *agents* en géographie se fondent sur trois dimensions (Torrens, 2015) : les deux premières concernent l'espace (*x,y*) alors que la troisième complexifie la structuration du modèle de simulation pour prendre en compte le temps. Pour Paul Torrens (2015), l'intérêt à porter à cette troisième dimension nécessite de se focaliser sur la qualification des objets, sur les informations géographiques disponibles, ainsi que sur le sens donné aux lieux et à l'espace. Cette troisième dimension implique aussi de s'arrêter sur les interactions entre les différents éléments du système mis en place (*cf.* section 6.2.2, p. 196).

## 5.3. La SMA pour la modélisation du trafic routier à partir des capteurs de trafic routier : le modèle SCAUP

Comme nous l'avons vu plus haut, la démarche de modélisation envisagée se construit à travers plusieurs modèles : le modèle conceptuel, le modèle de données et le modèle de simulation. L'objectif de cette section est d'offrir une représentation du trafic routier à partir des données issues des capteurs fixes dans une optique de simulation multi-agents. Cette section s'insère comme étant la première partie de la démarche de modélisation présentée au sein de la figure 2.1 (p. 128). Le modèle de données ainsi que le dispositif de quantification étant maintenant identifiés, le passage au modèle de simulation peut être envisagé. À travers le modèle SCAUP (*Simulation multi-agents à partir de CApteurs Urbains pour la Pollution atmosphérique automobile*) proposé dans cette thèse, nous proposons ici une démarche exploratoire autour d'un modèle de simulation multi-agents où les données de trafic sont directement issues des capteurs fixes (*cf.* Chapitre 3, p. 81).

Avant d'aborder frontalement la problématique de la simulation multi-agents du trafic routier à partir des comptages routiers, il est nécessaire d'identifier les différents éléments composant le système du trafic routier, autant au niveau de sa conceptualisation que de son fonctionnement. En suivant les trois grandes étapes de la démarche de modélisation (*cf.* section 4.1.1 p. 127), cette dernière section permettra de présenter les grandes lignes du modèle SCAUP mis en place dans le cadre de cette thèse.

### 5.3.1. **Une conceptualisation spatio-temporelle du trafic routier : quelles entrées ?**

#### a. ***Quelle conceptualisation du trafic routier ?***

Nous avons vu au cours de la partie 1.1.1 (p. 24) que l'approche prédominante pour la modélisation du trafic routier repose sur la formalisation de trois sous-systèmes urbains : *localisations ; pratiques et relations sociales ; transports et déplacements*. Cette approche apparait adéquate à mettre en œuvre lorsque l'on dispose de données permettant de relever les pratiques de mobilité quotidienne des individus, comme les EMD (Commenges, 2013). De surcroît, la mobilisation de ces données constitue un prérequis à la modélisation à quatre étapes. Néanmoins, dans notre cas, l'utilisation des données de capteurs réduit le système à étudier, dans le sens où ces données ne relèvent que les flux de trafic routier en différents points de l'espace (*cf.* Chapitre 3, p. 81). En effet, il n'est pas possible de rattacher aux données de capteurs des informations qualifiant les déplacements d'un ou plusieurs individus, de même qu'il n'est pas possible de rattacher les

comptages à une origine ou à une destination. Ainsi, il n'est pas nécessaire de reconstruire l'ensemble de ces sous-systèmes et la problématique du travail présentée s'insère plus spécifiquement au sein d'un seul sous-système. C'est pourquoi nous avons fait le choix de focaliser nos recherches sur le *sous-système des déplacements et transports,* en nous concentrant exclusivement sur le trafic routier. Cette approche réductionniste permet de conceptualiser ce travail autour de plusieurs *objets élémentaires* qui définissent le trafic routier (*cf.* figure 2.13, ci-dessous) : **l'infrastructure de transport** et les **véhicules routiers** qui définissent le trafic routier (*cf.* section 2.1.1, p. 56). Le dispositif de quantification du trafic routier a été présenté au sein du chapitre 3 (p. 81) au cours duquel nous avons pu convenir des *objets* mesurés par les capteurs fixes à travers la temporalité inhérente à ce système de mesure (les rythmes journaliers, hebdomadaires et annuels). En revanche, les véhicules routiers sont implicitement comptabilisés sous la forme d'un débit routier (nb de véhicules/quart d'heure).

Il faut alors convenir d'une génération des véhicules à partir du dispositif de quantification mis en place. On soulève, par ailleurs, un problème éminemment géographique, afin de passer d'une information ponctuelle à une information linéaire, ou en d'autres termes, d'une vision que l'on peut qualifier d'eulérienne à partir des comptages routiers à une vision lagrangienne des véhicules en déplacement sur le réseau routier. La formalisation du modèle conceptuel dans un contexte multi-agents permet de se focaliser sur les composants élémentaires du trafic routier comme illustré en figure 2.13, en vue d'affecter les comptages routiers sur le réseau routier. Il se conceptualise autour de deux *objets* :

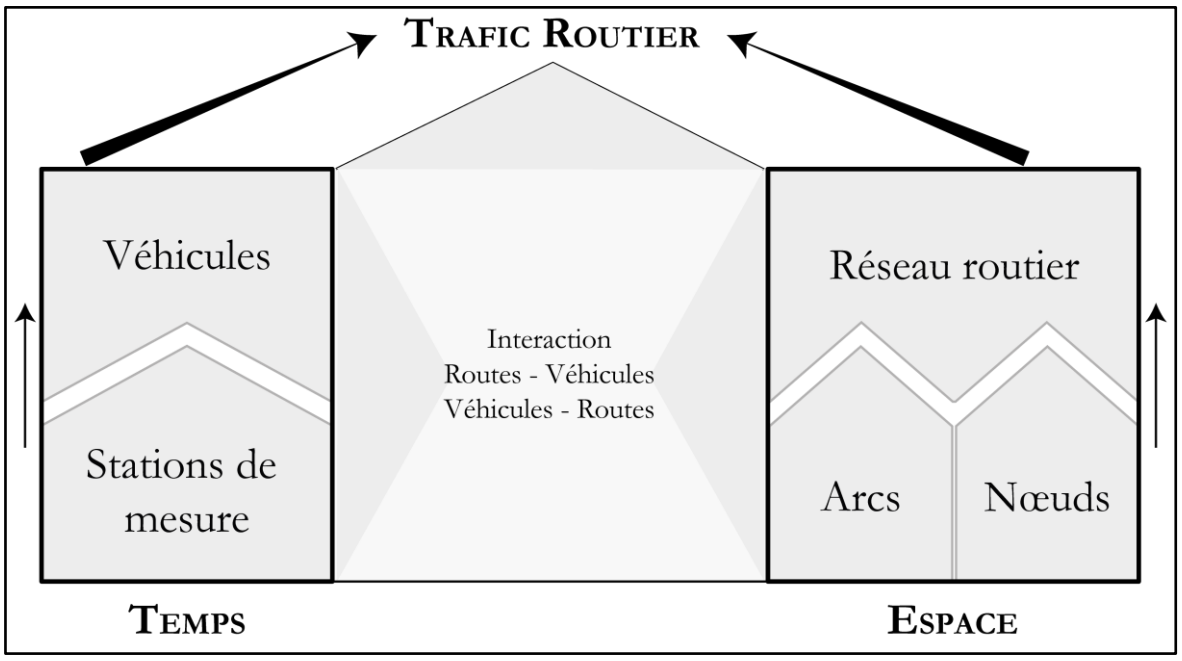



**Figure 2.13 : La conceptualisation du trafic routier sous le modèle SCAUP**

Dans cette acception, les *objets* à identifier et à construire sont donc de deux natures. Conceptuellement, le premier *objet* est une composante spatio-temporelle du système et correspond aux véhicules en déplacement sur une infrastructure de transport. Les véhicules sont des objets éminemment spatio-temporels qui se formalisent discrètement au sein des comptages routiers sous la forme d'un débit de véhicules recensé à chaque quart horaire. Ce sont des *objets mobiles* (Cheylan, 2007 ; Langlois, 2006), se manifestant au sein de la base de données construite à

partir des données de comptage (*cf.* Chapitre 3, p. 81). Une fois l'emplacement des véhicules replacé dans leur contexte de comptage, ceux-ci correspondent aux *agents* à générer pour la construction du trafic routier. Le second *objet* est une composante spatiale et fixe (invariante) du système à construire, et correspond au réseau routier qui constitue l'*environnement* dans lequel le trafic routier s'inscrit.

Enfin le trafic routier se construit conceptuellement par l'interaction de ces deux *objets* élémentaires (*cf.* figure 2.13) : sans route et véhicules, il ne peut pas y avoir de trafic routier. Sous cet angle ils se qualifient autant dans l'espace et dans le temps par les véhicules en mouvement, que sur les infrastructures routières. Par conséquent, c'est par l'interfaçage de ces *objets* au sein du modèle SCAUP que le trafic routier pourra être produit. Le caractère génératif des SMA (Epstein, 1999) apparait ici essentiel en vue de produire un trafic routier par l'interaction entre, non seulement, les véhicules comptabilisés par les stations (*les agents*), mais aussi, le réseau routier (*l'environnement*). L'ensemble de ces éléments en interaction doit nous permettre ensuite d'observer les flux de trafic routier émergent sur l'ensemble du réseau routier d'une ville ou agglomération.

### b. *Inscrire les SMA dans la donnée : l'approche « deepening »*

L'intérêt de l'approche initiée se situe dans la relation entre le dispositif de quantification et celui de modélisation, en vue de conserver le signal issu des comptages routiers, et donc de garder au maximum ce signal sans en déformer ses valeurs initiales. C'est aussi pour cela que nous privilégions une approche générative du trafic routier à partir des objets élémentaires du trafic routier identifiés plus haut. Les données à recueillir sont comme nous l'avons vu plus haut spatiales et temporelles. Par analogie, cette structuration est aussi reprise dans la figure 2.14 qui représente le système d'intégration des données de comptages routiers, au sein du modèle SCAUP, conçu avec la plate-forme de simulation multi-agents GAMA (*cf.* section 6.1.2, p. 187) :

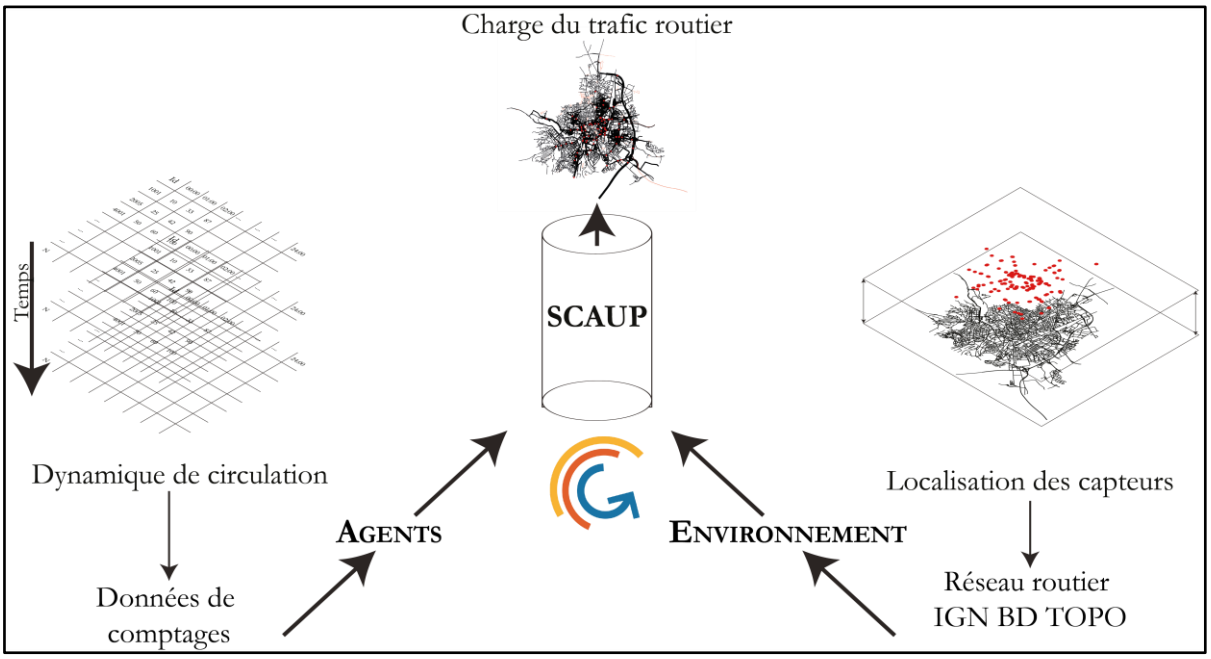



**Figure 2.14 : Schéma d'intégration des données au sein du modèle SCAUP**

Sous cette seconde forme, deux modules peuvent être transposés au modèle conceptuel présenté en figure 2.13 (p. 166) :

- **les agents**, qui reflètent les véhicules à créer via les données de comptage et dont l'objectif vise à initialiser la dynamique du trafic routier à partir des stations de mesure
- **l'environnement** du système multi-agents, qui permet de localiser le réseau routier ainsi que les capteurs de trafic sur l'espace urbain.

Dans ce sens, en rapprochant le dispositif de quantification et le dispositif de production du trafic routier (*cf.* figure 2.10, p. 158), la relation entre données de comptage et simulation du trafic routier est mise en avant de la démarche proposée. L'approche de simulation envisagée dans SCAUP s'inscrit dans l'optique d'une approche dite de « *deepening* »[96], qui « *peut contribuer à obtenir des résultats de simulation qui sont plus proches des observations et donc de l'objectif visé* » (Hassan *et al.*, 2010b, p. 2) :

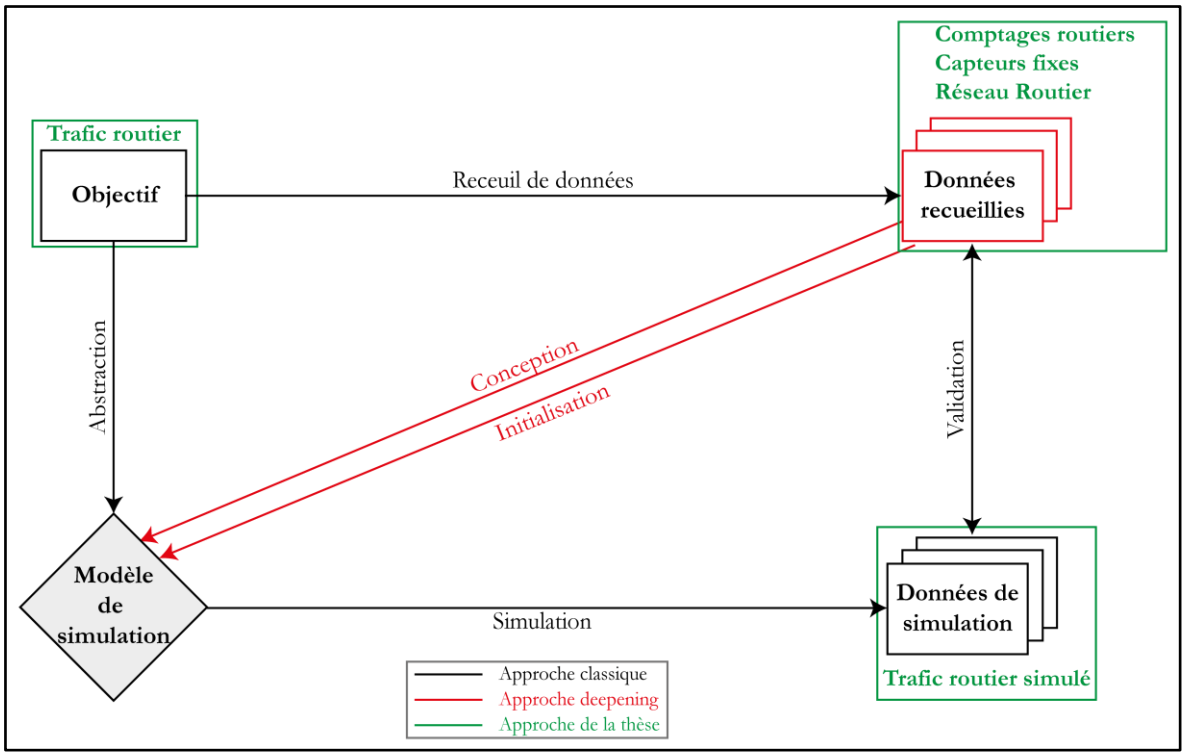



**Figure 2.15 : Les différentes approches de modélisation en simulation multi-agents : de l'approche déductive à l'approche inductive (source : Hassan *et al.*, 2010a, p. 7)**

Cette approche se distingue de la simulation multi-agents *classique*, qui s'inscrit dans une « *logique de simulation pour l'expérimentation scientifique* » (Hassan *et al.*, 2010a, p. 2), illustrée au sein de la figure 2.15 (en noir). Le point de départ de la démarche de simulation repose sur une approche déductive fondée sur les données recueillies. Toutefois, ces dernières sont déconnectées du processus de conception et d'initialisation du modèle de simulation. En effet, celles-ci n'interviennent qu'à la fin du processus de simulation en vue de valider les données de simulation (*cf.* figure 2.15). Dans cette optique, la construction du modèle de données à mobiliser peut être réalisée après la simulation. Alors que dans le cas d'une approche « *deepening* » (illustrée en rouge) la place accordée aux données dans le processus de construction du modèle est plus importante et la collecte des données doit être réalisée au préalable de la construction du modèle de simulation. En effet, les données interviennent à chaque étape du processus de simulation, autant

[96] Nous avons souhaité conserver l'anglicisme de « *deepening*, » car il apparait difficile de convenir d'une définition dans la langue française de ce terme, qui se traduit par *approfondissement*.

pour sa conception et son initialisation, que pour sa validation. La mise en place du modèle de simulation est davantage pilotée par les données de terrain que dans une approche classique ou déductive (Hassan *et al.*, 2009, 2010a, 2010b).

Dans cette optique, nous proposons la mise en place d'un modèle de SMA du trafic routier guidé par les comptages routiers, présenté par les zones vertes de la figure 2.15. Par nature, nous avons vu que la simulation multi-agents permet d'envisager une logique inductive radicalement centrée sur l'individu et sur les données. Sa formalisation, ainsi que son fonctionnement *bottum-up*, permettent d'être plus directement relié aux données d'entrée du modèle (Hassan *et al.*, 2010a, 2010b). L'intérêt porté au modèle de données dans l'approche « *deepening* » permet aussi d'explorer les potentiels d'utilisation des outils de géosimulation pour la représentation des « *données empiriques existantes, pouvant être massives, hétérogènes et bruitées* » (Banos, 2013, p. 68). Néanmoins, avec cette approche, les questionnements autour de la validation du modèle de simulation se posent. En effet, les données servant autant à alimenter le modèle de simulation qu'à le valider, il faut être en mesure de fournir des données de contrôle n'ayant pas été utilisées lors de l'initialisation de la simulation. Les défis d'une approche « *deepening »* se situent aussi au niveau de l'exploration détaillée des éléments liés à l'évaluation et à la validation du modèle de SMA, que nous aborderons au cours du chapitre 7 (p. 223). Par ailleurs, nos réflexions s'inscrivent, dans un dernier temps, sur le formalisme des approches en SMA, et plus particulièrement en géographie, en se focalisant sur les formalismes des modèles de SMA : les approches KISS et KIDS.

### c. *L'approche expérimentale en SMA : du KISS au KIDS*

Au sein de la communauté de la simulation multi-agents, deux grandes approches de modélisation se distinguent (Edmonds et Moss, 2004 ; Banos et Sanders, 2012 ; Bouquet *et al.*, 2015 ; Delay, 2015) : la modélisation dite KISS (*Keep It Simple Stupid*), dont l'objectif est de présenter une vision simple et parcimonieuse d'un système à étudier, et la modélisation dite KIDS (*Keep It Descriptive Stupid*), qui vise quant à elle à apporter une vision descriptive et complexe d'un système à étudier. En d'autres termes, les modèles KISS sont des modèles de compréhension qui « *visent à simplifier au maximum le modèle afin de construire un environnement contrôlé intelligible et focalisé sur la dynamique étudiée du système* », alors que les modèles KIDS sont des modèles descriptifs qui « *visent à décrire le système dans sa réalité la plus fine* » (Bouquet *et al.*, 2015, p. 33).

Pourtant la réalité du cheminement d'exploration des modèles montre à plusieurs reprises que le cloisonnement entre KISS et KIDS n'est pas si hermétique : « *aucun modèle n'est tout à fait KISS et complétement KIDS. La propension des deux varie en fonction des situations modélisées et de la connaissance empirique qu'en ont les modélisateurs/observateurs* » (Delay, 2015, p. 101). Bruce Edmonds et Scott Moss (2004, p. 143) ajoutent que « *ni l'approche KISS, ni l'approche KIDS n'est toujours la meilleure et qu'un mélange complexe des deux approches sera souvent approprié* ». Pour mieux appréhender ces éléments en fonction de cette distinction (KISS-KIDS), Arnaud Banos et Lena Sanders (2012) proposent une grille de lecture pour comparer et replacer les différents modèles utilisés par les géographes. Cette grille de lecture nommée le « *fer à cheval* » se construit selon deux axes, en vue de distinguer les modèles, non seulement selon leur niveau d'abstraction, mais aussi, selon leur degré de parcimonie (*cf.* figure 2.16, p. 170) :

- Le premier axe illustre le niveau d'abstraction (stylisé-particulier), qui oppose « *d'un côté les modèles reposant sur des informations complètes et fines au niveau spatial* [particulier]*, de l'autre les modèles dits "concepts driven*" [stylisé]*, en général plus parcimonieux sur le plan des variables mobilisées* » (Banos et Sanders, 2012, p. 53).
- Le second axe illustre le degré de parcimonie avec lequel le modèle est mis en place, en distinguant les plus parcimonieux (KISS) des plus descriptifs (KIDS).

En croisant ces deux composantes, les auteurs replacent quatre pratiques de modélisation (*cf.* figure 2.16) : des modèles les plus théoriques (A) aux modèles les plus descriptifs, où les données expérimentales prennent une place importante. Pour illustrer ces quatre pratiques, les auteurs reprennent la formalisation du modèle gravitaire : « *si la "masse" d'une ville peut être approchée par sa population par exemple, alors on doit pouvoir prédire les flux de population entre deux villes à partir de leur population respective et de la distance géographique les séparant. Cet énoncé, simple et stylisé, nous place dans le quadrant A du diagramme. L'ajustement statistique du modèle mathématique afférant aux données empiriques nous entraîne vers le quadrant D. En revanche, l'enrichissement de ce modèle « pauvre sur le plan théorique », par exemple par l'introduction « d'opportunités intermédiaires » ou « d'effets de barrières » relève du quadrant B, et l'ajustement empirique des modèles ainsi spécifiés amène un positionnement dans le quadrant C* » (Banos et Sanders, 2012, p. 54). À travers l'exemple du modèle gravitaire, on comprend que les approches ne s'opposent pas et sont plutôt d'ordre complémentaire (Mathian et Sanders, 2014). Ici, on comprend bien le rôle que les données prennent pour positionner les modèles sur l'axe d'abstraction, le passage aux données de terrain permettant de se positionner à différents niveaux du cadran :

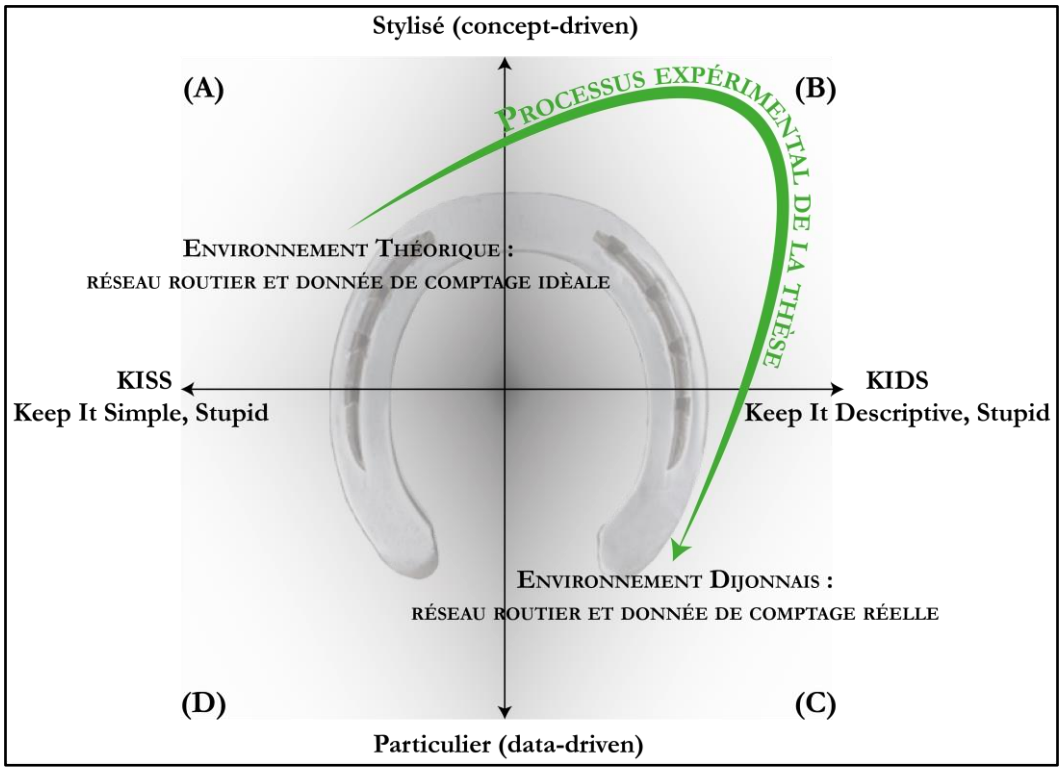



**Figure 2.16 : Le « *fer à cheval* » : une grille de lecture à deux entrées pour positionner les modèles de simulation multi-agents (source : Banos et Sanders, 2012, p. 54)**

Il faut noter que notre parti pris, du point de vue de la modélisation et de sa formalisation s'insère exclusivement dans la catégorie des modèles KISS. Le glissement du positionnement KISS vers

KIDS s'opère plutôt du point de vue de la simulation par le passage progressive d'un environnement simple vers un environnement « réel ». En effet, comme nous l'avons abordé plus haut (*cf.* section 1.1.1, p. 24), aborder la modélisation du trafic routier dans une perspective KIDS nécessiterait de prendre en compte un plus grand nombre *d'objets* caractérisant l'espace urbain et le trafic routier. C'est pourquoi, pour l'expérimentation, notre parti pris KISS-KIDS se situe plutôt au cœur de la simulation. En effet, en intégrant tout d'abord, un ensemble restreint d'éléments pour la SMA du trafic routier, nous nous positionnons dans la catégorie des modèles KISS ; toutefois, en nous focalisant sur le réseau routier et les comptages routiers sur un espace urbain réel, nous nous positionnons au sein des modèles KIDS. C'est pourquoi la démarche de modélisation que nous proposons vise à nous entourer de deux types de modèles, en faisant glisser progressivement le plan d'expérimentation, illustré par la flèche au sein de la figure ci-dessus, du modèle de simulation de la première catégorie (A) à la quatrième (C), sans en modifier le fonctionnement interne, mais plutôt en y insérant la réalité géographique d'un espace urbain.

Le premier modèle KISS permet de tester la dynamique du modèle SCAUP sur un réseau routier théorique auquel ont été associées des données de comptage « types »[97], permettant d'observer le cycle journalier du trafic routier (matin, midi, soir), ainsi que les mêmes caractéristiques attributaires permettant de qualifier le réseau routier (hiérarchie, vitesse…). Cette construction simple, illustrée ci-dessous (à gauche de la figure 2.17), permet de tester et d'expérimenter la mise en place du modèle de simulation sur un nombre restreint de routes (48 tronçons) et de stations de comptage (15 capteurs), auxquelles sont associées les données moyennes journalières des comptages routiers réels.

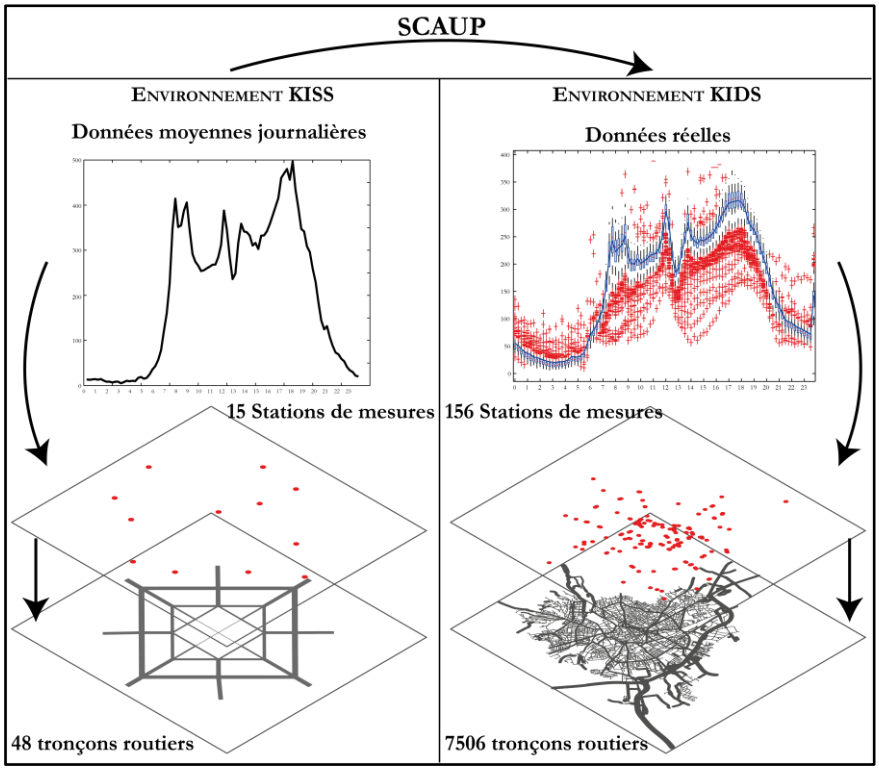



**Figure 2.17 : Du modèle KISS au modèle KIDS : application du cadre expérimental de SCAUP à la ville de Dijon**

[97] Nous avons fait le choix d'y insérer des valeurs de trafic caractérisant la courbe matin, midi et soir, en utilisant la moyenne d'un ensemble de capteurs routiers, qui a été ensuite affectée à des capteurs fictifs sur le réseau théorique.

Une fois le modèle KISS exploré, le glissement du premier modèle au second s'opère par la prise en compte du contexte urbain au complet, en y insérant son réseau routier, ainsi que l'ensemble des données de comptage recueillies (*cf.* figure 2.17). En conséquence notre parti pris respecte les préconisations mises en avant par Bruce Edmonds et Scott Moss (2004, p. 142) : « *le but de simplifier les modèles* [KISS] *ne doit pas se faire dans l'objectif que le modèle complexe* [KIDS] *conserve l'ensemble des comportements, mais plutôt dans le but de n'en conserver que les comportements pertinents* ». C'est bien dans cette acception que nous avons choisi d'orienter notre processus de recherche expérimentale autour du fer à cheval.

La complexification du modèle mis en place sur un espace urbain plus vaste tient plutôt compte d'un *environnement* de simulation complet, en employant l'ensemble des capteurs et des données associées au sein du dispositif de quantification. Cette seconde phase expérimentale sera testée à travers la ville de Dijon en nous entourant des données issues des capteurs électromagnétiques présentées plus haut (*cf.* Chapitre 3, p. 81), et des bases de données géographiques issue des nomenclatures nationales (IGN) et internationales (OpenStreetMap). Notons tout de même que nous aurions pu sélectionner un autre espace test du moment que l'on dispose des données de comptages routiers de l'espace d'étude cible. Dans cette optique, le cadre expérimental envisagé sur la ville de Dijon peut être réalisé sur un tout autre espace urbain, dans la mesure où l'on dispose au préalable de données issues des comptages routiers. En effet, comme nous le verrons le modèle KISS et le modèle KIDS possèdent les mêmes caractéristiques de fonctionnement interne, leur différenciation s'opérant plutôt par l'intégration d'un environnement urbain au sein du système multi-agents.

### 5.3.2. Le fonctionnement du modèle SCAUP : le rôle structurant des infrastructures de transports pour le déplacement des véhicules

Nous nous arrêtons maintenant sur les hypothèses de fonctionnement établies en vue de produire un trafic routier émergent à partir des comptages routiers au sein du modèle SCAUP. Son fonctionnement tel que nous l'envisageons vise à régir le déplacement des véhicules routiers sur les infrastructures de transports. En reposant sur une approche générative et sur les comptages routiers, nous faisons l'hypothèse que le fonctionnement de SCAUP repose en partie sur une interaction hiérarchique majeure au sein de laquelle le réseau routier porte et supporte l'ensemble des caractéristiques pour appréhender les déplacements des véhicules en milieu urbain.

#### a. *Une réinterprétation de l'affectation et l'écoulement des véhicules : une problématique spatio-temporelle*

Comme nous l'avons vu précédemment, simuler le trafic routier nécessite de s'arrêter sur deux éléments qui définissent les modèles de simulation du trafic routier : **l'affectation et l'écoulement** (*cf.* section 4.3, p. 141). Lorsque l'on aborde les problèmes de l'affectation, la question centrale tourne autour des chemins et des itinéraires qui seront choisis et empruntés par les véhicules pour se déplacer (Henn, 2001). Alors que pour l'écoulement on s'intéresse plutôt à la manière dont les véhicules vont se déplacer sur la route (Bourrel, 2003). En d'autres termes, deux éléments ressortent (Buisson, 1996): l'affectation est une problématique spatiale qui permet de répartir un nombre de véhicules sur un plus grand ensemble d'infrastructures de transport ; alors que l'écoulement est, quant à lui, une problématique plutôt temporelle permettant de régir la

façon avec laquelle les véhicules arpentent les routes. Dans une approche classique de simulation du trafic routier, tous les logiciels de simulation se fondent sur ces deux modules, auxquels se rajoute un troisième module permettant de représenter le réseau routier (Buisson, 1996) : le **modèle d'affectation** ; le **modèle d'écoulement** ; et la **représentation du réseau routier**. Ce dernier module contient l'ensemble des caractéristiques permettant de décrire le réseau routier, qui se conceptualise autour de deux éléments (*cf.* figure 2.13, p. 166) : les arcs auxquels on associe les informations sur le nombre de voies, les sens de circulation, la vitesse de déplacement… ; et les nœuds qui permettent d'identifier les liens de passage d'un arc à un autre. L'ensemble de ces caractéristiques étant le plus généralement disponible et appréhendé par l'intermédiaire d'un SIG au sein duquel on peut générer l'ensemble des informations relatives aux infrastructures de transports.

En faisant une analogie aux problématiques d'affectation et d'écoulement, nous mettons en lumière les enjeux liés au fonctionnement de SCAUP qui doit permettre, d'une part, de définir la direction des flux de véhicules recensés par les stations de comptage ; et d'autre part, d'en appréhender leurs déplacements et mouvements sur les infrastructures routières. L'ensemble doit donc se caractériser autant dans sa dimension spatiale que dans sa dimension temporelle :

- **Spatiale** : se formalisant comme une problématique d'affectation en généralisant l'information sur l'ensemble des infrastructures d'un espace urbain. Portée par l'interaction entre réseau routier et véhicules.
- **Temporelle** : se formalisant comme une problématique d'écoulement du trafic routier en vue de reproduire le signal du trafic routier par la reproduction du rythme journalier du trafic routier, dont la dynamique est portée par le dispositif de quantification.

Si nous ne pouvons appliquer les éléments d'affectation et d'écoulement dans leur acception la plus générale, notre approche repose sur une hypothèse au centre de laquelle le réseau routier joue un rôle fondamental pour mettre en mouvement les véhicules. C'est pourquoi, dans un second temps, nous nous focaliserons sur la notion de *réticularité* qui permet d'appréhender la morphologie et les caractéristiques circulatoires des infrastructures de transport pour qualifier le trafic routier.

### b. *Comment caractériser la morphologie du réseau routier ?*

Ici, nous nous focalisons plus spécifiquement sur les « *réseaux d'infrastructures*[98] » (Gleyze, 2005, p. 76). En géographie des transports, la notion de *réticularité* se définit comme la « *capacité d'une structure quelconque à s'organiser et à fonctionner en réseau. L'organisation et le fonctionnement réticulaire des réseaux de transport contribuent à permettre et à faciliter l'acheminement des personnes et des marchandises à la surface du globe* » (Bavoux *et al.*, 2005, p. 82). En d'autres termes, la notion se rapporte à l'ensemble des propriétés qui se rattache autant à la morphologie qu'au fonctionnement du réseau. Sous ce terme, le réseau dans son sens large, ainsi que le réseau routier, n'existe que s'il répond à des besoins de mobilité, de communication ou d'échange. Il permet d'apporter un regard sur la configuration et la morphologie que le réseau routier a afin d'organiser et de gérer les flux de circulation. Dans ce sens, le réseau routier présente non seulement des caractéristiques

[98] Les réseaux d'infrastructures sont « *ancrés matériellement sur le territoire afin de donner prise à l'action physique des phénomènes* » (Gleyze, 2005, p. 76) ; en d'autres termes, la notion se réfère à tous les moyens *physiques* permettant de véhiculer autant l'énergie ou l'information, que les biens ou les personnes.

morphologiques mais aussi des caractéristiques « *circulatoires* [99]» (Genre-Grandpierre, 2000), influençant autant l'organisation que la répartition des flux de circulation.

En nous intéressant à la notion de *réticularité*, nous souhaitons mettre en avant une piste de réflexion dans le cadre de ce travail, en vue de dépasser les contraintes liées à la manipulation des matrices O-D. En effet, pour mettre en mouvement le trafic routier, il est nécessaire d'identifier dans quelle(s) direction(s) et comment les véhicules routiers se déplacent sur les infrastructures de transport. Dans le cas de l'absence d'information de type O-D, il faut être en mesure d'identifier et de distinguer les routes les plus empruntées de celles qui le sont moins, en se focalisant plutôt sur la forme et les caractéristiques du réseau routier que sur les flux de déplacements, comme il en serait d'usage dans une approche classique.

Pour cela nous nous inspirons des travaux de Cyrille Genre-Grandpierre (2000, 2010, p. 96) : « *Il a été montré, notamment pour la fourmilière urbaine où la mobilité est forte et les matrices de déplacements complexes, que le pattern des flux observés dépend plus de la structure des réseaux de transport que des matrices origine-destination des déplacements* […] ; *il en ressort que c'est la structure du réseau qui dicte les trajets effectués et qui va au final dessiner la carte des flux* […]. *Certains réseaux (les plus hiérarchisés fonctionnellement) auront intrinsèquement tendance à conserver les flux sur les voies les plus rapides. En effet, les usagers, à la recherche de la minimisation du temps de transport, sont conduits à emprunter ces voies rapides, en se détournant des trajets plus directs géométriquement, mais plus lents. Avec la même matrice origine-destination, des réseaux plus homogènes fonctionnellement tendront au contraire à davantage répartir les flux sur l'ensemble des voies du réseau, causant ainsi moins facilement des problèmes de congestion* ». Les propos de Cyrille Genre-Grandpierre permettent d'envisager une alternative à la manipulation d'une matrice O-D, en portant non plus le regard sur les individus en déplacement (les flux), mais plutôt sur le réseau routier (la structure). Cela induit nécessairement un changement d'échelle, dirigé plutôt vers des attributs macroscopiques en vue de caractériser l'*offre* de transport. Dès lors, comme le relève l'auteur, le réseau routier est un élément structurant des déplacements, en particulier pour les déplacements routiers (Genre-Grandpierre, 2000, 2010), mais aussi pour d'autres modes de déplacement comme la marche à pied (Foltête *et al.*, 2008). Ici notre parti pris est de nous intéresser à la morphologie géographique du réseau routier en tant que tel, permettant d'en identifier le fonctionnement *circulatoire*, mais aussi les caractéristiques de déplacement des véhicules routiers.

### c. *La représentation du réseau routier : un élément fondamental du fonctionnement spatio-temporel de SCAUP*

Nous privilégions des informations qui sont plus facilement manipulables par tous et ne nécessitant que très peu d'actions de l'utilisateur pour construire son réseau routier. C'est pourquoi, nous nous sommes plutôt tourné vers les caractéristiques directement incluses dans des référentiels nationaux (IGN) ou internationaux (OpenStreetMap). En outre, ces informations permettent de caractériser la forme des réseaux urbains en fonction de la typologie des routes, permettant ensuite de les hiérarchiser et d'en appréhender la configuration générale.

[99] En parallèle, il existe de nombreux indicateurs pour caractériser la structure des réseaux routiers en vue d'en appréhender les fonctionnalités circulatoires (Gleyze, 2005 ; Ducruet, 2010a, 2010b ; Lhomme, 2012). Ces indicateurs sont issus de la théorie des graphes (Mathis, 2003), dont la géographie des transports s'est largement entourée (notamment les travaux de Kansky). Néanmoins, comme le souligne Cyrille Genre-Grandpierre (2010, p. 91), « *qu'ils soient locaux ou globaux, les indices de description de l'accessibilité sont exprimés sous forme d'une valeur globale, le plus souvent une moyenne, dont on ignore la composition* ».

Cyrille Genre-Grandpierre (2000, p. 31) illustre bien ces propos avec la vitesse réglementaire : « *la hiérarchie des vitesses entre les voies départementales, nationales et autoroutières, qui induit des parcours s'éloignant de la ligne droite et "impose" le rabattement sur les voies les plus rapides pour réaliser des chemins plus longs en distance mais plus rapides en temps, marque la primauté de la valeur vitesse (du temps) sur les autres valeurs, notamment environnementales* ». Ici, l'auteur distingue deux typologies du réseau routier, celle qui est rattachée à la vitesse réglementaire et une typologie plutôt fonctionnelle, voire institutionnelle, permettant de distinguer les routes de moindre niveau hiérarchique (départementales) par rapport à celles de niveau hiérarchique plus important (les autoroutes). Allant dans ce sens, nous poussons cette réflexion en posant l'hypothèse que la hiérarchie du réseau routier offerte à travers divers référentiels (BD TOPO, OpenStreetMap…) permet d'approcher les caractéristiques circulatoires du réseau routier, en vue d'identifier les routes qui seront privilégiées par les véhicules lors du processus de simulation multi-agents. Trois référentiels ont été explorés sur le réseau routier de la ville de Dijon (*cf.* figure 2.18 et figure 2.19).

Le premier référentiel est issu de l'attribut « Importance » inclus nativement au sein de la BD TOPO. Cet attribut *« matérialise une hiérarchisation du réseau routier fondée, non pas sur un critère administratif, mais sur l'importance des tronçons de route pour le trafic routier. Ainsi, les valeurs "1", "2", "3", "4", "5" permettent un maillage de plus en plus dense du territoire. Le graphe des éléments appartenant à un degré (autre que le plus bas) et aux niveaux supérieurs est connexe* ». (IGN, 2010, pp. 18–19). À travers cette définition, on retrouve les éléments qui permettent d'appréhender les caractéristiques circulatoires du réseau routier, dont l'illustration est proposée en figure 2.18. On distingue différents niveaux de routes, illustrés par deux couronnes de boulevards sur la ville de Dijon : les liaisons départementales assurent les grandes connexions Nord-Sud et Est-Ouest de l'agglomération ; alors que les liaisons interurbaines et urbaines assurent une desserte au sein de l'agglomération. Le second référentiel concerne les vitesses réglementaires. Il correspond à un élément déterminant pour, d'une part, définir la vitesse de déplacement des véhicules sur les tronçons, et d'autre part, hiérarchiser le réseau routier. Notons ici que les vitesses réglementaires ne sont pas incluses dans le référentiel de la BD TOPO (IGN, 2010). C'est pourquoi nous avons utilisé le module de génération du réseau routier associé au logiciel MOBISIM (Antoni et Vuidel, 2010 ; Lunardi, 2013), permettant, entre autres, d'estimer les attributs de vitesse réglementaire automatiquement à partir du référentiel « type de voie » de la BD TOPO comme illustré en tableau 2.1 (ci-dessous) :

| Type de voie issue de la BD TOPO | Vitesse réglementaire estimée |
|---|---|
| Autoroute | 130 km/h |
| Quasi-autoroute | 110 km/h |
| Bretelle | 50 km/h |
| Route à 2 chaussées | 90 km/h ou *25 km/h* |
| Route à 1 chaussée | 90 km/h ou *25 km/h* |

**Tableau 2.1 : Relation entre le type de voie issu de la BD TOPO et l'estimation des vitesses réglementaires intégrée à MOBISIM (Antoni et Vuidel, 2010)**

Les vitesses réglementaires sont estimées en tenant compte de l'attribut « *Type de voie* » issu de la BD TOPO (IGN, 2010), auquel est associée la limitation de vitesse réglementaire appliquée en France (130 km/h, 110 km/h, 90 km/h et 50 km/h). Une cinquième limitation de 25 km/h est

définie lorsque les routes sont situées à moins de 100 mètres des zones d'habitation, et où le trafic est alors assimilé à un trafic *pulsé*, c'est-à-dire soumis aux feux tricolores et à des arrêts multiples, par exemple à des intersections (Lunardi, 2013). Les attributs de vitesse sont une information essentielle du modèle SCAUP car ils permettent à la fois de proposer une hiérarchisation du réseau routier de la ville de Dijon (*cf.* en haut de la figure 2.19) et d'appréhender la vitesse de déplacement des véhicules sur le réseau routier (*cf.* section 6.2.2, p. 196). C'est une information non négligeable qui est portée par le réseau routier et qui permettra d'appréhender autant la dynamique spatiale (direction des véhicules) que temporelle (vitesse de déplacement) du modèle.

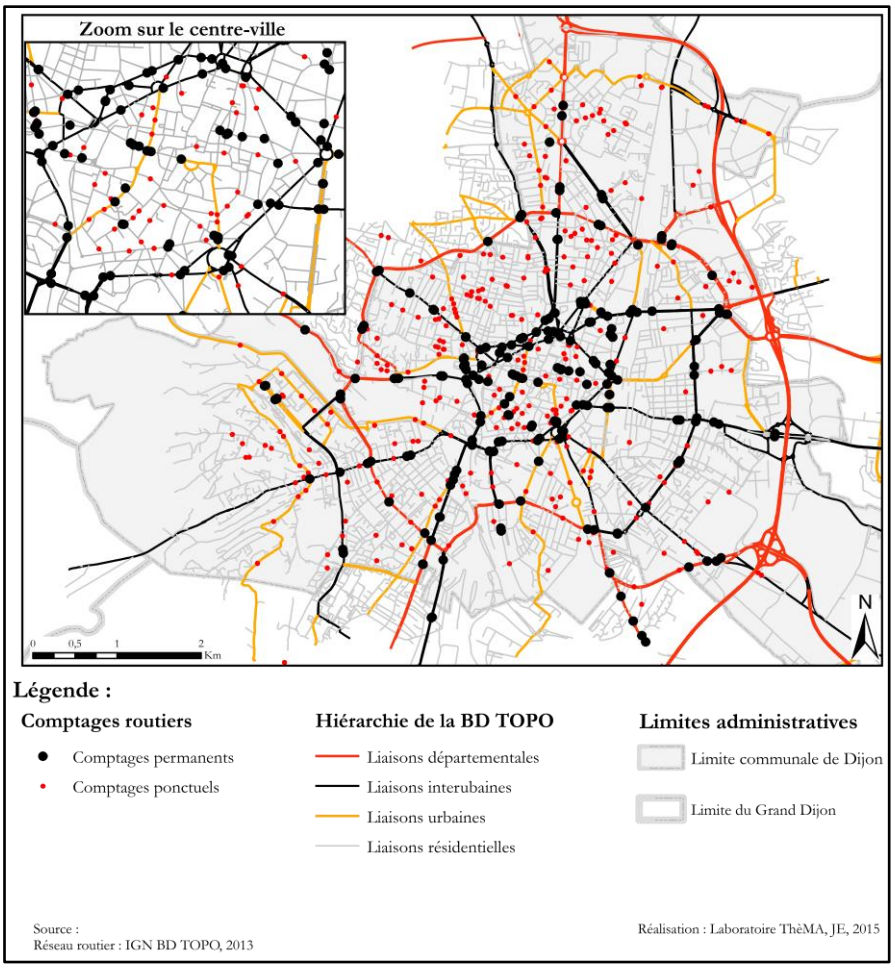



**Figure 2.18 : Hiérarchies du réseau routier issues de la BD TOPO pour la ville de Dijon**

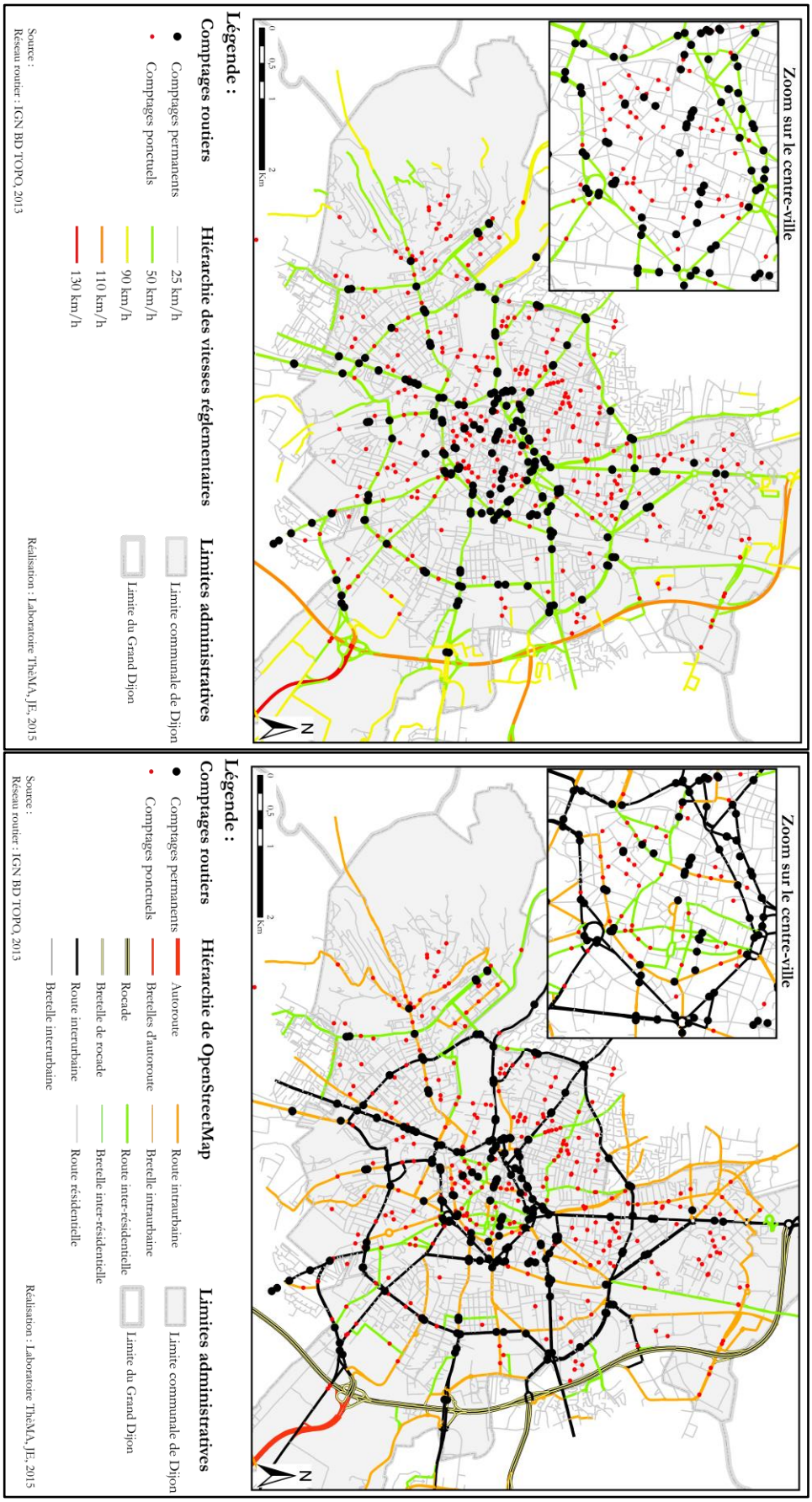



**Figure 2.19 : Hiérarchies du réseau routier retenues sur la ville de Dijon (en haut : les vitesses réglementaires ; en bas : la hiérarchie issue de OpenStreetMap)**

La figure 2.18 permet de distinguer les réseaux traversant l'espace urbain entre les routes présentant des liaisons d'importance départementale, et les liaisons interurbaines et urbaines. Cependant, à l'extérieur du réseau urbain, la différenciation entre les réseaux autoroutiers et la rocade n'apparait pas aussi nettement que sur les deux cartes présentées ensuite. La figure 2.19 (en haut) permet d'observer la répartition des vitesses réglementaires sur le territoire de la ville de Dijon ; toutefois, si l'autoroute ainsi que la rocade sont mises en évidence (à l'Est de la zone), cette illustration ne permet de faire ressortir nettement les routes traversant l'espace urbain qui apparaissent trop homogènes pour les tronçons définis à 50 et à 25 km/h. De plus, le référentiel de la BD TOPO est applicable à l'ensemble du réseau routier français, mais est plus difficilement manipulable pour d'autres pays.

C'est pourquoi, dans le souci d'obtenir des informations plus facilement exploitables sur un grand nombre de villes, nous nous sommes dans un troisième temps tourné vers les informations associées au référentiel OpenStreetMap (en bas de la figure 2.19). Notons ici que les attributs OpenStreetMap ont été affectés manuellement sur la géométrie du réseau routier de la BD TOPO afin de pouvoir comparer les trois référentiels de type de voies. Le référentiel du réseau routier associé à OpenStreetMap est défini par l'attribut « *highway* »[100]. Il est d'autant plus intéressant que les types de voirie sont associés à un référentiel d'équivalence internationale permettant d'appliquer cette typologie à toutes les villes mondiales[101]. En passant par cette information, 11 typologies de voies sont définies, que l'on peut, néanmoins, réduire à 6 types de routes, les 5 autres correspondant aux bretelles permettant d'accéder au réseau qui lui est associé. Le réseau routier est illustré en haut de la figure 2.19 : d'une manière générale, il permet d'observer les mêmes caractéristiques que la BD TOPO, en apportant une précision plus fine à l'intérieur des zones résidentielles pour les routes inter-résidentielles qui n'apparaissent pas au sein des précédents référentiels. Cette typologie permet d'avoir un réseau routier mieux hiérarchisé, où l'on peut observer une information plus hétérogène à l'échelle des quartiers résidentiels de la ville de Dijon.

Enfin, si nous avons présenté les éléments permettant de caractériser le réseau routier, il faut mettre en lumière la manière dont ces informations permettront de définir les axes routiers qui seront privilégiés par les véhicules routiers. Que les mesures soient permanentes ou ponctuelles, en nous reposant sur les figures précédentes (*cf.* figure 2.18 et figure 2.19), on peut observer que l'ensemble des typologies est couvert par au moins un comptage routier. En croisant d'une part les informations relatives à la fréquentation des axes (les comptages), et d'autre part la typologie des axes, on peut obtenir une image de la fréquentation type d'une catégorie de voirie, permettant de définir un pourcentage d'utilisation des différents types de voiries. Cette information permettra de définir les tronçons routiers qui seront privilégiés par les véhicules en tenant compte des valeurs de fréquentation par type de route que nous détaillerons au cours de la section 6.2.2 (p. 196). À noter que les informations relatives à la hiérarchie du réseau routier que nous avons retenues pour le fonctionnement de SCAUP sont présentées au sein de la figure 2.19 (ci-dessus).

Pour finir, nous soutenons l'idée que le réseau routier joue un rôle essentiel dans le processus de simulation pour introduire et appréhender les comportements des véhicules sur le réseau routier.

[100] Descriptif de la voirie sous OpenStreetMap : http://wiki.openstreetmap.org/wiki/FR:Voirie.
[101] Référentiel d'équivalence international de l'attribut « Highway » issus de OpenStreetMap : http://wiki.openstreetmap.org/wiki/FR:Highway:International_equivalence.

Il permettra de porter l'ensemble des informations qui permettront de définir non seulement la dynamique spatiale du système multi-agents, mais aussi sa dynamique temporelle en intégrant les informations de vitesse de déplacement des véhicules (*cf.* figure 2.19). À cet effet, l'affectation ainsi que l'écoulement des véhicules, sont, dans un premier temps, envisagés comme des informations macroscopiques de notre système au sein duquel les agents véhicules pourront parcourir l'ensemble du réseau routier en tenant compte d'une représentation fine du réseau routier. Dans cette perspective, le réseau routier n'est pas seulement un support, il est aussi un élément déterminant pour le fonctionnement du modèle SCAUP.

#### 5.3.3. **Les sorties pour l'évaluation de la PAA : les enjeux d'une approche intégrée**

Enfin après avoir vu les principaux éléments relatifs à l'abstraction et au fonctionnement du modèle SCAUP, nous nous intéressons à la dernière phase de la démarche de modélisation : l'interprétation. Nous nous arrêtons plus spécifiquement sur les éléments relatifs aux sorties du modèle (outputs) qui sont de deux natures : 1. générer un trafic routier émergent pour une journée de comptage routier sous la forme d'une carte de charge pour une journée de simulation ; 2. proposer une estimation des niveaux de PAA en mobilisant les sorties du modèle de simulation aux ASSQA. Le modèle SCAUP doit alors nous permettre d'atteindre trois objectifs :

- il doit être en mesure de **généraliser l'information issue des comptages routiers** sur un plus grand ensemble de tronçon routier à l'échelle de la ville de Dijon ;
- il doit être en mesure de **reproduire la dynamique journalière du trafic routier** en permettant d'observer le cycle type d'une journée identifié plus haut (matin, midi, soir) ;
- il doit être en mesure d'**alimenter le modèle d'estimation de la PAA** utilisé par les AASQA.

C'est au niveau de cette dernière composante que se situe la finalité de SCAUP. L'objectif est tout d'abord, de faire une évaluation précise de la PAA et pour cela de disposer de données relatives à l'intensité des flux de circulation routière autant dans l'espace que dans le temps pour une journée de comptage. Pour y parvenir, nous avons fait le choix de contourner l'usage des matrices O-D (*cf.* figure 2.6, p. 142). Dans le cas de la matrice technique exposée précédemment, les dispositifs techniques fonctionnent au sein d'un système que l'on qualifiera d'endogène (*cf.* figure 2.8, p. 146). Toutefois, pour le travail présenté ici, les dispositifs de modélisation et d'évaluation sont couplés et fonctionnent plutôt de manière exogène (*cf.* figure 2.9, p. 147).

S'inscrire dans cette démarche impose de se formaliser au standard de données utilisé par les AASQA, portant autant sur la nature des variables que sur la finesse des résultats. Le processus d'estimation de la PAA doit nous permettre de passer des données de trafic routier simulé à une estimation des niveaux de PAA sur l'espace de la ville de Dijon. D'une manière générale, la méthode d'estimation des polluants atmosphériques tient compte de l'intensité de l'activité humaine ($A$), comme la production énergétique, le chauffage résidentiel ou l'activité industrielle, à laquelle est appliquée un coefficient multiplicateur ($F$) pour estimer la quantité de polluant émis ($E$) par cette activité. Le calcul des émissions de polluants atmosphériques peut être formalisé sommairement par la formule suivante (PCIT, 2012) :

$$E_{s,a,t} = A_{a,t} * F_{s,a}$$

**E** : émission relative à la substance « *s* » et à l'activité « *a* » pendant un temps « *t* »
**A** : intensité de l'activité « *a* » pendant un temps « *t* »
**F** : facteur d'émission relatif à la substance « *s* » et à l'activité « *a* »

Cette formulation permet de prendre en compte les variations temporelles d'une activité (*t*) mais aussi ses variations dans l'espace (SETRA, 2009). De surcroît, elle permet par sa formulation sommaire de comprendre les apports de notre démarche, qui vise à affiner l'estimation des flux de trafic routier (*A*), plutôt que l'ensemble de la chaîne d'estimation en se focalisant plutôt sur les émissions (*E*). Pour le calcul des émissions de PAA, les AASQA mobilisent le logiciel CIRCUL'AIR développé par l'ASPA[102], qui est utilisé par l'ensemble des AASQA afin d'estimer les émissions de polluants atmosphériques automobiles sur la base de la méthodologie européenne dite COPERT[103] (Ntziachristos et Zissis, 2009 ; AEE, 2014).

CIRCUL'AIR permet d'estimer, pour chaque axe routier, les émissions de polluants atmosphériques rattachées au trafic routier. Nous explorerons plus en détails ces différents éléments au cours du Chapitre 8 (p. 263). La méthodologie COPERT permet de définir les facteurs d'émissions (*F*), afin de « *convertir des quantitatives d'activité (ici des données relatives aux caractéristiques du trafic automobile) en émission de polluants* » (SETRA, 2009, p. 2). Elle se fonde sur une estimation des polluants émis à l'échappement des véhicules en fonction de la vitesse de circulation des véhicules, ainsi que du type de véhicule, comme nous l'illustrons ci-dessous pour les oxydes d'azote ($NO_x$). À travers cette illustration, on remarque que les émissions de $NO_x$ prennent en compte les évolutions technologiques des véhicules en fonction non seulement des normes européennes (normes EURO), mais aussi des types de véhicules et de leur vitesse comme nous le présenterons plus en détail au cours du Chapitre 8 (p. 263).

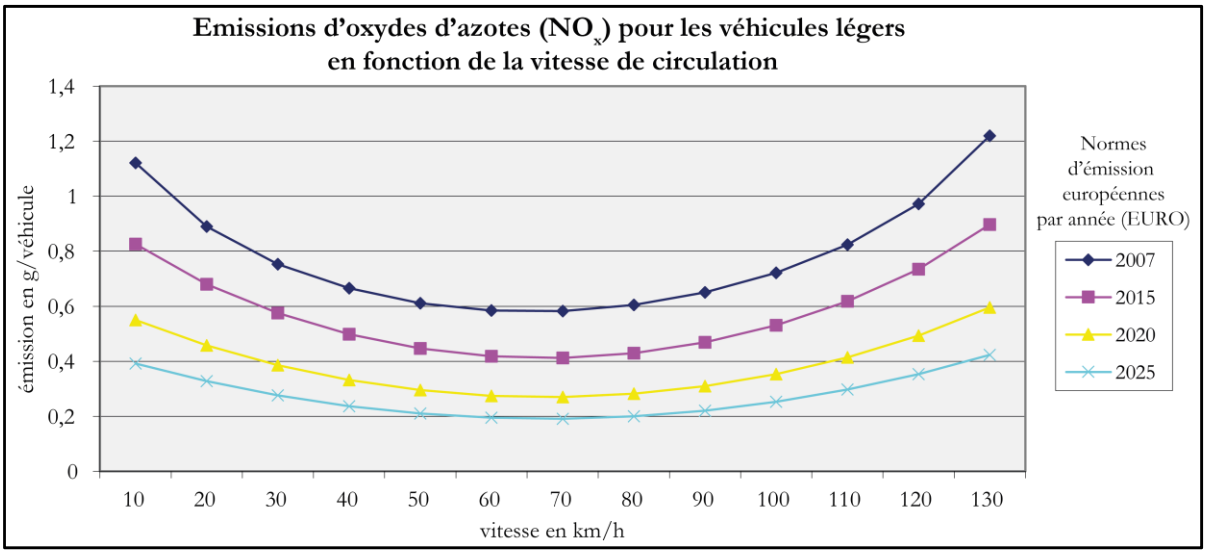



**Figure 2.20 : Illustration du facteur d'émission depuis la méthodologie COPERT pour l'estimation des oxydes d'azote (source : SETRA, 2009)**

[102] ASPA : Association pour la Surveillance et l'Étude de la Pollution Atmosphérique en Alsace.
[103] COPERT (COmputer Programme to calculate Emissions from Road Transport) est un logiciel libre téléchargeable sous le lien suivant : http://emisia.com/products/copert-4/download.

Pour cela, le logiciel CIRCUL'AIR intègre les informations relatives à la circulation routière sous la forme d'un trafic agrégé comme le TMJA (Trafic Moyen Journalier Annuel) ou le TMH (Trafic Moyen Horaire) affecté à chaque tronçon routier. Les sorties du modèle SCAUP doivent donc permettre d'obtenir une carte de charge du trafic routier par tranche de temps, qui sera ensuite mobilisée pour alimenter le modèle d'estimation de la PAA. Les apports de notre approche se situent ici au niveau de la capacité de SCAUP à fournir des données de circulation plus rapidement qu'en utilisant une approche de modélisation classique ; en effet, les données étant réagrégées, on aura tendance à lisser les apports spatio-temporels du modèle de simulation du trafic routier. De plus, la méthodologie COPERT se base sur les vitesses moyennes par voie de circulation. Il faut relever que d'autres outils, comme HBEFA[104], permettent quant à eux d'intégrer une approche plus fine des conditions de circulation lors du calcul des émissions de PAA (André *et al.*, 2013). Néanmoins, pour ce travail, nous avons fait le choix d'utiliser les mêmes approches que celles citées dans le guide défini par le Pôle de Coordination national sur les Inventaires d'émission Territoriaux (PCIT, 2012) et utilisé par les AASQA locales françaises.

[104] HBEFA (Handbook Emission Factors for Road Transport) : http://www.hbefa.net/e/index.html.

## Conclusion :

La géosimulation, et plus spécifiquement les modèles de simulation multi-agents (SMA), permettent de renouveler la manière dont les modèles de trafic routier sont conçus et construits. Ces outils sont aujourd'hui de plus en plus exploités pour la modélisation du trafic routier à un niveau spatial et temporel plus fin. S'inscrire dans le paradigme multi-agents apparait pouvoir apporter une autre réponse en vue d'estimer les niveaux de PAA en milieu urbain. L'approche que nous proposons s'inscrit dans une démarche de modélisation, qui s'est d'abord attachée à la constitution du modèle de données à travers le dispositif de quantification, puis à sa formalisation à travers un modèle conceptuel. L'ensemble de cette approche s'inscrit dans le cadre conceptuel de la *géocomputation* défini précédemment. Les SIG ont été mobilisés pour construire le modèle de données et les informations relatives aux agents et à l'environnement de la SMA, qui est ensuite relayé par les outils informatiques et statistiques pour la simulation du trafic routier (Openshaw et Abrahart, 2000).

L'objectif final du modèle de SMA du trafic routier SCAUP (*Simulation multi-agents à partir de CApteurs Urbains pour la Pollution atmosphérique automobile*) vise à fournir des valeurs de trafic routier pour ensuite en estimer les émissions atmosphériques à travers la PAA. Ainsi, s'inscrire dans une démarche de modélisation nécessite de mettre en place un processus de réflexion qui se caractérise par un ensemble interdépendant de modèles, allant de sa conceptualisation à sa finalité. Si nous nous rattachons au processus de modélisation abordé plus haut (*cf.* figure 2.1, p. 128), ce n'est qu'une fois la première phase d'abstraction établie que la mise en mouvement du modèle à travers la simulation peut être envisagée. Cette seconde phase vise à reproduire les dynamiques de la circulation routière en milieu urbain (Champion, 2003 ; Ferber, 2006).

En s'inscrivant dans les paradigmes multi-agents, cette seconde phase vise à modéliser les processus sous-jacents au trafic routier par l'interaction de deux composants élémentaires : les véhicules et l'infrastructure de transports. En outre, l'approche que nous proposons s'entoure de différents éléments conceptuels qui permettent d'inscrire notre travail dans une optique *deepening* (Hassan *et al.*, 2010a, 2010b). Cette expérimentation s'opère par un glissement progressif du protocole expérimental de SCAUP, d'un positionnement KISS vers une application KIDS (Edmonds et Moss, 2004). Ainsi dans cette optique, l'intégration progressive d'un *environnement* urbain complet permettra de tester les capacités opérationnelles d'une approche guidée par les comptages routiers pour l'évaluation de la PAA.

Pour finir, ce travail s'inscrit dans une optique expérimentale, dont les applications opérationnelles seront testées sur la ville de Dijon. Toutefois, l'application de SCAUP peut être transposée sur d'autres espaces urbains du moment que nous disposons d'informations issues des comptages routiers. En effet, toutes les villes ne disposent pas forcément de données permettant de construire des matrices O-D ; par contre, elles bénéficient très largement de comptages routiers qui sont parfois relevés manuellement ou automatiquement. L'approche expérimentale que nous proposons vise par ailleurs à s'inscrire dans une perspective de généralisation de l'approche expérimentale à d'autres espaces urbains que celui de la ville de Dijon.

# Chapitre 6. EXPÉRIMENTATION ET MISE EN MOUVEMENT DU TRAFIC ROUTIER PAR LA SIMULATION MULTI-AGENTS

Après avoir exposé les différents éléments conceptuels et les hypothèses, liées autant à la construction qu'au fonctionnement du modèle SCAUP, nous présentons au cours de ce chapitre les éléments de réponse mis en place pour mettre en mouvement le trafic routier à partir du modèle SCAUP sur notre terrain d'expérimentation, la ville de Dijon. La simulation, et plus spécifiquement la simulation multi-agents, ouvre la voie à l'expérimentation sous la forme d'un *laboratoire virtuel*, par l'intermédiaire d'une approche soit déductive soit inductive ou abductive (*cf.* 1.2.1, p. 36). Quelle que ce soit l'approche choisie, l'expérimentation par la simulation implique un usage central de l'ordinateur : à ce stade, on parlera plus aisément de simulation numérique, positionnant ainsi les outils associés à la *géocomputation* dans ce champ.

Il reste à préciser une dernière spécificité liée à la démarche géographique : si le modèle est une représentation simplifiée de la réalité (Haggett, 1965), il en constitue néanmoins une formalisation qu'il ne faut pas négliger dans une optique géographique, puisque ces modèles se caractérisent plus particulièrement par leur ancrage spatial et territorial (Daudé, 2005). Une entrée de lecture de la démarche du modélisateur en géographie peut se synthétiser autour de plusieurs éléments clés (Banos, 2013) :

- Étudier un système réel ;
- Les modèles doivent mieux s'enraciner dans la donnée ;
- La formalisation mathématique ne constitue pas l'unique solution ;
- Construire un modèle pour réaliser des expérimentations ;
- S'inscrire dans une démarche de simplification et de parcimonie.

L'ensemble de ces arguments nous amène à focaliser nos investigations sur les différents éléments constituants le SCAUP, en s'intéressant autant aux agents et à leur environnement, qu'à leurs interactions (Demazeau, 2003 ; Marilleau, 2006). Rappelons ici, que la démarche entreprise est exploratoire et vise à reproduire autant la dynamique spatiale que temporelle du trafic routier en milieu urbain, en mobilisant la richesse des comptages routiers. L'application du protocole expérimental de SCAUP vise à passer d'un modèle simple KISS (*Keep It Simple Stupid*), en vue de mettre en place les différents éléments constitutifs du modèle (Edmonds et Moss, 2004), pour ensuite passer à un modèle plus complet KIDS (*Keep It Descriptive Stupid*), en y intégrant l'ensemble de l'environnement de la ville de Dijon. Le passage de l'un à l'autre se combine au caractère empirique de la démarche initiée en vue de positionner le processus d'expérimentation pour l'estimation de la PAA (Banos et Sanders, 2012).

La construction du modèle de simulation envisagé nous amène dans un premier temps à effectuer un état de l'art des outils de simulation multi-agents, afin de choisir l'instrument de recherche ad hoc pour simuler le trafic routier. Ensuite, nous nous attèlerons à décrire l'ensemble du modèle SCAUP, en nous appuyant sur la grille de lecture VOYELLES définie par Yves Demazeau (2003). Enfin, nous illustrerons l'apport des différentes fonctionnalités du modèle de simulation, en nous arrêtant sur les différents tests de sensibilité réalisés lors de sa mise en place.

## 6.1. Les plates-formes de simulation multi-agents :

On compte un grand nombre de plates-formes permettant la mise en œuvre d'un modèle de simulation multi-agents. Tout comme nous avons séparé de la modélisation à quatre étapes les logiciels de modélisation du trafic comme DAVIS, VISUM, … (*cf.* encart 4, p. 136), nous séparons la simulation multi-agents des plates-formes de simulation multi-agents. La simulation multi-agents constitue un cadre de modélisation, et il existe une grande diversité de plates-formes permettant de définir, construire et simuler des systèmes multi-agents sous la forme de boîte à outils[105]. Le choix de l'une ou l'autre de ces plates-formes tient compte autant de la capacité de chacun à manipuler des langages plus ou moins complexes que de la complexité du modèle que l'on souhaite mettre en œuvre.

### 6.1.1. **Quelle classification des plates-formes de simulation multi-agents ?**

Toutes les plates-formes de SMA sont construites sur la même base de structure interne par l'intermédiaire d'un langage de programmation orienté-objet (POO), qu'on peut toutefois distinguer en trois catégories (Taillandier *et al.*, 2014 ; Bouquet *et al.*, 2015 ; Langlois *et al.*, 2015) :

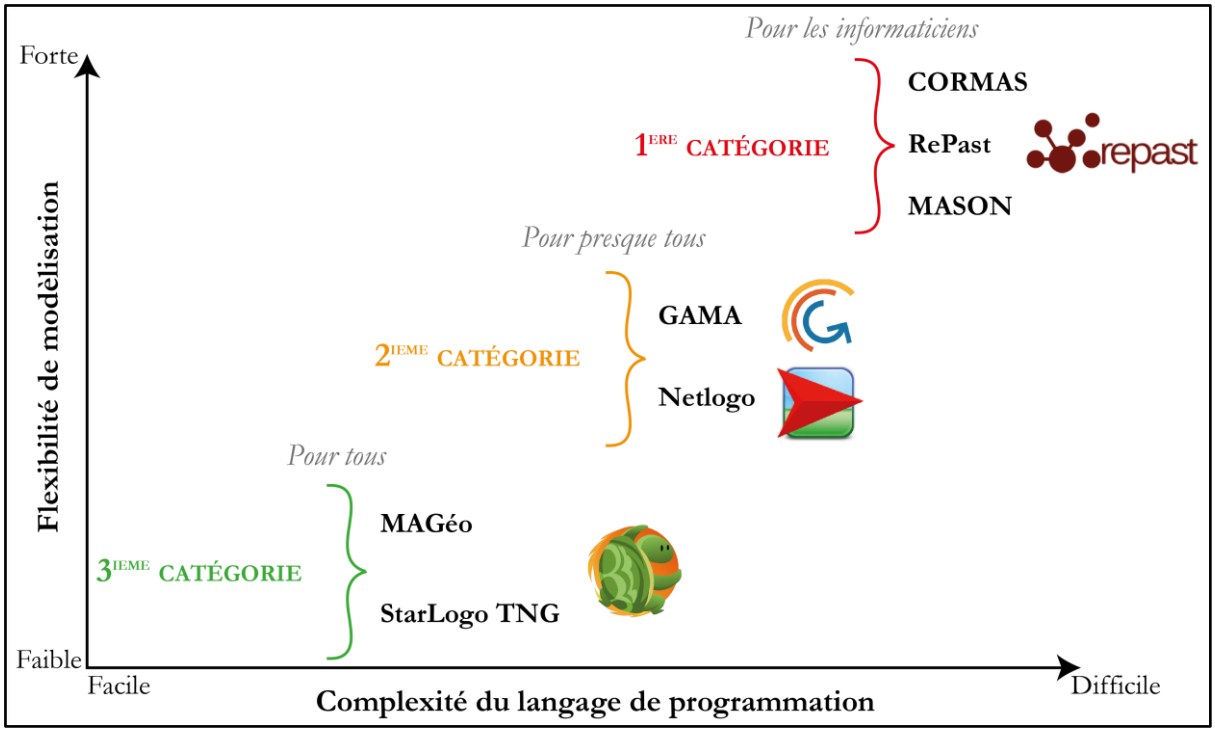



**Figure 2.21 : Une classification des plates-formes de simulation multi-agents (Source : North et Macal, 2007 et illustration d'après http://www.mcs.anl.gov/~leyffer/listn/slides-06/MacalNorth.pdf, diapo 26)**

[105] Différentes listes peuvent être accessibles :
Sur le site de Paul Torrens consacré à la géosimulation (12 plates-formes de SMA recensées) : http://www.geosimulation.org/geosim/abms.htm.
Sur le site OpenABM Consortium pour une liste plus exhaustive (33 plates-formes de SMA recensées) : https://www.openabm.org/page/modeling-platforms.
Sur Wikipédia pour un tableau comparatif des différentes plates-formes (86 plates-formes de SMA recensées) : https://en.wikipedia.org/wiki/Comparison_of_agent-based_modeling_software.

Ces trois catégories peuvent être distinguées selon deux axes illustrés en figure 2.21 (ci-dessus). L'axe des abscisses illustre la complexité du langage de programmation intégré au logiciel allant des plus difficiles, plutôt destinés aux informaticiens et nécessitant des connaissances solides en programmation informatique et algorithmique, aux plus faciles, plutôt destinés à tout public et intégrant un langage graphique dit *user-friendly*, ne nécessitant que peu de compétences en programmation (Langlois *et al.*, 2015). L'axe des ordonnées, quant à lui, tient compte de la flexibilité de la modélisation proposée, permettant d'illustrer la capacité des plates-formes à offrir des modèles plus complexes et plus riches aux modélisateurs. Par conséquent, en s'appuyant sur la figure ci-dessus, le positionnement et le choix de l'une ou l'autre des plates-formes de SMA s'effectuent en fonction des deux axes, en recherchant un juste compromis entre la complexité du langage de programmation et les besoins et compétences du chercheur pour mettre en place un modèle plus ou moins complexe (North et Macal, 2007).

La première catégorie regroupe l'ensemble des plates-formes utilisant un langage de programmation générique (C++, Visual Basic, Java, Python…). C'est particulièrement le cas de plates-formes comme MASON[106] ou Repas[107] utilisant toutes les deux un langage *Java* (Minar *et al.*, 1996 ; North *et al.*, 2013). Comme illustré, cette catégorie est plus particulièrement destinée aux informaticiens et nécessite des connaissances solides en programmation informatique. Elles présentent une flexibilité et un pouvoir de modélisation plus important que les autres catégories, procurant alors plus de liberté au modélisateur pour créer et développer des modèles très complexes. Au sein de cette catégorie, on peut noter qu'il existe des outils opérationnels construits exclusivement à partir de plates-formes informatiques sans nécessairement faire appel à une plate-forme de SMA. C'est le cas du modèle MOBISIM qui a été développé à travers la plate-forme informatique Java afin de bénéficier de la souplesse de modélisation la plus grande possible (Antoni et Vuidel, 2010).

La deuxième catégorie regroupe l'ensemble des plates-formes utilisant un langage de modélisation dédié et intégré à la plate-forme, comme pour NetLogo[108] ou GAMA[109] (Tisue et Wilensky, 2004 ; Taillandier *et al.*, 2014). En intégrant un langage complet de modélisation et un environnement de développement intégré (IDE), cette catégorie permet de construire des modèles relativement rapidement et facilement. C'est ce qui explique probablement le succès d'une plate-forme comme NetLogo, qui regroupe une forte communauté de chercheurs, comme en témoigne la grande richesse des librairies qui lui sont associée[110]. Néanmoins, si le langage reste foncièrement didactique, il demande quelques compétences qui ne sont pas toujours à la portée de tous ; c'est pourquoi une troisième catégorie de plates-formes existe, dont la vocation se destine plutôt à tous (Langlois *et al.*, 2015).

Enfin, la troisième catégorie englobe les plates-formes qui utilisent un langage de modélisation via une interface graphique pour construire un modèle multi-agents comme MAGéo[111] ou StarLogo TNG[112] (Resnick, 1996 ; Langlois *et al.*, 2015). Elles fournissent, comme illustré en figure 2.21,

[106] Page d'accueil de MASON : http://cs.gmu.edu/~eclab/projects/mason/.
[107] Page d'accueil de RePast : http://repast.sourceforge.net/.
[108] Page d'accueil de NetLogo : https://ccl.northwestern.edu/NetLogo/.
[109] Page d'accueil de GAMA : https://github.com/gama-platform.
[110] Librairies de NetLogo : http://ccl.northwestern.edu/NetLogo/models/.
[111] Page d'accueil de MAGéo : https://sites.google.com/site/mageosim/.
[112] Page d'accueil de StarLogo TNG : http://education.mit.edu/portfolio_page/starlogo-tng/.

« *un langage de manipulation des connaissances qui soit à la fois rapidement compréhensible, facilement mobilisable, tout en offrant une grande richesse combinatoire. Cependant ce qu'on gagne d'un côté – par exemple, le tout graphique comme StarLogo TNG –, on perd sur les autres – faire un modèle un peu complexe devient très rapidement fastidieux* » (Langlois *et al.*, 2015, p. 3). La construction d'un modèle ne se fait pas nécessairement en passant par le *code*, mais plutôt sous la forme d'un modèle conceptuel graphique constitué de différentes entités reliées entre-elles par différents types de relations.

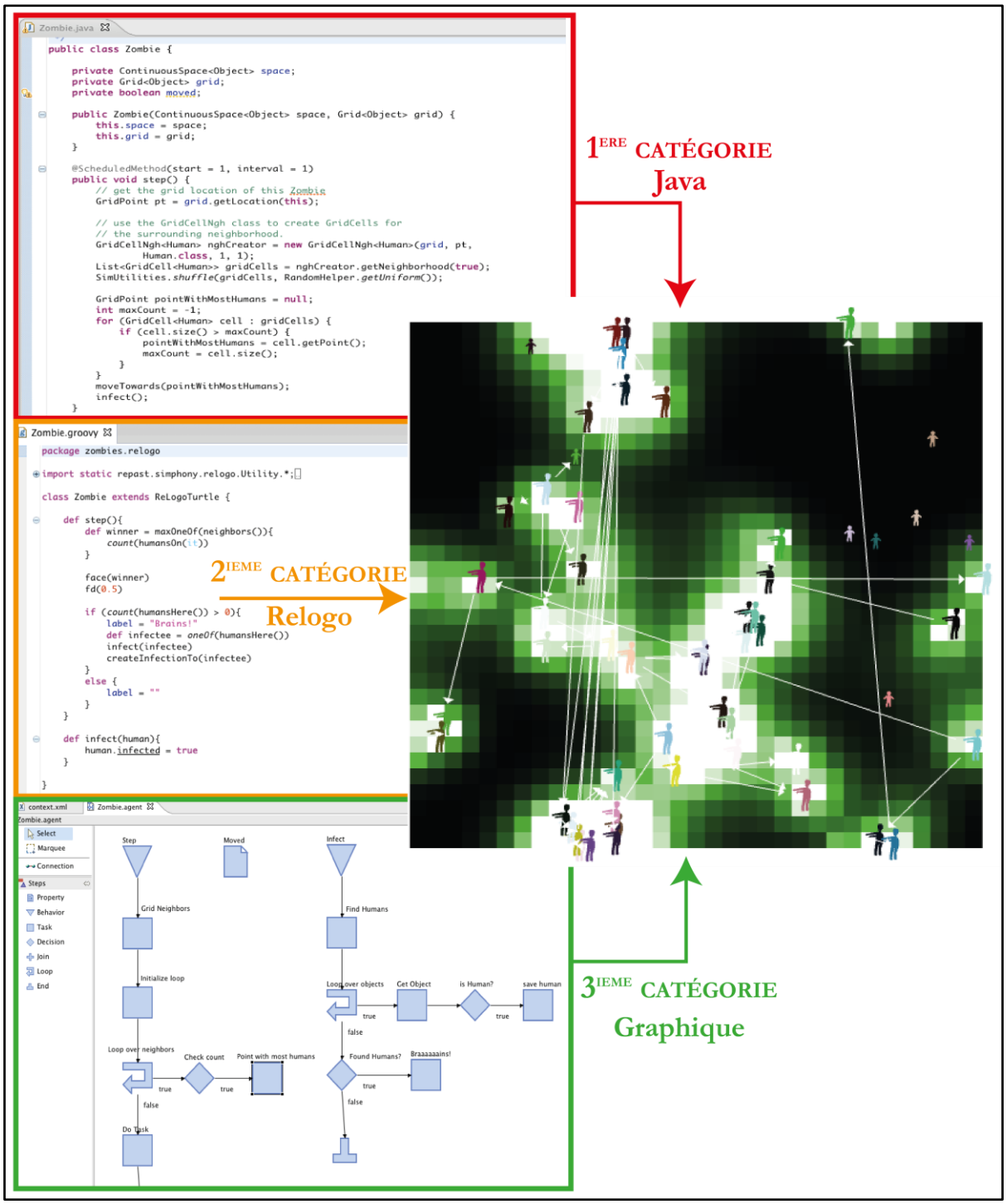



**Figure 2.22 : Trois manières de définir un modèle SMA avec RePast : l'exemple du modèle « Zombie » (Source des images : [http://repast.sourceforge.net/screenshots.php](http://repast.sourceforge.net/screenshots.php))**

Le choix de retenir une plate-forme doit alors tenir compte de cette typologie, ainsi que de la capacité du modélisateur à être plus ou moins à l'aise avec un langage informatique. Cependant, il faut relever que, dans des contextes interdisciplinaires, des langages comme le formalisme UML peuvent faciliter la communication entre le thématicien et l'informaticien. Le langage UML, qui permet de spécifier, visualiser, construire et documenter les éléments d'un modèle, est un standard (formel et normalisé) pour la modélisation des objets à la fois parlant et compréhensible par et pour tous (Ferber, 2006). Ce formalisme constitue une passerelle pour adopter des pratiques de modélisation en se positionnant dans la première catégorie des plates-formes de SMA lorsque l'on ne peut construire un modèle multi-agents par soi-même. Toutefois, le développement des plates-formes des deuxième et troisième catégories a permis de participer à une « *libération des pratiques de modélisation et de simulation en sciences humaines* » (Banos, 2013), où chacun peut construire un modèle : « *les chercheurs en SHS en général, et en géographie en particulier, ont en effet aujourd'hui les moyens d'accéder à un niveau d'autonomie et d'initiative inégalé dans le domaine du traitement des données et de leur modélisation* » (Banos, 2013, p. 76). De plus, il faut noter que l'on assiste actuellement à un développement de plates-formes intégrant différentes manières de définir des modèles agents (Bouquet *et al.*, 2015), comme illustré en figure 2.22 avec la plate-forme RePast qui permet d'intégrer à la fois un langage Java (en haut), Relogo[113] (au centre) et une modélisation graphique (en bas).

Si les plates-formes de SMA présentent toutes chacune leurs avantages et leurs inconvénients, il faut néanmoins relever que NetLogo, créé en 1999 (Tisue et Wilensky, 2004), semble faire consensus par sa simplicité d'utilisation et d'accessibilité pour de nombreuses applications (Bouquet *et al.*, 2015), dont la géographie est aussi partie prenante. Par exemple, de nombreuses applications sur les phénomènes spatiaux ont été mises en place via cette plate-forme au sein du réseau thématique de Modélisation multi-agents Appliquée aux Phénomènes Spatiaux[114] (MAPS). Les plates-formes développées après NetLogo, s'en inspirent et en reprennent les concepts (Bouquet *et al.*, 2015). C'est le cas de la plate-forme GAMA (*Gis & Agent-based Modelling Architecture*), qui constitue la plate-forme que nous avons choisie d'utiliser au cours de ce travail.

### 6.1.2. La plate-forme de simulation multi-agents GAMA[115] (*GIS & Agent-based Modelling Architecture*)

La plate-forme GAMA est développée depuis 2007 au sein de l'Unité Mixte Internationale de Modélisation Mathématique et Informatique des Systèmes Complexes (UMMISCO). Cette plate-forme de simulation multi-agents, qui est en accès libre, se situe dans la continuité des plates-formes de la deuxième catégorie (*cf.* figure 2.21, p. 184). En effet, elle se construit sur la base d'un langage de modélisation orienté objet qui lui est propre (Taillandier *et al.*, 2014) : le GAML (*GAma Modelling Language*). En parallèle, il est aussi possible de formaliser un modèle de simulation multi-agents par l'intermédiaire d'un éditeur graphique, au même titre que des plates-formes comme MAGéo, ou RePast (*cf.* figure 2.22 (en bas), ci-dessus). D'une manière générale,

[113] Le langage Relogo correspond à l'utilisation du langage utilisé sous NetLogo (2e catégorie) sous la plate-forme RePast : http://repast.sourceforge.net/docs/ReLogoGettingStarted.pdf.
[114] Liste des fiches pédagogiques du réseau MAPS : https://maps.hypotheses.org/production-pedagogique-de-maps/modeles-mapsiens.
[115] La version de GAMA utilisée au cours de ce travail est la version 1.6.1 ; à noter qu'une version 1.7 est disponible depuis le mai 2016.

l'espace de travail de GAMA se présente sous forme d'une interface graphique structurée en plusieurs sous-ensembles, permettant non seulement de construire et de définir le modèle, mais aussi de suivre autant le processus de construction que le processus de simulation du modèle (*cf.* figure 2.23) :

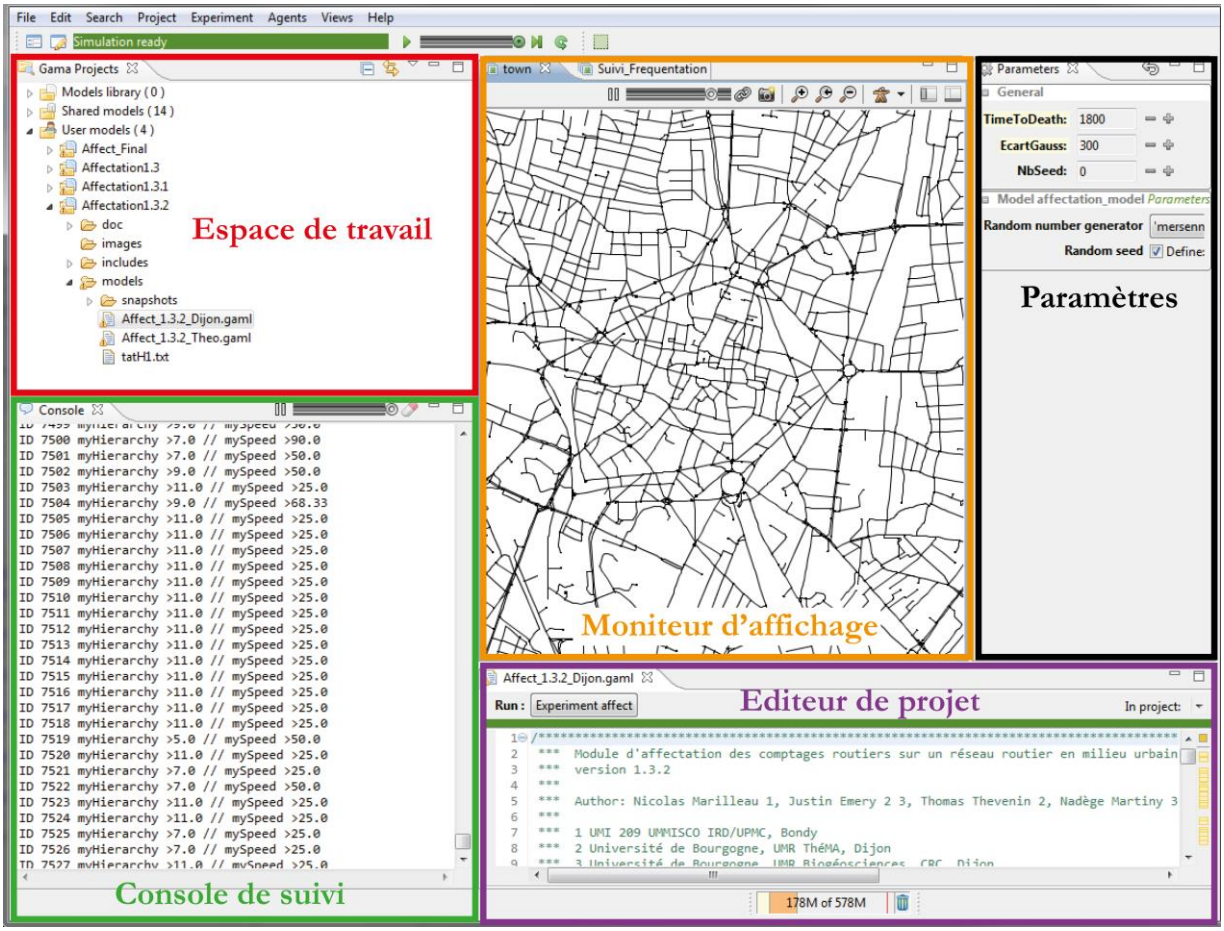



**Figure 2.23 : Interface graphique et environnement de travail sous la plate-forme GAMA**

L'ensemble du modèle se formalise sous la forme d'un projet, se composant de plusieurs éléments associés à l'interface graphique de la plate-forme, présentée ci-dessus :

- D'un **espace de travail** (en rouge) permettant d'accéder à un ensemble de librairies ainsi qu'à plusieurs modèles d'illustrations (*Models library, Shared models*). C'est aussi ici que sont stockés les projets (*User models*) permettant d'accéder à l'ensemble des attributs et données nécessaires en entrée et en sortie lors de la construction d'un projet de simulation multi-agents (*doc, images, includes, models).*
- D'un **moniteur d'affichage** (en orange) permettant d'observer l'évolution d'une simulation par l'intermédiaire de graphiques et de moniteurs définis par l'utilisateur.
- D'un contrôle des **paramètres** (en noir) définis par l'utilisateur et utilisés lors de l'initialisation du modèle de simulation. Cette partie de l'interface permet, par exemple, d'inspecter et de paramétrer le nombre d'agents au démarrage d'une simulation.
- D'une **console de suivi** (en vert) permettant de suivre en détail l'évolution générale d'un modèle à chaque cycle du processus de simulation.

- D'un **éditeur de projet** (en violet) permettant l'écriture d'un projet au format GAML[116]. Par l'intermédiaire d'un environnement de développement intégré (IDE), l'écriture d'un projet est facilitée en affichant en direct les erreurs et les différents statuts pour la compilation finale d'une simulation (Taillandier *et al.*, 2014). Au niveau de la structuration du langage de modélisation, l'éditeur de projet de GAMA présente de nombreuses similitudes avec NetLogo[117] en se construisant sous la forme de trois blocs au sein de l'éditeur de projet (Taillandier *et al.*, 2014) :
    - 1. Un **bloc global** décrivant les caractéristiques et les dynamiques générales du modèle représenté ;
    - 2. Des **espèces** qui permettent de définir les caractéristiques, comportements et représentations visuelles de chacun des agents créés dans le projet selon un formalisme « comportemental », où les agents agissent selon des réflexes (*reflex*) (Cura et Tannier, 2015) ;
    - 3. Les **expérimentations** dans lesquels sont définies les paramètres en entrées (*input*) de même que les affichages et sorties (*output*) pour l'exécution du modèle.

Ces différents arguments tendent à montrer les points communs qu'entretient GAMA avec les plates-formes de la seconde catégorie. Néanmoins, le choix de mobiliser GAMA se justifie sur plusieurs points qui en font sa spécificité par rapport à d'autres plates-formes de SMA (Taillandier *et al.*, 2014 ; Bouquet *et al.*, 2015) :

- Elle se situe dans la deuxième catégorie des plates-formes de SMA, permettant de faciliter l'expression et la compréhension du formalisme d'implémentation par l'intermédiaire du langage GAML.
- Elle permet la construction de modèles complexes, ce qui la distingue quelque peu des plates-formes de deuxième génération. En effet, « *contrairement à NetLogo qui se limite à la construction de modèles simples, GAMA permet la construction de modèles très complexes, aussi riches que ceux construits par un informaticien à partir d'outils comme RePast* » (Taillandier *et al.*, 2014, p. 1). Par exemple, il est possible de définir des agents comme étant composés d'autres agents, ce qui n'est pas possible avec NetLogo (Bouquet *et al.*, 2015). Dans ce sens, nous pouvons plutôt positionner GAMA dans un contexte original, en la situant à l'interface des plates-formes de première et deuxième catégories (*cf.* figure 2.21, p. 184).
- Elle permet d'intégrer des données géographiques avec une plus grande facilité que NetLogo. Elle se distingue par l'attention portée à la modélisation de l'environnement à partir des données SIG, que ce soit pour des données géographiques vectorielles ou raster.
- Elle comprend un ensemble d'opérateurs spatiaux naturellement associés à l'intégration des données géographiques. Ces opérateurs permettent de fournir des fonctionnalités proches d'un SIG comme la définition de requêtes spatiales, de tests géométriques… (Taillandier *et al.*, 2014).

[116] Liste exhaustive du langage GAML : https://github.com/gama-platform/gama/wiki/G__Index.

[117] Comparativement, la construction d'un modèle NetLogo se structure en quatre blocs relativement similaires à GAMA (Amblard *et al.*, 2015) : 1. La définition des variables globales ; 2. La définition d'un état initial et des variables globales ; 3. La définition des comportements des agents et de l'environnement ; 4. La définition des sorties du modèle.

- Elle a déjà fait l'objet d'utilisation avancée dans différents domaines, montrant ainsi sa capacité opérationnelle pour des questionnements liés à diverses applications, portant autant sur l'épidémiologie (Projet AEDESS : Maneerat et Daudé, 2013), le transport routier en milieu urbain (Projet MOSAIIC : Czura *et al.*, 2015), la mobilité quotidienne (Projet MIRO : Banos *et al.*, 2010), que la géohistoire (Projet TransMonDyn : Cura et Tannier, 2015).

Pour nuancer nos propos, on peut relever quelques défauts de la plate-forme GAMA : par exemple, en ce qui concerne les changements de quelques fonctionnalités ou la simplification du langage GAML entre les différentes versions de la plate-forme GAMA[118]. Ces modifications illustrent surtout la récente mise en place de la plate-forme de SMA. Enfin, il faut noter que le positionnement de GAMA à l'interface des plates-formes de première et deuxième catégories s'avance à la fois comme un atout, mais aussi comme un inconvénient qui se reflète par une complexité du langage GAML, plus difficile à appréhender que, par exemple, le langage natif de NetLogo.

Dans la suite, nous avons privilégié la plate-forme de simulation multi-agents GAMA[119] pour la constitution du modèle SCAUP (*Simulation multi-agents à partir de CApteurs Urbains pour la Pollution atmosphérique automobile*). De plus, l'insertion autant des données géographiques que des opérateurs spatiaux fournit une plus-value pour une gestion avancée des couches géographiques issues des SIG. D'autant qu'aborder les problématiques de la modélisation urbaine à travers un usage couplé des SIG et des SMA constitue une opportunité pour aborder les questions liées aux sciences de la complexité (Crooks, 2015). C'est aussi dans ce sens que les SIG et les SGBD ont été mobilisés, autant pour la constitution du dispositif de quantification que pour la reproduction des agents et de l'environnement du modèle de simulation multi-agents. Par ces derniers propos, on retrouve ainsi l'ancrage thématique de ce travail au sein de la *géocomputation* (*cf.* section 1.3, p. 45).

## 6.2. Formalisation du modèle de simulation multi-agents : du projet à l'expérimentation de SCAUP

L'enjeu du modèle SCAUP est de quantifier le trafic sur l'ensemble des tronçons d'un réseau routier à partir d'un signal ponctuel et épars, issu des données de comptage, pour ensuite estimer les émissions de PAA. Les SMA permettent d'inscrire des agents au sein d'un environnement dans lequel ils agissent, et interagissent entre eux mais aussi avec leur environnement. Par ailleurs, ils présentent l'avantage d'offrir un lien direct avec les SIG permettant, d'une part, d'inscrire des agents au sein d'un environnement géographique et, d'autre part, de reproduire la mobilité de ces agents par la description de leur comportement dans un environnement (Marilleau, 2006).

Au cours de cette section, nous nous appuierons sur la figure 2.24 qui décrit la structure complète du modèle SCAUP au format UML. Les « capteurs » et les « véhicules » correspondent aux agents, tandis que le « réseau routier » correspond à l'environnement spatial dans lequel s'inscrivent les capteurs et les véhicules. Nous décrirons la structure et la dynamique du modèle

[118] Documentation détaillées sur les évolutions et changements des différentes versions de GAMA : https://github.com/gama-platform/gama/wiki/G__OlderVersions.

[119] Du côté humain, un critère collaboratif avec Nicolas Marilleau, qui participe et contribue actuellement au développement de la plate-forme GAMA, a aussi conditionné notre choix de se tourner vers cette plate-forme de SMA.

selon l'approche VOYELLES de Yves Demazeau (2003), permettant de décrire la structure et le fonctionnement d'un modèle de simulation multi-agents en cinq étapes : Agents, Environnement, Interaction, Organisation, Utilisation (AEIOU). Toutefois, notons que nous n'aborderons pas le concept d'organisation (O), car le modèle que nous proposons pour la simulation du trafic routier ne présente pas d'organisation sociale entre les « agents ».

La structure générale de SCAUP, présentée en figure 2.24 (ci-dessous), fait directement référence au modèle conceptuel, ainsi qu'au schéma d'intégration des données, présentés précédemment (*cf.* figure 2.13, p. 166 et figure 2.14, p. 167). C'est pourquoi, sous la figure 2.24, on retrouve bien deux grands éléments, qui sont, à gauche, les agents du modèle, et à droite, leur environnement. On peut noter que les relations d'interdépendance entre ces deux éléments suivent la définition du trafic routier en distinguant les véhicules, d'une part, et le réseau routier, d'autre part (respectivement à gauche et à droite de la figure 2.24). Les caractéristiques spatio-temporelles du système de trafic routier sont illustrées par la relation entre les « capteurs » et les « véhicules » implantés sur le « réseau routier » :

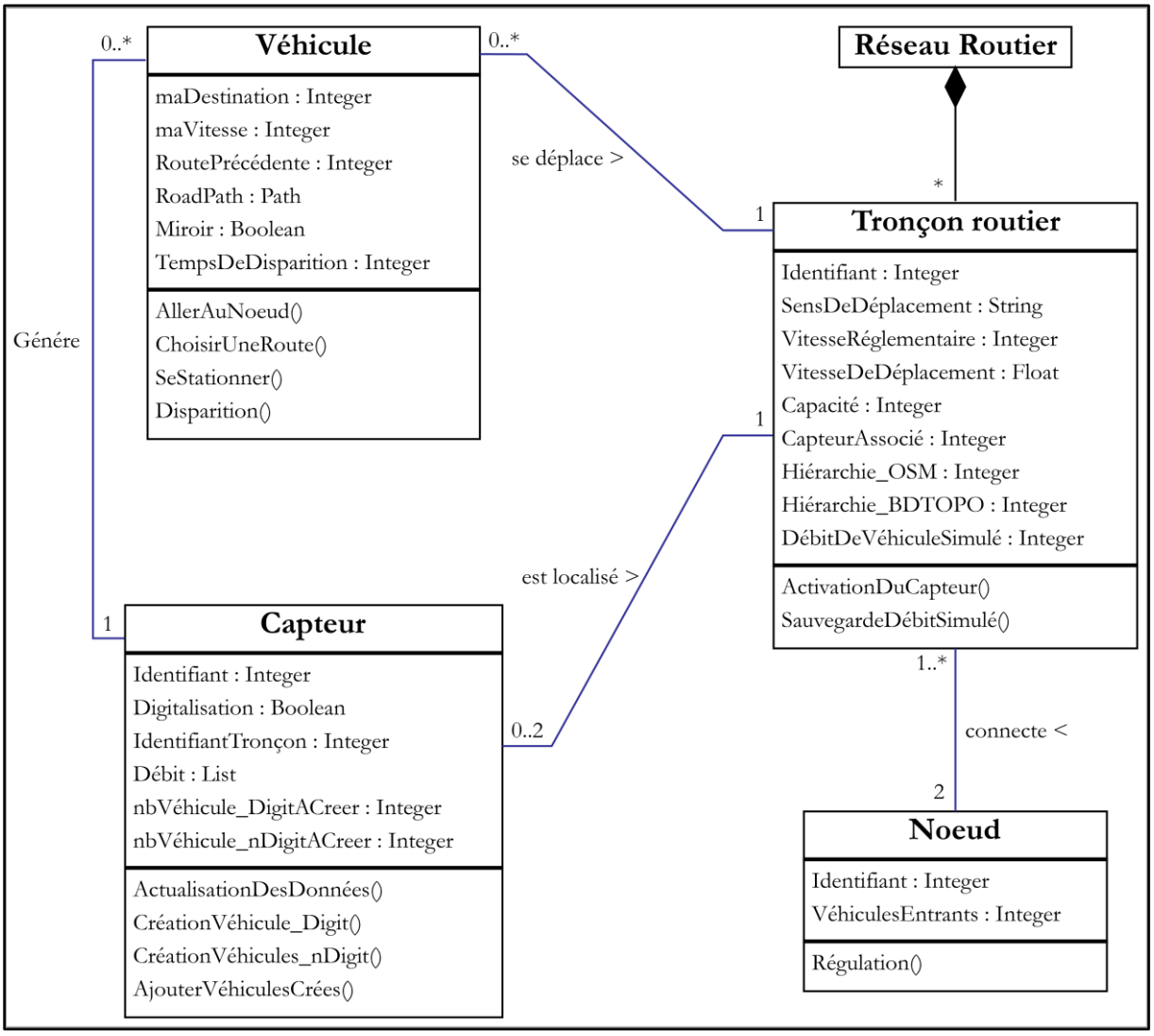



**Figure 2.24 : Structuration UML du modèle d'affectation des comptages routier au sein de SCAUP : du capteur au trafic routier**

### 6.2.1. **Agents et environnement : spatialisation et initialisation du système multi-agents SCAUP**

L'ensemble du corpus de données à simuler se constitue d'une partie que l'on peut qualifier de *spatiale*, le réseau routier, et d'une autre partie plutôt rattachée au *temps*, les véhicules en déplacement. Nous avons vu précédemment que la dynamique de la circulation routière est portée par les données de comptage. En revanche, elles ne permettent pas, en l'état, d'identifier le positionnement des véhicules comptabilisés, et donc de pouvoir produire un trafic routier émergent à une échelle individuelle. Il est nécessaire de resituer les véhicules dans leur contexte spatio-temporel, en les associant, dans un premier temps, aux différentes stations de mesure pour en appréhender la dynamique temporelle, puis, dans un second temps, en les rattachant au réseau routier pour en appréhender le déplacement dans l'espace.

#### a. *Du capteur au véhicule sur la route : formalisation des agents*

Le dispositif de quantification mobilisé a été présenté au cours de la partie 3.2.3 (p. 97). Il se base sur l'utilisation d'un réseau de 208 capteurs électromagnétiques disséminés sur le réseau routier de la commune de Dijon (*cf.* figure 1.38, p. 108). Une fois l'état des lieux du réseau de capteurs établi, les données ont été stockées au sein du logiciel PostGIS, en vue non seulement d'appréhender la localisation des stations sur le réseau routier, mais aussi de rattacher à chacune d'entre elles les données de comptage qui lui sont associées. C'est au sein du dispositif de quantification que la dynamique du système multi-agents est portée (*cf.* figure 1.40, p. 111). En revanche, il reste à identifier les véhicules au sein du dispositif de quantification. En effet, les véhicules se présentent sous la forme d'un lot de véhicules par quart d'heure (débit) ; il faut donc les générer à partir des données de débit rattachées à chaque capteur de trafic.

La relation entre les capteurs et les véhicules s'opère à partir des données de comptage, permettant ainsi d'initialiser le modèle de simulation à partir du dispositif de quantification du trafic routier (*cf.* figure 2.24). Le capteur est employé comme un *générateur* des véhicules (*cf.* figure 2.24). En effet, c'est à partir de la relation entre non seulement la base de données de comptages et le capteur, mais aussi avec les valeurs de débit qui lui sont associées, que les véhicules peuvent être formalisés et ensuite appréhendés individuellement dans le simulateur multi-agents. En conséquence, en étant localisés sur le réseau routier, les capteurs sont des générateurs de trafic routier, et ils insèrent les véhicules sur le réseau routier en respectant les données de comptage réelles qui leur sont associées. Les données étant au quart d'heure, pour affiner la résolution temporelle du modèle de simulation à la minute, la création des véhicules est équitablement répartie sur chaque quart d'heure ; c'est-à-dire qu'ils sont insérés de manière régulière tout au long d'un quart d'heure : par exemple, pour un débit compté de 15 véhicules, 1 véhicule est généré toute les minutes au lieu de 15 véhicules dès la première minute.

Dans ce sens, l'ensemble des caractéristiques de génération d'un véhicule est porté par le capteur, auquel est associé l'ensemble des informations relatives au comptage, ainsi que la route associée au capteur, pour en définir le sens de comptage (et donc de génération) des véhicules (*cf.* figure 1.40, p. 111). Ces informations permettent d'initialiser un véhicule pour qu'il puisse ensuite être replacé correctement dans son contexte (*digit* pour un même sens de comptage et de digitalisation du réseau routier, et *ndigit* pour un sens inverse, *cf.* figure 1.32, p. 101). En générant les véhicules de cette manière, le modèle de simulation et sa dynamique sont non seulement initialisés, mais

aussi portés par les données de comptage, s'inscrivant alors dans l'optique de SMA « *deepening* » (Hassan *et al.*, 2010a, 2010b)[120] :

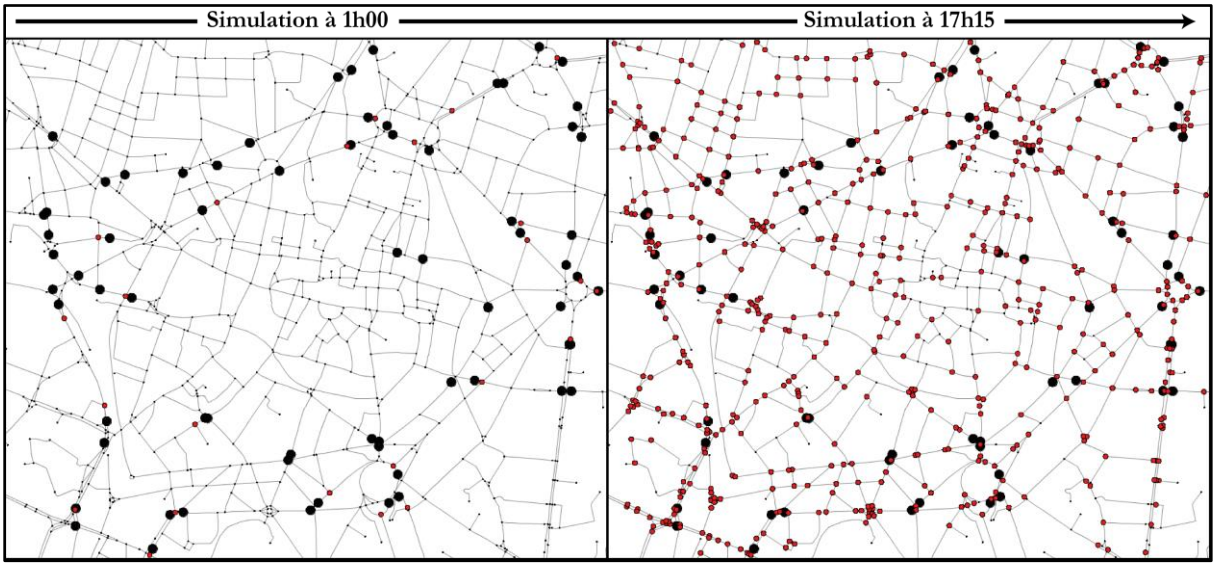



**Figure 2.25 : Génération de l'environnement, des agents capteurs (en noir) et des agents véhicules (en rouge) sur le réseau de la ville de Dijon**

Les véhicules, une fois créés, se déplacent de tronçon en tronçon, en tenant compte des interactions locales avec leur monde simulé (voirie et autres véhicules). Ils participent ainsi, par leur comptage sur chaque tronçon, à quantifier le trafic sur l'ensemble du réseau. Cependant, avant d'aborder les éléments relatifs au déplacement des véhicules, il faut les replacer dans l'environnement dans lequel ils s'inscrivent : le « réseau routier ».

### b. *La génération du réseau routier : l'environnement des agents*

En parallèle des agents du système, le second *objet* qui caractérise le trafic routier est le réseau routier. Il se constitue de l'ensemble des infrastructures permettant le déplacement des véhicules routiers dans l'espace et exclut, par nature, l'ensemble des réseaux pédestres ou non carrossables de la zone d'étude. Dans une perspective de modélisation, la description du réseau routier nécessite d'intégrer trois types d'informations pour qualifier et décrire les caractéristiques de la route (Cabrera Delgado et Bonnel, 2012) :

- i. la description de la structure entre sections et intersections ;
- ii. les propriétés des infrastructures (longueur, sens, vitesse…) ;
- iii. la définition des relations entre ces éléments et les flux de déplacements.

C'est pourquoi pour la constitution du réseau routier, nous nous sommes inspirés de la théorie des graphes, qui se définit comme « *un ensemble fini de points appelés sommets ou nœuds et un ensemble de relations entre ces points appelées arcs ou tronçons. La théorie des graphes porte essentiellement sur l'existence des relations entre les sommets* […]. *La théorie des graphes ne traite que des relations entre des éléments définis et en nombre limité* » (Mathis, 2003, p. 21). À travers cette théorie des graphes le réseau routier se compose de deux éléments :

[120] En d'autres termes, la relation entre le dispositif de quantification et celui de modélisation est directe, dans le sens où ils sont créés en adéquation avec les données de comptage recueillies sur la ville de Dijon.

- Les **intersections**, qui se constituent de nœuds plus ou moins simples représentant une relation causale d'entrée et/ou de sortie
- Les **axes de circulation**, qui sont représentés sous la forme d'arcs orientés, de tronçons représentant les sens de circulation et les propriétés caractérisant chacun des axes de circulation

Le réseau routier se formalise ainsi par ces deux éléments : les *nœuds* et les *tronçons routiers*. Au final, le **réseau routier** est représenté par l'ensemble des nœuds reliés par des arcs orientés, où chacune des relations au sein du réseau est décrite par le lien entre les nœuds et les arcs pour parcourir cet arc : par exemple, par la relation « de… à » (*From…To*). Dans ce sens, la notion de *réseau*, routier ou non, se fonde sur la prise en compte des relations (Lhomme, 2012, p. 106) : « *il n'y a réseau que s'il existe au moins une relation, voire une interrelation, entre des éléments disjoints* ». Ce n'est donc que s'il y a une relation entre les intersections (les *nœuds*) et les axes de circulation (les *arcs*) que le réseau routier peut être défini. Notons ici que ces informations de relation entre les différents axes de circulation de la ville de Dijon sont aussi appréhendées par les informations relatives à la hiérarchie du réseau routier permettant de conceptualiser les caractéristiques circulatoires du réseau routier (*cf.* section 5.3.2.c, p. 174).

Au même titre que les capteurs (illustrés en figure 2.24, p. 191), ce sont les « tronçons routiers » qui supportent l'ensemble des attributs décrivant l'infrastructure et les informations relatives aux déplacements des véhicules. Les « nœuds », quant à eux, permettent de représenter les relations de connectivité et de connexité du réseau routier en lien avec le sens de déplacement des véhicules routiers. Le « réseau routier » a été réalisé à partir de la BD TOPO de l'IGN, dans une perspective d'interfaçage avec le formalisme employé pour l'estimation des émissions de PAA (PCIT, 2012). En outre, il s'insère aussi dans une optique de généralisation du travail présenté sur d'autres terrains d'études, la base de données de l'IGN étant accessible pour l'ensemble des villes françaises : « *la BD TOPO se veut être le référentiel topographique de base pour toutes les applications d'aménagement et de gestion de l'espace, de l'échelle de la commune à celle du département* »[121].

Toutefois, les attributs natifs de la BD TOPO ne sont pas suffisants dans une perspective de simulation du trafic routier, les informations relatives à la vitesse de déplacement ou à la capacité des tronçons routiers ne permettant pas de décrire avec finesse les caractéristiques de déplacement des véhicules. Pour combler ces lacunes, le module de génération automatique du réseau routier associé au logiciel MOBISIM a été mobilisé (*cf.* encart 5, p. 195). Dès lors, les manipulations du réseau routier sont réduites à une activité de vérification, de validation et de test des attributs routiers à l'aide d'un SIG.

[121] Descriptif du contenu de la BD TOPO de l'IGN : http://professionnels.ign.fr/sites/default/files/DC_BDTOPO-2-1.pdf.

***Encart 5 Le module de génération du réseau routier de MOBISIM :***

Le modèle MOBISIM développé au laboratoire ThéMA (Antoni et Vuidel, 2010) contient, en amont de la simulation, un module de génération automatique du réseau routier permettant de générer automatiquement la géométrie du réseau et les caractéristiques de celui-ci à partir de la BD TOPO de l'IGN (Lunardi, 2013). La mobilisation de cet outil permet de dispenser « *l'utilisateur des manipulations en amont nécessaires au bon fonctionnement du réseau. Les attributs de géométrie, d'identifiants, de sens de circulation, de vitesse, de capacité… sont automatiquement générés* » (Lunardi, 2013, p. 25). Ce module permet d'une part d'obtenir un réseau opérationnel en vue de modéliser le trafic routier pour d'autres applications, et d'autre part de générer un ensemble d'informations complémentaires pour qualifier les tronçons routiers de toute la France. Le passage par le module de génération permet d'obtenir des informations complémentaires, comme illustré et décrit dans le tableau de synthèse ci-dessous[122] :

| Attributs | Origine de l'attribut | Caractéristiques et/ou calcul |
|---|---|---|
| **Nature** | BD TOPO | Attribut permettant de distinguer la nature des routes (Autoroute, Quasi-autoroute, Bretelle, route à 2 chaussées, route à 1 chaussée,…) |
| **Largeur** | BD TOPO | Largeur de la chaussée (d'accotement à accotement) exprimée en mètres. |
| **Importance** | BD TOPO | Hiérarchisation du réseau routier fondée, non pas sur un critère administratif, mais sur l'importance des tronçons de route pour le trafic routier |
| **Pente** | BD TOPO et MOBISIM | Altitude calculé par MOBISIM en fonctiondes attributs Z_INI et Z_FIN de la BD TOPO. La pente est exprimée en pourcentage par la relation suivante : $Pente_{\%} = \frac{Z_{fin} - Z_{init}}{longueur} \times 100$ |
| **Sens de circulation** | BD TOPO et MOBISIM | Sens 'inverse' et 'direct' de la BD TOPO redéfinit en trois catégories en fonction de la digitalisation du réseau pour MOBISIM : 'FromTo' (sens identique),'ToFrom' (sens inverse), 'Both' (double sens) |
| **Sinuosité** | MOBISIM | Sinuosité rapport entre la distance euclidienne et la distance réelle du tronçon ou 1 = un tronçon rectiligne et 0 un tronçon sinueux $Sinuosité = \frac{Distance\ euclidienne\ du\ tronçon}{longueur\ réelle\ du\ tronçon}$ |
| **Capacité** | MOBISIM | Nombre de véhicules par heure sur le tronçon correpsondant au débit maximal de véhicules supporté par le tronçon. La capacité est définie selon les caractéristiques du tronçon (nombre et largeur des voies), de la vitesse de circulation et de l'importance du tronçon |
| **Vitesse réglementaire** | MOBISIM | Est définie en fonction de la nature des voies et des caractéristiques des zones d'habitation (cf. tableau 2.2) |
| **Vitesse réelle** | MOBISIM | Vitesse réglementaire pondérée en fonction de la pente, de la sinuosité du tronçon et des coefficients d'ajustement α et β par la relation suivante : $Vitesse_{réelle} = Vitesse_{régl.} \times e^{\frac{-\alpha(1-sinuo)-(\beta\frac{pente}{100})}{(\alpha+\beta)}}$ |

**Tableau 2.2 : Caractéristiques générales des attributs du réseau routier (source : Lunardi, 2013) et http://thema.univ-fcomte.fr/mobisim/)**

[122] L'ensemble des caractéristiques précises des attributs de la voirie routière est détaillé dans le mémoire de Nicolas Lunardi (2013, pp. 25–30), ou sur le site officiel du modèle MOBISIM de 2012 (http://thema.univ-fcomte.fr/mobisim/images/documents/MobiSim_05_Module_D02_Graphes_reseaux.pdf).

### 6.2.2. Mise en mouvement du modèle de trafic routier : formalisation des interactions au sein du modèle SCAUP

Les véhicules sont des agents réactifs dont le but est de se déplacer au sein du réseau routier. Le déplacement des véhicules est régi par différentes interactions, qui s'opèrent autant au niveau des agents entre eux qu'en fonction du réseau routier au sein duquel ils se déplacent. Pour illustrer la dynamique du modèle SCAUP, nous nous focaliserons au cours de cette section sur les différents niveaux d'interaction qui permettent de mettre en mouvement les véhicules pris individuellement sur le réseau routier. Ces interactions sont de deux natures :

- Entre les agents (capteurs et véhicules), pour initialiser la création des véhicules
- Avec l'environnement (nœuds et tronçons), pour appréhender le déplacement et la dynamique du trafic routier.

#### a. *Interaction agent-agent : initialisation des véhicules routiers*

Le premier niveau d'interaction s'opère entre les différents agents du système : les capteurs et véhicules, qui visent à insérer sur le réseau routier les agents véhicules, qui se déplaceront ensuite en toute autonomie. La figure 2.26 décrit à travers un diagramme d'activité UML le processus de génération des véhicules à partir des capteurs et des données qui leur sont associées.

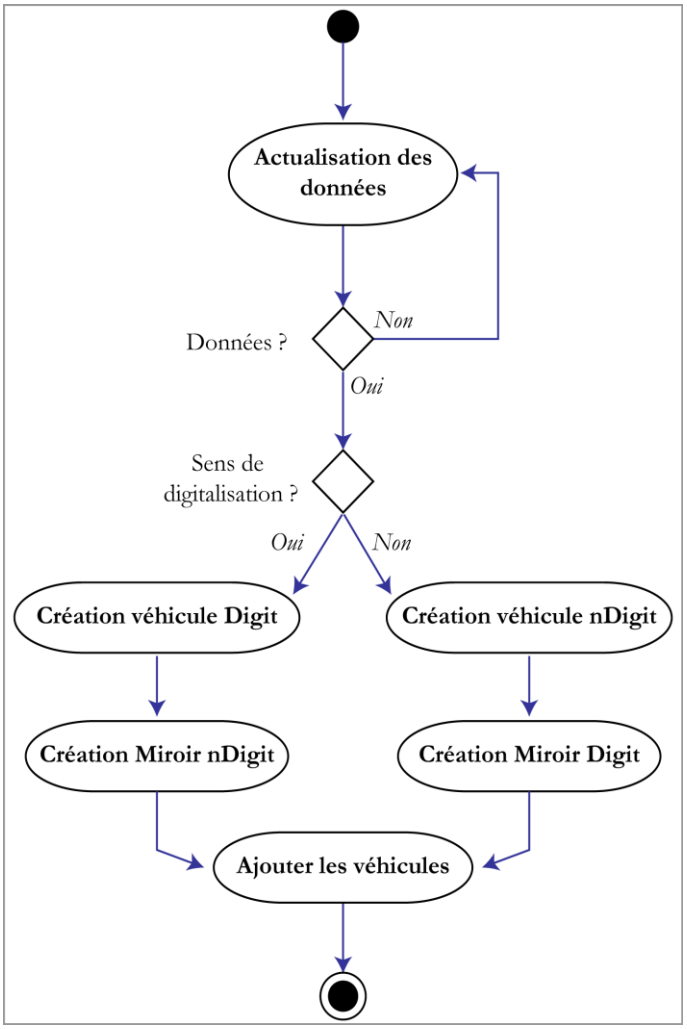



**Figure 2.26 : Formalisation UML de la génération des véhicules à partir des capteurs**

Le capteur permet par les informations qui lui sont affectées (*cf.* figure 2.24, p. 191) d'ajouter les véhicules dans leur contexte, par exemple, en les replaçant dans leurs sens de comptage en tenant compte de la digitalisation du réseau routier (digit et ndigit). Cette approche permet d'injecter à partir des données de comptage les véhicules à chaque pas de temps et au fil de l'eau. De cette manière, le trafic routier est généré sur les axes en amont, par un processus d'accumulation des véhicules, au cours du processus de simulation, en tenant compte des données de comptage réelles. Afin d'éviter des doubles comptages sur les tronçons routiers, lors de la génération d'un véhicule, les capteurs émettent aussi dans le sens inverse de génération d'un véhicule un agent dit « miroir » (*cf.* figure 2.24). La création des « miroirs » permet d'éviter qu'un véhicule ne soit compté plusieurs fois sur une section équipée

d'une station de mesure. En effet, les véhicules qui sont comptabilisés en un point de mesure ont été comptés auparavant par un autre point de mesure. Lorsqu'un véhicule est inséré dans le réseau par un capteur, son « miroir » est créé dans sa direction inverse (*cf.* figure 2.26). Ensuite, quand un agent « miroir » rencontre un agent véhicule, comme illustré ci-dessous, le couple (miroir/véhicule) se neutralise au moment de leur *collision*, faisant alors le lien entre deux points de comptage comme illustré ci-dessous[123]. En d'autres termes, la génération des « miroirs » permet de disposer d'une variable de régulation du trafic routier, permettant de contrôler le nombre de véhicules présents et entrant sur un tronçon possédant une mesure. Une illustration de cette dynamique est proposée ci-dessous :

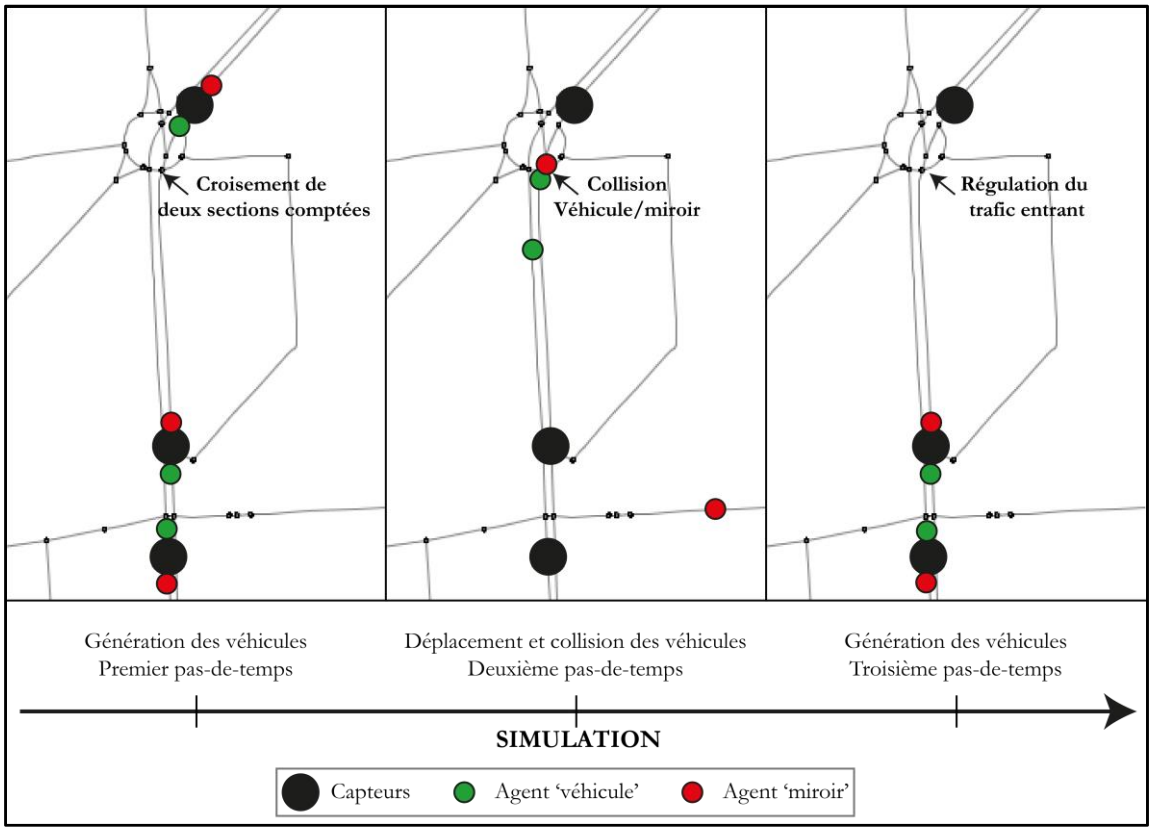



**Figure 2.27 : Illustration de la génération et de l'action des agents « miroir » sur un tronçon routier pour la régulation du trafic routier**

Pour finir, les interactions entre les agents du système sont de plusieurs ordres : la première relation s'opère entre les capteurs et les véhicules par une relation de dépendance avec la base de données de comptages routiers en vue d'initialiser le modèle de simulation multi-agents. Le second degré d'interaction s'opère entre les véhicules et les « miroirs » en vue de contrôler les flux de véhicules passant sur les infrastructures mesurées. Notons ici que nous avons fait le choix d'adopter une approche globale et expérimentale de simulation du trafic routier en milieu urbain à partir des capteurs de trafic routier. Ainsi, les interactions entre les véhicules (variables d'accélération et de décélération, distances de sécurité entre les véhicules se déplaçant sur les infrastructures…) ne sont pas intégrées. Nous reviendrons sur ces éléments au cours du dernier chapitre de cette thèse.

[123] Notons ici que par simplification les « véhicules miroirs » sont générés au même moment que les véhicules. Toutefois, les « miroirs » devraient être insérés à t-1 de la génération des véhicules. En effet, les véhicules déjà comptabilisés en aval des capteurs sont déjà passés en amont d'un tronçon routier.

### b. *Interaction agent-environnement : appréhender le déplacement des véhicules sur le réseau routier*

Après sa création par un agent capteur, le véhicule se déplace en choisissant sa trajectoire et ses actions, en suivant un schéma d'interaction stimuli-réponse illustré en figure 2.28 (ci-dessous). Il se déplace en fonction des caractéristiques du réseau routier afin de non seulement choisir une route, mais aussi pour définir sa vitesse de déplacement sur le tronçon routier où il se situe en fonction de l'attribut « vitesse réelle » (*cf.* encart 5). Nous avons vu précédemment que l'essentiel des interactions entre les « agents » est porté par les capteurs. Ici, pour ce second niveau d'interaction, c'est le réseau routier en tant qu'environnement qui porte, en très grande partie, les informations nécessaires au déplacement des véhicules :

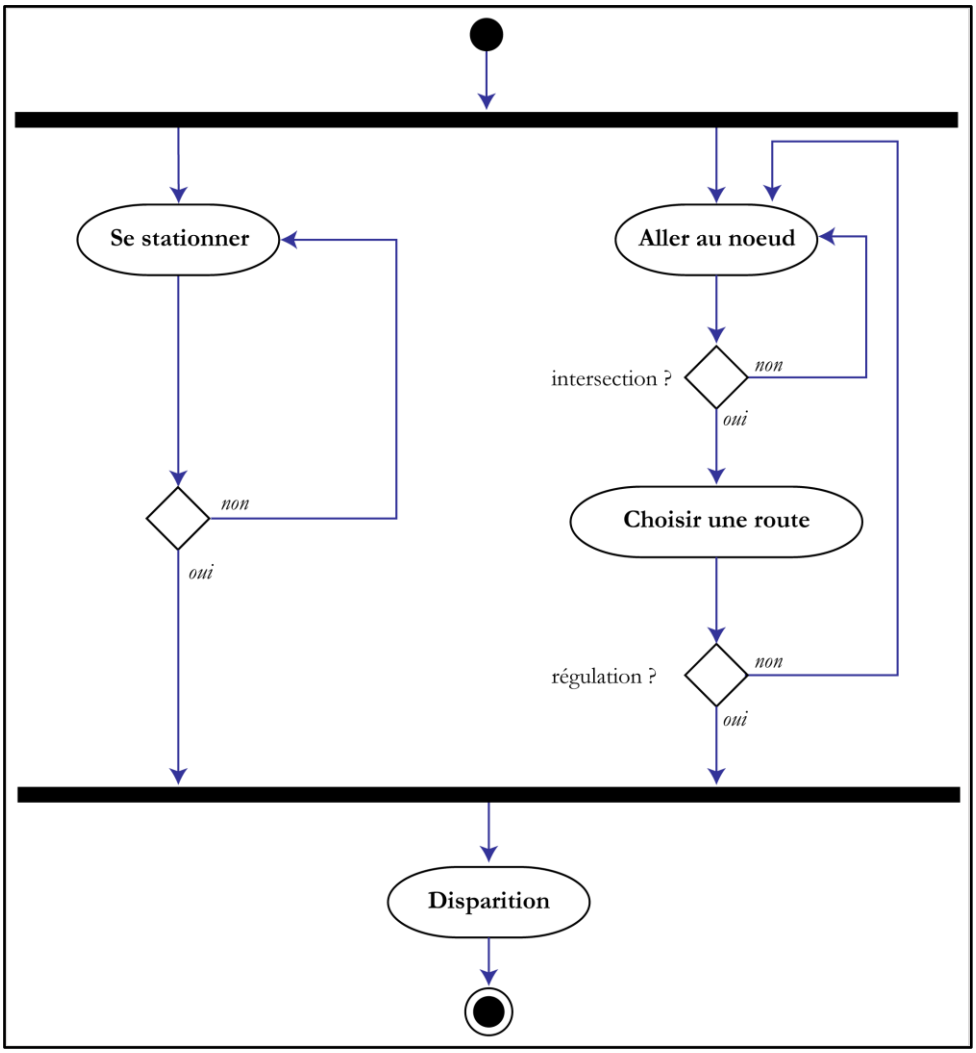



**Figure 2.28 : Formalisation UML de la dynamique de déplacement des véhicules sur le réseau routier**

De plus, la trajectoire des véhicules ne pouvant être appréhendée par la définition d'itinéraires origine-destination, nous avons fait le choix d'établir un modèle par non-énumération des itinéraires en privilégiant un mouvement des véhicules de « nœud à nœud » (de carrefour vers un carrefour). En conséquence, le positionnement d'un véhicule à une intersection détermine le prochain tronçon (connexe à l'intersection) sur lequel le véhicule va se déplacer. Le choix du

prochain tronçon s'opère sur la base d'un critère tenant non seulement compte des caractéristiques de fréquentation des axes mesurés, mais aussi des caractéristiques circulatoires du réseau routier (*cf.* c, p. 169). Le choix d'une route plus que d'une autre a été établi en croisant ces deux informations : les comptages et la hiérarchie du réseau routier.

Pour cela l'ensemble des valeurs de débit collectées a été localisé pour toutes les stations de la ville de Dijon (ponctuelles et permanentes). Cette carte (*cf.* figure 2.29) permet d'observer les flux moyens de trafic routier pour 11% des tronçons routiers de la ville de Dijon (870 comptages sur 827 tronçons). Si cette couverture reste faible, à part pour l'autoroute, l'ensemble des comptages permet d'obtenir au moins une valeur de fréquentation par type de route sur l'espace de la ville de Dijon (des voies résidentielles aux voies rapides). Ce premier traitement permet d'obtenir une image synthétique de la circulation routière sur la ville de Dijon :

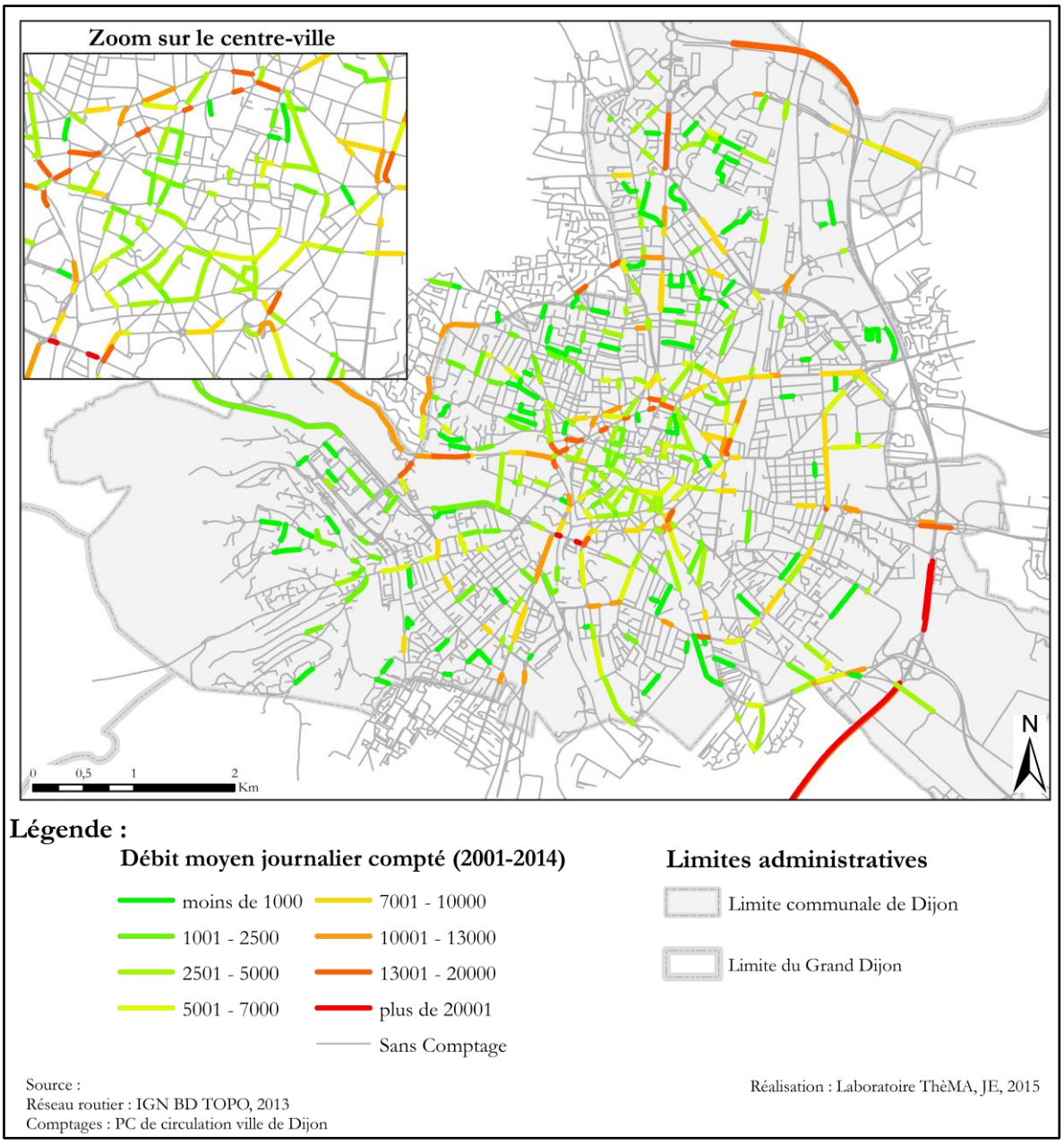



**Figure 2.29 : Synthèse des débits moyens journaliers comptabilisés par toutes les stations de comptages routiers (870 stations) pour la période 2001-2014**

Ces valeurs synthétiques de fréquentation des routes sont ensuite croisées avec les caractéristiques du réseau routier. Ce second traitement permet d'apporter un regard plus fin sur la fréquentation des routes en fonction de deux typologies : la vitesse réglementaire et la hiérarchie OpenStreetMap (*cf.* figure 2.19, p. 177). En croisant ces informations (fréquentation et hiérarchie), on peut alors affecter à chaque typologie d'axe une information relative à sa fréquentations « type » sous la forme d'une part de trafic, comme illustré ci-dessous :

| Vitesse | Nombre total de tronçon | Nombre de tronçon mesurés | Couverture des comptages | Débit moyen compté | % Fréquentation globale |
|---|---|---|---|---|---|
| **25 km/h** | 5604 | 575 | **10,26** | 3 792,7 | **8,66** |
| **50 km/h** | 1581 | 237 | **14,99** | 9 074,5 | **20,72** |
| **90 km/h** | 273 | 9 | **3,30** | 6 752,0 | **15,42** |
| **110 km/h** | 38 | 6 | **15,79** | 24 179,3 | **55,21** |
| **130 km/h** | 10 | x | x | x | x |
| **Total** | 7506 | 827 | **11,02** | 43 798,5 | **100** |

**Tableau 2.3 : Synthèse des comptages routiers en fonction de la vitesse réglementaire**

| Hiérarchie OSM | Nombre total de tronçon | Nombre de tronçon mesurés | Couverture des comptages | Débit moyen compté | % Fréquentation globale |
|---|---|---|---|---|---|
| **Autoroute** | 6 | x | **x** | x | **x** |
| **Rocade** | 40 | 6 | **15,00** | 24 179,3 | **48,87** |
| **Route interurbain** | 1120 | 175 | **15,63** | 10 043,4 | **20,30** |
| **Route intraurbaine** | 1076 | 163 | **15,15** | 6 137,1 | **12,40** |
| **Route inter-résidentielle** | 515 | 90 | **17,48** | 6 437,1 | **13,01** |
| **Route résidentielle** | 4605 | 393 | **8,53** | 2 684,4 | **5,43** |
| **Lien** | 144 | x | x | x | x |
| **Total** | 7506 | 827 | **11,02** | 49 481,3 | **100** |

**Tableau 2.4 : Synthèse des comptages routiers par type de route (hiérarchie OpenStreetMap)**

Si la couverture du réseau routier par les stations de comptage n'est pas parfaite, on observe que ces informations permettent d'avoir une image de la répartition de chacun des axes routiers en tenant compte des valeurs *in situ*. D'une manière générale, en s'appuyant sur ces deux tableaux de synthèse, on remarque que plus un axe présente une hiérarchie importante, plus il concentre une part importante du trafic routier. Que ce soit pour les vitesses réglementaires ou pour OpenStreetMap, le trafic routier aura tendance à se reporter des axes de plus faible hiérarchie vers ceux présentant une hiérarchie plus forte. Ces informations confortent nos hypothèses sur les caractéristiques circulatoires et la manière avec laquelle le réseau routier organise et hiérarchise les flux de circulation routière en ville (*cf.* section 5.3.1.c, p. 169). C'est pourquoi nous avons fait le choix de mobiliser ces valeurs de fréquentation en vue de définir les routes qui sont privilégiées par les véhicules lors de la simulation.

Le passage de ces valeurs de fréquentation en variable de choix est appréhendé par la définition d'une matrice de transition, sur laquelle chaque véhicule va choisir une nouvelle route en fonction non seulement de sa position relative au nœud, mais aussi du choix qui lui est donné par ces valeurs de fréquentation (*cf.* tableau 2.3 et tableau 2.4). Ainsi, la probabilité qu'un véhicule soit à la position $i$ à un nœud suivant ($n+1$), sachant qu'il était à la position $j$ au nœud précédent, est

donnée par la relation suivante :

$$P(X_{n+1} = i | X_n = j) = p_{ji}$$

La matrice de transition ($A$) est composée de l'ensemble des $p_{ji}$, dont les valeurs varient entre 0 et 1 en fonction des pourcentages définis ultérieurement (*cf.* tableau 2.3 et tableau 2.4). Lorsqu'un véhicule arrive à un nœud, il tient compte des différentes routes qui sont à sa disposition ; par exemple, à une intersection simple où seulement deux routes sont présentes, cela implique que tous les $p_{ji}$ sont nuls sauf deux valeurs $p_{jj}$ qui définissent que le véhicule conserve son choix précédent et $p_{ji}$ qui définit que le véhicule modifie sa route. L'ensemble de ces valeurs est stocké sous la matrice de transition $A$, regroupant l'ensemble des valeurs de fréquentation sous la forme suivante, où la somme des valeurs sur une ligne est égale à 1 pour toutes les routes *j* :

$$A = \begin{pmatrix} p_{11} & p_{12} & \dots \\ p_{21} & p_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix}$$

L'utilisation de la matrice de transition permet d'éviter que les véhicules modifient leur position constamment, et elle présente l'opportunité de mieux régir leur déplacement en tenant compte de la position relative du véhicule sur sa route initiale ($p_{jj}$). Dans ce sens, un véhicule aura tendance à mieux conserver sa position sur une route de même hiérarchie, au lieu de modifier constamment son choix pour une route d'une autre hiérarchie. Enfin, le choix des routes est une variable stochastique, car à chaque nouveau choix le véhicule tire au sort le choix d'une nouvelle route en fonction des valeurs définis par $p_{ji}$. L'ensemble des choix routiers est établi à une échelle macroscopique en se focalisant sur les caractéristiques circulatoires du réseau routier, mais aussi en tenant compte des spécificités locales apportées par les comptages *in situ*. Une fois que le véhicule a choisi sa route en fonction de la matrice de transition, celui-ci continue son déplacement vers un nouveau nœud en vue de choisir une nouvelle route. Ce processus est répété à chaque fois que le véhicule rencontre un nouveau « nœud ».

### c. *Disparition et stationnement des véhicules lors de la simulation*

Enfin, le mouvement des véhicules est contraint par leur position sur la route, mais aussi par leur temps de déplacement depuis leur génération. En effet, le trafic routier est produit par l'accumulation des véhicules sur les tronçons routiers lors de la simulation. Dans ce sens, le nombre de véhicules s'accumule très rapidement sur le réseau routier, alors que dans la réalité, on peut sortir de l'espace de la ville de Dijon pour se rendre vers d'autres lieux, on se stationne,… Il apparait nécessaire de gérer le temps de résidence (de déplacement) des véhicules sur le réseau routier. Il faut alors gérer ces disparitions lors du déplacement des véhicules au cours de la simulation. En effet, il est très rare de se déplacer plus d'une heure en ville.

Afin de gérer ce processus cumulatif, deux issues sont possibles (*cf.* figure 2.28, p. 198) : soit les véhicules disparaissent aux marges du réseau routier de la ville de Dijon lorsque aucun nœud n'est présent lors de leur déplacement (*disparition*) ; soit ils se déplacent pendant un temps de déplacement maximum à l'intérieur du réseau urbain (*stationnement*). Ce dernier cas peut être assimilé à un *stationnement*, alors que le premier cas se rattache plutôt au trafic de transit qui traverse l'espace de la ville de Dijon. Dans ce sens, la *disparition* permet au véhicule de sortir du réseau routier lorsque celui-ci arrive à la marge de l'espace d'étude. Notons ici, que l'assimilation

au stationnement ne tient pas compte des places de stationnement disponible en ville. Ce paramètre n'est pas établi spatialement mais plutôt temporellement par l'intermédiaire de la courbe illustrée en figure 2.30. Pour cela, nous nous sommes inspiré des travaux de Manuel Appert et Laurent Chapelon (2002, p. 14) sur la modélisation des chaines de transport, où le temps de recherche d'un stationnement est défini par un critère temporel, fixé à 15 minutes en centre-ville : « *ces temps sont calibrés à partir de reconnaissance de terrain et sont directement intégrés aux calculs* ». Dans notre cas, la reconnaissance sur le terrain de ces temps de recherche apparait très complexe par rapport à notre espace d'étude, c'est pourquoi le temps de stationnement est une variable sensible du modèle SCAUP, que nous explorerons plus finement lors de l'expérimentation du modèle de simulation multi-agents. Toutefois, nous avons défini un temps de stationnement par défaut à 30 minutes comme valeur de départ lors des tests de simulation.

L'insertion de ce paramètre temporel permet de limiter le nombre de véhicules présents sur les infrastructures de transport. Comme les véhicules ne stationnent pas tous au même moment, le temps de stationnement est défini par l'intermédiaire d'une loi gaussienne dont les paramètres sont établis manuellement. Mais les véhicules vont aléatoirement se stationner en fonction de l'espace défini par cette courbe (*cf.* figure 2.30). Par exemple, pour une durée de déplacement de 15 minutes et un écart-type de 5 minutes, le stationnement d'un véhicule sera sélectionné aléatoirement en fonction de l'espace défini par la courbe gaussienne (illustrée en gris ci-dessous). Chaque véhicule stationnera à des intervalles de temps différents, comme illustré : le *véhicule n° 1* se stationnera à partir de 11 minutes, alors que le *véhicule n° 4* se stationnera au bout de 21 minutes de déplacement. Par conséquent, chaque véhicule stationnera individuellement, évitant alors que les véhicules générés à un même moment ($t_0$) par différents capteurs ne disparaissent tous à un même instant ($t_{30}$).

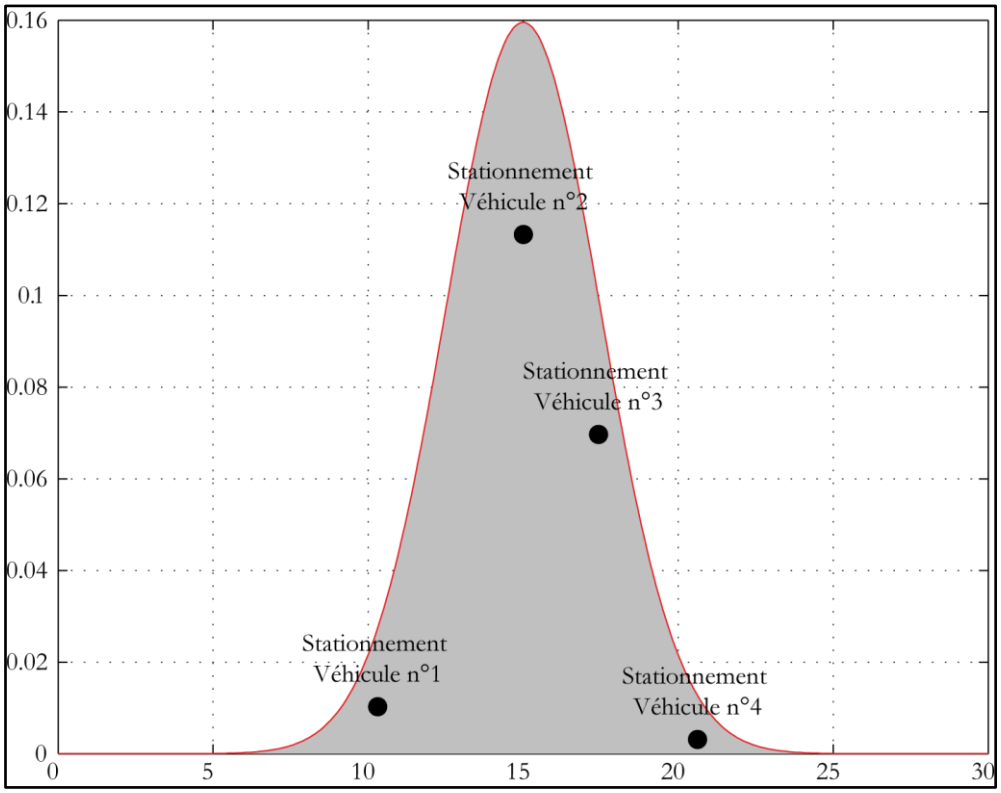



**Figure 2.30 : Paramétrage du stationnement des véhicules au cours de la simulation selon une loi gaussienne (temps de déplacement de 15 minutes pour un écart de 5 minutes)**

Enfin, notre approche étant guidée par les données de comptage, les véhicules en stationnement ou en sortie du réseau routier ne sont pas remis en jeu lors du processus de simulation. Nous considérons ici que la réapparition d'un véhicule se fera de manière discrète à travers les comptages routiers. Par exemple, un véhicule comptabilisé à 8 h 00 par une station de comptage à la sortie de la ville de Dijon pourra être comptabilisé de nouveau par une station située, elle, en entrée de la ville à 18 h 00. Cela est, par exemple, le cas des flux de trafic routier domicile-travail se dirigeant à l'extérieur du terrain d'étude le matin et revenant lors de la fin de journée.

### 6.2.3. **Le suivi du processus de simulation par l'utilisateur**

Cette dernière section vise à apporter un dernier regard sur les relations que l'utilisateur entretient avec le modèle de simulation multi-agents en s'appuyant sur la dernière composante de l'approche VOYELLES (U comme Utilisateur) (Demazeau, 2003). L'intégration de l'utilisateur dans la description du modèle de SMA permet de le positionner comme un acteur actif du modèle de simulation. Ici, nous nous arrêtons sur les éléments permettant à l'utilisateur non seulement de suivre, mais aussi de contrôler le bon fonctionnement du modèle de simulation. La simulation du trafic routier avec SCAUP est établie toutes les minutes, afin de reproduire les mouvements de circulation routière sur l'ensemble du réseau pendant 24h. L'observation du trafic routier s'opère sur chacun des axes du site d'étude en mesurant le nombre de véhicules passant toutes les 15 minutes. Ces données peuvent être visualisées à chaque pas de temps, directement sur le réseau routier de la ville de Dijon ou extraites au sein d'un tableur de données. Ces différents éléments permettent de suivre le comportement du modèle au cours du processus de simulation, duquel une illustration est proposée en figure 2.31, pour plusieurs pas de temps de la simulation. Le suivi du modèle de simulation est constitué de trois éléments permettant à l'utilisateur d'observer le comportement du modèle. En plus de ces trois illustrations, l'utilisateur peut préalablement définir une interface graphique en vue de suivre l'évolution d'un seul axe défini par l'utilisateur. Ce suivi local présente les mêmes caractéristiques que le suivi global présenté en bas de la figure 2.31 :

- Un suivi général du processus de simulation par l'intermédiaire d'une représentation cartographique en deux dimensions (*cf.* figure 2.31, en haut). Au sein de cette interface, on peut suivre le déplacement des véhicules pris individuellement, ou à l'échelle de la ville de Dijon. Il est possible d'observer le stationnement des véhicules par l'intermédiaire d'une grille comptabilisant le nombre de véhicules stationnés en chacun des points de la grille.
- Un suivi de la fréquentation moyenne des axes sous la forme d'une carte en trois dimensions (*cf.* figure 2.31, au milieu). Avec cette représentation, la hauteur de l'axe, de même que la variation des couleurs (du vert au rouge), permettent d'observer la fréquentation moyenne des axes au cours du processus de simulation.
- Un suivi global de la fréquentation des axes routiers à travers trois graphiques permettant d'observer (*cf.* figure 2.31, en bas): la fréquentation moyenne sur l'ensemble du réseau routier (en haut) ; le cumul de véhicules simulés par axe (en bas à droite) ; la relation entre le débit simulé rapporté à la capacité de l'axe, sous la forme d'une courbe débit-capacité[124] (en bas, à gauche).

[124] Cette dernière représentation vise à représenter la courbe débit/capacité (*cf.* encart 1, p. 68). Toutefois, la friction entre les véhicules routiers n'étant pas prise en compte, cette courbe présente une forme linéaire.

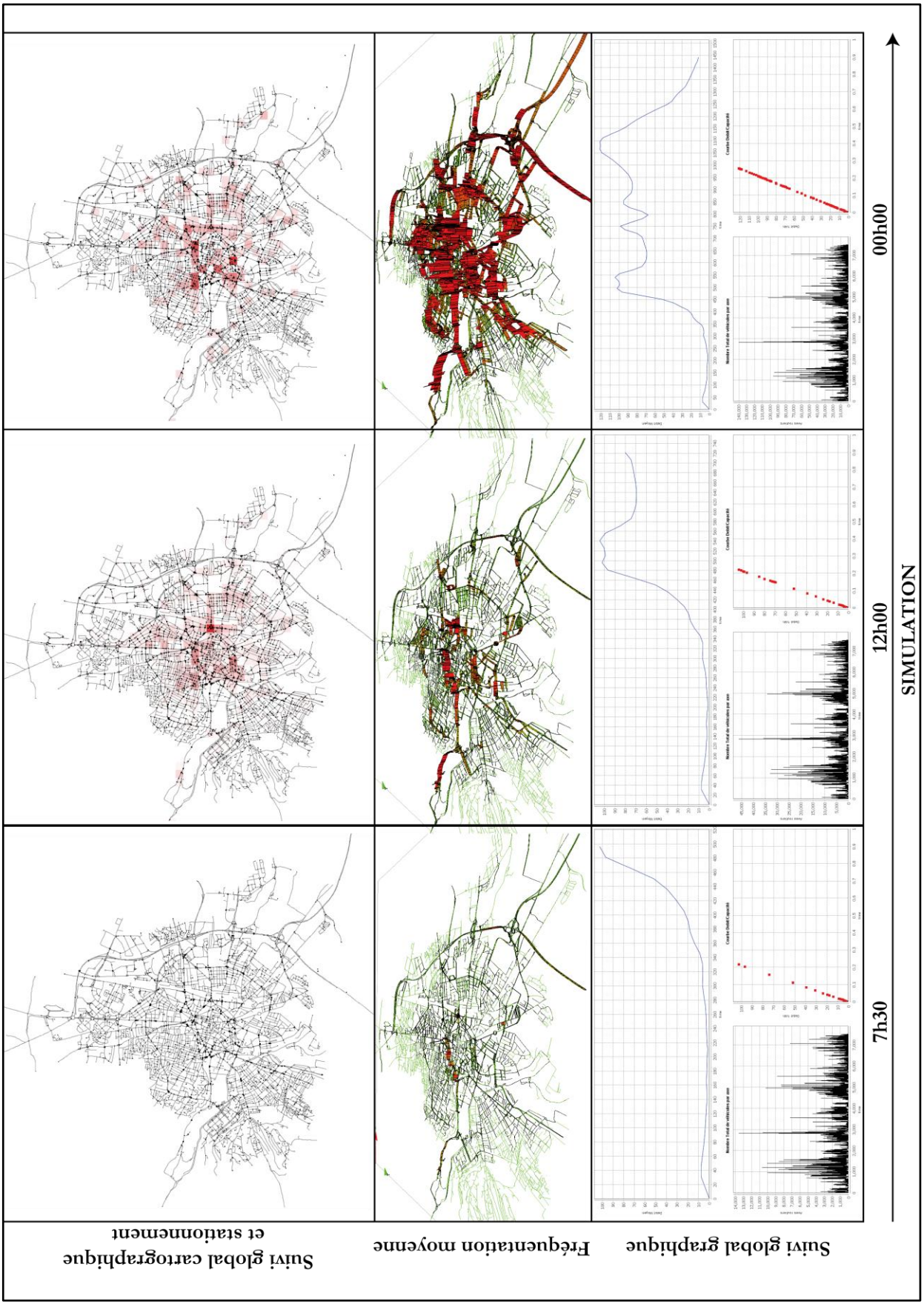



**Figure 2.31 : Illustration des interfaces utilisateur lors du processus de simulation à trois horaires différentes (7 h 30, 12 h 00 et 00 h 00) sur la ville de Dijon**

## 6.3. Expérimentation et plan d'expérience du modèle SCAUP :

### 6.3.1. **Une démarche expérimentale de simulation abductive**

Nous avons présenté l'ensemble des éléments nécessaires à la constitution du modèle SCAUP, or, la réalité de son processus de construction est bien plus tortueuse et abductive. En effet, les paramètres mis en avant précédemment ont été insérés progressivement lors de l'expérimentation du modèle de simulation qui a été réalisée pas-à-pas. Au cours de cette section nous présenterons le plan d'expérience et d'exploration qui nous ont permis d'aboutir à une simulation du trafic routier sur la ville de Dijon. Deux types d'environnement de modélisation ont été mis en place pour l'expérimentation de SCAUP (*cf.* section 5.3.1.c, p. 169). Un premier dit *théorique* (KISS) en mobilisant un réseau routier simple permettant de tester les capacités du modèle à produire un trafic routier émergent à partir de comptages routiers. Et un second modèle dit *réel* (KIDS) qui nous permettra de tester la dynamique du précédent modèle dans des conditions plus réelles via la ville de Dijon (*cf.* Chapitre 7, p. 223). Le plan d'expérience vise à étudier la capacité de l'approche KISS à produire un trafic routier émergent sur un ensemble restreint de tronçons routiers et de stations de mesure (*cf.* figure 2.17, p. 171). Cette première phase expérimentale vise à mettre en place les bases du fonctionnement du modèle SCAUP qui sera ensuite utilisé et testé sur la ville de Dijon. La synthèse de la première phase d'expérimentation est illustrée au sein du schéma ci-dessous :

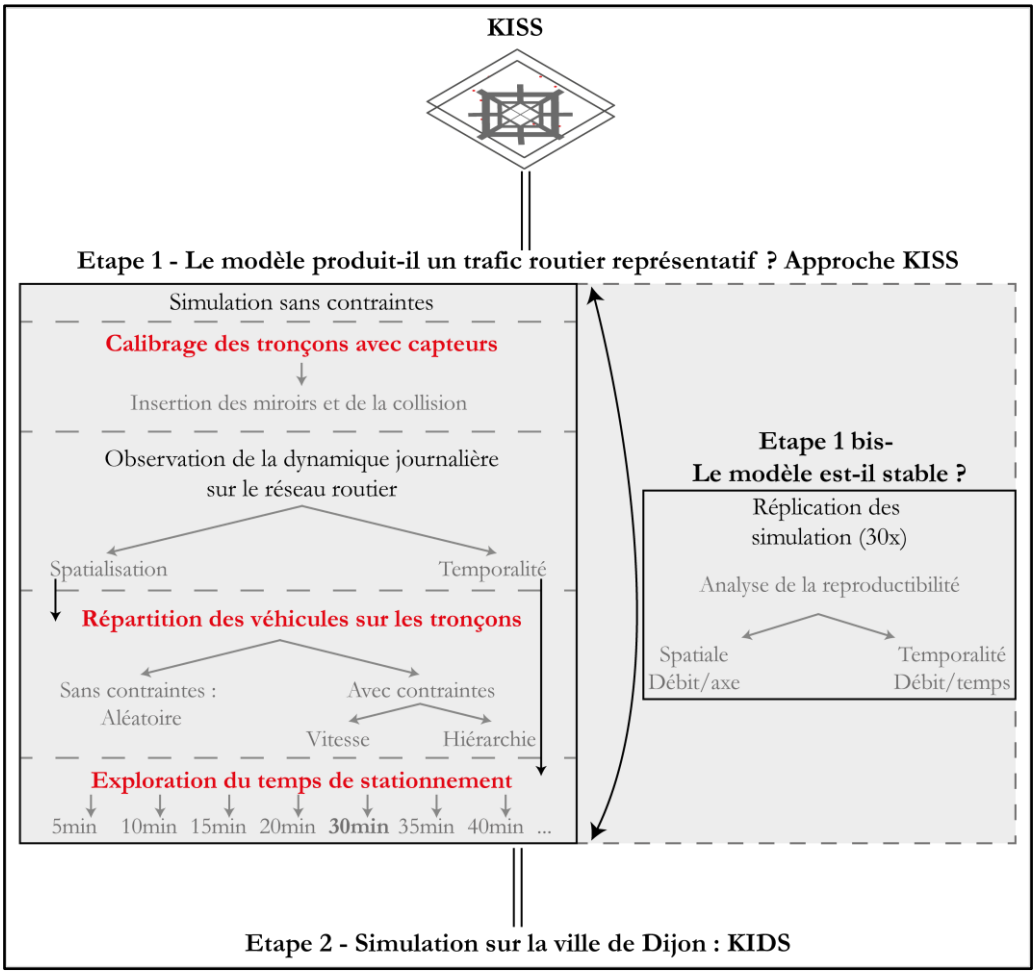



**Figure 2.32 : Plan d'expérience du modèle SCAUP : du réseau théorique à la ville de Dijon**

La mobilisation de ce premier environnement simple permet d'explorer bien plus rapidement la capacité et la sensibilité de SCAUP à produire le signal de trafic routier issu des capteurs électromagnétiques pour une journée de simulation (24 h). En effet, le temps de simulation de l'environnement KISS est estimé à 3 minutes pour 24 h de trafic routier simulé[125]. Le plan d'expérience s'organise en deux étapes parallèles (*cf.* figure 2.32, ci-dessus) :

- une première visant à **observer la capacité du modèle à produire un trafic routier émergent** à partir des stations de comptage, en insérant au fur et à mesure des nouvelles fonctionnalités.
- une seconde visant plutôt à **apprécier la capacité du modèle à reproduire un signal cohérent** entre les différentes simulations, en étudiant sa reproductibilité sur le plan spatial et sur le plan temporel.

La première phase expérimentale vise à tester la capacité du modèle à simuler le trafic routier à partir des comptages sans y introduire de contraintes. À ce niveau, les véhicules générés par les capteurs se déplacent aléatoirement sur les tronçons routiers pendant 24h. Ensuite, nous nous sommes focalisé sur les tronçons routiers équipés de capteurs, en vue de les calibrer aux comptages qui leurs sont associés. En effet, à ces lieux stratégiques, le nombre de véhicules générés doit être strictement similaire au comptage qui lui est associé lors du processus de simulation. Une fois les axes équipés de capteurs calibrés en concordance avec les données de comptage, le plan d'expérimentation vise à insérer progressivement différents paramètres permettant d'appréhender les déplacements des véhicules au niveau spatial et au niveau temporel ; par exemple, en contraignant le déplacement des véhicules à travers la matrice de transition des choix routiers, ou en gérant leur temps de déplacement sur les tronçons routiers. Nous décrirons plus en détail ces différents éléments afin d'exposer les résultats les plus caractéristiques du modèle SCAUP présenté ici (illustré en rouge en figure 2.32).

À chacune des sous-étapes, un processus d'analyse de la stabilité du modèle a été établi en répliquant (30 fois) chacun des blocs du plan d'expérimentation illustré en figure 2.32 (*cf.* section 6.3.2, ci-après). Pour analyser la stabilité et la sensibilité du modèle, un indice de reproductibilité a été mis en place pour observer spatialement et aussi temporellement la capacité du modèle de simulation à reproduire le signal global d'une simulation, rapportée à l'ensemble de ces réplications. Cet indice de reproductibilité[126], utilisé en climatologie pour analyser la sensibilité des modèles de climat à reproduire des phénomènes climatiques (Crétat *et al.*, 2011 ; Pohl et Douville, 2011 ; Pohl, 2015), a été pour notre travail réinterprété pour analyser la sensibilité du modèle de simulation du trafic routier. Il présente l'opportunité d'interpréter le signal des différentes simulations (les membres) par une analyse de leur variance, et donc *a fortiori* d'apprécier la stabilité du modèle. En d'autres termes, la reproductibilité est calculée comme un ratio *signal sur bruit*, entre la variance d'un membre et la variance totale de tous les membres (30 membres) :

$$\text{Reproductibilité} = \frac{\text{Variance moyenne de l'ensemble (1 membre)}}{\text{Variance totale (N membres)}} * 100$$

[125] Comparativement, le modèle KIDS prend bien plus de temps, avec une moyenne de simulation étant estimée à 1 h 30, 30 fois supérieure au temps de simulation du modèle KISS.

[126] À noter qu'au cours de ce travail le terme de reproductibilité sera employé au sens de cet indicateur statistique, et non pas au sens informatique du terme qui quant à lui est plutôt utilisé dans le domaine de la recherche reproductible, qui vise plutôt à passer d'un cas particulier à un cas plus général (Marilleau, 2006), faisant référence à la reproductibilité du calcul ou à sa reproductibilité sous d'autres logiciels…

L'indice de reproductibilité sera, par la suite, exploité pour analyser spatialement et temporellement les données de simulation. Pour l'analyse spatiale, l'indice est calculé en se focalisant sur le débit simulé par tronçon routier ; alors que pour l'analyse temporelle nous nous focalisons plutôt sur le débit de véhicules simulé par quart d'heure. En d'autres termes, l'indice spatial permet d'observer la fréquentation d'un axe, alors que l'indice temporel permet d'observer la capacité du modèle à reproduire la dynamique journalière du trafic routier (matin, midi et soir). Ainsi un signal est dit reproductible lorsque le même signal de trafic routier est observé entre chaque simulation (indice de 100 %). À l'opposé, la non-reproductibilité du modèle se caractérise par une forte différence entre les simulations face à la variance totale de l'ensemble (indice de 0 %). Une fois la capacité du modèle SCAUP à simuler et à reproduire un trafic routier émergent testés sur un environnement KISS, il sera ensuite appliqué à l'échelle de la ville de Dijon (*cf.* étape 2 en figure 2.32). Cette première phase d'expérimentation n'introduit pas les tests de validation du modèle de simulation, en confrontant les données simulées aux données de terrain, pour deux raisons : tout d'abord, la mise en application de l'environnement de simulation KISS vise à tester chacune des hypothèses liées au fonctionnement de SCAUP. Ensuite, nous considérons que la validation des données de simulation ne peut être réalisée sur un réseau dont on ne peut en apprécier la réalité, comme c'est le cas avec ce premier test. C'est pourquoi les éléments de validation des données de simulation de SCAUP seront appliqués et présentés sur le réseau réel de la ville de Dijon dans un second temps (*cf.* Chapitre 7, p. 223).

### 6.3.2. **« Répliquer, répliquer et répliquer »** [127]

« *Étant donné que presque toutes les simulations ne peuvent se prêter à une analyse formelle, la seule façon dont elles peuvent être vérifiées s'opère via l'expérimentation de simulations en cours d'exécution. Si nous voulons être en mesure de faire confiance aux simulations que nous utilisons, nous devons les répliquer indépendamment du reste. Une simulation sans réplication n'est pas digne de confiance – ne comptez pas sur ses résultats, ils sont presque certainement faux* » (Edmonds et Hales, 2003, p. 15). Cette citation illustre un écueil important lors de l'exploration des modèles multi-agents, qui montre que sans *réplication* il ne peut y avoir de confirmation ou de rejet des résultats issus du plan d'expérience (Bommel, 2009). Elle constitue aujourd'hui une étape indispensable, voire de base (Amblard *et al.*, 2006), lors de la mise en œuvre d'un modèle multi-agents. Malgré tout, cette étape est souvent sous-estimée, voire occultée, dans la littérature. Dans son sens large, l'objectif de la *réplication* vise à s'assurer que les résultats issus de l'expérimentation ne sont pas une coïncidence (Wilensky et Rand, 2007). Néanmoins, à travers la littérature mobilisée, il s'avère que le concept de *réplication* couvre deux sens qui sont complémentaires (Edmonds et Hales, 2003 ; Amblard *et al.*, 2006 ; Wilensky et Rand, 2007 ; Bommel, 2009 ; Banos *et al.*, 2015).

La première définition est dite « *indépendante du simulateur* » (Bommel, 2009, p. 15), c'est-à-dire que le même modèle est réécrit, voire ré-implémenté (Edmonds et Hales, 2003), en un autre langage de programmation que celui initial (par exemple, en GAMA puis en Netlogo). Dans ce sens, la réplication vise à se prémunir des biais qui sont inhérents à une plate-forme de SMA mobilisée, plutôt qu'au modèle construit. Sous cette acception, la réplication vise à avoir un modèle lisible à travers un modèle conceptuel complet, « *c'est-à-dire suffisamment explicite pour pouvoir être compris,*

[127] D'après l'article de Bruce Edmonds et David Hales (2003) intitulé : Replication, Replication and Replication: Some Hard Lessons from Model Alignment, *Journal of Artificial Societies et Social Simulation 6(4).*

*discuté et critiqué par tout un chacun* » (Bommel, 2009, p. 15). Cette première définition vise avant tout à vérifier le fonctionnement d'une plate-forme de simulation (Bommel, 2009).

C'est pourquoi, pour ce travail, nous nous appuierons plutôt sur la seconde définition pour illustrer le travail mis en œuvre au cours de l'étape 1-bis associée au plan d'expérience (*cf.* figure 2.32, ci-dessus). C'est un processus qui par opposition à la première définition, est *interne au simulateur*. Sous sa deuxième acception, la *réplication* est synonyme de *répétition*, dans le sens où elle vise à répéter et paramétrer un certain nombre de simulations (*réplication*) pour chaque configuration d'un plan d'expérience. Sous cet angle, la réplication vise à reproduire de façon indépendante *N* fois un même paramètre en utilisant la même plate-forme tout en en changeant la graine (*informatique*) de manière aléatoire, pour obtenir un nombre de simulations aléatoires suffisant (Wilensky et Rand, 2007). En d'autres termes, la réplication vise à observer la reproductibilité des expériences menées en s'assurant que les données de simulation ne présentent pas un caractère aléatoire trop important. En effet, « *pour des valeurs identiques des paramètres, les variables observées peuvent avoir des valeurs différentes (du fait de la stochasticité présente dans le modèle), et exécuter plusieurs fois la simulation pour chaque combinaison de valeurs choisie permet d'obtenir des résultats plus robustes* » (Banos *et al.*, 2015, p. 134). En d'autres termes, la *réplication* vise ici à tester la stochasticité d'un modèle de SMA. Ainsi, il est nécessaire de trouver un juste compromis entre un nombre de réplications maximal et un temps de calcul minimal à mettre en œuvre pour l'exploration d'un modèle, en vue de distinguer les biais qui relèvent plutôt de l'aléatoire de ceux qui relèvent du modèle en soi. Pour illustrer nos propos, la figure 2.33 représente les valeurs de débit simulées par l'exploration pas-à-pas de 1 à 1000 réplications (en abscisse) pour un paramètre donné[128] :

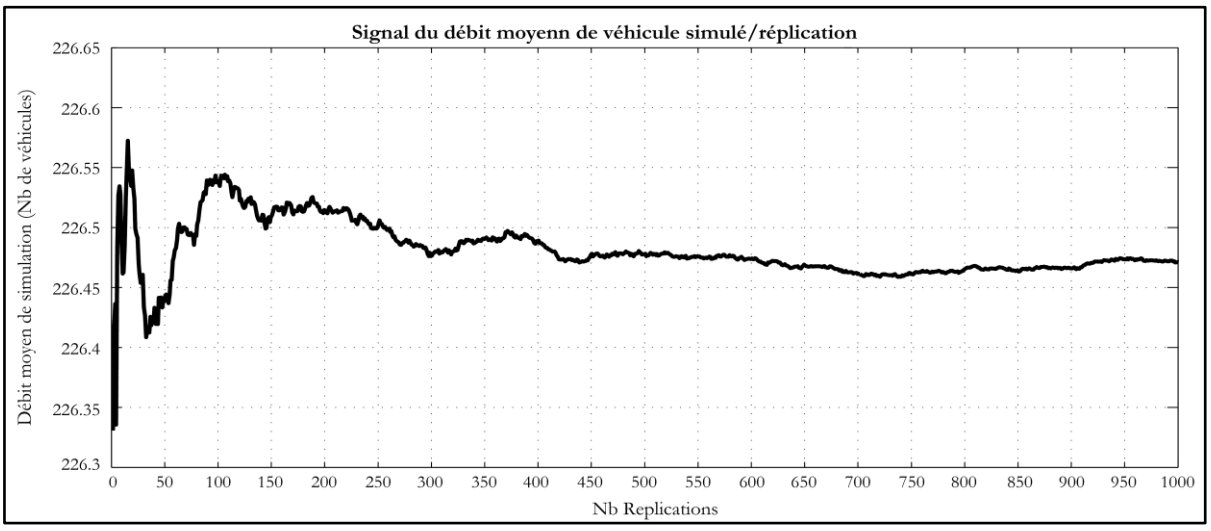



**Figure 2.33 : Signature de la réplication de SCAUP pour un paramètre de l'approche KISS (1000 réplications pas-à-pas pour un débit moyen de véhicules simulés)**

Chaque valeur en abscisse représente un membre indépendant de N réplications, l'ensemble des valeurs est ensuite calculé en observant la moyenne des véhicules simulés et en la rapportant au nombre de réplications réalisées. Sous cette forme, on observe qu'à partir de 100 réplications les

[128] Nous avons sélectionné le paramètre de déplacement des véhicules sans contraintes pour un stationnement de 30 minutes. À noter, que la sélection d'un paramètre quelconque donnerait des tendances relativement similaires à la figure 2.33 et à la figure 2.34.

valeurs de débit simulé entre chaque membre vont présenter un bruit moins important (*cf.* figure 2.33). Au-delà, ce signal a d'ailleurs tendance à s'atténuer et à se lisser entre chaque membre à partir de 300 réplications. Ainsi, plus on réplique un même paramètre, plus la signature stochastique des données de simulation entre chaque membre va s'atténuer. Et donc, a fortiori, les données de trafic routier présenteront un caractère qui est bien moins aléatoire à partir de 100 réplications. Par conséquent, on constate ci-dessus qu'au-delà de 100 réplications le modèle SCAUP est relativement stable. La figure 2.34 conforte ces observations, en représentant la variance des données de simulation entre chaque membre répliqué. On observe alors une tendance plus nuancée que précédemment. En effet, si la variance des différents membres augmente sensiblement au départ, celle-ci a plutôt tendance à augmenter moins rapidement, voire à se stabiliser à une variance de près de 11 véhicules après 30 réplications (variance de 10,4). En d'autres termes, lorsque le nombre de réplications est supérieur à 30 réplications, on observe que la variance intergroupe se stabilise, puis que celle-ci n'évolue que très faiblement à mesure que le nombre de réplications augmente :

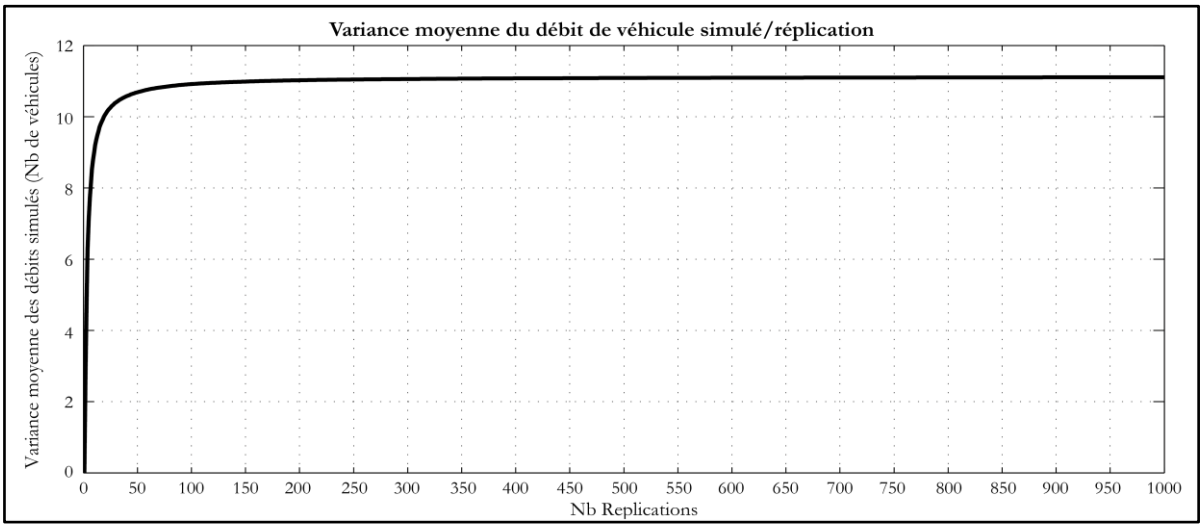



**Figure 2.34 : Analyse de la réplication de SCAUP par le signal de variance moyen de débit de véhicules simulés (1000 réplications)**

Ainsi, en nous appuyant sur ces deux dernières illustrations, nous avons souhaité mettre en avant l'importance de la *réplication interne* d'un modèle de SMA lors de la mise en place d'un plan d'expérimentation. La mise en œuvre systématique de ce travail, réalisé parallèlement, est un prérequis indispensable en vue de bien appréhender les biais qui sont inhérents à la stochasticité d'un modèle de SMA (Amblard *et al.*, 2006 ; Banos *et al.*, 2015). Dans ce sens, on préserve aussi l'interprétation des résultats en évitant d'y associer des biais inhérents au système multi-agents. C'est pourquoi, nous avons considéré un minimum de 30 réplications lors de la mise en œuvre du plan d'expérience.

Pour finir, il faut relever que l'application de cette procédure nécessite un temps de formalisation et de calcul non négligeable, qui peut s'avérer très vite chronophage. Toutefois, ces approches se démocratisent actuellement grâce au développement croissant d'outils informatiques plus performants. Etienne Delay (2015, p. 106) illustre ces propos : « *si aujourd'hui il est possible sur des ordinateurs de bureau de paralléliser 8 à 10 simulations, il n'est pas loin le temps où nos processeurs n'étaient dotés que d'un seul cœur. Parallèlement le nombre d'équipements de calcul intensif s'est largement démultiplié ces*

*dernières années, que ce soit sur des clusters universitaires, ou les VO (Virtual Organization) ; qui lancent de manière coopérative des calculs plus lourds* ». Ainsi, la réalisation d'un plan d'expérience plus lourd est de plus en plus aisé, autant au niveau individuel que collectif. Ces développements se sont aussi accompagnés de la création d'outils spécifiques permettant de faciliter la mise en œuvre et la réalisation des plans d'expérience, ainsi que de leurs réplications. C'est particulièrement le cas d'un outil comme OpenMole (Reuillon *et al.*, 2013), ou d'outils intégrés aux plates-formes de SMA, comme le B*ehaviorSpace* de Netlogo (Banos *et al.*, 2015).

### 6.3.3. **Calibrage et ajustements des axes équipés de capteurs : contrôler les flux de trafic entrants et sortants du réseau routier**

Les premiers tests de simulation ont eu pour objectif de calibrer les axes équipés de stations de mesures avec les données qui leur sont associées. Cette analyse, réalisée parallèlement, vise à calculer les indices de reproductibilité du modèle sans paramétrage pour les 30 réplications (*cf.* figure 2.35). La synthèse de ces premiers résultats montre que le modèle de simulation du trafic routier présente des caractéristiques intéressantes avec une reproductibilité moyenne de 95,3 % sur le plan spatial (en haut) et de 94,6 % sur le plan temporel (en bas). De surcroît, en s'appuyant sur la figure 2.36 (p. 212), on peut observer que la dynamique de la circulation est correctement simulée en reproduisant la courbe intra-journalière du trafic routier.

De manière plus détaillée, à travers la figure 2.35, on observe que le modèle présente une reproductibilité que l'on peut qualifier de stable, autant sur le plan spatial (indice allant de 99 %[129] à 100 %) que temporel (de 80 % à 100 %). Spatialement, l'indice montre que la fréquentation des axes routiers reste relativement similaire (en termes de variance) entre les différentes réplications. Dans ce sens, les véhicules simulés sur les axes routiers entre chaque membre présentent des valeurs de fréquentation qui sont reproduites dans plus de 99 % des simulations répliquées. Ainsi, le caractère stochastique apparait prendre peu de place dans le processus de simulation, entre les différentes réplications (*cf.* section 6.3.2, p. 207). Malgré tout, on observe que les valeurs de reproductibilité se répartissent de manière hétérogène sur les axes routiers, sans présenter de logique spatiale. Ce constat est à nuancer dans le sens ou les valeurs inférieures à 99 % (en rouge) présentent un taux de reproductibilité qui demeure de bonne qualité, en étant toujours supérieur à 97 % de reproductibilité.

Sur le plan temporel, on peut voir que plus le temps de simulation avance, plus l'indice de reproductibilité s'approche de 100 %. On observe cependant deux périodes distinctes : la première correspond à l'initialisation et au lancement de la simulation ; alors que la seconde représente les périodes de fréquentation. L'augmentation de l'indice de reproductibilité tend à illustrer le processus d'accumulation progressif des véhicules lors du processus de simulation. Pour ces deux périodes, on observe un différentiel entre la nuit (de 20 h à 6 h, avec une faible fréquentation) et la journée (de 6 h à 20 h, avec une plus forte fréquentation), ce qui fait naturellement diminuer l'indice de reproductibilité en termes de variance entre ces deux moments de la simulation (*cf.* en bas de la figure 2.36). Ce dernier résultat montre que les conditions initiales de simulation divergent jusqu'à ce que le trafic routier augmente en début de journée ; ensuite,

[129] La valeur minimale est de 97% pour l'ensemble des paramètres testés, ce qui constitue pour cette dernière classe un indice de reproductibilité de bonne qualité.

l'ensemble des simulations apparait stabilisé en termes de fréquentation temporelle, illustrant la capacité du modèle à reproduire la dynamique intra-journalière du trafic routier. De surcroît, l'indice de reproductibilité permettant d'observer l'amplitude des courbes intra-journalières entre chaque membre, on peut en conclure que pour toutes les réplications le phasage des courbes matin, midi et soir est appréhendé d'une façon similaire. Ce dernier constat est à relier à la génération des véhicules. En effet, les conditions d'initialisation et de simulation du modèle étant reliées à la dynamique des comptages, les données de simulation suivent forcement la forme générale de la courbe intra-journalière (*cf.* figure 2.36). On peut alors en conclure qu'au niveau temporel le modèle conserve et reproduit correctement le signal issu des comptages routiers lors du processus de simulation :

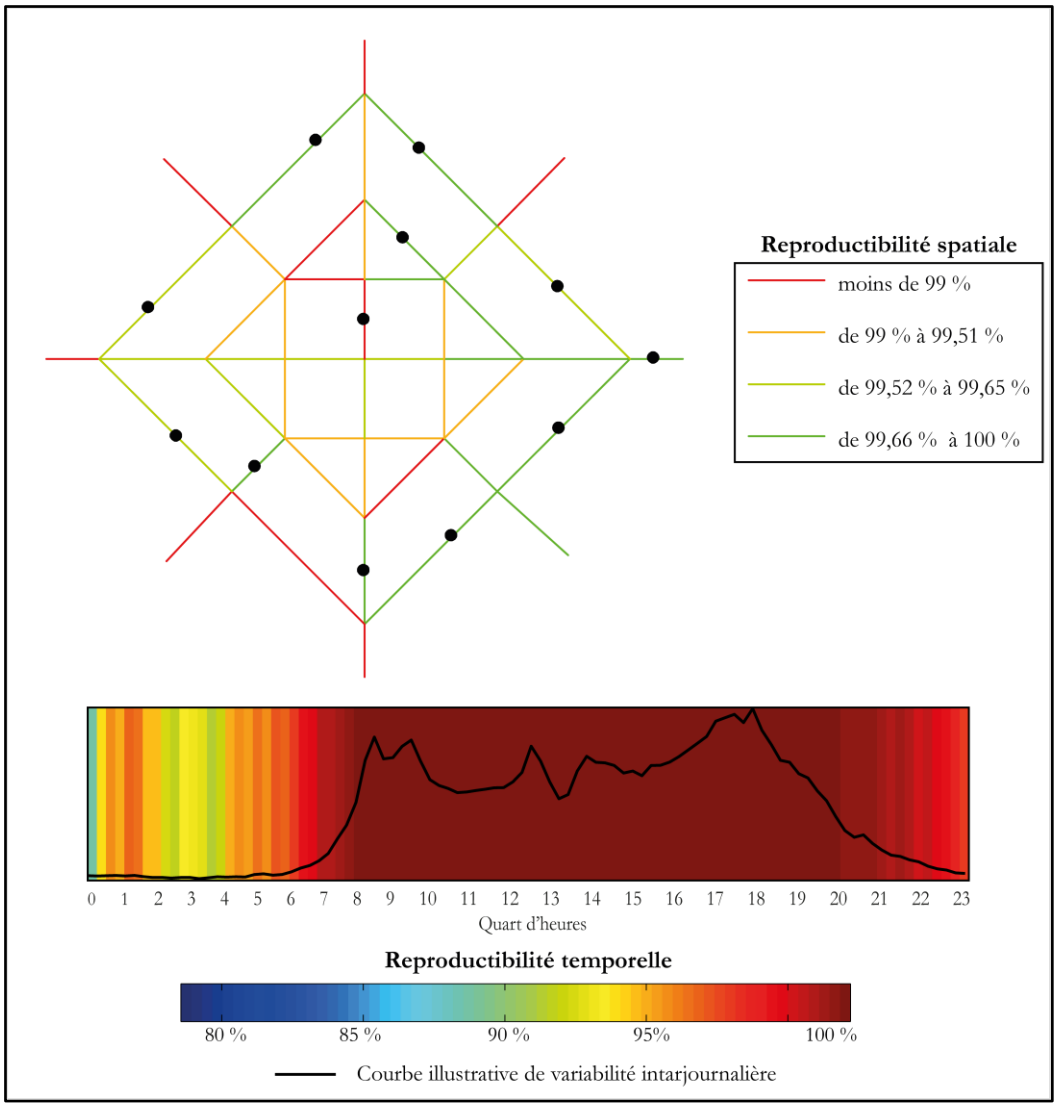



**Figure 2.35 : Indice de reproductibilité moyen (30 réplications) spatial (en haut) et temporel (en bas) du modèle de simulation sans paramétrage, ni contraintes de déplacement**

Malgré tout, si la stabilité du modèle de simulation est relativement bonne (proche de 100 %), les premiers tests de simulation (sans aucun paramétrage) ont eu tendance à montrer que le nombre de véhicules simulés sans paramètre est 5 fois plus important que le nombre de véhicules compté sur un axe équipé d'un capteur fixe (illustré par la courbe noire en figure 2.36). Ce premier

constat montre qu'il est essentiel de mieux contraindre le nombre de véhicules entrants à l'échelle des tronçons équipés de ces capteurs. Il s'agit en effet de conditions de départ essentielles afin de forcer le modèle à respecter les valeurs de comptage. À ce niveau, nous considérons que le calibrage des axes à la mesure est essentiel pour l'initialisation ainsi que pour l'insertion des véhicules routiers sur un plus grand ensemble d'axes routiers.

Ainsi le calibrage des axes routiers s'est avéré nécessaire et a été réalisé pas à pas en analysant la capacité du modèle à conserver la mesure exacte associée à un axe équipé (illustrée en tirets bleus la figure 2.36). Chacune des courbes illustre l'introduction d'un nouveau paramètre lors du processus de simulation : en noir, la simulation est réalisée sans hypothèses et sans contraintes sur le déplacement des véhicules ; en vert, les agents « miroir » (collision) de même que la « disparition » des véhicules sont introduits ; en rouge, l'insertion du temps de « stationnement » des véhicules[130].

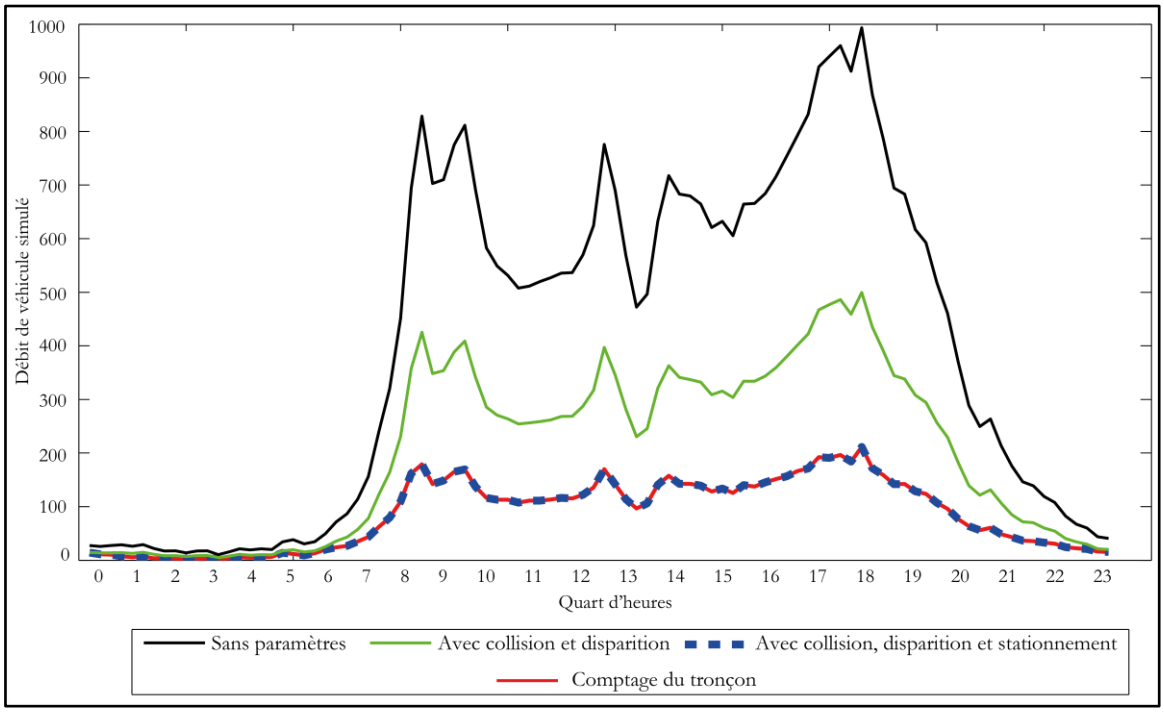



**Figure 2.36 : Calibrage des axes possédant des stations de mesure (moyenne des 15 axes équipés de stations)**

En s'appuyant sur la figure 2.36, on peut observer que les axes mesurés ont été calibrés en ajustant les valeurs de débit simulé à leurs valeurs de référence (les comptages). L'insertion des « miroirs » et de la « disparition » permet non seulement de diminuer de deux fois le débit de véhicules simulé en maîtrisant le nombre de véhicules entrant sur un tronçon (les « miroirs »), mais aussi ceux sortant du réseau routier, en permettant au véhicule de sortir du système routier par sa disparition en marge du réseau routier.

Enfin, l'introduction du temps de stationnement, dans un dernier temps, permet de diminuer de nouveau de moitié le débit de trafic entrant sur le réseau routier. La combinaison de ces trois paramètres permet d'ajuster les valeurs de trafic simulées sur les valeurs de trafic observées à

[130] Le paramètre par défaut est défini à 30 minutes de déplacement pour 24 h de simulation (*cf.* section 6.2.2.c, p. 193).

travers le dispositif de quantification. Ainsi, pour chaque axe équipé de capteur de trafic, le degré de liberté du modèle de simulation est contraint, en laissant peu de place au modèle, pour y introduire des incertitudes à l'échelle de ces tronçons lors de la simulation. On retrouve bien ici les grandes lignes de l'abduction, où chaque test a permis d'observer de nouvelles tendances, pour ensuite arriver à un calibrage adéquat des axes équipés de capteurs, illustré par la courbe en pointillés bleus (*cf.* figure 2.36).

### 6.3.4. **L'impact de l'insertion des choix routiers : hiérarchiser les flux de trafic routier**

Le choix des routes lors du processus de simulation de SCAUP vise à hiérarchiser la fréquentation des voiries en fonction de leurs caractéristiques circulatoires. Pour observer l'introduction de ce paramètre, deux choix routiers ont été établis : un premier dit *aléatoire*, où le mouvement et la direction des véhicules ne sont pas contraints ; un second dit *hiérarchique*, où les véhicules choisissent une nouvelle route en fonction de la matrice de transition préalablement définie (*cf.* section 6.2.2.b, p. 198). Par souci de comparabilité, les deux tests ont été réalisés avec les mêmes conditions initiales, en insérant les paramètres de collision, de disparition et de stationnement (défini par défaut à 30 minutes). La synthèse de ces tests de simulation par typologie est présentée en figure 2.37 (ci-dessous), où chaque courbe correspond au débit simulé pendant 24 h :

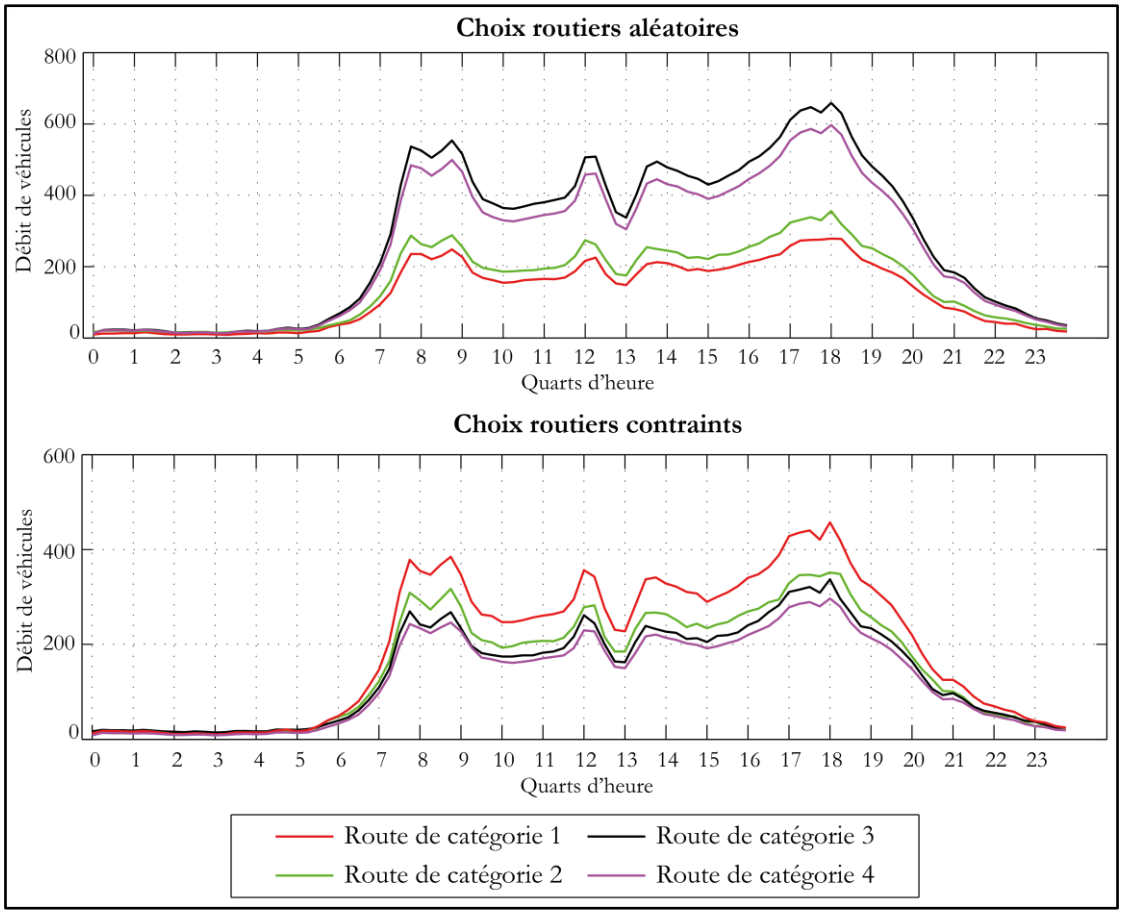



**Figure 2.37 : Simulation du trafic routier sans (en haut) et avec (en bas) contraintes des choix routiers (moyenne des débits simulés pour les 30 réplications) : hiérarchiser la fréquentation des axes routiers**

Le premier test avec choix aléatoire des routes montre que la fréquentation des routes diffère entre les routes des plus faibles catégories et celles qui concentrent le plus de trafic. En somme, on ne retrouve que peu de cohérence entre les axes devant concentrer, d'une part, le plus de véhicules routiers (*catégories 1 et 2*), et ceux devant concentrer, d'autre part, le moins de véhicules (*catégories 3 et 4*). On observe d'ailleurs en figure 2.37 (en haut) un résultat contre-intuitif pour le premier test de simulation, où les axes de plus faible catégorie (3 et 4) concentrent trois fois plus de véhicules que les autres typologies. L'insertion des choix routiers illustrée en bas de la figure 2.37 montre que la contrainte liée aux préférences des véhicules permet de hiérarchiser les axes fréquentés par les véhicules, en allant des axes de plus forte catégorie vers ceux de plus faible catégorie. De surcroît, le nombre de véhicules fréquentant les axes est plus homogène que lors du test précédent.

Par conséquent, le modèle SCAUP apparait sensible à l'insertion de la matrice de transition (*cf.* section 6.2.2.b, p. 198) dans le processus de simulation les véhicules ne se déplacent plus aléatoirement et leurs mouvements sont maitrisés par la définition des routes à privilégier. Comme illustré ci-dessus, les débits simulés respectent plus rigoureusement la typologie circulatoire du réseau routier, dans le sens, où le débit simulé croît sur les axes possédant les caractéristiques les plus importantes par rapport à ceux qui ont les potentiels circulatoires les plus faibles (catégories 3 et 4). La figure 2.37 permet d'observer la sensibilité de SCAUP lors de l'insertion des choix routiers à travers la matrice de transition. Les contraintes associées aux déplacements des véhicules permettent de mieux appréhender leur destination. Malgré tout, l'affectation des choix routiers (nœud à nœud) augmente sensiblement le temps de simulation passant de 3 minutes à 6 minutes sur le réseau routier théorique, soit 100 % de temps de simulation supplémentaire. En effet, à chacun des nœuds, chaque véhicule interroge d'une part les routes qui sont à sa disposition, et d'autre part la matrice de transition pour choisir la route qui lui est la plus adéquate ; c'est pourquoi le temps de simulation augmente parallèlement à l'introduction de la matrice de transition.

### 6.3.5. **Exploration du temps de stationnement : comment gérer le nombre de véhicules présents lors du processus de simulation ?**

Enfin le dernier paramètre caractérisant le modèle SCAUP nous amène à nous focaliser sur le temps de stationnement des véhicules lors du processus de simulation (*cf.* 6.2.2.c., p. 201). Défini par défaut à 30 minutes, il apparait toutefois important d'en étudier l'impact sur le trafic routier qui en résulte par la suite. C'est pourquoi un dernier test sur le modèle de simulation a été mis en place afin d'explorer pas-à-pas le temps de stationnement, en le faisant progressivement varier par pas de 5 minutes, pour trois conditions initiales de simulation illustrées ci-dessous (en bleu, sans paramétrage lié aux déplacements des véhicules[131] ; en rouge, avec paramétrages et déplacement aléatoire des véhicules ; en vert, en intégrant tous les paramètres de déplacement des véhicules ainsi qu'en contraignant le déplacement des véhicules sur la hiérarchie du réseau routier). Si le nombre de véhicules simulé augmente sensiblement à mesure que le paramètre de stationnement évolue (en abscisse de la figure 2.38), on observe qu'à partir de 30 minutes le débit de véhicules simulé n'augmente plus aussi sensiblement, et a même tendance à plafonner.

[131] C'est-à-dire sans l'insertion des paramètres de collision, disparition et stationnement ; tout en intégrant un mouvement totalement aléatoire des véhicules sur les tronçons routiers.

On observe d'ailleurs que lorsque le temps de stationnement des véhicules n'est pas établi (*illimité* en figure 2.38), le débit simulé est proche de celui simulé à 30 minutes de stationnement. Ce résultat est assez surprenant, mais illustre bien le caractère abductif de la démarche initiée, où les résultats que nous trouvons par simulation sont parfois contre-intuitifs et surprenants par rapport à ce à quoi nous en attendions au départ. Dès lors, le paramètre de stationnement a un impact sur le nombre de véhicules simulés, jusqu'à un certain moment où son apport apparait relativement limité. Toutefois, notons ici que le réseau routier sur lequel les véhicules se déplacent est de taille réduite (48 axes), ce qui peut aussi avoir une conséquence lors du processus de déplacement et de disparition des véhicules positionnés en marge du réseau routier. À titre indicatif, pour tous les paramètres testés, le taux d'évolution du débit moyen simulé entre un stationnement défini à 15 minutes et un stationnement de 30 minutes est de l'ordre de 20 % ; alors que le taux d'évolution du débit moyen simulé entre un stationnement défini à 30 minutes et un de 45 minutes est de l'ordre de 5 % quel que soit le paramètre testé. Par conséquent, plus le temps de simulation augmente plus le taux d'évolution du débit diminue.

Ce premier état des lieux permet d'une part de justifier notre choix de considérer un temps de stationnement par défaut de 30 minutes (*cf.* section 6.2.2.c., p. 201). En outre, cela nous permet d'envisager plus aisément l'exercice de simulation de SCAUP sur la ville de Dijon (KIDS) qui sera défini à la suite. En effet, ces résultats tendent à nous montrer que la définition d'un temps de stationnement supérieur à 30 minutes n'améliorera que très peu les résultats. Il apparait alors plus intéressant pour la suite de se focaliser sur des temps de simulation inférieurs à cette borne.

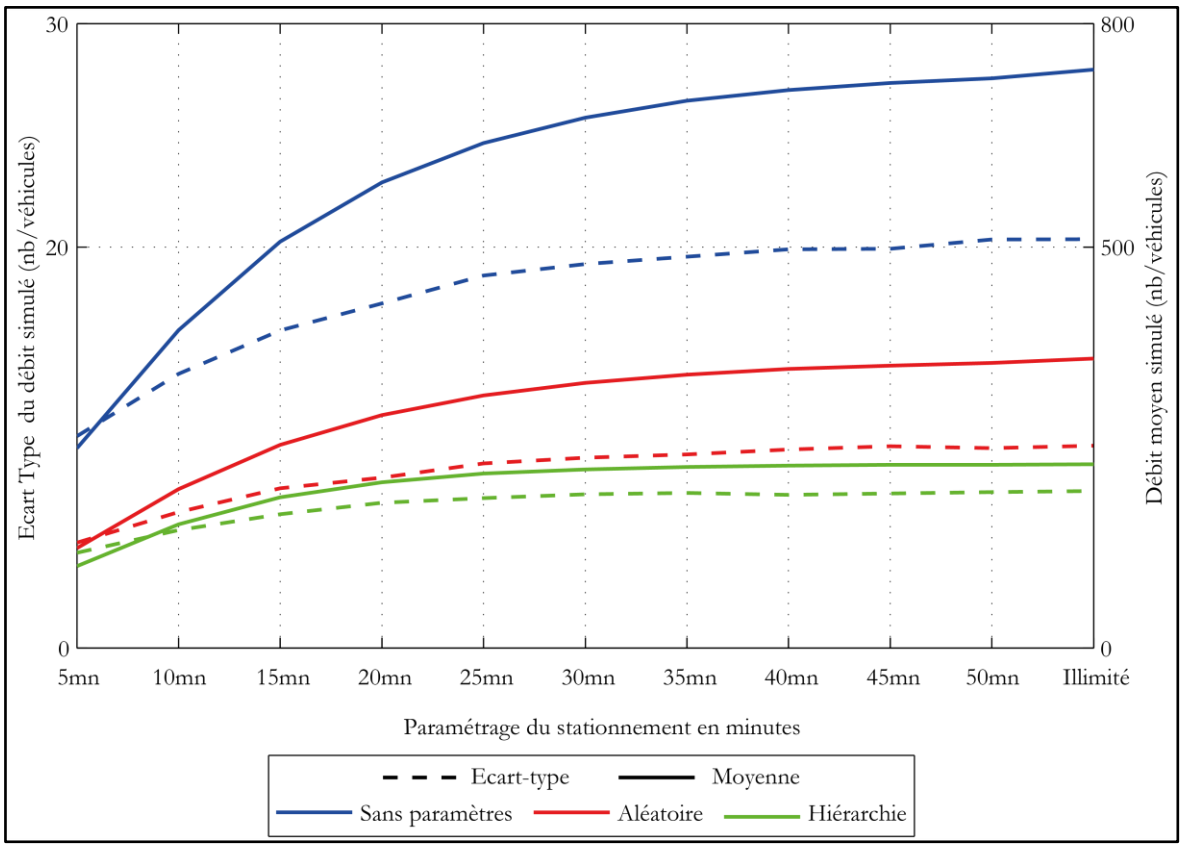



**Figure 2.38 : Variation du temps de stationnement (en abscisse) pour trois conditions initiales de simulations (en ordonnées, à droite, est présenté l'écart-type moyen, ainsi que la moyenne des débits simulés, à gauche, pour les 30 réplications)**

Enfin, la figure 2.38 apporte une seconde information relative au paramètre des choix routiers. En effet, si on observe la courbe de choix hiérarchique (en vert) par rapport aux deux autres conditions initiales de simulation, on peut voir que l'écart-type reste très faible par rapport aux deux autres conditions initiales de simulation. Le débit moyen de véhicules simulés est également deux fois moins important que lors d'un mouvement totalement aléatoire des véhicules. Ce dernier résultat tend à montrer que les débits de trafic routier se répartissent plus uniformément sur l'ensemble des tronçons que lors des deux autres conditions initiales de déplacement des véhicules. Dans ce sens, le modèle montre de nouveau sa sensibilité à l'insertion de la matrice de transition, qui permet de mieux appréhender la répartition spatiale des véhicules sur le réseau routier.

Pour conclure, ces trois dernières sections ont permis d'illustrer les principales fonctionnalités du simulateur multi-agents de trafic routier SCAUP. En s'appuyant sur un environnement de simulation simplifié (KISS), cette première phase d'expérimentation a permis d'explorer et d'exposer l'approche de simulation guidée par les données de comptages routiers. En effet, les mesures sont mises en avant à chaque étape du processus et jouent un rôle primordial pour la mise en œuvre de notre approche ; non seulement, pour l'initialisation et la génération des véhicules sur le réseau routier par la relation de dépendance entre les capteurs et les données qui leur sont associées ; mais aussi lors de la définition des préférences de choix routiers à travers la mise en œuvre de la matrice de transition. En effet, pour cette dernière, nous avons exploité l'ensemble des valeurs de comptage recueillies (ponctuelles et permanentes) afin de définir la fréquentation « type » des tronçons en fonction de leurs caractéristiques (*cf.* tableau 2.4, p. 200). Par ailleurs, ces données ont aussi été utilisées pour le calibrage des tronçons routiers équipés de capteur, en vue d'obtenir des conditions d'observation adéquates aux valeurs de terrain relevées sur ces différents points du réseau routier.

## CONCLUSION DE LA DEUXIÈME PARTIE

Cette seconde partie a permis d'explorer les enjeux liés au passage de la donnée d'observation à la donnée simulée à travers le « *dispositif de modélisation* » (Commenges, 2013). En nous focalisant sur la matrice technique et les relations que celle-ci entretient avec le dispositif de quantification et le dispositif d'évaluation, nous avons mis en avant les enjeux d'une approche guidée par les données. Ce glissement de l'observation vers la simulation nécessite de repenser les outils de simulation du trafic routier pour être en adéquation avec les objectifs énoncés au cours de la première partie. En effet, la modification de l'un des dispositifs modifie l'ensemble de la matrice technique ; ainsi, mobiliser les données issues des capteurs au détriment des enquêtes O-D, de même que s'inscrire dans une optique d'évaluation de la PAA, nous incite à repenser l'ensemble des dispositifs associés à la matrice technique.

De ce point de vue, nous proposons de nous rapprocher, via une approche *bottom-up*, du véhicule en tant que tel pour reproduire la circulation urbaine, en faisant appel à la Simulation Multi-agents (SMA) plutôt qu'en se basant sur les approches plus habituellement utilisées dans les études d'impact du trafic routier. En effet, la SMA apparait appropriée à la démarche proposée, parce qu'elle se prête très bien à la représentation des mobiles, en étant plus intimement liée aux SIG et aux données (Marilleau, 2006, Taillandier et al, 2014). Ainsi, il apparait intéressant de mobiliser les données issues des capteurs électromagnétiques en s'appuyant sur les outils de SMA pour la simulation du trafic routier.

Cette partie a, en outre, permis de mettre en avant les enjeux liés au positionnement des données de capteurs pour la modélisation du trafic routier. Cette première lecture nous a permis d'aborder les limites de l'approche *classique* non seulement en vue de faire ressortir les problématiques liées aux impacts environnementaux du trafic routier, mais aussi dans sa capacité à se rapprocher des données pour l'initialisation de la démarche de simulation. La mise en place d'un modèle est une démarche de construction de la connaissance qui se construit en plusieurs étapes pour convenir, conceptualiser, et simuler les différents éléments d'un système. Cette grille de lecture a permis de construire une approche expérimentale de simulation du trafic routier, à partir des comptages routiers, en les positionnant au centre du processus de modélisation, afin de dépasser les limites liées à la constitution d'une matrice origine-destination en proposant une approche de simulation expérimentale, à travers le modèle développé au cours de ce travail : SCAUP (*Simulation multi-agents à partir de CApteurs Urbains pour la Pollution atmosphérique automobile*). Au sein de SCAUP, les données de comptages routiers interviennent à chaque étape du processus de modélisation, que ce soit pour la construction du dispositif de quantification de celui de modélisation, ou de celui d'évaluation.

De plus, ce positionnement nous inscrit bien au sein des outils de la *géocomputation*, permettant alors de disposer d'un ensemble d'instruments de recherche conçu pour passer de la donnée d'observation en différents points de l'espace à la donnée simulée sur l'ensemble d'un réseau routier. Les approches génératives, au premier rang desquels la SMA se positionne, présentent un potentiel d'autant plus intéressant, pour non seulement se rapprocher du véhicule (spatialement), mais aussi en vue d'en construire un modèle tourné vers les données (*data driven*) dans une optique de SMA : l'approche *deepening* (Hassan *et al.*, 2010a, 2010b). La mise en place du modèle

SCAUP suit un raisonnement abductif dans le sens où l'exploration des différentes sorties nous a permis de mettre en place progressivement les différents paramètres à travers une première application sur un trafic routier virtuel. À travers l'application d'un environnement de simulation KISS, cette première phase expérimentale a permis de mettre en lumière l'ensemble des éléments liés autant à la structure qu'à la dynamique de SCAUP. Au cours du dernier chapitre, nous avons abordé le processus à la fois expérimental et exploratoire de celui-ci. En décomposant les données issues des capteurs, l'approche présentée permet d'envisager la production d'un trafic routier émergent par l'intermédiaire d'un modèle de simulation multi-agents. L'application de SCAUP vise à disposer de données de simulation pour une journée complète de comptage (24 h), à une granularité fine (la minute). Pour cela, nous nous sommes focalisé sur une vision simplifiée du système de transport, en nous inscrivant dans un seul des sous-systèmes urbains (Masson, 2000) : les déplacements. Cette formulation simple nous a permis de construire un système de trafic routier autour des deux composants élémentaires qui définissent le trafic routier : les véhicules et la route.

Le processus exploratoire et expérimental que nous proposons à travers le modèle SCAUP s'inscrit dans une optique opérationnelle sur la ville de Dijon, pour ensuite alimenter le logiciel d'estimation de la pollution atmosphérique (CIRCUL'AIR) exploité par l'AASQA locale bourguignonne (ATMOSF'AIR BOURGOGNE). Cependant, avant d'en aborder les caractéristiques méthodologiques, une question de fond est soulevée quant à notre capacité à valider un modèle initialisé à partir des données qui vont ensuite être aussi utilisées pour la validation du contexte dijonnais. On se retrouve confronté à une approche hermétique, où les données vont avoir un lien non seulement avec les *inputs* mais aussi avec les *outputs* du modèle de simulation SCAUP. Dès lors, il faut être en mesure de proposer une démarche de validation de celui-ci, en vue d'être en adéquation avec cette contrainte et avec le positionnement adopté d'une simulation guidée par les données.

La troisième partie de ce travail poursuit cette réflexion avec un passage au contexte local de la ville de Dijon, en vue d'expérimenter et de valider le modèle en proposant une démarche originale de validation des données de simulation. Les enjeux de la simulation du trafic routier issue de SCAUP se situent en amont du processus d'estimation de la PAA. En effet, avant d'être en mesure de fournir des valeurs relatives à la PAA, il faut déjà être en mesure de quantifier le nombre de véhicules routiers présents au sein du réseau routier. Nous aborderons ensuite les éléments relatifs au passage des données de simulation de SCAUP vers les données environnementales d'émissions atmosphériques dues au trafic routier en milieu urbain.

# PARTIE 3 -

# DU TRAFIC ROUTIER À LA POLLUTION ATMOSPHÉRIQUE AUTOMOBILE : VALIDATION, ÉVALUATION ET PERSPECTIVES


INTRODUCTION DE LA TROISIÈME PARTIE ..... 221

CHAPITRE 7. SIMULATION DU TRAFIC ROUTIER SUR LA VILLE DE DIJON : DE LA SIMULATION À L'ÉVALUATION ..... 223

*7.1. Exploration du modèle SCAUP à travers le cas de la ville de Dijon ..... 225*

*7.2. Analyse préliminaire de l'approche expérimentale de simulation du trafic routier ..... 235*

*7.3. Comment choisir la meilleure simulation : de la simulation à la réalité ..... 246*

*Conclusion : ..... 260*

CHAPITRE 8. UN DISPOSITIF D'ÉVALUATION ENVIRONNEMENTALE : DU TRAFIC ROUTIER AUX ÉMISSIONS DE POLLUANTS ATMOSPHÉRIQUES AUTOMOBILES ..... 263

*8.1. Les enjeux de la spatialisation des émissions de polluants atmosphériques automobiles ..... 265*

*8.2. Données d'entrée et méthodologie de calcul des émissions avec CIRCUL'AIR ..... 271*

*8.3. Spatialisation de la pollution atmosphérique automobile sur la ville de Dijon ..... 281*

*Conclusion : ..... 298*

CHAPITRE 9. LES PERSPECTIVES DE DÉVELOPPEMENT AUTOUR DE LA MATRICE TECHNIQUE : DE LA QUANTIFICATION À L'ÉVALUATION ..... 301

*9.1. Les réflexions autour du dispositif de quantification pour alimenter et façonner le dispositif de modélisation ..... 303*

*9.2. Une approche de modélisation SCAUP à affiner ..... 311*

*9.3. S'inscrire dans une démarche intégrée d'évaluation de la PAA en milieu urbain : de l'émission à l'imission ..... 320*

CONCLUSION DE LA TROISIÈME PARTIE ..... 333

# INTRODUCTION DE LA TROISIÈME PARTIE

La visualisation des données observées, calculées ou simulées est un terrain d'étude largement utilisé pour l'analyse et l'interprétation des phénomènes urbains et de leurs impacts environnementaux. Ses domaines d'applications sont multiples allant des besoins pour caractériser les phénomènes climatiques en milieu urbain, les impacts de la pollution urbaine, l'ilot de chaleur urbain, aux déplacements des individus dans l'espace (CGDD, 2012). Pour ce travail, les enjeux liés aux impacts du trafic routier sur la qualité de l'air à travers l'étude de la PAA constituent un enjeu autant social et politique qu'environnemental (OMS, 2006 ; AEE, 2011 ; Primequal-Predit, 2012). En effet, la mise en place des inventaires territoriaux de la pollution atmosphérique s'inscrit aujourd'hui dans ce cadre, visant non seulement à mieux identifier mais aussi à mieux gérer et contrôler les émissions atmosphériques à l'échelle des territoires. Dès lors disposer de données issues de modèle de simulation permet de fournir un ensemble d'informations pour caractériser le trafic routier et ses impacts environnementaux.

Après nous être intéressé au dispositif de quantification au cours de la première partie (pp. 17-118), dont les caractères intrinsèques positionnent ce dispositif en amont du dispositif de modélisation que nous avons abordé au cours de la deuxième partie (pp. 119-218), cette dernière partie vise à s'intéresser à l'aval du dispositif de modélisation, en s'intéressant au dispositif d'évaluation qui lui se situe en amont du dispositif de modélisation. En d'autres termes, après s'être intéressé aux *IN,* puis à la démarche de modélisation SCAUP, nous nous focaliserons ici aux *OUT* qui caractérisent le dispositif d'évaluation. Cependant, sous cette terminologie, la notion d'évaluation sera employée dans un sens plus général que la définition donnée par Hadrien Commenges (2013). C'est pourquoi au cours de cette partie nous nous focaliserons non seulement sur l'évaluation des capacités du modèle SCAUP à reproduire un trafic routier émergent, mais aussi à mettre en avant la démarche de collaboration avec ATMOSF'AIR BOURGOGNE, en vue de passer d'une information sur la circulation routière en milieu urbain à une estimation des émissions de pollution atmosphérique associées à ce secteur : la pollution atmosphérique automobile (PAA).

Les questionnements liés à l'évaluation d'une approche guidée par les données nécessitent de s'arrêter sur la manière avec laquelle la relation entre les capteurs et le modèle a été entreprise au sein du modèle SCAUP. La force temporelle des capteurs électromagnétiques ainsi que leur positionnement spatial fournissent des valeurs de référence en différents points du terrain d'étude pour évaluer et valider les données de trafic routier simulé. En effet, les interrogations sur la validation et l'évaluation des SMA guidées par les données demeurent un questionnement de fond (Hassan *et al.*, 2010a, 2010b ; Banos, 2013), et ce d'autant plus lorsque l'on aborde les problématiques des mobilités intraurbaines (Banos et Thévenin, 2008). C'est une partie à laquelle il faut porter une attention particulière pour juger de l'opérationnalité d'un modèle (Bonnafous, 1990 ; Bonnel, 2002). En effet, juger de l'opérationnalité d'un modèle nécessite d'aborder plusieurs questionnements portant autant sur la *pertinence du modèle*, et la *cohérence du modèle*, que sur sa *mesurabilité* (Bonnafous, 1990): Le modèle est-il conforme à ce que l'on croit de la « réalité » : *pertinent* ? Le modèle répond-il aux objectifs ciblés : *cohérent* ? Le modèle est-il mesurable et valable avec des données extérieures : *mesurabilité* ?

C'est pourquoi, selon nous, l'ensemble de ces questionnements liés à l'opérationnalité s'insère dans les problématiques associées au dispositif d'évaluation. Ces questionnements peuvent paraître discutables, cependant un modèle qui est déjà capable de reproduire des situations observées, ou tout du moins en partie, est-il opérationnel pour autant ? Nous supposons donc que pour juger le modèle SCAUP comme opérationnel, il est nécessaire de remplir plusieurs conditions, répondant aux questions suivantes : le modèle reproduit-il des tendances et des mesures de circulation routière observées en différents points du domaine d'étude ? Le modèle reproduit-il la dynamique du trafic routier par la reproduction de la courbe matin/midi/soir ? Le modèle permet-il d'alimenter les outils d'évaluation de la pollution atmosphérique d'origine automobile ?

Notre démarche étant guidée par les données (*data-driven*), elle s'appuie sur une modification du statut des données de comptages routiers, vers une information simulée sur le réseau routier, pour *in fine* constituer une connaissance sur les émissions atmosphérique à l'échelle de la ville de Dijon Cette démarche s'appuie sur les compétences de l'AASQA ATMOSF'AIR BOURGOGNE pour le calcul des émissions de PAA à partir des données de trafic routier simulé issues de SCAUP. L'objectif de cette dernière partie vise aussi à mettre en avant les apports opérationnels de l'approche présentée dans cette thèse. En effet, le passage à la PAA ouvre ce travail vers des applications opérationnelles, en proposant la mise en œuvre d'un modèle simple pour obtenir des informations sur les émissions de polluants en différents points du réseau routier. C'est pourquoi un regard particulier sera apporté à la chaîne de travail mise en place pour alimenter les outils de calcul des émissions de polluants atmosphériques.

Dans cette partie, le terme d'évaluation portera sur les deux dernières étapes du protocole expérimental réalisé à au cours de cette thèse. C'est pourquoi, dans un premier temps, nous nous attacherons à évaluer et à qualifier l'expérimentation de SCAUP réalisée sur la ville de Dijon en se focalisant sur les données de trafic routier simulé. Ces données étant utilisées en entrée du calcul des émissions de PAA, ce n'est que par la suite que nous nous intéresserons à l'évaluation environnementale des émissions de PAA sur la ville de Dijon. Cette dernière étape a été mise en œuvre au cours d'une collaboration avec ATMOSF'AIR BOURGOGNE, qui dispose des compétences et des outils nécessaires aux calculs des émissions de PAA. C'est pourquoi nous nous intéresserons à tous les éléments qui se situent en aval du dispositif de modélisation expérimenté sur la ville de Dijon, c'est-à-dire autant au niveau du trafic routier simulé qu'aux estimations des émissions de polluants d'origine automobile. Dans un dernier temps, nous nous arrêterons sur les perspectives de développement du modèle SCAUP, ainsi que sur les éléments de réflexion de la démarche de modélisation initiée dans ce travail. C'est pourquoi nous reviendrons sur les trois dispositifs de quantification, en remettant en perspective l'ensemble de la démarche de modélisation du trafic routier entreprise.

# Chapitre 7. SIMULATION DU TRAFIC ROUTIER SUR LA VILLE DE DIJON : DE LA SIMULATION À L'ÉVALUATION

Le passage d'un environnement de simulation KISS vers KIDS vise à s'assurer que les données de trafic routier simulées le sont en adéquation avec les données d'observation. Cette partie de « *construction de la preuve* » (Fouillé *et al.*, 2012, p. 199) est importante, car comme le soulèvent Laurent Fouillé *et al.* (2012, p. 199) : « *commenter les résultats d'un travail de modélisation en évitant soigneusement la discussion sur les jeux de données entrantes, et leur condition de production revient à réciter un mantra et se convaincre de sa véracité au seul titre de sa quantification* ». Ici c'est bien la problématique d'*évaluation* du modèle de trafic routier qui est soulevée et qui constitue la dernière phase du processus de modélisation (*cf.* figure 2.1, p. 128).

Il est à noter que le terme d'évaluation englobe non seulement les éléments liés à la validation du modèle, mais aussi ceux liés au fonctionnement interne du modèle sans en aborder sa vérification[132] (Biernat *et al.*, 2015). C'est pourquoi, par la suite, nous engloberons sous la même acception les termes de validation et d'évaluation. En d'autres termes, le concept d'évaluation est plus vaste et ne vise pas seulement à valider, mais aussi à s'assurer de l'*opérationnalité* (Bonnafous, 1990) du modèle de SMA. S'inscrire dans cette optique d'*évaluation* apparait aussi être une réponse à la validation des systèmes multi-agents car comme le soulèvent Frédéric Amblard *et al.* (2006, p. 103) : « *Certes la comparaison des résultats de simulation avec des jeux de données empiriques constitue un exercice important qui s'inscrit dans le processus de modélisation. Mais est-ce suffisant pour conclure à la validité du modèle ?* ». De surcroît, il apparait qu'il n'existe pas encore d'outils fiables et utilisés par tous pour la validation des simulations multi-agents (Amblard *et al.*, 2006).

Ces questionnements prennent encore plus de sens lorsque l'on s'inscrit dans une démarche guidée par les données : comment assurer une démarche de validation qui assure l'autonomie de l'échantillon d'évaluation qui est utilisé au cours du processus d'initialisation et lors de la simulation ? Quelles sont les forces et les faiblesses du modèle SCAUP dans un contexte réel ? Quels sont les impacts d'une modification d'un ou de plusieurs paramètres sur l'ensemble des données simulées ?

L'approche *deepening* (Hassan *et al.*, 2010a, 2010b) soulève encore plus fortement ces questionnements. En effet, comme le relèvent Samer Hassan *et al.* (2010a, 2010b), dans une démarche de SMA classique la validation peut seulement être cantonnée à une étape de vérification de la théorie. Alors que dans la perspective d'une approche guidée par les données, le processus d'évaluation doit être opéré avec des données qui n'ont pas été incluses en entrée du modèle de SMA. Il faut donc être en mesure de s'inscrire dans un aller et retour entre ce qui relève de la simulation et ce qui relève de l'observation. La validation croisée est issue de la l'analyse des données de laquelle une multitude d'approches découle (Biernat *et al.*, 2015). Ici, l'une des difficultés est de proposer une validation croisée synthétique du modèle de simulation, tout en proposant une démarche iconographique de la validation d'un modèle. L'enjeu de l'approche présentée vise à tester les capacités du modèle SCAUP, afin de généraliser les données

[132] Par vérification, nous entendons l'ensemble des tests qui vise à éprouver mathématiquement le fonctionnement interne du modèle de SMA que nous n'aborderons pas au cours de cette thèse.

de comptages routiers sur un espace d'étude expérimental : la ville de Dijon. Cette démarche s'inscrit non seulement dans une perspective opérationnelle pour l'évaluation de la PAA en milieu urbain, mais aussi dans une optique de généralisation du dispositif de modélisation, du moment que nous disposons des données de comptages routiers et du réseau routier d'un espace urbain quelconque. Dans ce sens, le fonctionnement interne de SCAUP n'est pas modifié, et le glissement de l'approche KISS vers l'approche KIDS nécessite d'insérer les éléments nécessaires à son initialisation à travers un environnement de simulation plus complet. Au cours de ce chapitre, l'approche expérimentale du modèle de simulation vise à tester ces capacités opérationnelles de SCAUP à travers les données recueillies sur la ville de Dijon : le réseau routier, les capteurs électromagnétiques et le dispositif de quantification

Par ailleurs, notre travail visant à sélectionner les « meilleures » valeurs de trafic routier simulé pour l'évaluation de la PAA, il faut être en mesure de proposer une évaluation des sorties issues du modèle qui soit intuitive, synthétique, et communicable à différents tiers, pour éviter toute ambiguïté, non seulement sur les hypothèses et les méthodes sous-jacentes au fonctionnement de SCAUP, mais aussi sur ce qui est correctement simulé et sur ce qui ne l'est pas. Au cours de ce chapitre, nous explorerons ces différents écueils pour évaluer les capacités du modèle SCAUP à produire et généraliser un trafic routier émergent à partir d'un système de capteurs fixes. Dans un premier temps, nous exposerons l'ensemble du plan d'expérience élaboré sur la ville de Dijon en nous focalisant sur les paramètres du modèle explorés et sur, les données initialisées au sein du système multi-agents, puis en soulevant les différents enjeux liés à l'évaluation d'un modèle alimenté par des données de comptage, par le biais d'une méthode de validation croisée. Ensuite, nous nous focaliserons sur le fonctionnement interne de SCAUP, en relevant l'ensemble des éléments liés aux simulations expérimentées par le biais d'une analyse de reproductibilité du modèle. Enfin, nous nous focaliserons sur la relation entre données d'observation et simulations en vue de sélectionner les meilleures données issues de SCAUP pour l'étude de la pollution atmosphérique d'origine automobile. En mobilisant différentes approches issues de la statistique classique, de l'ingénierie du trafic et de la climatologie, cette dernière section permettra d'évaluer et de sélectionner les sorties du simulateur de trafic routier.

## 7.1. Exploration du modèle SCAUP à travers le cas de la ville de Dijon

Après avoir établi le modèle « théorique » en vue de le tester et de valider les principales fonctionnalités de SCAUP, la seconde étape du processus expérimental de simulation vise à passer à l'application d'un cas pratique à travers la ville de Dijon (*cf.* figure 2.16, p. 170). L'application de ce second test expérimental offre non seulement l'opportunité d'évaluer le modèle, et sa capacité à produire un trafic routier émergent sur un réseau routier plus vaste, mais aussi de valider les valeurs de trafic routier simulé avec les valeurs d'observation recueillies sur l'espace d'étude via les capteurs électromagnétiques. L'ensemble du processus expérimental de SCAUP est exposé ci-dessous et se situe dans la continuité du plan d'exploration présenté plus haut (*cf.* figure 2.32, p. 205).

### 7.1.1. **Plan d'expérience de SCAUP sur la ville de Dijon : l'approche KIDS**

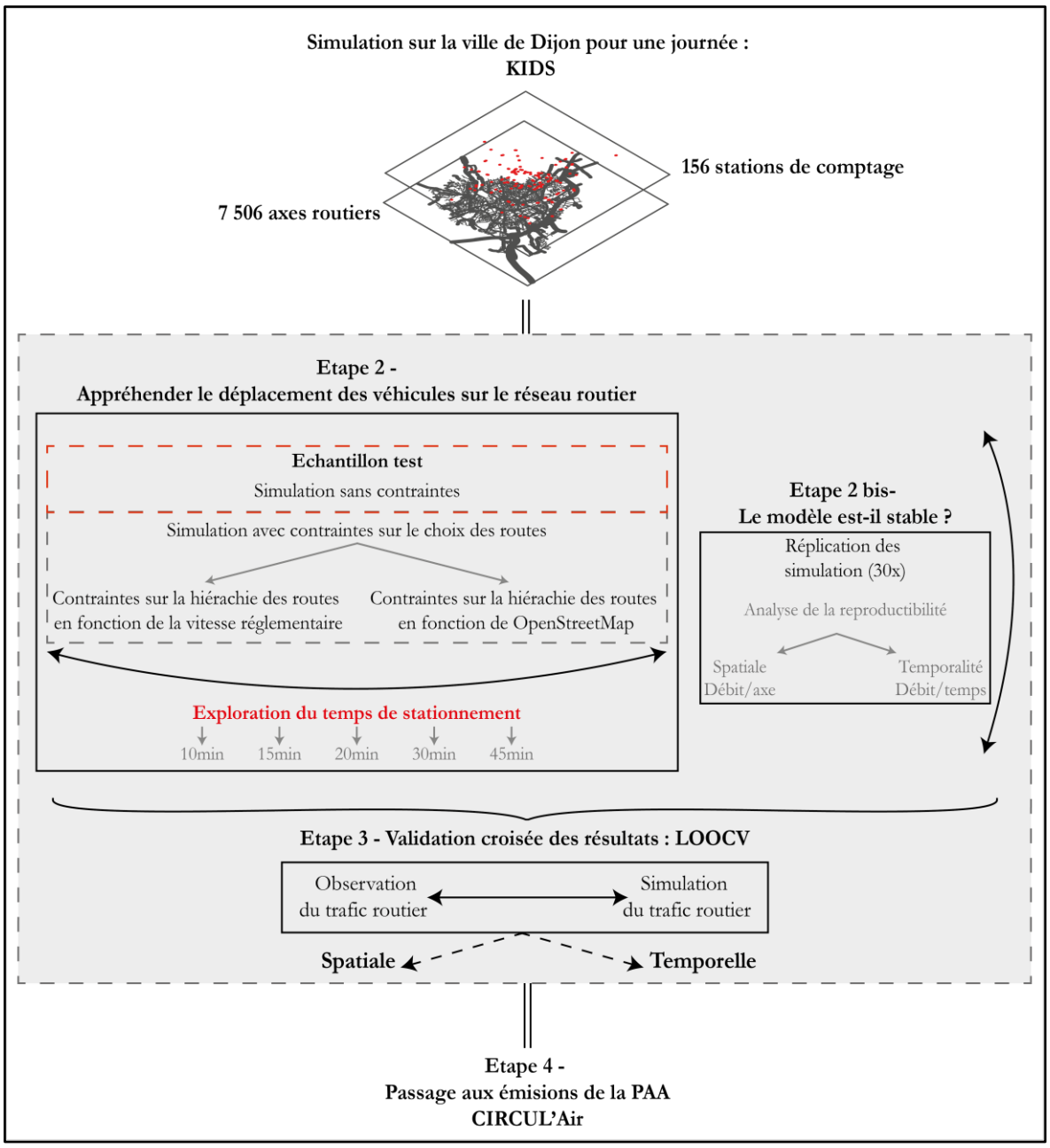



**Figure 3.1 : Plan d'expérience du modèle SCAUP sur la ville de Dijon**

Tout d'abord, le processus d'exploration, présenté ci-dessus, se situe dans la continuité du plan d'exploration du modèle KISS exposé en figure 2.32 (p. 205). Parallèlement à chaque test, comme pour la phase précédente, une analyse de reproductibilité de SCAUP est établie pour apprécier la capacité de celui-ci à être stable ou non (*étape 2 bis* en figure 3.1). Rappelons ici que les indicateurs de reproductibilité permettent de déterminer la variance d'une simulation face à l'ensemble des autres simulations. Nous mobilisons ici une seule et même matrice de transition afin de bénéficier des mêmes conditions initiales de simulation. En outre, les valeurs définies au sein de la matrice de transition ne sont pas explorées car d'une part celles-ci sont définies en fonction des valeurs empiriques recueillies au cours de ce travail (*cf.* section 6.2.2.b, p. 198), et d'autre part, car son exploration pas-à-pas nécessiterait un temps de simulation important. À titre indicatif l'exploration de la matrice de transition basée sur OpenStreetMap nécessiterait de définir $11^{11}$ possibilités[133]. C'est pourquoi, nous exploiterons une seule et même matrice de transition permettant de bénéficier des mêmes conditions initiales de simulation.

Concernant les phases d'exploration et de sensibilité du modèle SCAUP (*étape 2* en figure 3.1), deux paramètres ont été retenus pour *appréhender le déplacement des véhicules sur le réseau routier* :

- Le type de **déplacement des véhicules**, soit ils se répartissent aléatoirement sur les routes, soit ils se déplacent en tenant compte de la hiérarchie du réseau routier (*cf.* section 6.3.4, p. 213). Le paramètre *aléatoire* fera office d'échantillon test pour l'analyse des données de trafic routier simulé. Il permet, en outre, d'apprécier les résultats en s'assurant d'analyser ce qui tient plutôt compte des paramètres testés et ce qui tient plutôt compte du fonctionnement interne de la plate-forme de SMA.
- Le **stationnement des véhicules** permet de définir un temps maximum de déplacement des véhicules sur le réseau routier. Les tests de simulation visent à faire varier ce temps de résidence, en explorant ses effets pour chaque paramètre de déplacement, de 10 à 45 minutes. Conformément au résultat exposé en figure 2.38 (p. 215), nous n'avons pas engagé d'exploration fine du temps de stationnement au-delà de 30 minutes. Toutefois, le paramétrage d'un temps de stationnement défini à 45 minutes a tout de même été envisagé, car nous travaillons sur un environnement de simulation plus vaste que celui établi précédemment avec le modèle KISS.

En somme, la combinaison de l'ensemble de ces tests permet d'explorer 15 conditions initiales de simulation du trafic routier (*cf.* tableau 3.1, ci-dessous). Chacun des paramètres testés est établi pour une journée de 24 h au pas de temps minute. Chaque simulation prend en moyenne près de 1 h 30[134] et a été effectuée à l'aide des ressources de calcul haute performance (HPC) du mésocentre de l'Université de Bourgogne. Par ailleurs, ces ressources ont été utilisées car elles permettent de faciliter un calcul parallélisé et distribué des différentes expériences du modèle, qui

[133] Soit près de trois milliards (285 311 670 611) de possibilités pour explorer la matrice de transition rattachée à la hiérarchie OpenStreetMap. La matrice de transition relative à la hiérarchie des vitesses réglementaires est plus réduite avec $5^5$ (3 125) possibilités.

[134] Le temps de simulation s'accroît lorsque le déplacement des véhicules est contraint en fonction du réseau routier. L'insertion de la matrice de transition ainsi que du temps de stationnement augmentent sensiblement le temps de simulation. À titre, d'exemple, les simulations établies pour un temps de stationnement de 45 min (en violet au sein du tableau 3.1) s'opèrent sur une durée moyenne de simulation de 8 h, soit un temps de simulation 5 fois plus important que pour les paramétrages aléatoires.

n'aurait pu être envisageable sur un seul ordinateur[135]. Par souci de clarté et de lecture, le tableau 3.1 ci-dessous présente les codes couleurs ainsi que les appellations qui seront utilisés au cours de cette partie en vue d'exposer les résultats du modèle SCAUP. Par simplification, les paramètres de choix routiers seront codés en « *Aléatoire (R pour Random)* » lorsque aucune contrainte n'est associée aux mouvements des véhicules, en « *Hiérarchie (H pour Hierarchy)* » lorsque le déplacement est contraint en fonction de la typologie du réseau routier issue de OpenStreetMap (*cf.* en bas de la figure 2.19, p. 177), et en « *Vitesses (S pour Speed)* » (*cf.* figure 2.18, p. 176) lorsque le déplacement est contraint en fonction de la vitesse réglementaire. À noter que les résultats des expériences réalisées avec SCAUP seront présentés en s'appuyant sur la nomenclature présentée ci-dessus tout au long des développements suivants :

| | **Temps de stationnement** | **10 min** | **15 min** | **20 min** | **30 min** | **45 min** |
|---|---|---|---|---|---|---|
| **Choix routiers** | **Aléatoire (R)** | R10 | R15 | R20 | R30 | R45 |
| | **Hierarchie (H)*** | H10 | H15 | H20 | H30 | H45 |
| | **Vitesse (S)** | S10 | S15 | S20 | S30 | S45 |
| * Nomendature de la hiérarchie des routes issue de OSM | | | | | | |

**Tableau 3.1 : Nomenclature des acronymes et des codes couleur des expériences réalisées lors de l'expérimentation du modèle SCAUP sur la ville de Dijon**

La dernière phase du plan d'expérimentation du modèle (*étape 3* en figure 3.1) vise à s'assurer de la qualité des données de trafic routier simulées à partir de SCAUP en les confrontant aux valeurs d'observation issues des capteurs de la ville de Dijon. Toutefois, comme nous l'avons relevé, la validation d'un modèle de simulation initialisé avec ces mêmes données nécessite d'adapter le cadre de validation du modèle en s'assurant de l'indépendance des échantillons d'observation et de simulation. Cette problématique constitue d'ailleurs un véritable enjeu lors de la mise en œuvre d'une démarche guidée par les données (Hassan *et al.*, 2010b).

L'approche envisagée pour la validation des résultats de SCAUP vise à s'entourer des méthodes de validation croisée (Stone, 1974), et plus spécifiquement la méthode de validation croisée dite « *Leave One Out Cross Validation (LOOCV)* », dont nous explorerons plus en détails les éléments conceptuels et méthodologiques au cours de la section 7.1.3 (p. 231). Cette dernière étape a pour finalité de non seulement valider les données de simulation, mais aussi de faciliter la sélection des meilleures données de trafic simulé pour ensuite alimenter le modèle d'émission des polluants atmosphériques automobiles (PAA) appliqué par l'AASQA[136] de la ville de Dijon : CIRCUL'AIR.

[135] Présentation du centre de calcul du mésocentre de l'université de Bourgogne : https://haydn2005.u-bourgogne.fr/dsi-ccub/spip.php?article959, ainsi qu'une illustration photographique du cluster de calcul : https://haydn2005.u-bourgogne.fr/dsi-ccub/spip.php?article162.
[136] AASQA : Associations Agréées pour la Surveillance de la Qualité de l'Air.

### 7.1.2. Expérimentation de SCAUP sur la ville de Dijon : données d'entrée, disponibilités et initialisation de la simulation multi-agents

La mise en place du terrain d'expérimentation est réalisée via le système de gestion de base de données mis en place pour manipuler les données de comptages routiers (*cf.* section 3.2.3, p. 97). Rappelons que l'insertion du terrain de la ville de Dijon au sein de SCAUP s'insère dans un positionnement KIDS (*cf.* figure 2.17, p. 171). Mobilisés en amont, les SIG ont permis de construire le monde virtuel dans lequel les agents seront ensuite simulés. Ces outils sont exploités en suivant les préconisations de Stan Openshaw et Robert J. Abrahart (2000) pour la *géocomputation*, dans le sens où les SIG sont utilisés en vue de produire et de créer les données nécessaires à l'initialisation du modèle SCAUP. La génération du terrain d'étude (environnement et agents) vise à insérer au sein du simulateur de trafic routier l'ensemble des données géographiques, ainsi que des données de comptage du trafic routier sur la ville de Dijon, nécessaires à la mise en œuvre de SCAUP : le réseau routier (arcs et nœuds), les stations de mesure et les données de comptage qui leur sont associées. L'ensemble du terrain d'étude utilisé est présenté au sein de la figure 3.2 (ci-dessous) :

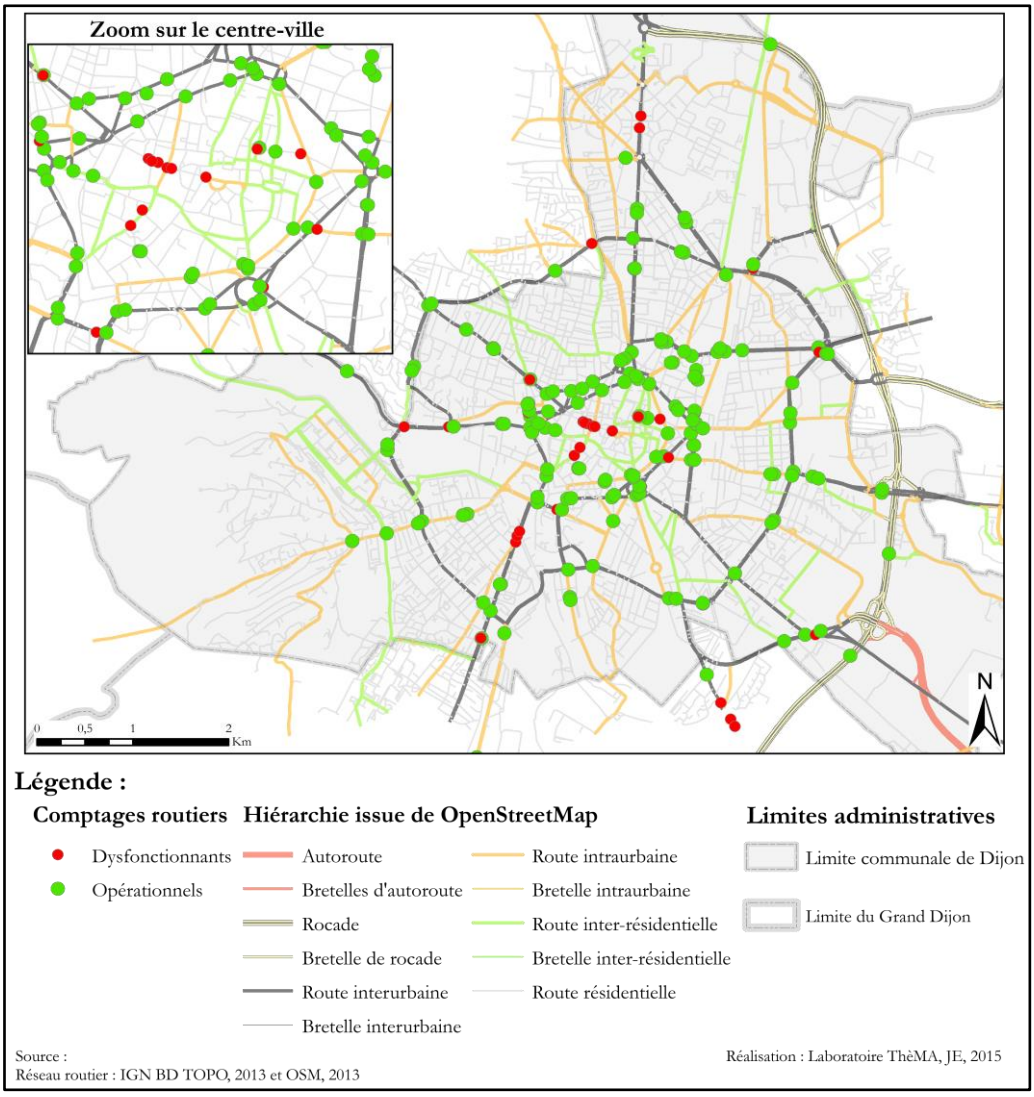



**Figure 3.2 : Les stations de mesure opérationnelles (156 postes) pour la simulation du trafic routier sur une journée représentative : le mardi 9 septembre 2008**

Comme précédemment vu (*cf.* encart 5, p. 195), le **réseau routier** a été créé à partir du module automatique associé au modèle MOBISIM. L'ensemble du terrain d'étude, qui correspond aux limites communales de la ville de Dijon, a été défini dans le SIG par un zonage autour duquel une zone tampon (*buffer*) de 500 mètres a été fixée afin d'englober l'ensemble des routes situées en marge de la limite communale (*cf.* figure 3.2, ci-dessus). Ceci est particulièrement utile dans certains cas, et notamment dans celui de la Rocade Georges Pompidou située à l'Est de la commune (*cf.* figure 3.2), dont le tracé n'est pas totalement inclus dans la commune dijonnaise. Une fois le domaine d'étude défini, l'ensemble du réseau routier est directement intégré dans la plate-forme de SMA GAMA. Rappelons que le réseau routier, duquel nous en avons présenté plusieurs fois les caractéristiques, contient l'ensemble des informations nécessaires au déplacement des véhicules sur le réseau routier, en intégrant autant les valeurs liées à la vitesse de déplacement des véhicules que l'ensemble des valeurs liées à la hiérarchie du réseau routier (OpenStreetMap, vitesse réglementaire, capacité…). Dès lors, chaque tronçon routier possède des caractéristiques qui lui sont propres et qui serviront aux *agents* « véhicules » pour réaliser leurs déplacements au cours du processus de simulation (*interaction agents-environnement*).

En ce qui concerne les **stations de comptage**, rappelons tout d'abord que les différents tests de simulation sont ici envisagés pour la simulation d'une seule journée de comptage (24 heures de comptage). Ainsi, au préalable, il a été nécessaire de sélectionner, puis d'identifier une journée de comptage à extraire parmi d'autres au sein de la base de données. Cette journée est considérée comme une journée « type » de la circulation routière en milieu urbain en s'appuyant sur les comptages routiers collectés. Pour cela, différents critères ont été établis en vue de sélectionner et d'extraire une journée de comptage représentative de la circulation routière intra-journalière sur l'ensemble des valeurs de trafic routier collectées entre 2001 et 2012 :

- 1. Bénéficier d'une journée de comptage représentative de la dynamique intra-journalière du trafic routier ;
- 2. Bénéficier d'un nombre de stations de mesure optimal pour la période sélectionnée, c'est-à-dire présentant un taux de disponibilité important ainsi que peu de dysfonctionnements
- 3. Bénéficier d'une couverture spatiale significative du réseau routier, autant sur les tronçons routiers ayant de fortes valeurs de fréquentation que sur les réseaux présentant une fréquentation plus faible.

L'ensemble de ces critères vise, tout d'abord, à utiliser les données de comptage sans en modifier leur nature, en utilisant exclusivement les données *brutes* issues du système de gestion de base de données (SGBD) mis en place. Le premier critère exclut par nature les WE, de même que les périodes de vacances, qui présentent des caractéristiques de fréquentation différentes d'une journée de comptage que l'on peut qualifier de *classique*[137]. La temporalité des données apparait primordiale pour ce travail ; dans ce sens, les valeurs de comptage au quart d'heure doivent respecter le signal intra-journalier qui a été mis en exergue à plusieurs reprises lors des sections précédentes. En effet, c'est une condition essentielle pour la simulation de la dynamique du trafic routier à partir du modèle SCAUP.

[137] En effet au cours de la section 3.3.3 (p. 106), nous avons vu que la circulation routière apparait rythmée et en phase avec les rythmes scolaires et de travail (routine du signal issue des comptages routiers).

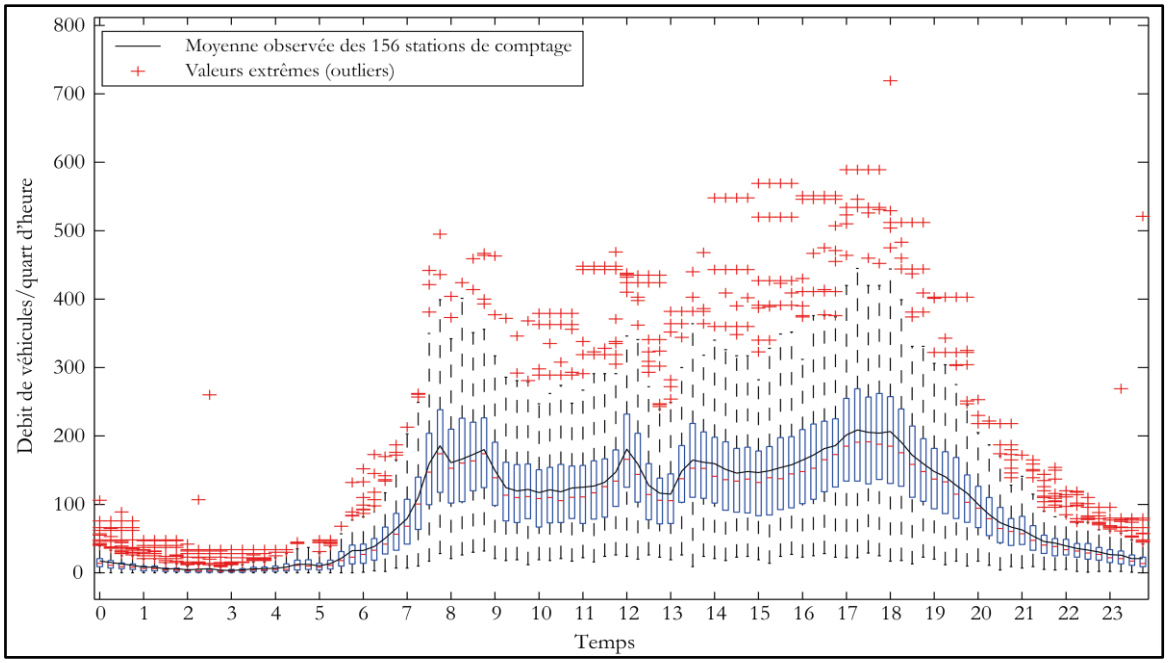



**Figure 3.3 : Observation du trafic routier pour les 156 stations de mesure opérationnelles : le mardi 9 septembre 2008**

Dès lors, comme illustré ci-dessus, notre choix s'est tourné vers un des jours de la semaine en privilégiant, au même titre que les EMD, une journée de référence qui est soit un mardi soit un jeudi (Thévenin *et al.*, 2007 ; Commenges, 2013). Les second et troisième critères de sélection sont intimement liés, dans le sens où plus il y a de stations disponibles, plus la couverture spatiale du réseau de mesure est significative. En s'appuyant sur le travail présenté en figure 1.39 (p. 109), la journée qui a été sélectionnée pour l'initialisation de SCAUP sur la ville de Dijon est le mardi 9 Septembre 2008. Le choix de retenir cette journée parmi d'autres se justifie par deux points :

- 1. C'est une journée caractéristique et représentative de la circulation routière sur l'ensemble de la période collectée (*cf.* figure 3.3, ci-dessus) ;
- 2. C'est une journée où les stations de mesure disponibles sont parmi les plus nombreuses (87 % de disponibilité des données).

L'extraction de cette journée de comptage est automatisée, en associant à chacune des stations localisées les valeurs de débit observées *via* le gestionnaire de base de données PostgreSQL. La chaîne de traitement mise en place présente des limites lorsque les données associées aux capteurs présentent des dysfonctionnements liés à des ruptures de comptage ou à des valeurs de débit égales à 0 (nulles) : l'application de la moyenne mobile sur des valeurs nulles renvoie naturellement à des valeurs nulles (*cf.* section 3.2.3, p. 97*)*. En outre, il est difficile de rattacher des valeurs nulles à un dysfonctionnement du capteur ou à une valeur de non-fréquentation de l'axe mesuré par la station. Ainsi, en s'appuyant sur la figure 1.39 (p. 109) 181 stations sont disponibles et opérationnelles (85% de disponibilité), mais en définitive, comme illustré (en rouge) au sein de la figure 3.2, 25 stations ont été exclues, soit car celles-ci sont situées hors de la zone d'étude définie *a posteriori*, soit car elles présentent des dysfonctionnements. C'est pourquoi, seulement 156 stations de mesure ont été intégrées au modèle SCAUP : cela signifie que ces stations de comptage, ainsi que les comptages associés, ne présentent aucun dysfonctionnement ou valeurs nulles pour la journée du 9 septembre 2008.

### 7.1.3. Les enjeux liés à une approche guidée par les données : les questionnements liés à la validation et à l'évaluation du modèle SCAUP

Lorsque Samer Hassan *et al.* (2010b, p. 4) abordent les concepts des SMA guidées par les données, ceux-ci soulignent que lors de l'approche *deepening*, « *les données utilisées dans la validation ne doivent pas être les mêmes que celles utilisées lors des étapes précédentes (initialisation, simulation)* ». Dans cette acception, il est important de s'assurer du caractère disjoint des observations utilisées pour construire le modèle de simulation et des observations qui seront confrontées aux résultats de la simulation. Comparativement à une approche de simulation classique, qui peut être validée par la théorie ou par les données (*cf.* figure 2.15, p. 168), lors d'une approche *deepening* les données jouent un rôle plus prépondérant à chaque étape du processus de modélisation. C'est pourquoi s'inscrire dans une optique guidée par les données nécessite un effort particulier lors de la phase de collecte, ainsi que lors de la phase de validation et d'évaluation du modèle de simulation multi-agents (Hassan *et al.*, 2010a, 2010b). Les auteurs y voient néanmoins une des limites quant à s'inscrire dans une optique *deepening*, car la phase de validation nécessite « *un coût élevé qui peut être difficile à justifier, en dépit de l'amélioration attendue des résultats* » (Hassan *et al.*, 2010a, p. 10).

#### a. *Les principes généraux de la validation des modèles guidée par les données*

En suivant les différentes préconisations mises en évidence par Samer Hassan *et al.* (2010a, 2010b), il faut être en mesure d'une part de proposer un système de validation non seulement adéquat à l'approche *deepening*, mais aussi adapté aux enjeux liés au modèle de SMA. Dans notre cas, ces enjeux sont liés à l'insertion des données de comptages positionnées au centre du processus de modélisation et de simulation du trafic routier. D'autre part, il faut être en mesure de fournir une démarche d'évaluation du modèle de SMA raisonnable en termes de coût (en temps et en ressources) par rapport à une approche de SMA classique.

Les concepts de validation apparaissent relativement évidents car ils visent à découper une base de données en différents « paquets » permettant non seulement d'initialiser et de simuler mais aussi d'évaluer le modèle de simulation ; toutefois « *il existe beaucoup d'alternatives permettant de sophistiquer cette approche : on parle alors de validation croisée* » (Biernat *et al.*, 2015, p. 166). Dans notre cas, le modèle SCAUP étant autant initialisé que validé à partir des stations de comptages routiers, la place accordée à celles-ci dans le processus de modélisation mérite de s'y arrêter pour en identifier les enjeux. En effet, les données interviennent à chaque étape du processus de simulation et on se retrouve confronté à enjeu important : comment conserver le maximum de stations de mesure disponibles tout en assurant une évaluation cohérente du modèle ? Effectivement, on se retrouve ici confronté à un modèle « *consanguin* » dans le sens où les données servant à l'initialisation et à la simulation du trafic routier doivent aussi être pris en compte tout au long de la démarche initiée et plus particulièrement lors de la phase de validation et d'évaluation des données de trafic routier simulé.

L'ensemble des enjeux de validation se rapproche des éléments conceptuels rattachés à la validation croisée. En effet, dans son acception la plus générale, la validation croisée vise à évaluer la qualité d'un modèle en évitant d'utiliser les mêmes données pour initialiser le modèle et

pour évaluer celui-ci (Stone, 1974 ; Biernat *et al.*, 2015) [138]. La figure ci-dessous illustre le principe général de la validation croisée en partant d'un même jeu de données. Cette phase de validation du modèle (Stone, 1974) vise à déconnecter les données d'observation (*validation*) des données de simulation (*entrainement*), pour ensuite les confronter en vue d'apprécier non seulement les qualités, mais aussi les défauts du modèle de simulation pour produire un trafic routier émergent à partir des comptages routiers :

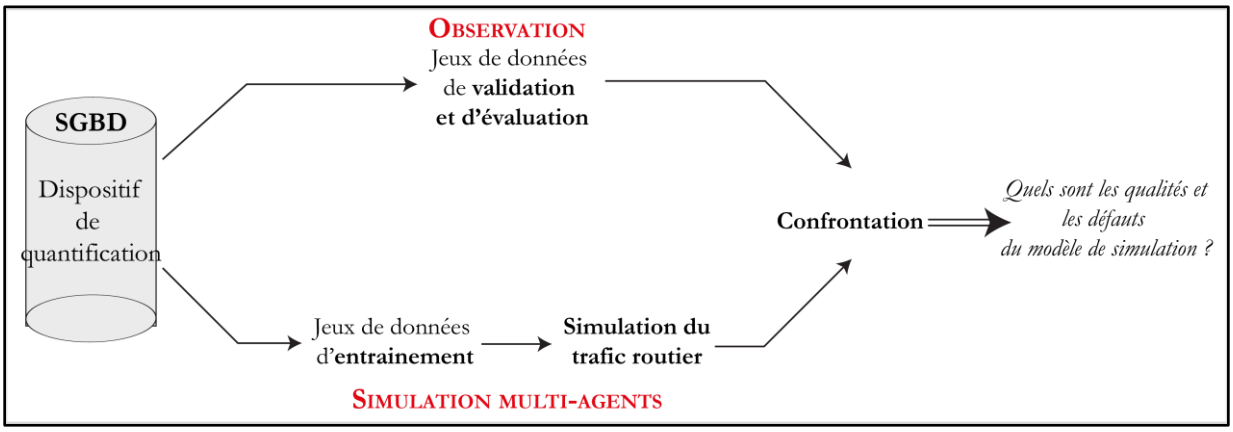



**Figure 3.4 : Principe général de la validation croisée : simulation et observation du modèle de simulation multi-agents du trafic routier**

Il existe plusieurs approches de validation croisée. Deux grandes catégories se distinguent en particulier : celles dite *exhaustives* qui visent à explorer l'ensemble des observations d'une même base de données ; et celles dites *non-exhaustives*, qui sont une « *approximation des méthodes exhaustives, mais nécessitent moins de temps de calcul* » (Biernat *et al.*, 2015, p. 167). En ce qui concerne les méthodes *non-exhaustives*, l'approche la plus connue est celle dite « *k-fold cross-validation* ». Elle vise à scinder la base de données en deux échantillons distincts en sélectionnant aléatoirement un premier échantillon de données qui sera exploité lors de la simulation (*échantillon d'entrainement*) ; tandis que le second échantillon est, quant à lui, conservé pour la phase de validation (*échantillon de validation et/ou d'évaluation*). C'est l'approche la plus simple à mettre en œuvre ; par exemple, on sélectionne aléatoirement 40 % des données pour la validation, et les 60 % restants seront utilisés pour la simulation du modèle. Cette méthode dite *non-exhaustive* permet la mise en œuvre rapide de la phase de validation. Elle se distingue des méthodes *exhaustives* qui visent à utiliser l'ensemble d'une base de données autant pour la validation que pour la simulation, dont les deux plus connues sont (Biernat *et al.*, 2015) :

- L'approche « *LKOCV* (*Leave-K-Out Cross-Validation*) » qui présente les mêmes principes que la méthode *k-fold cross-validation* en découpant le corpus de données en deux ensembles distincts. Cependant, dans ce cas de figure, l'ensemble des combinaisons possibles est exploré et répété en vue de couvrir l'ensemble des découpages possibles de la base de données. Par exemple en explorant l'ensemble des données par la mise en place d'un traitement par « fenêtre glissante » permettant alors de couvrir l'ensemble de la base de données, que ce soit pour la simulation ou pour la validation.

[138] Notons tout de même qu'il existe un grand nombre de méthodes de validation croisée ; on peut par exemple citer les méthodes de *bootstrap* qui se fondent plutôt sur les théories de rééchantillonage statistique.

- L'approche « *LOOCV* (*Leave-One-Out Cross-Validation*) » vise quant à elle à explorer l'ensemble de la base de données en ne retirant non pas un ensemble d'observations mais plutôt une seule observation au sein de la base de données (N-1). Le processus de validation est répété autant de fois qu'il y a d'observations comprises dans la base de données. Par conséquent, avec cette méthode, on conserve un maximum d'observations pour le processus de simulation.

Ces deux dernières approches, *exhaustives*, permettent d'exploiter l'ensemble des données à disposition, autant pour construire le jeu de données d'entrainement que celui de simulation. L'utilisation de ces méthodes nécessite toutefois un temps de traitement bien plus important que lors d'une validation *non-exhaustive*. Il est donc très souvent nécessaire de faire appel aux ressources HPC[139] pour réaliser ces traitements (Stone, 1974). Elles présentent une réelle force lorsque l'on s'inscrit dans une approche guidée par les données, car elles permettent non seulement de conserver un maximum d'observations pour la simulation, mais aussi de fournir une évaluation précise du modèle au niveau de toutes les observations collectées. C'est pourquoi, ici, la phase d'évaluation de SCAUP s'est plutôt orientée dans cette optique de validation croisée dite *exhaustive*, en vue d'exploiter toutes les stations disponibles du début à la fin de la démarche de modélisation guidée par les données entreprise dans cette thèse.

#### b. *La méthode LOOCV : tester le modèle de simulation avec l'ensemble du dispositif de quantification à disposition*

Nous avons choisi d'utiliser la méthode LOOCV, qui permet d'apporter une réponse opérationnelle et adéquate aux enjeux de l'approche *deepening* (Hassan *et al.*, 2010a, 2010b ; Banos, 2013). Plusieurs raisons justifient l'emploi de cette méthode de *validation croisée* pour ce travail :

- **Au niveau du traitement des résultats de simulation** : C'est une méthode exhaustive qui permet d'utiliser l'ensemble du dispositif de quantification pour l'évaluation du modèle. De surcroît, le modèle de simulation du trafic routier étant initialisé à partir des stations de mesure, il est primordial de bénéficier d'un maximum de stations de comptage lors du processus de simulation. En conséquence, chaque simulation est établie en utilisant N-1 stations de mesure, ce qui n'est pas le cas avec la méthode LKOCV.
- **Au niveau de l'évaluation** : Elle procure l'avantage de fournir une estimation sur l'ensemble des stations de mesure prises une à une. Une fois l'ensemble du plan d'expérimentation établi, la mise en place de cette méthode permet de constituer une information au niveau de chaque station. Dans ce sens, chaque indicateur est illustré au niveau d'un capteur, où les relations entre les données de simulation et les données d'observation (corrélation, erreur moyenne, écart-type…) sont appréhendées à une échelle individuelle.

Dans notre cas de figure, l'établissement de cette méthode consiste, comme présenté en figure 3.5, à effectuer l'apprentissage du modèle avec N-1 stations de mesure (soit 155 stations), et ensuite à l'évaluer avec une seule station de mesure, tout en répétant l'expérimentation autant de fois qu'il y a de stations de mesure (N fois, soit dans notre cas 156 fois). L'échantillon de validation qui sera ensuite extrait réunit l'ensemble des tronçons où les stations de mesure ne sont

[139] Notons que nous nous retrouvons de nouveau confrontés à un des éléments caractérisant la *géocomputation* : le calcul haute performance (HPC) comme nous l'avons vu plus haut : *cf.* section 1.3.3, p. 41.

pas intervenues lors du processus de simulation, comme présenté en rouge sur la figure 3.5. Si l'on reprend le plan d'exploration du modèle (*cf.* figure 3.1, p. 225) en y ajoutant l'ensemble du processus de validation croisée pour chaque exploration des paramètres, le nombre de simulations envisagé à partir de SCAUP est 156 fois plus important. En effet, comme nous l'avons soulevé, les principales contraintes pour la mise en œuvre de la LOOCV se situent au niveau des capacités de calcul et du temps de simulation, qui peut s'avérer très long. Toutefois, s'inscrire dans cette optique permet d'apporter une réponse opérationnelle et adéquate aux enjeux de l'approche.

Au final, l'application du plan d'exploration du modèle de simulation envisagé à l'échelle de la ville de Dijon nécessite l'application de 2 340 simulations, correspondant à l'ensemble des 15 combinaisons de paramètres à explorer (*cf.* tableau 3.1, p. 227) multiplié par le nombre de stations de mesures insérées dans le système multi-agents (*cf.* figure 3.2, p. 228). On comprend assez naturellement la limite de cette approche, qui nécessite un coût informatique important tant pour l'expérimentation du modèle de trafic routier que pour la mise en œuvre du plan d'expérience. À titre indicatif, une simulation sur une journée de comptage prenant en moyenne 1 h 30, le temps total de simulation nécessaire à la mise en œuvre complète du plan d'expérience du modèle SCAUP est de 3 510 heures de calculs cumulés[140]. Pour faciliter la construction ainsi que la mise en œuvre du plan d'expérience, nous nous sommes entouré du logiciel *OpenMole*, permettant de faciliter l'exécution du plan d'expérience et d'exploration d'un modèle de simulation (Reuillon *et al.*, 2013).

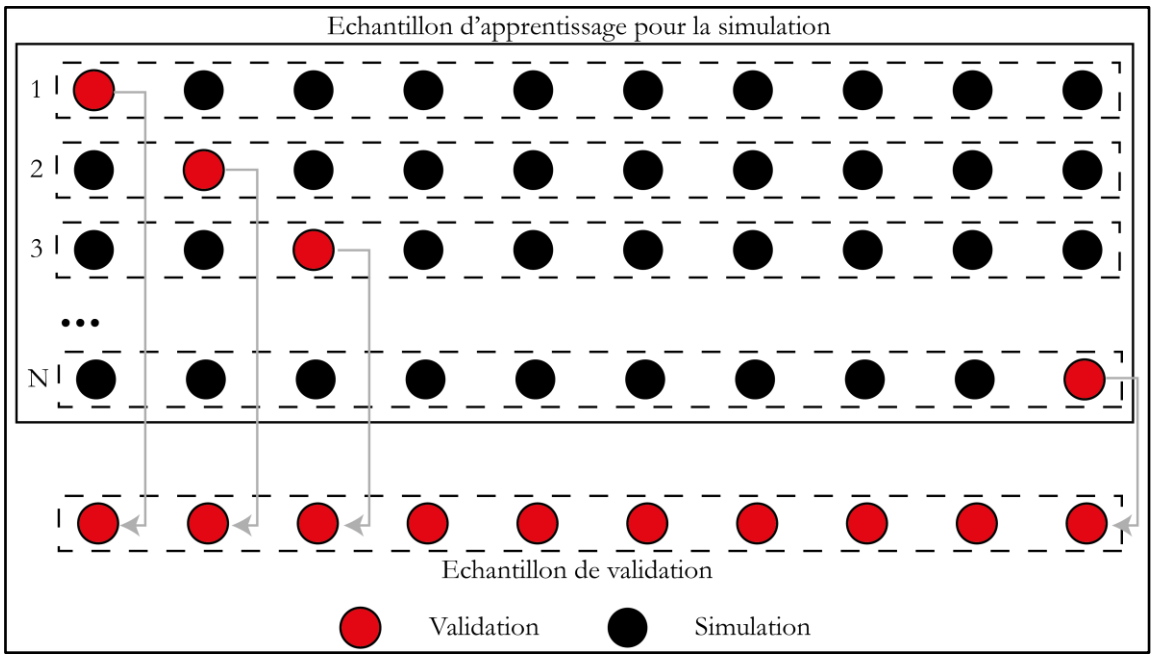



**Figure 3.5 : Principe de la validation croisée et d'échantillonnage d'apprentissage pour la simulation du trafic routier à partir des stations de mesure : l'approche *Leave One Out Cross Validation* (LOOCV)**

[140] Plus précisément, les 156 simulations * 15 paramètres correspondant à l'exploration de 2 340 tests pour une durée de simulation moyenne de 1 h 30 de simulation représentent 210 600 minutes de calculs (soit 8 775 jours ou 1 253 semaines). L'ensemble étant inexploitable par des ressources individuelles, les simulations ont été réalisées via les ressources de calcul HPC mises à disposition par le centre de calcul de l'Université de Bourgogne (CCUB) : une simulation établie sur un cœur avec une mémoire vive de 5 Go.

Si la capacité de calcul apparait être une condition *sine qua non* pour la validation croisée (LOOCV) de SCAUP, la répétition des simulations par bloc de N-1 stations procure parallèlement l'opportunité d'apprécier la stabilité de celui-ci. En effet, en reproduisant 156 fois chacun des paramètres de simulation, on peut aussi considérer que chaque simulation servira à étudier la capacité de SCAUP à reproduire des mêmes conditions de trafic routier. C'est pourquoi, les analyses préliminaires et les sorties associées à celui-ci seront analysées en appréciant la moyenne des 156 simulations établie lors du processus de validation croisée.

## 7.2. Analyse préliminaire de l'approche expérimentale de simulation du trafic routier

Les objectifs de l'approche expérimentale de SCAUP sur la commune de Dijon visent à produire un trafic routier émergent à partir d'une journée de comptage, tout en reposant sur le caractère génératif des systèmes multi-agents. Rappelons ici que les enjeux opérationnels de l'approche visent à généraliser les comptages issus des capteurs électromagnétiques sur l'ensemble des tronçons routiers de la ville de Dijon, pour ensuite en apprécier les niveaux de PAA. À cet aspect spatial se juxtaposent des objectifs temporels en vue de produire[141] la dynamique intra-journalière de la circulation routière. L'ensemble des tests de simulation décrits précédemment nous permet maintenant d'apprécier les capacités du modèle sur un cas d'étude non seulement expérimental, mais aussi pratique, appliqué à la ville de Dijon.

Pour la présentation des résultats, chaque analyse des résultats se fera en confrontant les simulations *tests* initialisées avec un mouvement aléatoire des véhicules sur le réseau (R) et les simulations tenant compte de la hiérarchie du réseau routier (H et S). Au cours des prochaines sections, nous illustrerons nos propos en sélectionnant les résultats les plus significatifs. Le lecteur est invité à consulter l'annexe cartographique (pp. vi - xxx), où l'ensemble des sorties cartographiques issues des différents tests d'exploration du modèle de SMA est exposé.

### 7.2.1. **Un trafic routier généralisé et simulé sur la ville de Dijon**

Tout d'abord, les résultats cartographiques des premiers tests de simulation pour le temps de stationnement défini à 15 minutes sont présentés pour les trois types de déplacement des véhicules : aléatoire (*R15* : figure 3.6, p. 236), hiérarchie (H 15 : en haut de la figure 3.7, p. 237) et vitesse (*S15* : en bas de la figure 3.7, p. 237). Ces premières sorties illustrent la capacité du modèle de trafic routier à généraliser sur l'ensemble du terrain d'étude les valeurs de comptages issues des 156 stations de comptage sélectionnées précédemment. Ces trois premiers exemples montrent, tout d'abord, la capacité du modèle à produire une carte de charge de la circulation routière à partir d'un processus génératif porté par la plate-forme de SMA. En effet, le premier constat de ces simulations montre que, pour l'ensemble des illustrations présentées, peu d'axes routiers présentent des valeurs nulles.

Le modèle apparait ensuite sensible à l'insertion d'une contrainte de déplacement des véhicules associée à la dimension circulatoire du réseau routier (Genre-Grandpierre, 2010). Effectivement, au niveau de la répartition spatiale des débits de véhicules, on peut remarquer que celle-ci n'est pas la même en fonction des contraintes associées aux déplacements des véhicules, comme

[141] Nous avons vu que la SMA permet non pas de reproduire mais bien de produire un trafic routier émergent.

représenté en figure 3.7 (p. 237). Lorsque les véhicules se déplacent aléatoirement sur le réseau routier, la répartition des flux de circulation est bien plus homogène que lorsque le déplacement des véhicules est contraint en fonction de la hiérarchie du réseau routier :

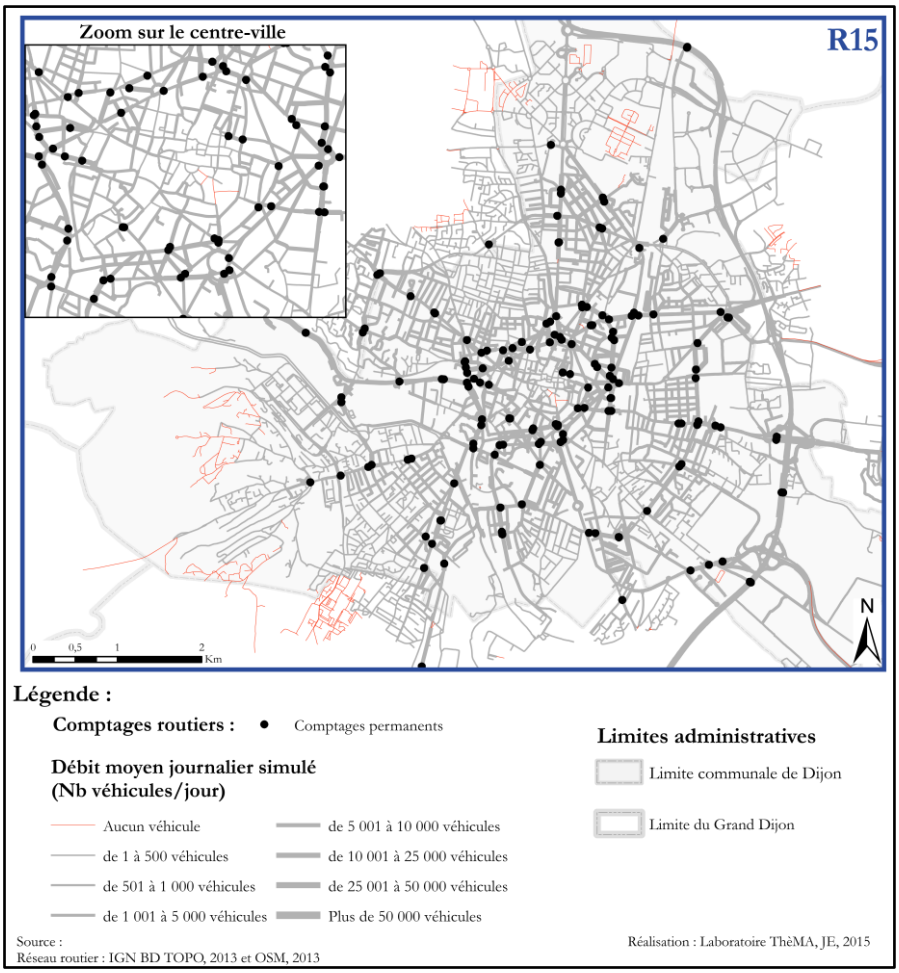



**Figure 3.6 : Débit moyen journalier pour le paramètre *R15* (moyenne journalière des 156 simulations en débit/jour)**

On remarque cela assez intuitivement en visualisant et en confrontant les différentes sorties cartographiques proposées ci-dessous (*cf.* p. 236 et p. 237). En effet, lors de la cartographie des flux de trafic simulés à partir de SCAUP, deux structures spatiales se discernent : lorsque les véhicules se déplacent aléatoirement (*cf.* figure 3.6), on relève que les flux de trafic routier se répartissent en ne présentant aucune structure ; alors que lorsque le mouvement des véhicules est contraint, la répartition des flux de trafic routier révèle la structure radioconcentrique du réseau routier de la ville de Dijon (*cf.* figure 3.7, p. 237). En effet, les flux majeurs de trafic routier se situent sur les principaux axes hiérarchiques, tandis que les flux mineurs se concentrent au niveau des axes résidentiels. Cet état des lieux est encore plus flagrant lorsque l'on se concentre sur les axes routiers non fréquentés (illustrés en rouge). Pour apporter un ordre de grandeur, le taux de route où aucun véhicule n'a été simulé est de 6 % lors du déplacement aléatoire des véhicules, de 22 % lorsque les déplacements sont contraints selon la hiérarchie (H), et de 17 % lorsque ceux-ci sont contraints en fonction de la vitesse (S).

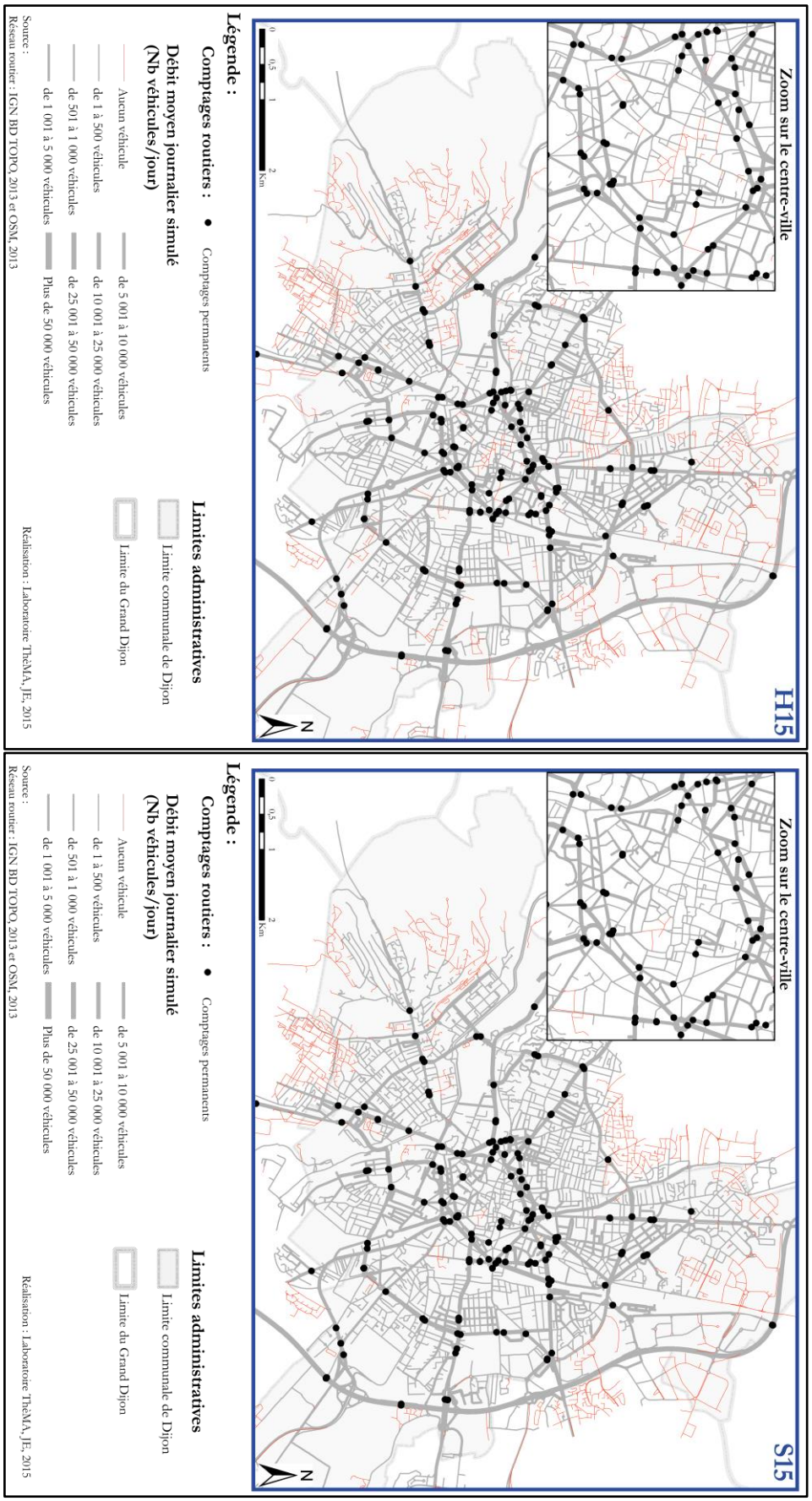



**Figure 3.7 : Débit moyen journalier pour le paramètre *H15* (en haut) et le paramètre *S15* (en bas) : moyenne journalière des 156 simulations en débit/jour**

Comme illustré en figure 3.8, on remarque que plus le temps de stationnement des véhicules est important, plus la part des axes non fréquentés diminue, et inversement, plus la fréquentation des routes augmente, plus le temps de stationnement augmente. De plus, on remarque que la part des axes non fréquentés par les véhicules lors de la simulation est aussi tributaire du paramètre de déplacement des véhicules. Le paramètre de déplacement des véhicules en fonction des contraintes locales impacte sensiblement la fréquentation des axes routiers résidentiels. En effet, on s'aperçoit que lorsque les véhicules se déplacent aléatoirement les axes les moins fréquentés se situent en marge du réseau routier (*cf.* figure 3.6, p. 236), alors que pour les deux autres cas figure 3.7, p. 237) ce constat est plus nuancé et les axes concentrant le moins de véhicules sont plutôt localisés à l'intérieur des quartiers. Si on se reporte à la figure 3.9 (ci-dessous), on relève que contraindre le déplacement des véhicules permet de mieux répartir les flux de trafic routier sur les différents axes routiers Ce bilan reflète l'apport de la matrice de transition en vue de mieux appréhender les contraintes de déplacement des véhicules sur les axes concentrant les potentiels de trafic routier les plus importants. Et cela plus particulièrement sur les axes interurbains et résidentiels, où l'on observe une tendance inversée de la fréquentation entre les paramètres de choix aléatoire et ceux de choix hiérarchique :

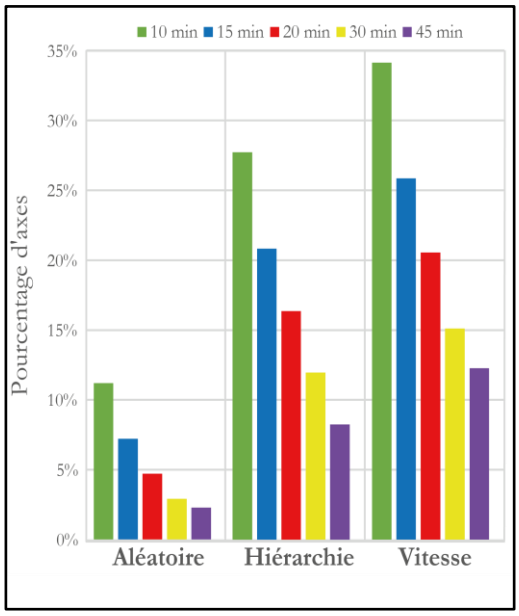



**Figure 3.8 : Pourcentage des axes sans véhicules en fonction du temps de stationnement**

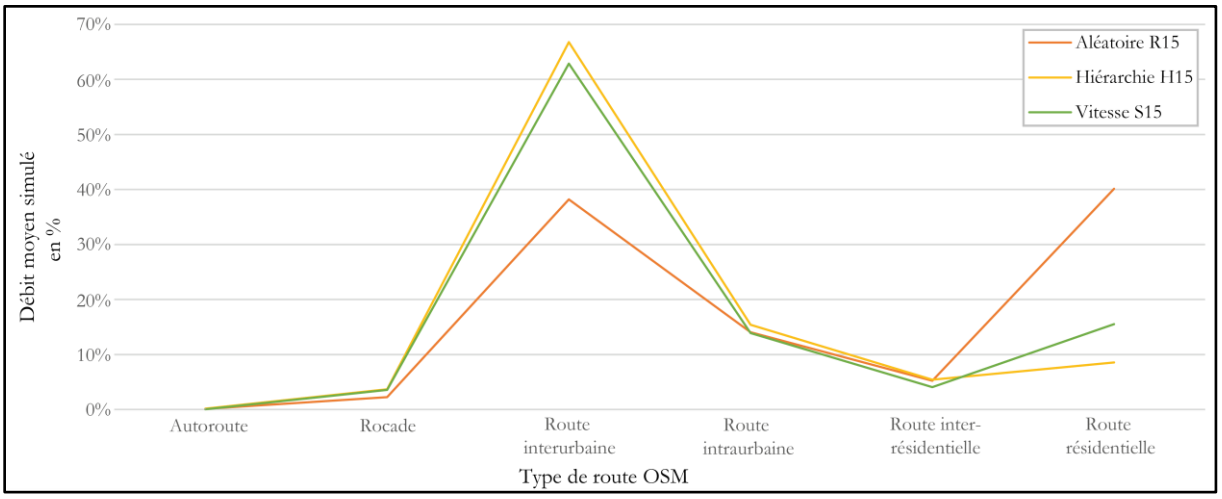



**Figure 3.9 : Pourcentage de fréquentation des axes en fonction de la typologie OpenStreetMap pour les simulations établies à 15 minutes de stationnement (moyenne des débits simulés pour les 156 simulations des paramètres *R15*, *H15*, *S15*)**

Ces premiers résultats sont encourageants et permettent d'illustrer la capacité de SCAUP à généraliser un trafic routier sur un ensemble d'axes routiers. Néanmoins, il demeure de nombreux écueils à soulever quant à ses capacités à produire des mêmes conditions de trafic routier, en appréciant les indicateurs de reproductibilité de celui-ci, autant dans sa dimension spatiale que temporelle.

### 7.2.2. La reproductibilité spatiale de SCAUP : comment le trafic routier est-il simulé ?

L'analyse de reproductibilité de SCAUP vise à interpréter l'emphase (les caractères de sensibilité, de stabilité et d'instabilité) d'un paramètre de simulation, en relations avec les données simulées, pour les 156 réplications du modèle de trafic routier (*cf.* section 6.3.1, p. 205). Dans ce sens, cette première analyse procure l'opportunité d'apprécier la stabilité du modèle, à savoir sa capacité à reproduire les mêmes conditions de circulation routière. En d'autres termes, plus l'indice de reproductibilité est proche de 100 % plus les trafics routiers simulés sont stables, et inversement lorsque celui-ci est proche de 0 %. En ce qui concerne la reproductibilité globale des paramètres explorés, on peut noter en s'appuyant sur la figure 3.10 deux tendances. La première est à rattacher au temps de stationnement des véhicules lors de la simulation, tandis que la seconde se rattache plutôt aux paramètres de déplacement des véhicules (R, H et S) :

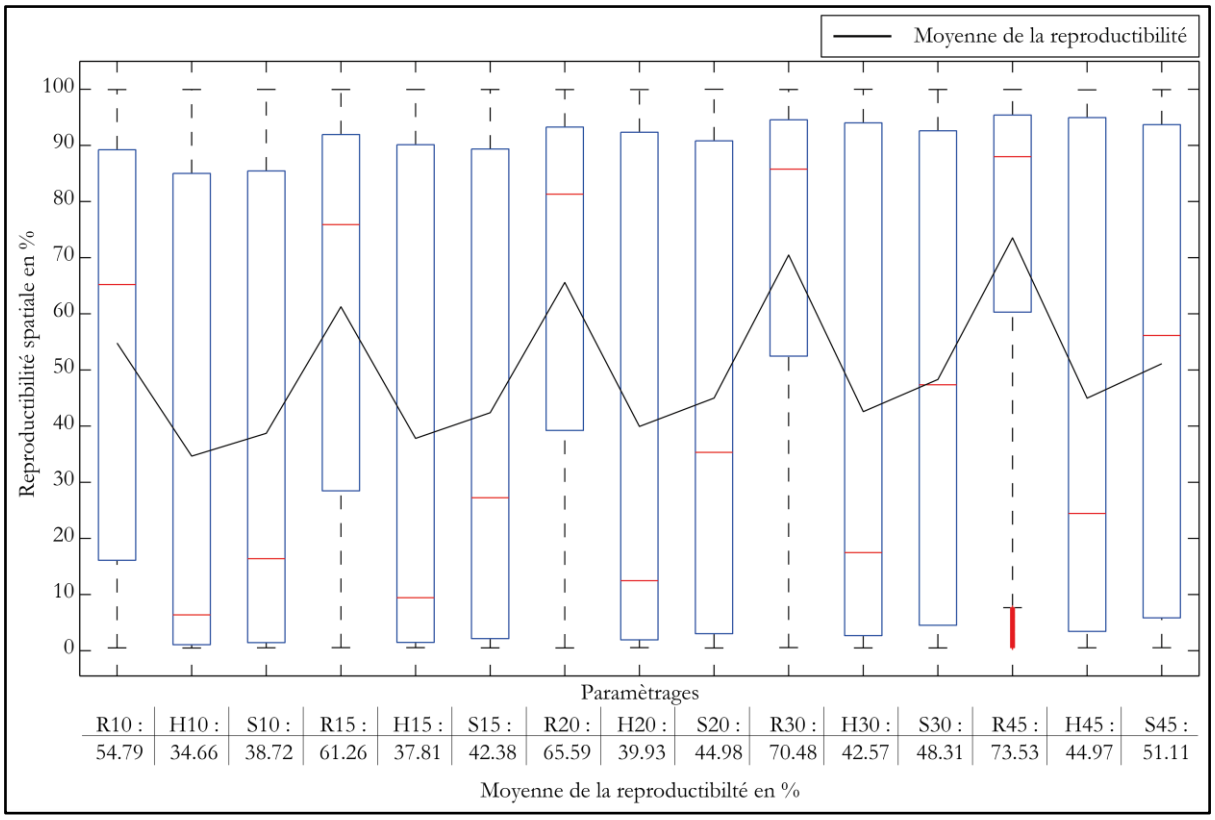



**Figure 3.10 : Répartition des indices de reproductibilité spatiale pour les 7 506 axes de la ville de Dijon sur l'ensemble des paramètres d'exploration**

Tout d'abord en ce qui concerne la reproductibilité moyenne[142], celle-ci évolue progressivement à mesure que le temps de stationnement des véhicules augmente (*cf.* figure 3.10). En effet, on constate que pour tous les différents tests de simulation, l'augmentation du temps de stationnement s'accompagne en parallèle d'un gain de reproductibilité d'environ 30 % lors du passage de 10 minutes à 45 minutes de stationnement. Ensuite, on remarque des différences de reproductibilité entre les trois paramètres de déplacement des véhicules (Aléatoire, Hiérarchie et Vitesse) que l'on peut relier aux sorties cartographiques présentées en figure 3.11 (ci-dessous). En

[142] Illustrée par la courbe noire en figure 3.10, la reproductibilité moyenne correspond à la moyenne des indicateurs de reproductibilité pour les 7 506 axes de la ville de Dijon.

effet, l'indice de reproductibilité moyen présente une bien meilleure qualité lorsque les déplacements des véhicules sont aléatoires (66 % en moyenne) que lorsque les déplacements sont contraints en fonction de la hiérarchie du réseau routier (en moyenne 40 % pour la hiérarchie et 45 % pour la vitesse). Cette affirmation se reflète au sein de la figure 3.10, où le positionnement du 1er quartile et de la médiane (illustrée en rouge) rattachés aux paramètres aléatoires (R) présente une structure générale plus resserrée que pour les deux autres paramètres (H et S).

Si ces résultats tendent à montrer une meilleure qualité des simulations lorsque le déplacement des véhicules est aléatoire, spatialement la figure 3.11 tend à nuancer ces premiers propos, lorsque l'on localise ces informations sur le réseau routier. Cette seconde illustration montre la répartition spatiale de l'indice de reproductibilité sur chacun des axes. En confrontant les deux cartes, le principal enseignement de ces deux sorties *brutes* tend à montrer que si la reproductibilité des simulations demeure relativement acceptable, on distingue deux structures spatiales : lors du déplacement aléatoire des véhicules (en haut de la figure 3.11), la répartition spatiale des indices présente une forme radioconcentrique dans le sens où plus on s'éloigne du centre-ville, plus l'indice de reproductibilité diminue ; alors que lorsque les déplacements des véhicules sont contraints (en bas de la figure 3.11), l'indice de reproductibilité tend plutôt à faire ressortir la structure hiérarchique du réseau routier.

À ce niveau on constate que les meilleurs taux de reproductibilité sont rattachés aux axes concentrant le plus de trafic routier, tandis que les taux les plus faibles s'observent sur les axes résidentiels ainsi qu'en marge du réseau. On observe alors un lien entre le paramétrage des choix routiers et les indices de reproductibilité, et donc *a fortiori* entre les indices de reproductibilité et les débits de véhicules simulés. Ce dernier constat s'illustre parfaitement lorsque l'on met en relation les données de trafic simulées (en abscisses) et les indices de reproductibilité (en ordonnées), comme ci-dessous (*cf.* figure 3.12). Il est à noter, que la même relation est retrouvée pour l'ensemble des 15 paramètres explorés sur la ville de Dijon (*cf.* annexe cartographique 2, p. xv). En d'autres termes, plus le débit simulé augmente plus la reproductibilité du modèle de simulation est bonne. En outre, on remarque un lien avec le positionnement des capteurs électromagnétiques situés sur les axes les plus performants. On observe que l'emplacement d'un capteur est aussi à relier aux trafics routiers simulés les plus importants et les plus reproductibles. Cet état des lieux tend à montrer que lorsque le trafic routier est important, le comportement du modèle est stable. Ceci s'explique en partie par la localisation des stations de comptages sur les axes routiers à forte circulation (*cf.* figure 1.28, p. 96), où le modèle est de fait calé sur ces points stratégiques. En revanche, pour les débits faibles, le modèle présente une composante chaotique plus importante. Le *bruit* important associé à ces débits est sans doute lié au grand nombre d'axes qui concentrent des débits faibles (*cf.* section 3.3 , p. 102).

En définitive, malgré son caractère stochastique le modèle SCAUP présente une réelle stabilité lorsqu'il faut produire un trafic routier émergent sur les axes les plus fréquentés du terrain d'étude. Notons que les incertitudes rattachées aux débits simulés sont le plus généralement reliées aux axes de plus faible catégorie, où aucune donnée d'observation n'a pu être intégrée.

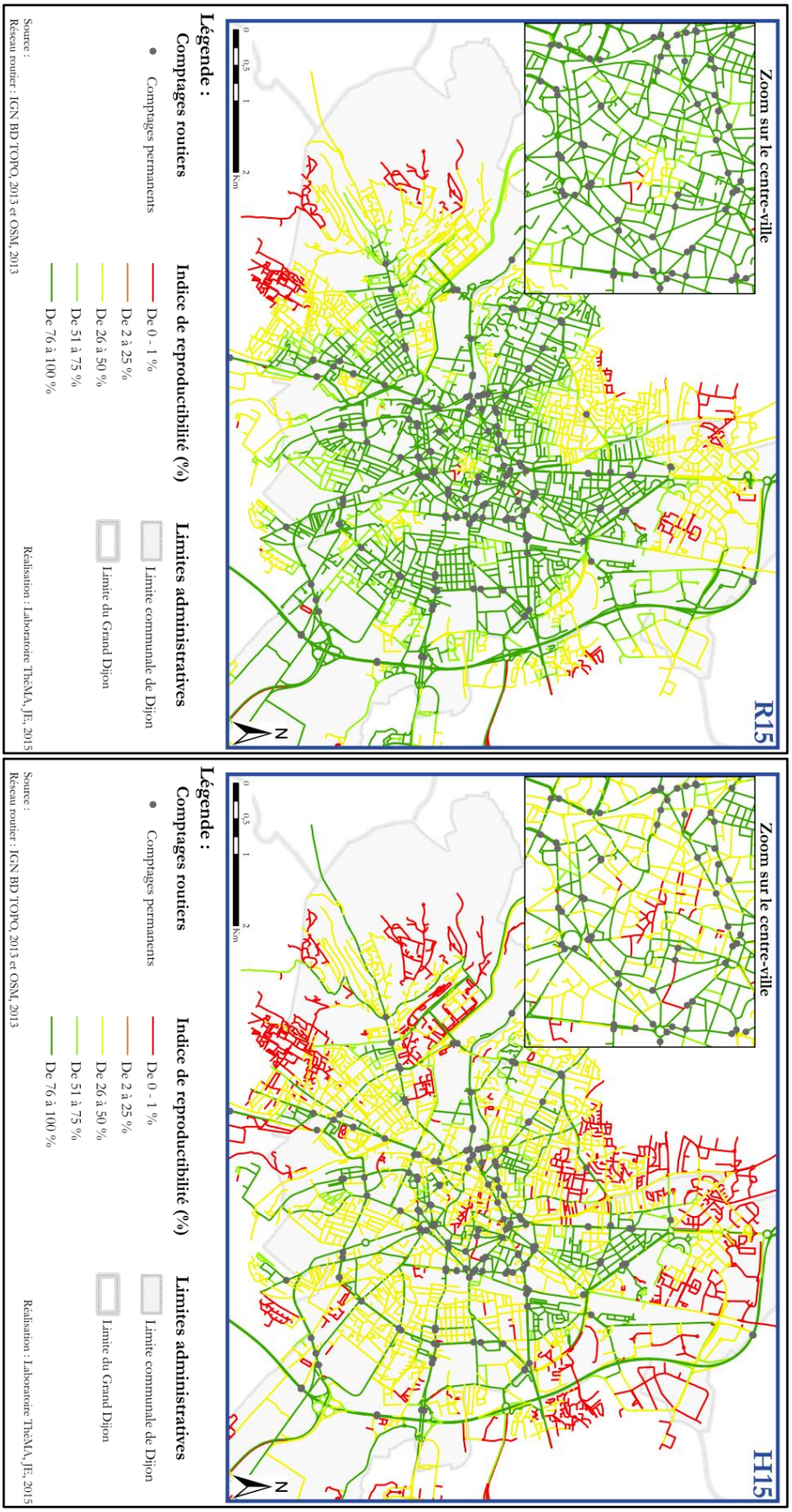



**Figure 3.11 : Reproductibilité spatiale des 156 simulations pour les paramètres de stationnement définis à 15 minutes, pour les choix routiers aléatoire (R : en haut) et hiérarchique (H : en bas)**

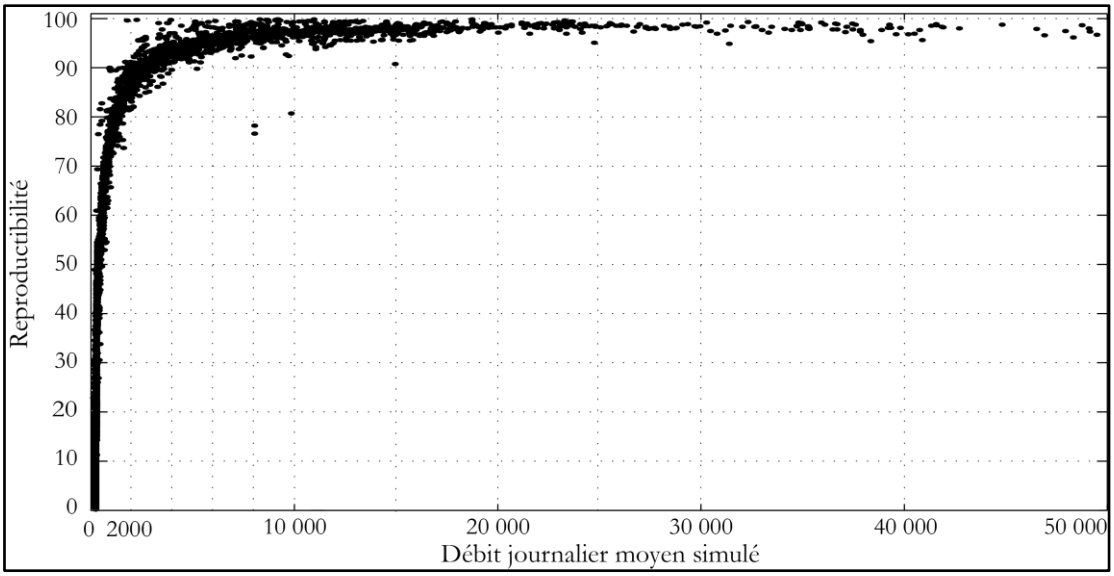



**Figure 3.12 : Relation entre le débit simulé et la reproductibilité pour chaque axe du domaine d'étude selon le paramètre *H15***

Comment interpréter ces différents constats ? Tout d'abord, on retrouve un résultat proche du constat établi lors de l'analyse de la reproductibilité du modèle KISS (*cf.* p. 210) : le modèle de simulation montre une relative stabilité lorsque le trafic simulé est important, et c'est particulièrement le cas sur les axes possédant une hiérarchie plus importante. En revanche, lorsque le trafic simulé est faible, voire très faible, le trafic routier simulé est plus aléatoire, voire disparate, et ne présente pas les mêmes caractéristiques d'une réplication à l'autre. Rappelons que les résultats présentés précédemment visent à apprécier les capacités et la sensibilité de SCAUP en se focalisant exclusivement sur les différentes simulations et leur emphase, comparées entre-elles. Pour le moment, ces premiers résultats, même s'ils permettent de faire un constat de l'opérationnalité de SCAUP au niveau de sa stabilité et de sa capacité à produire une carte de charge du trafic routier sur un plus grand ensemble de tronçons, ne permettent pas d'évaluer les estimations de trafic routier obtenues par la simulation au regard des données réelles.

### 7.2.3. **La fréquentation intra-journalière : analyse du signal temporel**

Ici, nous nous focaliserons sur les capacités du modèle SCAUP à conserver mais aussi à reproduire le signal temporel intra-journalier issu des stations de comptages routiers[143]. Tout d'abord, on constate que pour l'ensemble des tests de simulation réalisés, la moyenne de fréquentation au quart d'heure suit la forme de la courbe intra-journalière du trafic routier des données d'observation (*cf.* figure 3.3, p. 230). Que ce soit pour l'un ou l'autre des tests de simulation, la même tendance est produite, dans le sens où l'on observe une montée de la fréquentation routière lors des périodes de pointe (matin, midi, soir), et inversement lors des périodes creuses. Ce premier résultat montre la capacité du modèle à conserver, mais aussi à généraliser le signal issu des stations de comptages sur un plus grand ensemble d'axes routiers[144]

[143] Notons que pour bien appréhender les résultats de simulation au cours de cette section, seuls les axes ne possédant aucune station de comptage ont été sélectionnés : soit 156 axes/7 506 axes routiers de la ville. De même que les axes possédant des valeurs nulles (*cf.* figure 3.8, p. 230) ne sont pas pris en compte au cours de cette section.

[144] À savoir que les valeurs de trafic routier nulles ne sont pas prises en compte lors du calcul de la moyenne de la fréquentation routière. Inversement, si ces valeurs étaient prises en compte la tendance moyenne serait inversée par rapport résultats exposés en figure 3.13.

(*cf.* figure 3.3, p. 230). Il faut noter que les valeurs de débit présentées ci-dessous sont plus faibles que celles présentées pour les 156 stations comptages (*cf.* figure 3.3, p. 230) car nous avons considéré ici l'ensemble des axes routiers qui ne sont pas associés à un capteur.

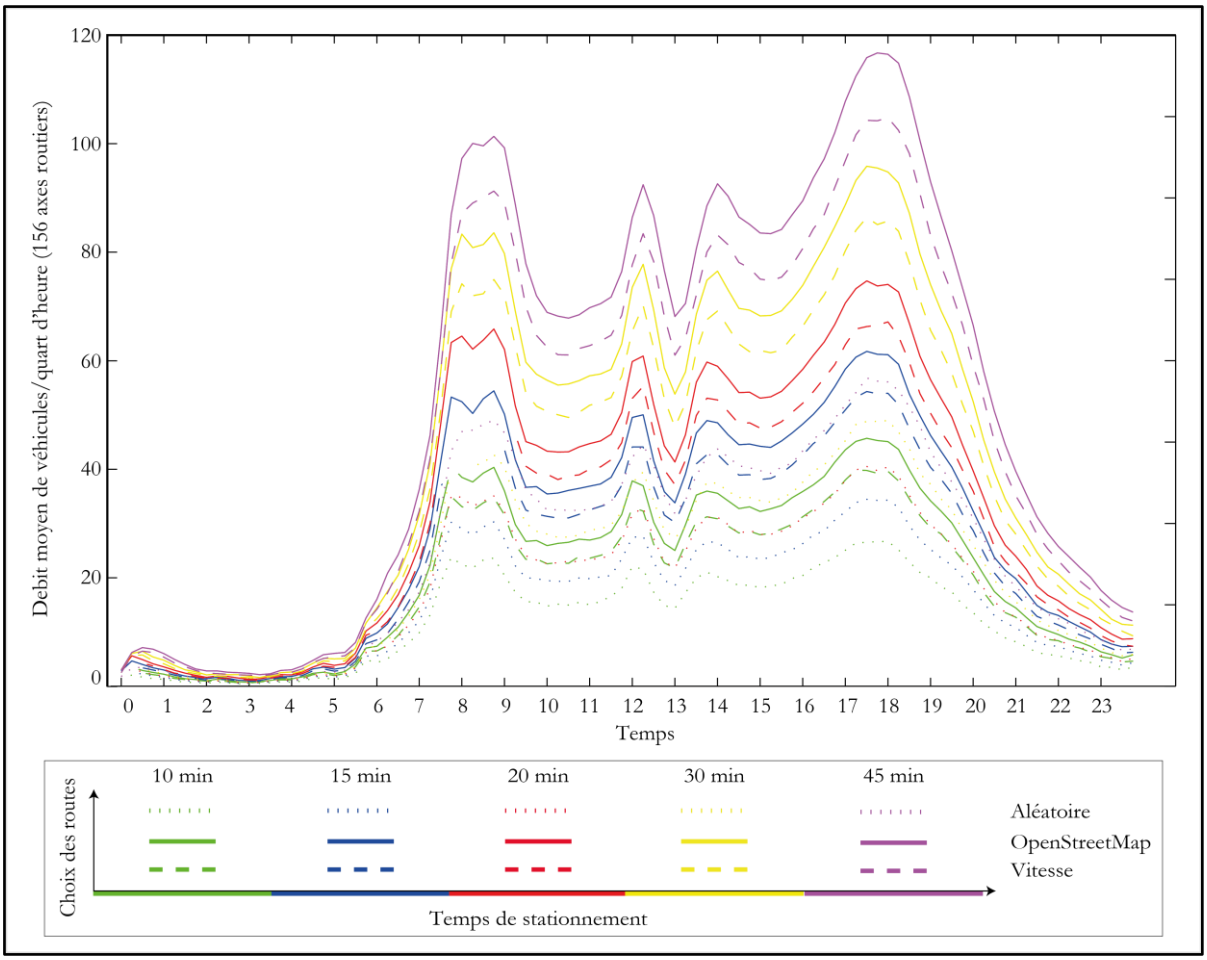



**Figure 3.13 : Moyenne de la fréquentation intra-journalière du trafic routier pour les différents paramètres de SCAUP explorés sur la ville de Dijon (hors axes possédant une station de comptage routier : 7 350 axes routiers)**

Dès lors, si la forme de la courbe apparait relativement identique pour l'ensemble des paramètres testés, les résultats se distinguent en termes de débit moyen simulé. De surcroît, on retrouve le même constat tel qu'exposé précédemment. En effet, en s'appuyant sur la figure 3.13, on observe que les débits simulés se distinguent, selon le temps de stationnement, qui fait augmenter progressivement le nombre de véhicules en déplacement à chaque pas de temps, et selon les paramètres de choix routiers, avec un débit simulé moyen qui est moins important lorsque les véhicules se déplacent aléatoirement (illustré en pointillés) que lorsque le mouvement de ceux-ci est contraint en fonction de la hiérarchie du réseau routier. À titre indicatif, la différence entre le débit simulé avec un déplacement aléatoire (R) et le débit simulé avec un déplacement hiérarchique (H) est de 6 % pour un temps de stationnement de 10 minutes, et de près de 20 % lorsque le stationnement est de 45 minutes.

Ce dernier constat s'explique assez naturellement lorsque l'on observe à travers la figure 3.6 (p. 236), pour le paramètre aléatoire (R), que le débit de véhicules simulé est relativement homogène sur l'ensemble de l'espace d'étude. À l'inverse, lorsque le déplacement des véhicules est contraint

par la hiérarchie du réseau routier, les débits simulés sont bien mieux hiérarchisés (*cf.* figure 3.7, p. 237). En d'autres termes, on retrouve pour ces deux contraintes de déplacement (H et S) des axes concentrant une part plus importante du trafic routier et d'autres où aucun véhicule n'est passé. En conséquence, comparativement au déplacement aléatoire des véhicules, la moyenne de la fréquentation routière aura plutôt tendance à augmenter lorsque le mouvement des véhicules est contraint par le réseau routier (*cf.* figure 3.13, ci-dessus). De ce fait, plus le temps de stationnement est important, plus les véhicules simulés se déplacent longtemps sur le réseau routier, augmentant naturellement les valeurs des débits moyens. À titre indicatif, en s'appuyant sur l'illustration ci-dessus, le débit moyen simulé à chaque pas de temps (le quart d'heure) est multiplié par trois entre un temps de stationnement défini à 10 minutes et un temps de stationnement de 45 minutes.

La figure 3.14 (ci-dessous) reflète quant à elle une seconde tendance en se focalisant plutôt sur une comparaison temporelle des réplications réalisées pour analyser l'emphase des différents tests de simulation à travers l'indice de reproductibilité[145]. Cette seconde analyse vise à s'assurer que SCAUP conserve bien le signal porté par les stations de comptage quel que soit le paramètre testé. En d'autres termes, cette analyse vise à observer si les courbes intra-journalières suivent toujours les mêmes temporalités. Si les valeurs de reproductibilité demeurent acceptables (supérieures à 77 %), à travers la figure 3.14, on constate tout d'abord deux périodes distinctes, entre le début de simulation (de 00:00 à 06:30) et le reste (de 06:30 à 23:45), et ensuite une différence selon que les véhicules se déplacent aléatoirement (R) ou qu'ils sont contraints (H et S). À travers cette illustration, on observe une différence, non plus à relier au temps de stationnement, mais plutôt au niveau des conditions de déplacement des véhicules (R, H, S).

La distinction entre les deux périodes, illustrée ci-dessous, parait être en adéquation avec le constat établi lors des tests sur l'environnement de simulation KISS présenté précédemment (*cf.* figure 2.35, p. 211). Ce premier élément montre bien l'impact que les stations de comptage ont sur la dynamique de SCAUP, dans le sens où ce sont les capteurs qui en portent la temporalité, en jouant un rôle non seulement sur les conditions initiales mais aussi sur la dynamique de simulation. Ce constat peu surprenant est à relier au dispositif de quantification mobilisé, dont la vocation vise à initialiser la dynamique de la circulation routière. Ici on retrouve le caractère empirique et inductif de la démarche de simulation entreprise. Effectivement, comme nous l'avons envisagé au cours des précédentes sections dans une optique de *data driven*, la temporalité et la dynamique de SCAUP reposent en très grande partie sur les données de débit mesurées par les capteurs électromagnétiques utilisés pour l'initialisation de celui-ci. Ainsi en insérant les véhicules en adéquation avec les données de comptages et au fil de l'eau, la dynamique de la circulation routière parait être correctement reproduite par le modèle, avec une période de « charge (00:00-06:30) » du réseau routier.

Pendant cette période, les débits de véhicules étant plutôt faibles, les conditions de reproductibilité sont plus aléatoires, et donc moins en phase que lors de la période où le réseau routier est chargé. Ce constat s'atténue lors de la seconde période (06:30-23:45). Néanmoins, on observe bien une alternance de la reproductibilité qui diminue entre les périodes de pics et de

[145] L'analyse de l'indice de reproductibilité pour la partie spatiale se focalisait sur le débit/véhicules axe par axe. Ici pour la partie temporelle, nous nous focalisons sur le débit de véhicules moyen des axes/pas-de-temps (le quart d'heure).

creux (*cf.* figure 3.14). Pour appuyer ces observations, on peut noter que le trafic routier simulé est d'autant plus reproductible que le nombre d'agents en déplacement sur le réseau routier est important. À titre indicatif et en s'appuyant sur la figure 3.12 (p. 242), on observe que dès lors que le nombre d'agents dépasse les 50 véhicules simulés par axe, l'indice de reproductibilité présente une bonne qualité. En d'autres termes, une *taille critique* de 50 véhicules simulés permet alors d'obtenir des données de simulations de relativement bonne qualité :

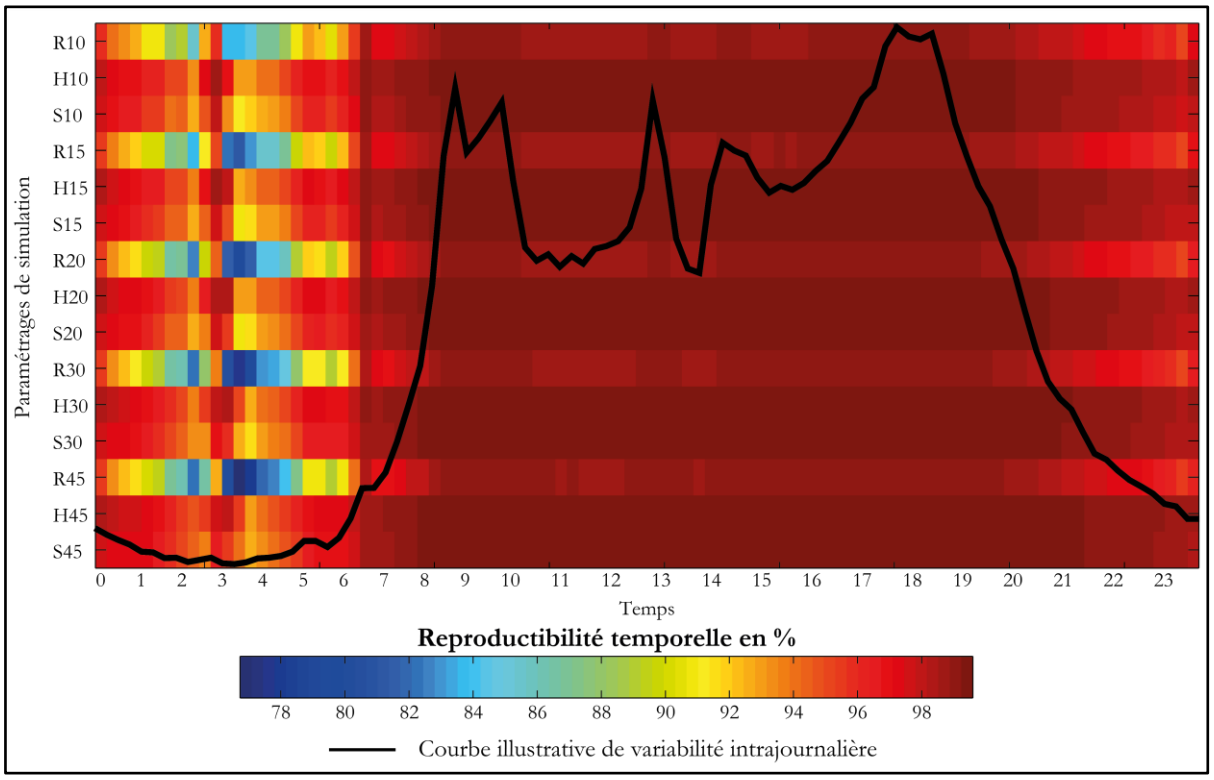



**Figure 3.14 : Reproductibilité temporelle moyenne des différents paramètres de simulation (en ordonnées) au pas de temps quart d'heure (en abscisses)**

Cette distinction entre réseau routier chargé et non chargé s'illustre parfaitement à travers notre échantillon test (R), lorsque les véhicules se déplacent aléatoirement sur l'espace d'étude. En effet, comme illustré en figure 3.14, on constate que le taux de reproductibilité est bien plus faible que lorsque le déplacement des véhicules est appréhendé en fonction de la hiérarchie du réseau routier (H et S). À ce niveau, le caractère aléatoire du modèle de simulation s'illustre autant au niveau spatial (*cf.* section 7.2.2, p. 239) que temporel. À l'inverse pour les paramètres H et S, lors de cette seconde phase la reproductibilité temporelle des simulations présente une stabilité de meilleure qualité que pour le paramètre aléatoire (R). En d'autres termes, chaque simulation établie avec ce caractère aléatoire (R) présente des caractéristiques qui lui sont propres, avec une temporalité qui est moins bien conservée qu'avec les autres paramètres (H et S), à savoir que les courbes intra-journalières ne commencent pas au même moment.

Cette seconde section nous a permis d'appréhender le fonctionnement interne de la simulation du trafic routier avec SCAUP, en nous focalisant dans un premier temps sur une comparaison des différents tests d'exploration présentés en tableau 3.1 (p. 227). À travers les indices de reproductibilité, cette seconde section a permis d'évaluer et de tester la sensibilité du modèle à produire un trafic routier émergent à l'échelle d'un territoire réel, en se focalisant non seulement

sur chacun des paramètres testés, mais aussi en confrontant l'ensemble des réplications entre-elles. En somme, cette analyse vise à apporter des éléments quant à la sensibilité du modèle SCAUP, en se concentrant spécifiquement sur son fonctionnement interne. De surcroît, la confrontation entre l'échantillon de simulation test (R) et les autres échantillons a permis de révéler les apports de la matrice de transition en vue de contraindre le déplacement des véhicules sur le réseau routier. Une fois l'état des lieux du fonctionnement interne à SCAUP mis en place, la troisième phase du plan d'exploration (*cf.* figure 3.1, p. 225) vise à valider et évaluer ses sorties en comparaison des valeurs de terrain observées.

## 7.3. Comment choisir la meilleure simulation : de la simulation à la *réalité*

L'ensemble de ce travail a pour ambition d'évaluer et de valider les estimations issues des différents tests de simulation explorés pour la journée du mardi 9 septembre 2008 ; pour ensuite sélectionner le(s) paramètre(s) de simulation qui s'approchent le plus de ce qui a été observé par les capteurs fixes de la ville de Dijon. À ce niveau, les questionnements tournent autour des méthodes et des indices de validation. Les résultats qui seront présentés ici sont effectués en adéquation avec la méthode de validation croisée LOOCV (*cf.* section 7.1.3.b, p. 233). Dans ce sens, les résultats seront exclusivement illustrés au niveau des stations (156 valeurs, *cf.* figure 3.5, p. 234), seuls endroits où les confrontations entre les données d'observation et les données de simulation issues SCAUP sont possibles pour réaliser une évaluation complète de son comportement.

### 7.3.1. L'analyse statistique classique pour mesurer les relations entre données de simulation et données d'observation

Lorsque l'on aborde les questionnements liés à la validation et à l'évaluation en statistique bivariée classique, les deux indicateurs qui viennent à l'esprit pour mesurer l'intensité d'une relation statistique entre deux variables quantitatives, telles que les valeurs observées, d'une part, et les valeurs simulées, d'autre part, sont le coefficient de corrélation (noté R) et le coefficient de détermination (noté $R^2$). Si ces indicateurs présentent de nombreuses limites et nécessitent, par exemple, de leur associer un test de significativité, ils font aujourd'hui largement consensus dans la communauté scientifique, non seulement de par leur simplicité de calcul, mais aussi d'interprétation. De surcroît, lorsque le $R^2$ est employé pour la validation de données de simulation face à des données d'observation, il permet de visualiser rapidement et synthétiquement les qualités ainsi que les défauts du modèle de simulation.

La figure 3.15 (ci-dessous) présente les coefficients de détermination (en ordonnées) calculés pour les 156 stations de comptage, en lien avec chacun des paramètres de simulation (en abscisse), sous la forme de *boxplots*. Cette figure permet d'observer la répartition de l'ensemble des $R^2$ pour chaque station de comptage[146]. Les débits simulés sont dans leur ensemble d'une très bonne qualité, et présentent pour l'ensemble des $R^2$ calculés une moyenne de 0,876[147]. De plus,

[146] Pour visualiser l'ensemble des valeurs du $R^2$, ainsi que les valeurs de débits associées à une représentation cartographique au niveau de capteurs, nous renvoyons le lecteur vers 0 (p. vii).

[147] Cet indicateur a peu de sens car il combine l'ensemble des $R^2$ représentés sous la figure 3.15 : il permet juste d'apporter une grandeur synthétique de l'ensemble des paramètres explorés.

pour chacun des paramètres testés, le $R^2$ moyen n'est jamais inférieur à 0,84 (*cf.* tableau sous la figure 3.15). Ce second constat s'illustre aussi sous la forme des boxplots, dont les frontières comprises entre le 1er et le 3e quartile s'échelonnent entre 1 et 0,8 de $R^2$. Malgré tout, à travers cette figure de synthèse, on identifie difficilement des différences entre les divers paramètres testés, ce premier bilan tendant à montrer que le trafic routier simulé présente, quel que soit le cas, de faibles différences par rapport aux données d'observation (en termes de variance et de signal) :

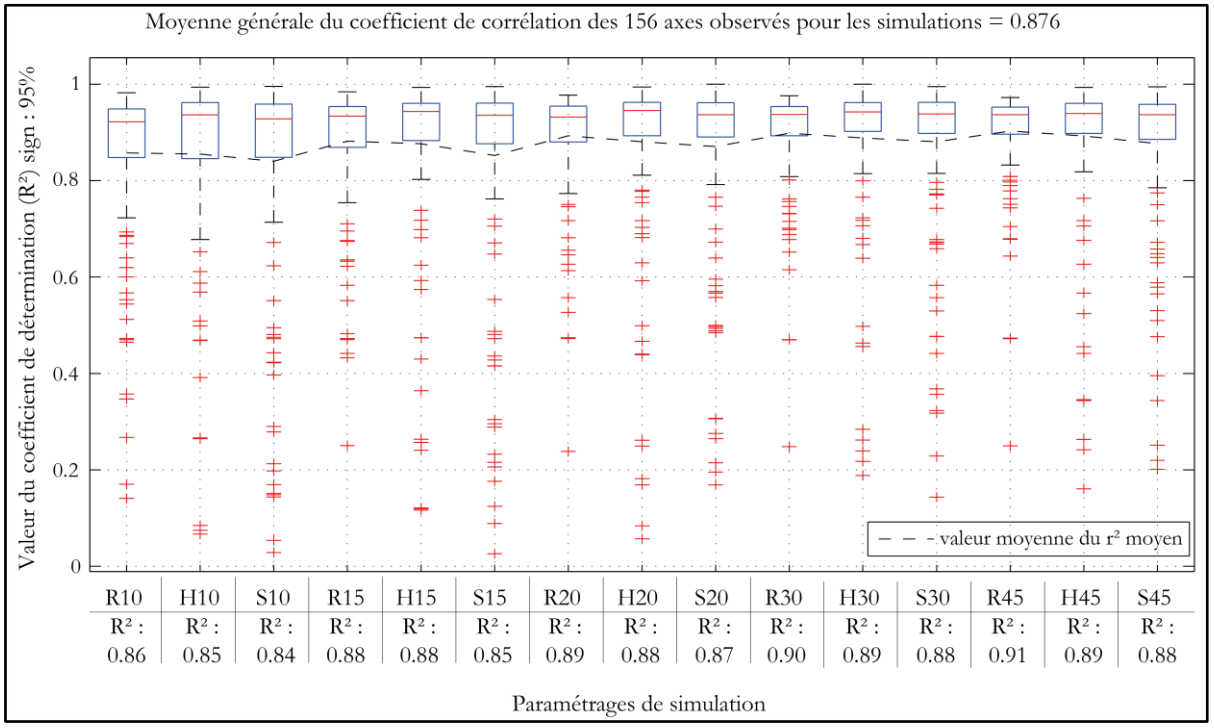



**Figure 3.15 : Coefficients de détermination ($R^2$) pour les 156 stations de comptage (en ordonnées) et pour l'ensemble des paramètres explorés (en abscisses) : le tableau associé à l'abscisse correspond à la moyenne des 156 coefficients de détermination calculés**

Il est alors nécessaire de bien séparer ces éléments en vue de mieux appréhender les sorties du modèle SCAUP. Pour cela, une analyse complémentaire a été mise en place en étudiant la relation directe entre les valeurs d'observation et les valeurs de simulation, autant dans l'espace (*cf.* figure 3.16, p. 248) que dans le temps (*cf.* figure 3.17, p. 249). En d'autres termes, l'aspect spatial vise à se focaliser sur le nombre total de véhicules passant en un point donné (le capteur électromagnétique), alors que l'aspect temporel vise plutôt à se focaliser sur le nombre moyen de véhicules présents par quart d'heure pour l'ensemble des capteurs électromagnétiques.

Cette seconde analyse permet d'établir un bilan plus contrasté des valeurs de trafic simulé, non seulement dans l'espace, mais aussi dans le temps. Avant tout, notons que pour l'ensemble des 15 paramètres testés, les relations entre débits simulés et débits observés présentent des formes très proches[148], c'est pourquoi pour cette section nous présentons les résultats pour un seul paramètre : *H15*. En effet, en s'appuyant sur ces résultats (p. 248), deux tendances ressortent : la première, spatiale (*cf.* figure 3.16, p. 248), montre que SCAUP a généralement tendance à sous-estimer le

[148] L'ensemble des graphiques de régression linéaire pour les 15 paramètres testés sont présentés en annexe 5 (p. ix). Pour cette section, seuls les résultats issus du paramètre *H15* sont présentés.

nombre de véhicules simulé présent en un point donné ; alors que la seconde tend à nous montrer qu'il reproduit correctement, voire parfaitement, la dynamique de la circulation routière mesurée au niveau des stations de comptages routiers.

Pour ce qui est de l'analyse spatiale (*cf.* figure 3.16, p. 248), on remarque qu'il n'existe pas de relation linéaire entre les valeurs d'observation et les valeurs de simulation. Cette tendance se caractérise au sein de la figure par une forte dispersion des points (correspondant aux stations) et un coefficient de détermination proche de 0 et significatif ($R^2$ de 0,064). La relation entre observation et simulation par station de comptage ne permet pas de faire ressortir une réelle tendance entre les valeurs de simulation et d'observation et donc de relation linéaire. Nous nous intéresserons maintenant au couple de données observation/simulation, et plus spécifiquement au nombre de stations où le trafic simulé est sous-estimé par rapport aux observations (en nombre de véhicules cumulés), comme synthétisé en tableau 3.2 :

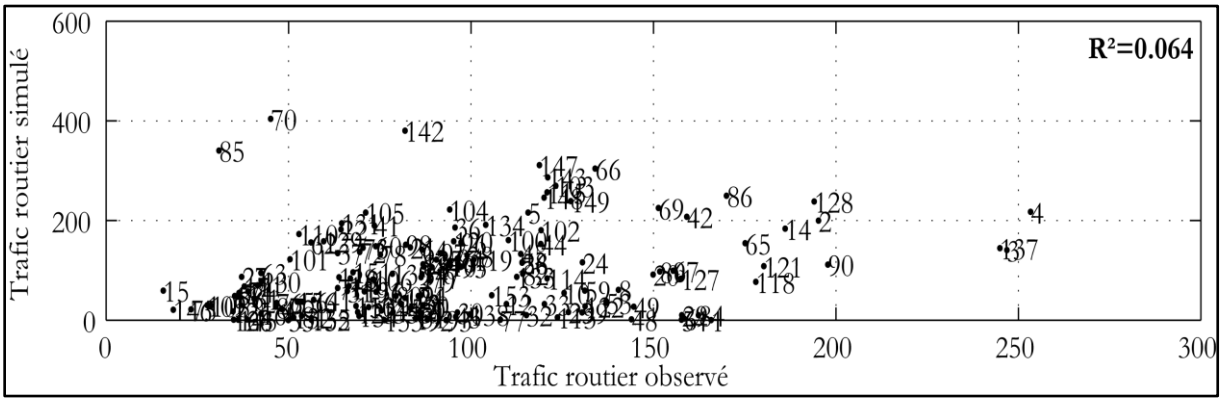



**Figure 3.16 : Relation entre le nombre de véhicules simulés et observés par station de comptage pour les 156 stations d'observation : paramètre *H15* (les points correspondent aux stations)**

On peut alors constater que le pourcentage de stations dont le cumul de véhicules simulés est inférieur au cumul de trafic observé varie entre 56,4 % et 65,4 % (*cf.* tableau 3.2). En somme, dans la majorité des cas (supérieurs à 50 %), le modèle de simulation a tendance à produire un trafic routier plus faible que celui observé au niveau des stations. À ce niveau, on observe que le temps de stationnement n'impacte que très peu les résultats, de même que le choix de déplacement des véhicules routiers ne permet pas d'observer de réelles tendances en vue de distinguer un test des autres. Ceci montre que les résultats tiennent plutôt compte de la matrice de transition que des paramètres sélectionnés pour l'application du plan d'expérience (*cf.* tableau 3.1, p. 227) :

| | R10 | H10 | S10 | R15 | H15 | S15 | R20 | H20 | S20 | R30 | H30 | S30 | R45 | H45 | S45 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| % de stations sous-estimées | 65,4% | 56,4% | 58,3% | 61,5% | 57,7% | 59,0% | 63,5% | 58,3% | 60,3% | 65,4% | 59,6% | 59,6% | 64,7% | 60,9% | 62,2% |

**Tableau 3.2 : Pourcentage de stations où le trafic routier est sous-estimé par SCAUP (nombre de stations où le trafic routier simulé est inférieur au trafic routier observé)**

Si la qualité du trafic routier simulé par SCAUP n'apparait pas complétement satisfaisante sur le plan spatial, *a contrario* sur le plan temporel une autre tendance ressort (*cf.* figure 3.17, p. 249) :

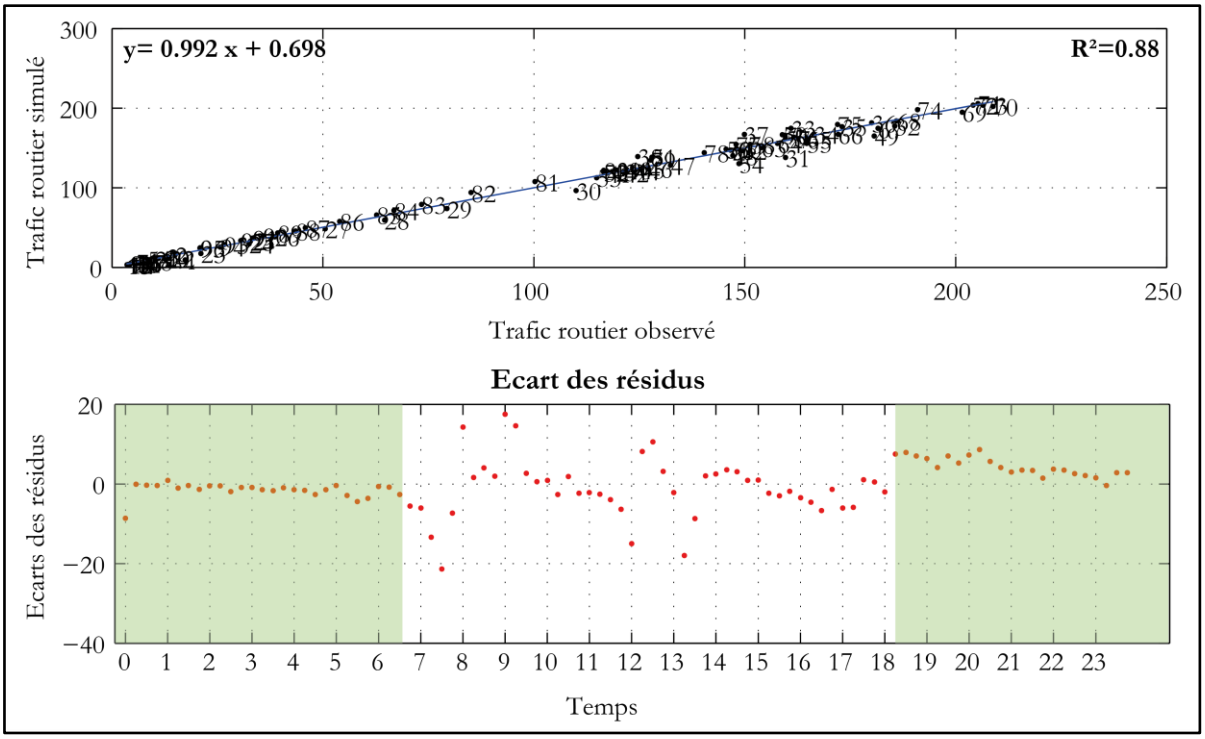



**Figure 3.17 : Régression linéaire (en haut) et analyse des résidus (en bas) en moyenne temporelle des 156 stations d'observation (Nb de véhicules moyen/Quart d'heure) : paramètre *H15* (les points correspondent au quart d'heure de la journée)**

On observe ci-dessus une relation linéaire entre les valeurs de trafic routier observées et simulées de très bonne qualité ($R^2$ de 0,88 avec un coefficient directeur proche de 1). Ce second état des lieux sur le plan temporel, c'est-à-dire en se focalisant sur le nombre de véhicules par quart d'heure apparait correctement produit par SCAUP. Cependant, en se focalisant sur l'analyse des résidus (en bas de la figure 3.17, p. 249), on remarque deux périodes distinctes, illustrées par les zones en verts :

- 1. les périodes de faible circulation (*de 00:00 à 06:00 et de 18:30 à 00:00*), où l'écart aux résidus est quasiment nul (proche de 0)
- 2. les périodes de forte circulation (*de 06:30 à 18:30*), où l'écart aux résidus est relativement plus important que pour la première période (supérieur ou inférieur à 0). Malgré tout, ce constat est à nuancer dans le sens où l'écart aux résidus demeure très faible et n'excède pas un écart de plus de 60 véhicules sur l'ensemble des paramètres explorés (*cf.* annexe 5, p. ix).

Au final, ce résultat montre que le modèle SCAUP a tendance à correctement simuler la dynamique de la circulation routière, quelles que soient les conditions de circulation (faible ou forte). Pour finir, on remarque deux résultats distincts lors de cette première phase d'évaluation des résultats de simulation : un premier constat sur le plan spatial qui s'avère relativement discutable, tandis que sur le plan temporel le second constat est bien plus acceptable et convenable pour l'ensemble des paramètres explorés. Par ailleurs, appréhender la validation du modèle de simulation du trafic exclusivement par cet indicateur (le $R^2$) ne permet pas de révéler l'hétérogénéité des différents tests de simulation mis en place, alors que les analyses précédemment exposées sur le fonctionnement de SCAUP tendent à nous montrer plusieurs

différences entre les tests établis. Ceci reflète bien notre besoin de combiner plusieurs indicateurs pour confronter observation et simulation. C'est pourquoi, nous nous sommes dans un second temps tourné vers un indicateur dont l'application est plus spécifiquement rattachée à l'ingénierie du trafic routier, en vue d'étudier le calage entre données de simulation et données d'observation (Barceló, 2010) : le GEH (initiale de Geoffrey E. Havers) .

### 7.3.2. Un indicateur synthétique issu de l'ingénierie du trafic routier pour l'évaluation des modèles de trafic : l'indicateur GEH

Le GEH est un indicateur adapté pour l'évaluation des modèles de simulation du trafic routier à une échelle microscopique (Barceló, 2010 ; Casas *et al.*, 2010), et donc adapté à notre problématique de travail avec le modèle SCAUP. Cet indicateur tire son nom de son créateur, Geoffrey E. Havers (GEH), qui le mit en place au cours des années 1970. C'est un indicateur spécifique au trafic routier, permettant d'observer à la fois les écarts absolus et relatifs entre le débit observé et le débit simulé à chaque station de mesure. Le calcul du GEH permet de fournir une vue d'ensemble de la qualité des simulations du trafic routier à chaque station de comptage routier (Barceló, 2010). La formule est la suivante, pour chaque station de comptage routier (ou $x_i$ correspond à la valeur de simulation et $y_i$ à la valeur d'observation) :

$$GEH_i = \sqrt{\frac{2(x_i - y_i)^2}{x_i + y_i}}$$

L'indicateur est ensuite interprété par rapport aux observations comme suivant : lorsqu'il est inférieur à 5 les estimations sont excellentes, lorsqu'il est situé entre 5 et 10 elles sont bonnes, et quand il est supérieur à 10 elles sont jugées comme moyennes. En somme, si les valeurs sont excellentes dans 85 % des cas, le modèle est jugé comme très acceptable (Barceló, 2010).

Le tableau 3.3 (ci-dessous) illustre le pourcentage des indicateurs GEH calculés (excellents, bons et moyens) pour les 15 paramètres explorés lors du plan d'expérience. En complément, la figure 3.18 ( p. 253) illustre les valeurs du GEH au niveau de chaque capteur pour le paramètre *R15* (en haut) et *H15* (en bas)[149]. En s'appuyant sur ces éléments, on observe plusieurs tendances. On remarque, tout d'abord, qu'aucun des tests de simulation explorés ne présente un pourcentage de *GEH excellent* dans 85 % des cas, permettant de considérer le modèle de SMA comme très acceptable. Toutefois, en excluant les simulations établies pour un stationnement de 45 minutes, on peut observer que les valeurs du GEH sont le plus généralement (dans 70 % des cas) *bonnes*, voire *excellentes* (*cf.* tableau 3.3). On observe d'ailleurs une répartition du GEH assez équilibrée, quel que soit le paramètre choisi.

Dans ce sens, les pourcentages de valeurs du GEH sont hiérarchisés selon la relation allant d'un indicateur Excellent à Moyen, en passant par Bon. Par ailleurs, dans la plupart des cas, le pourcentage de GEH excellent est supérieur à 40 %, ce qui est particulièrement le cas des paramètres *H10*, *S10*, *H15*. Ces constats généraux nous permettent de juger le modèle SCAUP comme relativement acceptable, même s'ils tendent à montrer que des améliorations peuvent être envisagées pour affiner les résultats de simulation.

[149] L'ensemble des valeurs du GEH cartographiées sont proposées en annexe cartographique 3 (p. xxiii).

C'est au niveau du temps de stationnement que l'on remarque le plus de différences : lorsque le temps de stationnement des véhicules est compris entre 10 minutes et 20 minutes, les tests de simulation qui présentent le pourcentage le plus faible de *GEH moyen* sont associés aux contraintes de déplacement des véhicules sur le réseau routier (*H10*, *S10*, *H15*, *S15*, *H20*, *S20*). Au-delà, c'est-à-dire lorsque le temps de stationnement est supérieur à 20 minutes, la tendance s'inverse et ce n'est que lorsque les véhicules se déplacent aléatoirement (R) que le taux de *GEH moyen* est le plus faible. Cette tendance s'observe d'autant plus lorsque l'on se focalise sur le temps de stationnement à 30 minutes. On observe à ce niveau que les pourcentages les plus acceptables sont rattachés au paramètre *R20*, alors que pour les autres paramètres la qualité du GEH a plutôt tendance à se dégrader.

| Stationnement 10 min | | | |
|---|---|---|---|
| Choix routiers | R | H | S |
| Excellent | 37,82 | 47,44 | 46,15 |
| Bon | 33,33 | 30,13 | 32,69 |
| Moyen | 28,85 | 22,44 | 21,15 |

| Stationnement 15 min | | | |
|---|---|---|---|
| Choix routiers | R | H | S |
| Excellent | 42,95 | 44,87 | 41,03 |
| Bon | 31,41 | 35,90 | 41,67 |
| Moyen | 25,64 | 19,23 | 17,31 |

| Stationnement 20 min | | | |
|---|---|---|---|
| Choix routiers | R | H | S |
| Excellent | 42,31 | 40,38 | 42,31 |
| Bon | 36,54 | 39,10 | 41,03 |
| Moyen | 21,15 | 20,51 | 16,67 |

| Stationnement 30 min | | | |
|---|---|---|---|
| Choix routiers | R | H | S |
| Excellent | 38,46 | 41,03 | 41,03 |
| Bon | 42,95 | 34,62 | 31,41 |
| Moyen | 18,59 | 24,36 | 27,56 |

| Stationnement 45 min | | | |
|---|---|---|---|
| Choix routiers | R | H | S |
| Excellent | 40,38 | 35,26 | 33,97 |
| Bon | 41,03 | 33,33 | 35,26 |
| Moyen | 18,59 | 31,41 | 30,77 |

| | Valeurs de classes |
|---|---|
| Excellent | GEH <5 |
| Bon | 5<GEH<10 |
| Moyen | GEH >10 |

**Tableau 3.3 : Pourcentage des indicateurs du GEH compris entre 0 et 5 (excellents), entre 5 et 10 (bons) et supérieurs à 10 (moyens) pour l'ensemble des paramètres de simulation**

Ce premier tableau de répartition des valeurs du GEH permet de discriminer et d'affiner notre analyse en dégageant quelques tendances entre les différents tests de simulation. Parallèlement, il est important d'analyser les valeurs du GEH au niveau de chaque capteur car, à travers le tableau 3.3, on peut remarquer que la qualité du calage et de l'adéquation des données de trafic routier simulé et observé est de meilleure qualité lorsque le mouvement des véhicules est contraint en fonction de la hiérarchie du réseau routier. La figure 3.18 (p. 253), qui représente la cartographie du GEH au niveau des stations de comptage, permet d'apporter des précisions sur la qualité des résultats de simulation. Nous comparons les simulations établies pour un temps de stationnement défini à 15 minutes, d'une part lorsque les véhicules se déplacent aléatoirement (*R15* en haut), et d'autre part lorsque les véhicules se déplacent en fonction de la hiérarchie du réseau routier (*H15* en bas).

Nous pouvons remarquer que les stations de comptage qui présentent le meilleur pourcentage de GEH *excellent* et *bon* ne présentent que peu de logique spatiale lorsque les véhicules se déplacent aléatoirement (*R15*). En revanche, lorsque le mouvement des véhicules est contraint par la hiérarchie du réseau routier (*H15*), on observe que les pourcentages de GEH de bonne qualité les plus importants sont le plus généralement situés sur les axes les plus significatifs de la ville de Dijon. En d'autres termes, en contraignant le mouvement des véhicules au niveau du réseau routier, on améliore l'adéquation entre trafic routier simulé et trafic routier observé sur les axes concentrant le plus de trafic routier. Ce dernier constat montre que lorsque les véhicules se déplacent en fonction des contraintes locales, et donc ne modifient pas leur positionnement aléatoirement, le trafic routier simulé présente une meilleure adéquation au contexte local. Dans ce sens, le modèle SCAUP apparait bien sensible à ce paramètre.

Pour terminer, l'évaluation des données de trafic simulé par SCAUP à travers le GEH nous a permis de distinguer différents éléments par rapport à l'analyse précédente (*cf.* section 7.3.1, ci-dessus), nous incitant d'exclure l'ensemble des simulations définies avec un temps de stationnement de 45 minutes (ce qui est conforme aux constats établis au sein de la figure 2.38). De surcroit, elle nous a permis de sélectionner quelques paramètres tels que *H10*, *S10* ou *H15*. En revanche, au même titre que l'analyse par régression linéaire présentée précédemment (*cf.* section 7.3.1, p. 246), de nombreux résultats semblent assez similaires. Par ailleurs, l'indicateur GEH étant purement empirique, comme le soulève Jaume Barceló (2010, p. 47) « *il n'y a ni théorie derrière la méthode, ni de critères théoriques pour déterminer les seuils de valeurs qui sont purement empiriques et donc basées sur la pratique* », il est à manipuler avec une certaine prudence lors de l'évaluation d'un modèle de trafic routier.

Finalement, en exploitant les indicateurs présentés jusqu'ici (coefficient de détermination et GEH), il est difficile de conclure à une évaluation précise du comportement de SCAUP, de même qu'il est difficile de choisir avec certitude les simulations optimales pour l'évaluation de la pollution atmosphérique automobile (PAA). C'est pourquoi, nous nous sommes tourné vers un troisième indicateur synthétique pour l'évaluation des performances du modèle SCAUP, à travers le diagramme de Taylor.

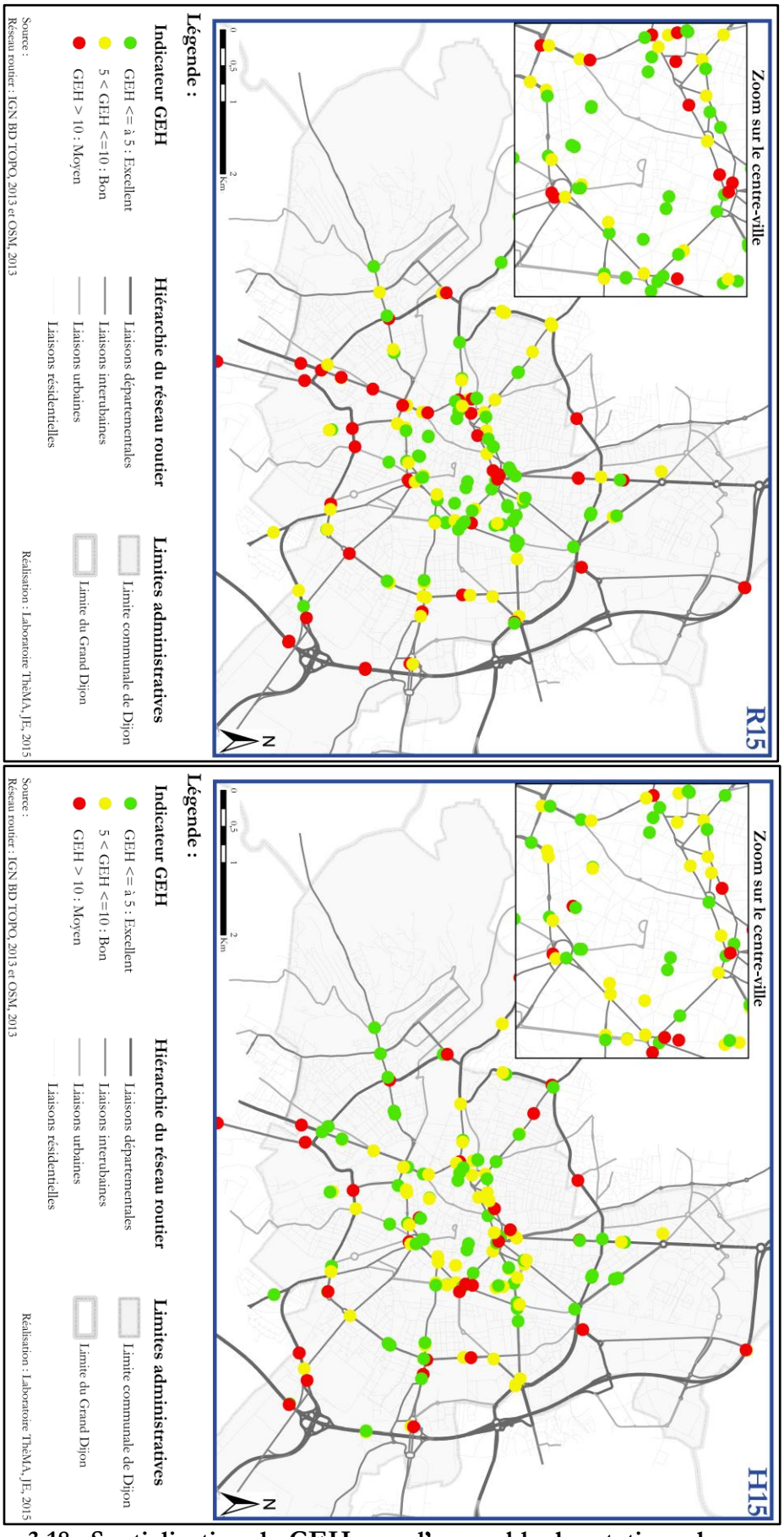



**Figure 3.18 : Spatialisation du GEH pour l'ensemble des stations de comptage**

### 7.3.3. **Un indicateur synthétique pour l'évaluation des modèles : le diagramme de Taylor (2001)**

Le diagramme de Taylor (2001) permet d'apporter une information synthétique pour visualiser et pour sélectionner les simulations qui se rapprochent le plus significativement des données d'observation en se focalisant autant sur les dimensions spatiales que temporelles du modèle de simulation. Ces questionnements sont depuis longtemps abordés en géographie physique par les climatologues lorsqu'il faut évaluer de la fiabilité d'un modèle numérique de climat face à un nombre réduit d'observations pour juger de la qualité des résultats de ce modèle. Dans un article fondateur intitulé « *Résumer les multiples aspects de la performance des modèles dans un seul diagramme* » [150] , Karl E. Taylor (2001, p. 7183) introduit ce problème pour les modèles numériques de climat : « *pour les modèles avec une multitude de variables et de multiples dimensions, la comparaison des champs simulés et observés devient impraticable, même si seulement une petite fraction de la sortie du modèle est considérée. Il est alors nécessaire soit de se concentrer sur certains aspects limités du système étudié, soit de mobiliser des indicateurs statistiques pour quantifier les correspondances entre ce qui est modélisé et ce qui est observé* ».

Ces questionnements de fond sont facilement transposables à notre travail, et d'une manière plus générale lors de la validation des modèles utilisés en géographie. De plus, rappelons que les méthodes d'évaluation et de validation des SMA constituent un enjeu de recherche important. Le diagramme de Taylor est une représentation graphique permettant de réaliser une étude comparative des données de simulation, en les positionnant dans un espace géométrique en fonction de leurs relations avec les données d'observation. En d'autres termes, le diagramme permet de résumer, sous une même représentation, le degré de correspondance entre les données de simulation et les données d'observation dites de référence (Taylor, 2001). De surcroît, en utilisant cette représentation synthétique, il est possible de positionner les différentes réplications d'un même paramètre[151]. Cela procure un atout considérable lors de l'interprétation des résultats de simulation, qui sont nombreux, en représentant sous une même figure un grand nombre d'informations, liées autant à l'analyse de sensibilité du modèle (réplication, réponse du signal…) qu'aux éléments rattachés à l'évaluation et à la validation de celui-ci.

Au sein de cette section, nous exploitons le diagramme de Taylor en vue non seulement d'évaluer et de valider les données de trafic routier simulé à partir de SCAUP, mais aussi de sélectionner les sorties les plus caractéristiques pour ensuite estimer les niveaux de PAA à l'échelle de la ville de Dijon. À travers ce diagramme, illustré en figure 3.19 (p. 256), le positionnement des données de simulation (en couleurs) est obtenu en tenant compte de trois indicateurs statistiques :

- 1. **L'écart-type** des échantillons mobilisés (en noir) permet de définir le positionnement radial des données de simulation (en couleurs) et des données d'observation (en noir) ;
- 2. **Le coefficient de corrélation** avec les observations (en bleu) permet de définir le positionnement azimutal des différents points ;
- 3. **L'erreur quadratique moyenne centrée** (RMS, en vert) permet de définir la distance entre les observations et les simulations proportionnellement aux valeurs de la RMS (qui définit la magnitude d'erreur constante entre simulation et observation).

[150] Le titre anglophone de l'article est : « *Summarizing multiple aspects of model performance in a single diagram* ».
[151] Dans le cas ici présent, les réplications sont toutes positionnées au même endroit, ce qui explique qu'on ne peut observer les réplications (156) associées aux paramètres testés.

L'interprétation des résultats s'opère en observant, tout d'abord, la position des données de simulation (en couleurs) face aux données d'observation (en noir) ; puis ensuite on interprète le positionnement relatif des données de simulation entre elles, en se focalisant sur leur position face aux trois indicateurs représentés.

Sur la figure 3.19, on observe d'abord des tendances conformes aux observations précédentes : sur le plan spatial (a), les résultats de la simulation (en couleurs) sont éloignés des valeurs observées ; alors que (b) sur le plan temporel les deux échantillons présentent de fortes similarités. Par ailleurs, au contraire des résultats présentés précédemment, on observe intuitivement que la projection des données de trafic simulé à travers le diagramme de Taylor apporte des informations complémentaires, puis supplémentaires, aux analyses précédentes (coefficient de détermination et GEH).

On remarque que la disposition générale des données de simulation, en a et b, suit la même tendance et présente plutôt une symétrie angulaire. Cette disposition générale est à rattacher à la valeur du coefficient de corrélation, qui impacte la disposition angulaire des données : sur le plan spatial (a), les simulations de trafic sont peu significatives, avec un coefficient compris entre 0 et 0,3 ; alors que sur le plan temporel (b), les valeurs du coefficient sont de très bonne qualité, avec des valeurs comprises entre 0,95 et 1. Ces résultats confortent le constat établi lors de l'application des régressions linéaires au cours de la section 7.3.1 (p. 246). En d'autres termes, si on ne simule pas correctement le bon nombre de véhicules passant sur le réseau routier (a), SCAUP reproduit correctement la dynamique intra-journalière du trafic routier portée par les capteurs (b). Ce premier résultat montre deux tendances :

- Les paramètres testés (stationnement et déplacement) impactent plutôt la dynamique portée par la SMA, alors qu'ils ne jouent un rôle que très minime sur la répartition des véhicules routiers. À ce niveau, l'amélioration de ces aspects apparait plutôt reliée aux paramètres de la matrice de transition (*cf.* section 6.2.2.b, p. 198).
- Dans l'optique d'une approche guidée par les données, la SMA conserve correctement la temporalité portée par les capteurs[152]. Cette tendance illustre la capacité du modèle à conserver et à généraliser les pulsations de trafic routier insérées par l'intermédiaire des capteurs électromagnétiques.

[152] Il faut néanmoins relever que l'analyse de la fréquentation intra-journalière est réalisée en observant la moyenne de la fréquentation par quart d'heure, cette analyse ayant tendance à lisser les valeurs absolues de fréquentation des axes routiers en annihilant le bruit observé sur l'ensemble des axes de l'espace d'étude. En d'autres termes, il est important de manipuler ces résultats avec prudence.

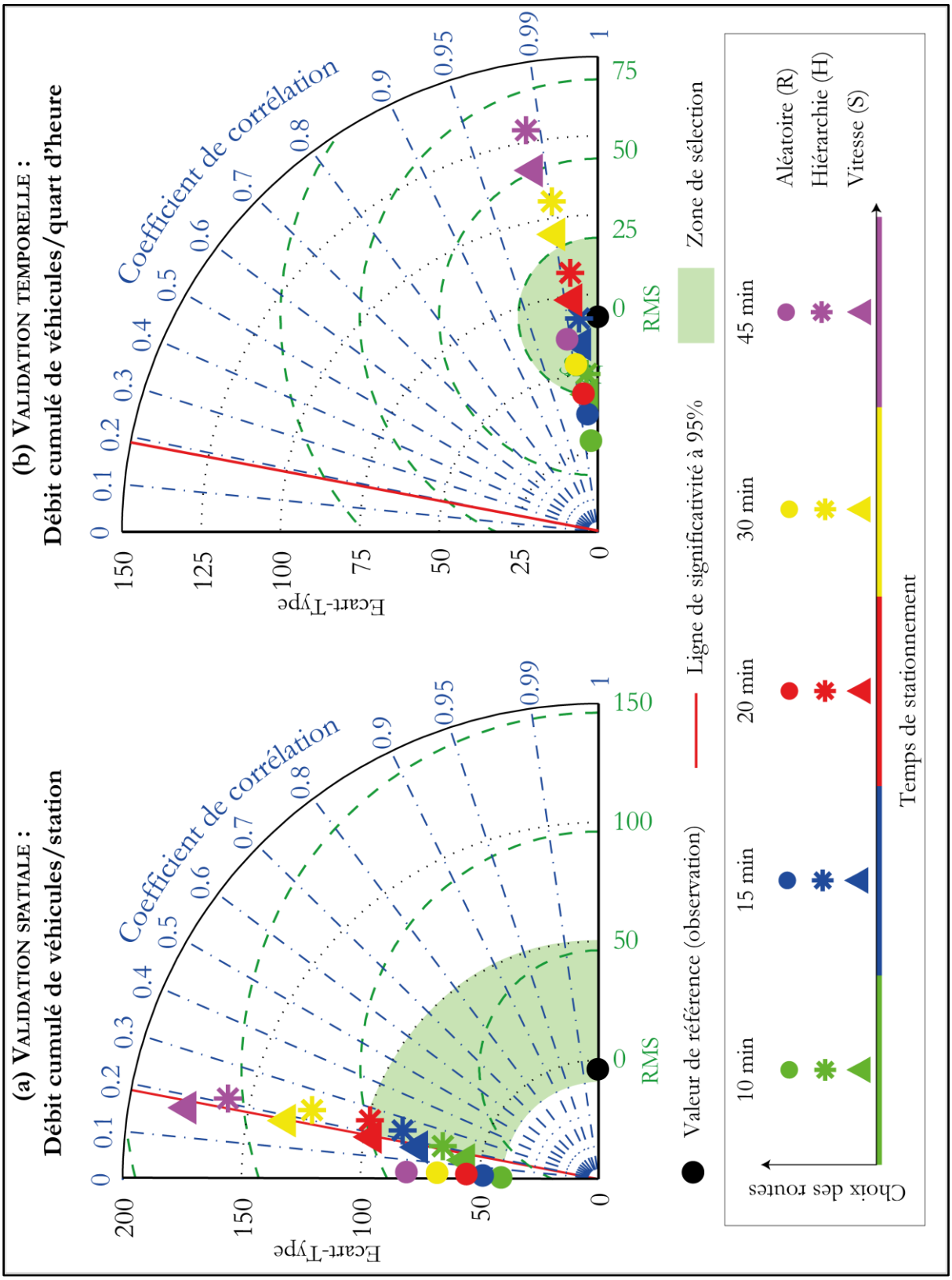



**Figure 3.19 : Diagramme de Taylor (2001) de la validation spatiale (a) et temporelle (b) du plan d'exploration du modèle SCAUP**

***La coordonnée radiale indique l'écart-type des débits routiers (véhicule/quart d'heure) ; la coordonnée angulaire correspond aux corrélations linéaires avec les observations ; la distance entre les observations et chaque membre est proportionnelle à la RMS calculée***

Si les valeurs du coefficient de corrélation ne montrent que très peu l'impact des paramètres sur les valeurs de trafic routier simulé, on remarque que ceux-ci impactent plutôt les valeurs de l'écart-type et de la RMS. Après avoir interprété la distance des points de simulation avec l'observation, ici, nous nous focalisons plutôt sur la répartition de ces points entre eux (en couleurs sous la figure 3.19). Au premier abord, on observe deux groupes qui se distinguent selon les paramètres de déplacement des véhicules :

- les simulations établies avec un mouvement aléatoires des véhicules (R), et qui constituent l'échantillon test de la démarche exploratoire
- les simulations établies avec une contrainte associée au déplacement des véhicules (H et S) ; au sein de ce groupe, on observe que chaque point de simulation suit pour un même temps de stationnement une logique de couple (H et S).

Ainsi, si les paramètres testés ne permettent que très peu d'améliorer la répartition des véhicules sur le réseau routier en termes de coefficient de corrélation, ils présentent néanmoins une sensibilité qui diffère des paramètres aléatoires, dans le sens où le second groupe se positionne au-delà de la ligne de significativité (illustrée en rouge en figure 3.19). Ces éléments nuancent quelque peu les propos généraux établis précédemment et on observe bien un « *raffinement* » des données de trafic simulé lorsque le déplacement des véhicules est contraint. Ce constat s'observe à l'intérieur du second groupe (entre les couples H et S). En effet, en s'appuyant sur la partie spatiale (a) de la figure 3.19, on observe une alternance de ces couples, qui passent la ligne de significativité pour les paramètres H, alors que pour le paramètre S (associé à la vitesse réglementaire) la significativité des points de simulation n'est pas acceptable. Dès lors, on observe bien une réponse des paramétrages de mouvement des véhicules sur le positionnement angulaire (coefficient de corrélation) des différents tests.

Cependant, si l'on s'intéresse maintenant au temps de stationnement, on observe que celui-ci est d'autant plus important que la distance entre simulation et observation est importante, et ce quelle que soit la représentation (*cf.* figure 3.19). Ce résultat n'est pas surprenant dans le sens où le temps de stationnement joue un rôle sur le nombre de véhicules se déplaçant en un instant donné sur le terrain d'étude. On note que pour le premier groupe (R) la dispersion des points est relativement minime, avec un écart-type ainsi qu'une RMS qui ne varient que très peu entre les différents points. Le second groupe (H et S) présente quant à lui une dispersion qui est de deux à trois fois plus importante que pour le premier groupe. On observe que la position des points de simulation pour chaque paramètre se disperse en fonction du temps de stationnement, c'est-à-dire que plus le temps de stationnement augmente, plus l'écart-type ainsi que la RMS augmentent proportionnellement. En d'autres termes, le temps de stationnement joue un rôle sur le nombre de véhicules présents sur les infrastructures, et impacte naturellement l'écart-type et la RMS. C'est-à-dire que le trafic routier sur l'ensemble des axes se disperse beaucoup plus autour de la moyenne en hiérarchisant et hétérogénéisant le trafic routier simulé sur les axes par rapport à moyenne (écart-type), augmentant proportionnellement l'erreur moyenne (RMS) du modèle. On observe d'ailleurs qu'à partir d'un temps de stationnement supérieur à 20 min, l'écart-type et la RMS augmentent significativement. Ces propos confortent l'analyse établie lors de l'expérimentation de SCAUP avec l'environnement de simulation KISS, dans le sens où un temps de stationnement supérieur à 30 minutes augmente l'écart-type et la RMS, ce qui impacte sensiblement le trafic routier simulé (*cf.* figure 2.38, p. 215).

L'ensemble des informations apportées par le diagramme de Taylor permet d'interpréter plus finement la sensibilité de SCAUP à produire un trafic routier émergent plus proche des observations *in situ*. Ces premiers constats illustrent bien les apports du diagramme pour la problématique de ce travail. À l'inverse des résultats précédents, on observe différentes tendances nous permettant d'évaluer plus finement le comportement du modèle à produire un trafic routier émergent depuis les données de comptages routiers. La représentation des données de simulation de SCAUP au sein du diagramme de Taylor (*cf.* figure 3.19) permet ainsi de mieux visualiser et donc de hiérarchiser les paramètres de simulation testés et expérimentés.

La dernière phase d'exploitation du diagramme vise à sélectionner les données de trafic routier les plus caractéristiques pour ensuite les utiliser dans une perspective environnementale : l'évaluation de la PAA. Pour cette partie, nous nous concentrerons sur la zone verte illustrée au sein de la figure 3.19. Cette zone a été définie en fonction du premier quartile des échantillons projetés, en tenant compte pour la partie spatiale (a) de l'écart-type (de 0 à 50), et pour le plan temporel (b) de la RMS (de 0 à 25). Les critères de sélection s'opèrent ensuite en trois temps, en se focalisant d'abord sur l'analyse spatiale (a), puis ensuite sur l'analyse temporelle (b), et enfin sur la combinaison des deux illustrations[153].

Dans un premier temps, pour la partie spatiale (a) différents critères ont été mis en place, au premier rang desquels l'écart-type (projection radiale) se présente, de prime abord, comme un critère important à prendre en compte. Les données de trafic routier simulé sont sélectionnées en excluant, tout d'abord, l'ensemble des simulations situées sous la ligne de significativité du coefficient de corrélation. On sélectionne ensuite les points de simulation situés dans la zone verte, c'est-à-dire ayant un écart-type compris entre 45 et 100 véhicules qui se rapproche de l'écart-type des valeurs observées. En appliquant ce second filtre, cela permet de sélectionner les simulations qui présentent des caractéristiques analogues ou proches (en termes d'écart-type) avec les données d'observation (46.22 d'écart-type en véhicules). Enfin, pour le troisième critère, on sélectionne les valeurs ayant la RMS la plus faible. En appliquant l'ensemble de ces filtres de sélection, seulement trois simulations répondent aux critères (*cf.* figure 3.19) : *H10, H15 et H20*.

Dans un second temps, sur la partie temporelle (b), le critère prédominant est la zone où la RMS présente les valeurs les plus faibles (*cf.* zone verte sous le figure 3.19). Ce choix s'explique aussi car le diagramme de Taylor sur cette partie présente une structuration idéale, avec un coefficient de corrélation de très bonne qualité, un écart-type proche des données d'observation (67.87 véhicules), et une RMS relativement faible, comprise entre 0 et 75 pour toutes les données de simulation[154]. Nous avons alors fait le choix de nous concentrer plus spécifiquement sur les magnitudes d'erreur comprises entre 0 et 25 (1$^{er}$ quartile des échantillons), afin de sélectionner les données de simulation les plus caractéristiques des données d'observation. En appliquant ce deuxième filtrage, aux simulations sélectionnées précédemment (*H10, H15 et H20*) se rajoutent les paramètres : *S10, S15 et S20*.

[153] À noter que les paramètres de choix aléatoire sont exclus naturellement de notre sélection. Pour rappel, ces données de simulation servent de référence pour appréhender les variations liées au modèle en soi, et celles liées à la mise en place des paramètres de stationnement et de répartition des véhicules routiers.

[154] La valeur de la RMS varie de 0 à l'infini. Elle n'est pas normée : 0 représente un cas idéal de validation, c'est-à-dire que la magnitude d'erreur d'une simulation est nulle.

En outre, à travers le paramètre *H15*, on peut identifier un positionnement quasiment idéal et parfait de ces données. En effet, la position des données de simulation est quasiment superposée aux données d'observation[155]. En d'autres termes, avec ce paramètre la dynamique portée par SCAUP est produite en adéquation aux données de comptage. Toutefois, cet élément de validation est à nuancer car le stationnement défini à 15 minutes est identique à la granularité temporelle des données de comptage (un comptage par quart d'heure), ce qui peut expliquer en partie son positionnement optimal. On retrouve ici les aspects inductifs et *deepening* de l'approche de simulation, dans le sens où les capteurs jouent un rôle prépondérant en supportant l'ensemble de la dynamique du système multi-agents de trafic routier. Néanmoins, ce résultat demeure intéressant car la dynamique de SCAUP est définie à la minute, de même que le temps de stationnement est défini à l'aide d'une fonction gaussienne (*cf.* figure 2.30, p. 202).

Enfin dans un troisième temps, la dernière phase de sélection des données de simulation s'effectue en croisant les données de simulation sélectionnées grâce aux deux filtres mis en place. Cette dernière phase est la plus simple car on sélectionne seulement les données de simulation concordantes aux deux filtrages, soit dans notre cas d'étude, les paramètres *H10, H15 et H20*. Ces trois paramètres sont ceux où les données de trafic routier simulé présentent le plus de similitudes avec les données d'observation, qui seront ensuite exploitées pour quantifier les émissions de PAA.

[155] C'est le cas idéal recherché lors de la validation des données de simulation à travers le diagramme de Taylor.

## Conclusion :

Les différents constats que nous avons pu établir sur les capacités du modèle SCAUP ont eu tendance à montrer les atouts d'une approche *deepening*, en démontrant ses capacités à conserver le signal temporel porté par les capteurs électromagnétiques. Par contre, nos difficultés se sont révélées lorsqu'il a fallu répartir ces débits observés sur l'ensemble des axes routiers. En somme, si la dynamique du système de trafic routier est relativement bien produite au sein du simulateur, les principaux écueils à soulever se situent plutôt au niveau de la répartition des véhicules sur le réseau routier. Ces résultats tendent à montrer les limites d'une approche de simulation du trafic routier sans matrice O-D permettant de correctement répartir les flux de trafic entre différentes zones. Toutefois si l'approche « matrice O-D » présente des avantages pour répartir les flux entre les différentes zones, elle présente des lacunes au niveau temporel, alors que notre approche, elle, présente plus particulièrement des forces au niveau de cette dimension. En d'autres termes, s'entourer des enquêtes O-D présente des avantages pour répartir les flux de trafic routier, mais des inconvénients pour appréhender la dynamique de la circulation routière.

À l'inverse, s'entourer des comptages routiers permet de mieux appréhender cet aspect temporel, et les limites se situent plutôt au niveau de la répartition spatiale des véhicules routiers. L'approche d'affectation associée à la matrice de transition, définie en fonction des caractéristiques circulatoires des infrastructures de transport (hiérarchie des axes) et des comptages routiers, apparait être une piste de travail intéressante pour affiner la répartition des véhicules au cours de la simulation, en vue de dépasser les limites intrinsèques aux enquêtes O-D. S'il demeure encore des résultats discutables quant à la répartition spatiale des véhicules, il semble néanmoins que visuellement on approche quand même la structure radioconcentrique du réseau routier dijonnais. De surcroît, tandis que l'approche par enquêtes O-D nécessite un grand volume de données, notre approche quant à elle présente un faible coût (en temps) pour obtenir une carte de charge du trafic routier. C'est aussi ici que se situe l'un des atouts du modèle SCAUP initialisé et calibré à partir des comptages routiers.

Toutefois, on comprend assez aisément les limites d'une approche guidée par les données, voire inductive, dans le sens où plus le nombre de capteurs est important plus la capacité du modèle à produire un trafic routier proche de la réalité du terrain est importante. C'est pourquoi la méthode de simulation a visé à toujours mobiliser un nombre optimal de capteurs pour l'initialisation de la simulation, en s'entourant de la méthode de validation croisée dite LOOCV (Leave One Out Cross Validation). Si cette approche présente des avantages en vue d'assurer de l'indépendance des échantillons de simulation et d'observation, il faut en revanche relever qu'il s'agit d'une méthode relativement gourmande en temps de calcul, qui n'est pas abordable par tous. En outre, ce chapitre a permis d'explorer les grands éléments d'évaluation du modèle SCAUP dans un contexte urbain réel, en mobilisant différentes approches issues autant de la statistique classique ($R^2$ et régression linéaire) et de de l'ingénierie du trafic (GEH), que de la climatologie (Diagramme de Taylor). Sur ce dernier élément, en nous appuyant sur le diagramme de Taylor nous avons pu voir que cette représentation permet, en comparaison avec les approches classiques, de mieux caractériser des ensembles de même nature, en vue de sélectionner les meilleures simulations. De surcroît, les diagrammes de Taylor présentent l'avantage de synthétiser des grands ensembles de données, comme dans notre cas, avec la représentation de plus de 2 300 tests de simulation, dont l'analyse par des outils classiques peut

apparaitre plus difficile, voire chronophage. Pour finir, en nous reposant sur cette représentation synthétique, nous disposons d'un outil qui peut justifier ou pas les différentes qualités du modèle SCAUP, en lui apportant une représentation iconographique relativement intuitive. L'ensemble de ce travail de sélection a aussi été réalisé afin de justifier et de communiquer à l'association de qualité de l'air locale (ATMOSF'AIR BOURGOGNE) les choix opérés pour la simulation du trafic, en sélectionnant les paramètres de simulation *H10, H15 et H20*. Cette partie de « *construction de la preuve* » (Fouillé *et al.*, 2012, p. 199) a été réalisée en vue de passer à une *évaluation environnementale* du trafic routier, à travers l'estimation de la pollution atmosphériques d'origine automobile (PAA).

Pour finir, il faut noter que le protocole d'*évaluation* présenté au cours de ce chapitre s'agence aussi comme une proposition opérationnelle pour l'*évaluation* des modèles de SMA, et plus particulièrement pour les approches de SMA guidée par les données, comme c'est le cas avec le modèle SCAUP. En effet, à travers la littérature, on peut remarquer que les démarches d'*évaluation* et d'*interprétation* des données de simulation issues des SMA constituent encore un enjeu de recherche important (Amblard *et al.*, 2006 ; Hassan *et al.*, 2010b ; Banos, 2013). Le diagramme de Taylor, en proposant une visualisation synthétique des résultats de simulation, apparait intéressant à exploiter. Effectivement, sous une même illustration il est possible de reporter de nombreux éléments *d'évaluation*, portant sur la sensibilité du modèle à différentes modifications (paramètres, environnement…), tout en y associant aussi les différentes réplications du modèle de SMA. À notre sens, cette représentation constitue une réelle force pour l'évaluation des modèles guidés par les données.

# Chapitre 8. UN DISPOSITIF D'ÉVALUATION ENVIRONNEMENTALE : DU TRAFIC ROUTIER AUX ÉMISSIONS DE POLLUANTS ATMOSPHÉRIQUES AUTOMOBILES

Les différents éléments d'évaluation présentés précédemment ont permis de faciliter la communication avec l'agence de qualité de l'air (ATMOSF'AIR BOURGOGNE), afin de choisir les données de trafic routier simulé à partir de SCAUP qui soient les plus proches de la « réalité » du territoire d'étude. Une fois ces choix opérés, la suite logique du travail vise à construire un dispositif d'évaluation de la pollution atmosphérique d'origine automobile (PAA), en s'inscrivant dans les référentiels mobilisés par les AASQA en vue de tester les capacités opérationnelles de notre approche. La continuité de ce travail vise maintenant à passer d'une information spatio-temporelle du trafic routier en milieu urbain à une connaissance des niveaux de PAA (*cf.* figure 1.6, p. 38). L'étude de la PAA demeure un enjeu important ; en effet, « *la question de l'impact des transports terrestres sur la qualité de l'air est au cœur des préoccupations des pouvoirs publics et des citoyens* » (Primequal-Predit, 2012, p. 9), comme les auteurs le relèvent en préambule de leur ouvrage de synthèse dédié aux dix années de recherches sur la pollution de l'air et les transports terrestres. Par exemple les actions mises en place lors des épisodes de mauvaise qualité de l'air vont tenter de réduire, voire de réguler temporairement les flux de trafic routier[156]. C'est typiquement le cas d'actions locales qui doivent viser à « *intensifier les contrôles de pollution des véhicules, limiter, voire interdire la circulation dans certains secteurs géographiques, comme les zones urbaines denses* […] ; *limiter le trafic routier des poids lourds en transit, voire les en détourner* […] ; a*baisser temporairement de 20km/h les vitesses maximales autorisées sur les voiries localisées dans la zone concernée par l'épisode de pollution, sans toutefois descendre au-dessous de 70 km/h* » (*Arrêté interministériel du 26 Mars 2014 relatif au déclenchement des procédures préfectorales en cas d'épisodes de pollution de l'air ambiant*, 2014).

L'ensemble de ces actions montre que la spatialisation des émissions de polluants atmosphériques est un travail primordial en vue de cibler les zones géographiques où des actions de réduction de la circulation routière doivent être menées. Pour les organismes en charge de la surveillance de la qualité de l'air (ASSQA), cette étape repose en partie sur l'élaboration d'un inventaire des émissions de polluants atmosphériques qui est ensuite cadastré, c'est-à-dire que les émissions sont spatialisées à l'échelle du territoire d'étude. Sa mise en œuvre s'inscrit au sein d'un référentiel national (PCIT : Pôle de Coordination des Inventaires Territoriaux) qui vise à formaliser et mutualiser leurs réalisations à l'échelle des territoires (PCIT, 2012). Ces inventaires territoriaux permettent ensuite d'alimenter les outils numériques de modélisation de la qualité de l'air, qui, eux, caractérisent les concentrations de polluants dans l'air (Primequal-Predit, 2012). Toutefois, le passage des données de trafic routier simulé à la spatialisation des émissions de PAA nécessite de faire appel à de multiples sources de données qui sont difficilement accessibles au grand public (méthodologie COPERT, Parc automobile, Démarrage à froid…). C'est pourquoi, le travail présenté au cours de ce chapitre a été réalisé en collaboration avec l'AASQA locale ATMOSF'AIR

[156] Le Monde du 8 Avril 2015 : « *Nouveau pic de pollution à Paris et nouvelle polémique* » : http://www.lemonde.fr/pollution/article/2015/04/08/nouveau-pic-de-pollution-a-paris-et-nouvelle-polemique_4611739_1652666.html#qMyCfrIVs4smmv6I.99.

BOURGOGNE[157]. En s'inscrivant dans ce contexte, comme nous l'avons vu plus haut (*cf.* section 5.3.3, p. 179), cela nécessite de nous conformer au formalisme employé par les AASQA en termes de données d'entrée à fournir. En revanche, cela permet de se placer dans une optique de généralisation et de tester l'opérationnalité de SCAUP en s'appuyant sur les référentiels utilisés sur l'ensemble du territoire français (PCIT, 2012), mais aussi européen (AEE, 2011, 2014).

Dans notre cas, le passage à la PAA s'opère en deux temps. La première phase permet de sélectionner les sorties du modèle SCAUP qui sont les plus significatives de ce qui est observé sur le terrain (*cf.* Chapitre 7, p. 223). Une fois cette première phase définie, la seconde vise à passer des données de simulation de SCAUP aux données d'émissions de PAA par l'intermédiaire du logiciel CIRCUL'AIR. Cette seconde phase d'*évaluation* s'inscrit dans une optique d'*évaluation environnementale* de la contribution du trafic routier par l'estimation des émissions de PAA à l'échelle locale de la ville de Dijon. Elle nécessite de soulever différents questionnements sur notre capacité à passer de la donnée de trafic routier simulé à la spatialisation de la PAA : dans quelles conditions le modèle de simulation du trafic routier permet-il d'alimenter les outils institutionnels de spatialisation de la qualité de l'air ? Quels en sont les apports et les limites pour la quantification des niveaux de PAA ? Quelles réponses opérationnelles peuvent être apportées aux AASQA pour une meilleure connaissance de la PAA ?

Dans un premier temps, nous reviendrons sur les enjeux de la spatialisation de la PAA en nous focalisant sur la spécificité du trafic routier pour le calcul de ces émissions ainsi que sur l'approche qui a été mise en œuvre en collaboration avec ATMOSF'AIR BOURGOGNE. Dans un second temps, nous présenterons l'ensemble des données d'entrée, et des calculs qui sont rattachés au fonctionnement du logiciel CIRCUL'AIR pour l'estimation des émissions PAA. Cette deuxième phase permettra de mettre en avant le processus d'interopérabilité réalisé entre le modèle SCAUP et le logiciel de calcul des émissions de PAA. Enfin, la troisième section de ce chapitre nous permettra de présenter les estimations de PAA calculées en les confrontant aux estimations réalisées par ATMOSF'AIR BOURGOGNE pour l'année 2008.

[157] Site officiel de ATMOSF'AIR BOURGOGNE : http://www.atmosfair-bourgogne.org/fr/accueil-3.html.

## 8.1. Les enjeux de la spatialisation des émissions de polluants atmosphériques automobiles

Avant toute chose, il est important d'apporter une précision sémantique sur le terme d'*émission de polluants atmosphériques*. En effet, l'*émission* se distingue de l'*immission*. Cette dernière permet de caractériser la qualité de l'air que l'on respire et s'exprime sous la forme d'une concentration de polluant dans l'air ambiant, dont la mesure la plus courante est un volume : le $\mu g/m^3$. De la concentration (*immission*), on distingue l'*émission* qui, elle, est plus directement reliée à une source émettant des substances polluantes, et dont la mesure la plus courante est une quantité : le kg, voire la tonne de polluants émis, et le ppb (partie par milliard). Au cours de cette thèse, nous n'abordons que les éléments liés aux émissions de PAA. Cette mesure ne permet pas d'estimer la qualité de l'air proprement dite, mais elle demeure, toutefois, une information qui lui est étroitement liée (Sportisse, 2007a). En revanche, elle permet de mieux identifier la contribution d'un secteur ciblé aux émissions de polluants à l'échelle locale.

Ce travail d'estimation de la PAA s'inscrit dans ce champ et se rattache aux inventaires des émissions de polluants atmosphériques territoriaux réalisés par les AASQA (PCIT, 2012). L'élaboration de ces inventaires vise à étudier les potentiels d'émissions sur un territoire, de manière à obtenir une vision cadastrée des émissions atmosphériques associées à différents secteurs d'activité humaine, le plus généralement pour une année complète. En d'autres termes, c'est une photographie des émissions « réelles » que l'on peut rencontrer dans l'espace urbain à un instant donné. La réalisation des inventaires territoriaux de la pollution atmosphérique s'inscrit dans le cadre de l'arrêté SNIEBA[158] qui vise à « *assurer la cohérence des résultats pour toutes les utilisations de ces données* [d'émissions]*, en particulier celles permettant de répondre aux obligations résultant des directives et décisions européennes ainsi que des conventions internationales et de leurs protocoles* » (*Arrêté du 24 août 2011 relatif au SNIEBA*, 2011, Article 1). C'est un des documents d'entrée des différents plans d'actions locaux comme les PPA[159] (Plans de Protection de l'Atmosphère), qui visent à offrir une image des émissions de polluants atmosphériques à l'échelle d'un territoire donné (PCIT, 2012). Leur élaboration relève de la compétence des AASQA locales qui sont en charge de la surveillance de la pollution atmosphérique à l'échelle des territoires.

### 8.1.1. **La spécificité du trafic routier : un secteur à part entière ?**

Pour commencer, il est important de positionner le trafic routier au sein de ces différents éléments généraux de spatialisation de la pollution atmosphérique à l'échelle locale. En somme, pourquoi le trafic routier tient une place si particulière au sein des cadastres d'émissions atmosphériques ? Quelles en sont les spécificités et les caractéristiques ?

D'après l'inventaire des émissions établi au niveau national (CITEPA, 2015), comparativement aux autres modes de transport (ferroviaire, maritime, aérien et fluvial), le transport routier est le mode de transport qui contribue à plus de 70% des émissions totales du secteur du transport (*cf.* figure 3.20, ci-dessous), alors que le secteur du transport n'est pas toujours le principal

[158] Arrêté du 24 août 2011 relatif au Système National d'Inventaires d'Émissions et de Bilans dans l'Atmosphère (SNIEBA) : https://www.legifrance.gouv.fr/eli/arrete/2011/8/24/DEVR1124328A/jo#.

[159] Le PPA est obligatoire pour les agglomérations de plus de 250 000 habitants et les zones présentant un risque atmosphérique important (comme les zones industrielles) : http://www.developpement-durable.gouv.fr/Presentation-des-plans-de.html.

contributeur des émissions représentées ci-dessous[160] (*cf.* à gauche) comme, par exemple, pour les émissions de Composés Organiques Volatils (COV) ou de monoxyde de carbone (CO) (CITEPA, 2015). On remarque tout de même qu'au sein de ce secteur, la contribution relative du transport routier est prédominante par rapport aux autres modes de transport, avec une contribution moyenne de 82 % des émissions des neufs polluants représentés au sein de la figure 3.20 (*cf.* à droite). De surcroît, si on se focalise, non plus sur le bilan des émissions atmosphériques à l'échelle de la France, mais plutôt à l'échelle des espaces urbains, la contribution du secteur des transports, et donc du trafic routier est bien plus importante (Primequal-Predit, 2012). Par exemple le secteur du transport contribue à plus de 65 % des émissions de $NO_x$ à cette échelle, alors qu'au niveau national sa contribution est toujours inférieure à 60 %.

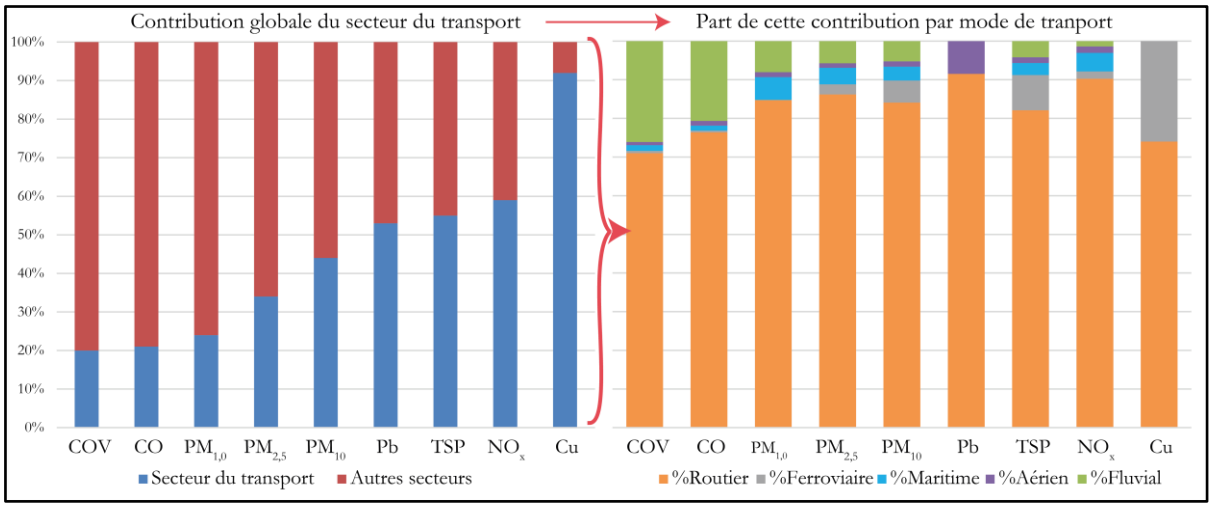



**Figure 3.20 : La contribution du secteur des transports (à gauche) et des différents modes de transport (à droite) aux émissions des principaux polluants atmosphériques en France (source : CITEPA, 2015, p. 172 : Émissions pour l'année 2013)**

Ces premiers éléments permettent de montrer pourquoi le trafic routier tient une place particulière au sein des cadastres d'émissions. Cependant, s'il est la principale source de la pollution atmosphérique de proximité en milieu urbain, sa spécificité au sein des cadastres d'émission tient aussi à son caractère spatio-temporel, sur lequel nous nous sommes déjà arrêtés longuement. D'une manière générale, on remarque que le trafic routier se formalise sous la forme d'entités linéaires diffuses. En effet, le trafic routier est une source spécifique qui est mobile et diffuse autant dans le temps que dans l'espace, ce qui le distingue des autres sources comme les industries ou les zones résidentielles qui sont moins volatiles et plus stables (PCIT, 2012). Le trafic routier est une source de pollution spécifique, car si grâce aux évolutions technologiques les véhicules routiers polluent moins individuellement, c'est plutôt la part croissante des véhicules routiers qui va jouer sur l'intensité des émissions atmosphériques (AEE, 2011). En d'autres termes, les véhicules polluent moins au niveau de leur échappement, mais ils sont plus nombreux, ce qui fait très peu diminuer la contribution relative du transport routier au sein des inventaires d'émissions nationaux et internationaux (CITEPA, 2015). De plus, l'insertion des nouvelles

[160] Ici nous n'avons sélectionné que quelques polluants pour lesquels la contribution du trafic routier est la plus significative. La liste détaillée des différents polluants où le secteur des transports contribue à plus de 5% des émissions est présenté dans l'inventaire national diffusé par le CITEPA (2015, pp. 172–200).

technologies comme les carburants sans plomb, le pot catalytique ou le *stop et go*[161], qui visent à limiter les émissions unitaires des véhicules, s'opère sur de longues périodes en raison de l'inertie du renouvellement du parc automobile (Bourdeau, 1997). C'est typiquement le cas de l'insertion des pots catalytiques qui a conduit à une diminution progressive des émissions de $NO_x$ (CITEPA, 2015, p. 173) : « *Globalement, le renouvellement du parc automobile et l'introduction généralisée des pots catalytiques sur les véhicules essence particuliers depuis 1993, et les véhicules particuliers diesel à partir de 1997, conduit à une réduction des émissions* [de $NO_x$] *du transport routier de 55 % entre 1990 et 2013, malgré une croissance du trafic de 34 %. La baisse des rejets de $NO_x$ devrait encore se poursuivre du fait du renouvellement des véhicules et de la pénétration accrue de véhicules aux normes plus performantes* ».

Au final, les émissions liées au trafic routier présentent une multitude d'éléments à prendre en compte, qui sont à relier autant aux évolutions technologiques et aux individus qu'à l'environnement. Elles apparaissent constituer un secteur à part entière. Cela se traduit aussi dans le guide méthodologique pour l'élaboration des inventaires territoriaux qui traite plus spécifiquement du secteur du transport routier (PCIT, 2012, pp. 165–180) : « *le calcul des émissions routières doit permettre une information spatialisée sur les voies, ou a minima à la commune pour des secteurs manquants de données. Ainsi l'évaluation fine du trafic routier est nécessaire* […], *approchée autant que possible voie par voie à partir du moment où des données de comptage et/ou une modélisation du trafic existent* ». Par cette citation, et suite aux différents propos rapportés plus haut, on comprend bien pourquoi les données issues des modèles de trafic sont « *à privilégier si elles existent* » (PCIT, 2012, p. 168) pour le calcul des émissions issues du trafic routier. C'est pourquoi ce calcul diffère des autres secteurs, qui reposent eux sur le référentiel OMINEA (2015), tandis que la méthodologie d'estimation des émissions de PAA repose plutôt sur la méthodologie COPERT (Ntziachristos et Zissis, 2009 ; AEE, 2014).

La méthodologie COPERT repose sur l'intégration d'un plus grand nombre de variables pour le calcul des émissions, comme la pente du tronçon routier, le parc automobile ou la vitesse de circulation moyenne (*cf.* figure 2.20, p. 180). Elle se fonde sur le calcul des émissions unitaires des véhicules en fonction des différents types de véhicules, des normes environnementales en vigueur (normes EURO[162]), de la motorisation des véhicules (diesel/essence), et bien sûr, des données de circulation routière. C'est pourquoi la ventilation des données de circulation en fonction du parc automobile est primordiale pour le calcul des émissions de PAA (Bourdeau, 1997 ; Hugrel et Joumard, 2004 ; Carteret *et al.*, 2014). Dans ce sens, le calcul des émissions rattachées à la PAA est quelque peu différent de la formule générale présentée plus haut (p. 180). En effet, si l'activité ($A_t$) est bien associée au débit de trafic routier sur les tronçons routiers pour un temps donné, les facteurs d'émissions quant à eux dépendent de plusieurs variables qui tiennent compte du parc automobile (défini en sous-groupes $K$) et des caractéristiques des infrastructures routières (la pente, la capacité et la longueur des tronçons routiers).

Le principe général des modèles d'émissions du trafic routier s'appuie sur la formule suivante (Chanut et Chevallier, 2012 ; PCIT, 2012) :

[161] Ce système permet de diminuer la consommation de carburant et *a fortiori* des émissions de polluants. Il permet de stopper momentanément le moteur d'un véhicule lorsque celui-ci est arrêté.
[162] Les normes EURO (normes européennes d'émissions) fixent les limites maximales d'émissions des polluants sur les véhicules neufs. Depuis septembre 2015, les véhicules neufs sont soumis à la norme EURO 6 : http://www.developpement-durable.gouv.fr/Depuis-le-1er-septembre-la-norme.

$$E_{s,t} = \sum_j A_{j,t} * (f_{ij,t} * n_{j,t}) \qquad avec\ N = \sum_j nj$$

Pour le calcul de l'émission d'un polluant $E_{s,t}$, cette formule tient compte d'un temps donné *t* d'une substance polluante *s*, pour un trafic routier $N$ dont on en distingue $K$ sous-groupes de $n_j$ ($j \in [1, K]$) de types de véhicules (catégorie, normes EURO,…) où chacun de ces sous-groupes a une loi d'émission définie au sein de la méthodologie COPERT $f_{ij}$ pour un polluant donné.

Dans cette optique, c'est une approche mixte qui est privilégiée, en mobilisant autant un mécanisme *top-down* en vue de ventiler les données du parc automobile qui sont disponibles à l'échelle nationale ou régionale, qu'un mécanisme *bottom-up* en ce qui concerne les valeurs de débit simulées au niveau des tronçons (PCIT, 2012 ; OMINEA, 2015). Malgré tout, il faut relever que ce calcul tient compte d'un débit de trafic routier agrégé par l'unité de temps à laquelle on souhaite étudier les émissions de PAA. Usuellement, la mesure des valeurs de débit est obtenue via le TMJA, ou le TMJH[163], auxquels on associe un parc automobile pour en décomposer les $K$ sous-groupes de types de véhicules.

### 8.1.2. **Présentation de la démarche d'évaluation environnementale de la PAA**

Cette section est consacrée au développement d'une chaîne de travail permettant d'estimer les émissions de PAA en s'appuyant sur une interopérabilité des données issues de SCAUP (*données de trafic simulé*), et du modèle de calcul CIRCUL'AIR utilisé par les AASQA, pour estimer les émissions de polluants rattachées aux infrastructures de transports. Les données en entrée de notre approche sont le trafic routier simulé par voie, dont nous avons sélectionné les paramètres de simulation les plus représentatifs au cours de la section 7.3.3 (p. 254). L'ensemble du travail d'estimation des émissions d'origine automobile (PAA) a été réalisé en collaboration avec l'AASQA ATMOSF'AIR BOURGOGNE. Pour cela, nous nous sommes appuyé sur leurs outils pour tester et expérimenter les potentiels opérationnels de notre approche au sein des cadastres d'émissions de la pollution atmosphérique classiques. L'objectif est de se rattacher aux formats de données habituellement utilisés par les AASQA, en fournissant au préalable les valeurs de trafic routier simulé à partir de SCAUP, pour les paramètres de choix routiers hiérarchiques (H) et pour trois temps de stationnement (10 minutes : *H10*, 15 minutes : *H15*, et 20 minutes : *H20*), conformément aux résultats présentés en figure 3.19 (p. 256). Le trafic routier simulé, nécessaire à l'application des calculs de PAA, impacte en partie leur calcul, et celui-ci s'insère dans un processus d'estimation nécessitant l'emploi d'un plus grand nombre d'informations

La figure 3.21 (ci-dessous) illustre la méthodologie de calcul des émissions du trafic routier suivie par les AASQA (en noir), ainsi que le processus méthodologique suivi au cours de ce travail (en vert). Ce logigramme a été élaboré à partir des informations fournies dans le guide méthodologique du PCIT (2012), qui constitue le document de référence méthodologique, en appui des AASQA, pour le calcul des émissions du trafic routier. D'une manière générale, le schéma permet de voir qu'il n'y a pas une seule méthode, mais bien plusieurs méthodes de calcul des émissions de la PAA, allant « *de la méthode la plus complète jusqu'à des méthodes beaucoup moins précises en cas d'absence de données* » (PCIT, 2012, p. 168). La figure 3.21 permet d'observer l'ensemble de ces méthodes, au sein desquelles on distingue deux étapes distinctes :

[163] TMJA : Trafic Moyen Journalier Annuel ; TMJH : Trafic Moyen Journalier Horaire.

- 1. pour **quantifier les volumes et vitesses moyennes de trafic sur les axes**
- 2. pour **calculer les émissions de PAA**

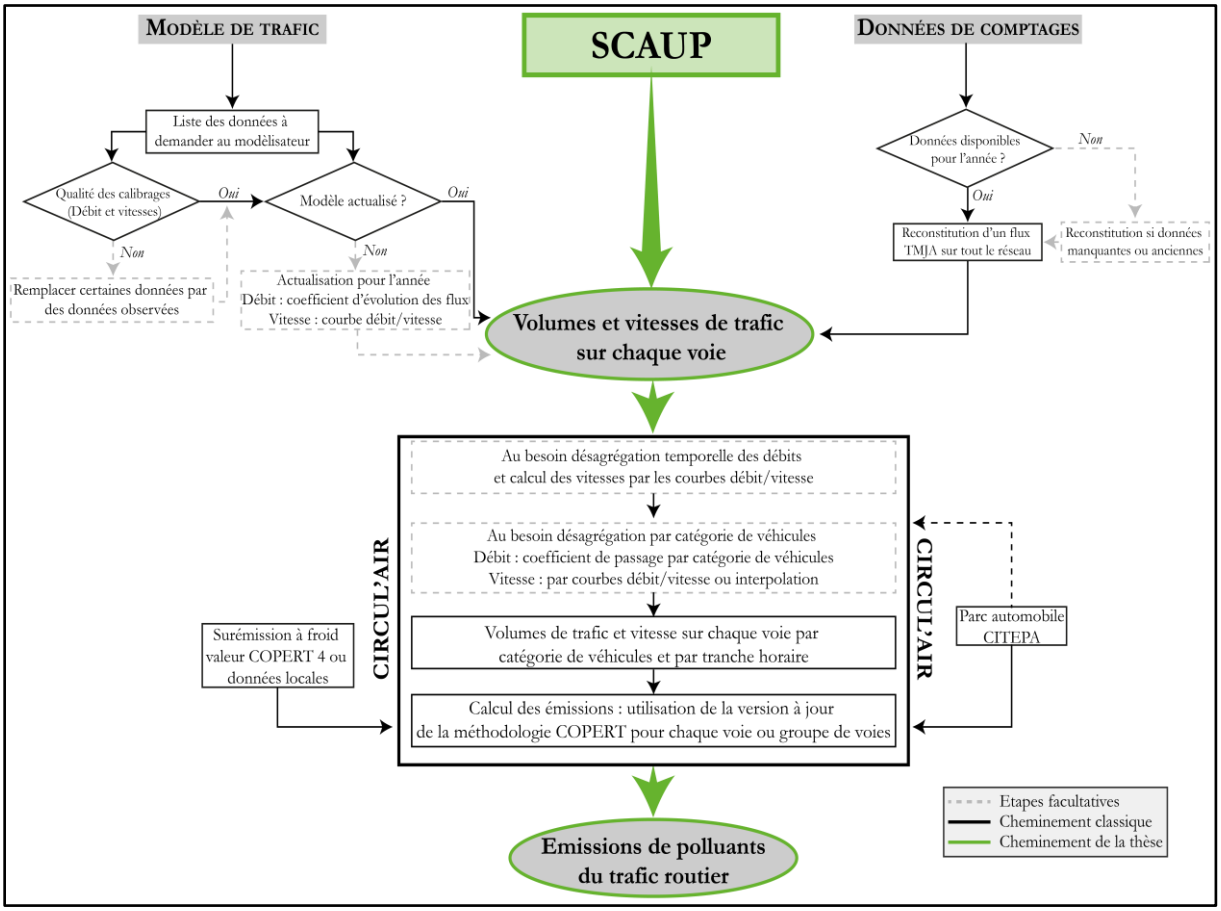



**Figure 3.21 : Diagramme des étapes de calcul des émissions du trafic routier (d'après PCIT, 2012, p. 168, figure 28)**

La première étape permet de **quantifier les flux de trafic routier**. Elle vise à obtenir et/ou calculer les volumes et la vitesse du trafic routier sur chaque voie du domaine d'étude. Dans le processus classique (en noir) suivi par les AASQA, deux itinéraires se distinguent pour quantifier le trafic routier :

- 1. soit en s'appuyant sur un modèle de trafic comme illustré à droite de la figure 3.21. Cette étape nécessite de faire appel au modélisateur du trafic en vue de s'assurer que les données de trafic soient calibrées et actualisées pour l'année considérées.
- 2. soit en s'appuyant sur les données de comptage comme illustré à gauche de la figure 3.21. À ce niveau, les données de comptage doivent être reconstituées sous la forme de TMJA sur l'ensemble du réseau routier, et nécessitent donc une reconstitution des données manquantes.

*In fine*, l'accès aux données de trafic routier est externalisé et doit se faire en *concordance* avec les services en charge de la gestion du trafic routier, qui disposent le plus généralement de ces informations (ASPA, 2005 ; PCIT, 2012). Cette étape qui est essentielle en vue d'alimenter les cadastres d'émissions de la pollution atmosphérique nécessite un travail de validation rigoureux en vue de calibrer et/ou de reconstituer les données de comptage sous la forme de TMJA. Pour la ville de Dijon, aucun modèle de trafic n'étant disponible, ATMOSF'AIR BOURGOGNE s'appuie

sur une estimation du trafic routier à partir des comptages routiers. Lorsqu'aucune donnée issue des modèles de trafic n'est disponible, ce qui est souvent le cas pour les villes petites et moyennes, le modèle SCAUP construit dans le cadre de ce travail permet alors d'apporter une première réponse aux AASQA, comme pour ATMOSF'AIR BOURGOGNE. C'est pourquoi nous positionnons notre approche à l'intersection du modèle de trafic et des données de comptage pour l'estimation des volumes de trafic routier, en proposant une troisième alternative à ces méthodes à travers l'utilisation du modèle SCAUP, comme illustré par le cheminement vert en figure 3.21. Par ailleurs, comparativement aux deux autres méthodes, l'approche SCAUP permet aussi, en facilitant le calcul des volumes de circulation routière, une réduction des coûts, autant en temps de collecte qu'en temps de traitement des données. En effet, pour la simulation du trafic routier à partir de SCAUP, il est seulement nécessaire de collecter les données et la localisation des stations de comptage, ainsi que le réseau routier de la zone étudiée, pour simuler le trafic routier. En nous appuyant sur la figure 3.21, nous situons l'apport du modèle à ce niveau (entre comptage et modèle), en vue de fournir plus directement des données de trafic routier en entrée du calcul des émissions de PAA.

La seconde étape permet de **calculer les émissions de PAA** (*cf.* figure 3.21). Cette étape vise à intégrer progressivement un ensemble de données relatives autant à la circulation routière qu'à la morphologie des infrastructures de transport routier (capacité, vitesse moyenne de circulation, longueur, pente…). En s'appuyant sur la figure 3.21, on remarque qu'à l'insertion des données de trafic routier se juxtaposent de nombreuses données d'entrée relatives au parc automobile, ainsi qu'à la méthodologie COPERT. L'ensemble de ces éléments d'entrée permet de mettre en œuvre le calcul des émissions de polluants, en s'appuyant sur la formule présentée plus haut (*cf.* section 8.1.1, p. 268). À ce niveau, pour le calcul des émissions de PAA, les AASQA s'appuient aujourd'hui sur un logiciel développé par l'Association de Surveillance de la Pollution en Alsace (ASPA), CIRCUL'AIR (ASPA, 2005). Ce logiciel, dont nous décrirons plus finement les données d'entrée et le fonctionnement au cours des prochaines sections, permet de faciliter la réalisation de ces étapes de calcul pour la mise en œuvre des calculs d'émissions de PAA à l'échelle des tronçons routiers.

Au final, l'interopérabilité entre les données de simulation issue de SCAUP et le logiciel CIRCUL'AIR s'opère par l'insertion des données géographiques de débits simulés (trafic routier/tronçon) sur l'ensemble du territoire de la ville de Dijon comme illustré en figure 3.22 (ci-dessous). Dès lors, il est important de mettre en avant l'ensemble des données d'entrée nécessaire à l'estimation des émissions de PAA, qui a été réalisée grâce à l'appui technique d'ATMOSF'AIR BOURGOGNE.

## 8.2. Données d'entrée et méthodologie de calcul des émissions avec CIRCUL'AIR

Le logiciel CIRCUL'AIR développé par l'ASPA[164] depuis les années 1990[165], est aujourd'hui accessible et utilisé par toutes les AASQA françaises. Il permet de faciliter autant l'intégration des données en entrée du logiciel que l'automatisation des calculs des émissions de PAA à l'échelle d'un territoire, et ce, par le biais de différentes étapes indiquées en figure 3.21. Il est à noter qu'il se base en très grande partie sur la méthodologie COPERT (Ntziachristos et Zissis, 2009 ; AEE, 2011), en y associant un ensemble de données d'entrée (détaillées ci-dessous) relatives au trafic routier et à sa fréquentation, au parc automobile et aux infrastructures de transports. En tenant compte des valeurs de débit, il permet la mise en place et l'estimation des émissions de PAA par le calcul de plusieurs étapes comme illustré en figure 3.21. Pour commencer, il faut noter que pour être en mesure d'interpréter des résultats qui n'ont pas été réalisés dans le même contexte que celui d'ATMOSF'AIR BOURGOGNE en 2008, il est important de décrire l'ensemble des données d'entrée qui ont été utilisées pour le calcul des émissions de PAA, et dont une synthèse est proposée en figure 3.22, qui sera reprise en détail au cours des sections suivantes. Cette figure illustre sous une forme procédurale le positionnement des données d'entrée, ainsi que le passage du modèle SCAUP au modèle CIRCUL'AIR pour le calcul des émissions de PAA :

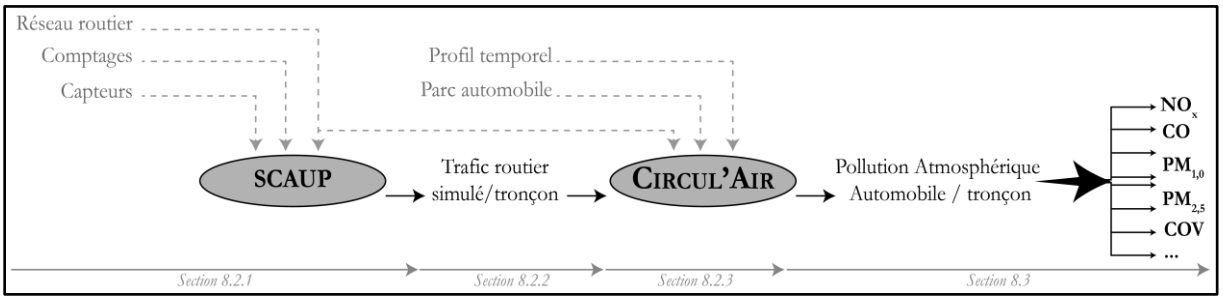



**Figure 3.22 : Données nécessaires et *workflow* pour le calcul des émissions des polluants atmosphériques automobiles**

### 8.2.1. La quantification du trafic routier : intégration des données issues de SCAUP

Pour cette première étape, comme nous l'avons illustré en figure 3.21, nous situons l'apport de notre approche à l'intersection des données de comptage et des modèles de trafic, dans le sens où ce sont bien les données de comptages routiers qui sont simulées par l'intermédiaire d'une plate-forme de SMA (*cf.* figure 3.21). Dans cette perspective, le modèle SCAUP permet de disposer de données de trafic routier simulé relativement rapidement pour une journée de comptage et par quart d'heure sur l'ensemble d'un territoire. C'est pourquoi, il n'est pas nécessaire de mettre en œuvre toutes les étapes d'estimation des volumes de trafic routier comme ce serait le cas avec des données issues des modèles de trafic classiques (*cf.* figure 3.21). L'accès aux valeurs de circulation routière et leur estimation en sont alors d'autant plus facilités, en s'appuyant exclusivement sur les

[164] ASPA : Association pour la Surveillance et l'Étude de la Pollution Atmosphérique en Alsace.

[165] Avec un certain regret, on retrouve très peu d'informations relatives au fonctionnement de CIRCUL'AIR ; pour la présentation et la description de ses différents éléments, nous nous appuyons sur l'expertise de ATMOSF'AIR BOURGOGNE et sur les documents suivants : (ASPA, 2005 ; Ntziachristos et Zissis, 2009 ; AEE, 2011 ; ASPA, 2012 ; PCIT, 2012 ; OMINEA, 2015)

données de simulation de SCAUP, qui comparativement au processus classique d'estimation des débits de circulation routière facilitent l'accès à des données de volumes de trafic routier sur chaque voie.

### a. *Les débits de trafic routier : une donnée d'entrée spatio-temporelle*

Les sorties de simulation de SCAUP sélectionnées sont les paramètres *H10*, *H15* et *H20* (*cf.* figure 3.19, p. 256), desquels nous avons extrait le trafic routier simulé au quart d'heure pour l'année 2008. Cependant, le logiciel CIRCUL'AIR fonctionnant avec des données agrégées de TMJA, les données de simulation (au quart d'heure) ont été réagrégées sous la forme d'un débit moyen journalier et horaire de simulation correspondant à la moyenne du trafic simulé pour une journée de comptage (*cf.* Chapitre 7, p. 223). En somme, deux types de données ont été pris en compte pour le calcul des émissions de la PAA : un trafic moyen horaire (TMH) et un trafic moyen journalier (TMJ). Ceux-ci sont ensuite spatialisés sur les tronçons utilisés pour la simulation du trafic routier, comme illustré pour le paramètre *H15* en haut de la figure 3.23 (ci-dessous).

En ce qui concerne les données de trafic routier exploitées par ATMOSF'AIR BOURGOGNE en 2008, aucun modèle de trafic n'étant disponible, les données de débits ont été définies et estimées en tenant compte des comptages routiers disponibles sur le territoire dijonnais (*cf.* logigramme de la figure 3.21, p. 269). L'ensemble des données pour l'année 2008 a été spatialisé en tenant compte du réseau routier issu du référentiel spatial de la BD CARTO de l'IGN[166], comme illustré en haut de la figure 3.23 (ci-dessous). Cependant en confrontant les deux réseaux routiers (*cf.* figure 3.23, ci-dessous), on remarque rapidement que comparativement au réseau routier utilisé pour SCAUP qui est issu de la BD TOPO (IGN, 2014), les réseaux routiers ne présentent pas la même résolution pour un même espace d'étude. C'est pourquoi, en s'appuyant sur le tableau 3.4 et la figure 3.23, il est nécessaire d'apporter un ordre de grandeur des différences entre les réseaux routiers exploités, car la BD CARTO exploitée par ATMOSF'AIR BOURGOGNE ne permet pas d'avoir une résolution aussi fine que la BD TOPO exploitée dans SCAUP. En effet, alors que la BD TOPO est adaptée pour l'analyse de données spatiales à des échelles spatiales allant de 1/5 000$^{e}$ à 1/50 000$^{e}$ (IGN, 2014), la BD CARTO présente une résolution moins fine, variant entre 1/50 000$^{e}$ et 1/100 000$^{e}$ (IGN, 2010). En outre, la BD CARTO constitue le référentiel national afin de « *faciliter le partage d'informations entre les différents acteurs d'une même collectivité territoriale, et sur lequel chaque utilisateur peut rattacher les données de son domaine thématique* » (IGN, 2010, p. 4). Malgré tout, si la finesse des réseaux routiers est différente, les référentiels étant tout deux établis par l'IGN ceux-ci présentent des caractéristiques communes quant aux attributs de hiérarchie du réseau routier sur lesquels nous nous appuierons par la suite. Mis à part pour les liaisons autoroutières, nous avons dénombré trois types de voies communes aux deux référentiels, et sur lesquels nous nous appuierons par la suite : *les liaisons locales, les liaisons principales et les liaisons régionales*[167].

[166] À noter que depuis 2010 ATMOSF'AIR BOURGOGNE exploite le réseau routier issu de la BD TOPO de l'IGN.

[167] Les attributs « Importance » de la BD TOPO (IGN, 2014) ont été mis en relation avec l'attribut « Vocation » de la BD CARTO (IGN, 2010) comme suit :

- BD TOPO « Importance » 2 = BD CARTO « Vocation » Liaisons principales
- BD TOPO « Importance » 3 = BD CARTO « Vocation » Liaisons régionales
- BD TOPO « Importance » 4 = BD CARTO « Vocation » Liaisons locales.

En s'appuyant sur le tableau de synthèse ci-dessus, on remarque que le réseau routier issu de la BD CARTO présente une longueur totale 80 % inférieure à celle du réseau routier de la BD TOPO. Néanmoins, on remarque que plus la hiérarchie du réseau routier est importante plus cette différence diminue, ce qui parait peu surprenant du fait de la résolution des deux réseaux routiers exploités. En d'autres termes, plus le réseau routier est d'importance (en termes de hiérarchie), plus la différence entre les deux référentiels est faible, comme synthétisé ci-dessous :

| Vocation Réseau | Longueur totale (km) | | |
|---|---|---|---|
| | BD CARTO | BD TOPO | Différence * |
| Ensemble | 132,70 | 698,19 | -80,99% |
| Liaisons locales | 10,41 | 57,29 | -81,82% |
| Liaisons principales | 39,20 | 90,92 | -56,89% |
| Liaisons régionales | 65,43 | 86,11 | -24,02% |

$$* \frac{(\text{BD CARTO} - \text{BD TOPO})}{\text{BD TOPO}}$$

**Tableau 3.4 : Comparaison de la longueur cumulée des réseaux routiers utilisés sur la ville de Dijon pour le calcul des émissions de la PAA**

Ces différences sont importantes et doivent être mises en avant dès maintenant, car celles-ci peuvent impacter sensiblement l'analyse des émissions de PAA qui sera présentée sur la ville de Dijon. Pour cela, il faut être en mesure de confronter les deux réseaux routiers sur une même base. C'est pourquoi nous avons privilégié une approche par carroyage en définissant un maillage d'une résolution de 100 mètres[168] (*cf.* figure 3.23). En utilisant cette information, nous sommes en mesure de confronter deux informations dont la résolution spatiale et l'emprise ne sont pas toujours identiques. En s'appuyant sur une analyse SIG simple (jointures spatiales) des valeurs d'émissions, nous sommes alors en mesure de mettre en relation les deux réseaux routiers. Dans ce sens, les données d'émissions de PAA seront toutes par la suite non seulement localisées au niveau des tronçons routiers de la ville de Dijon, mais aussi à l'intérieur de ce carroyage, où nous y associerons les émissions de PAA calculées par ATMOSF'AIR BOURGOGNE et celles calculées *via* SCAUP.

[168] Des tests ont été établis avec un maillage d'une résolution de 50m, or celui-ci ne permet de correctement associer les axes adjacents entre ceux localisé à partir de la BD CARTO et ceux localisés à partir de la BD TOPO au sein d'une même maille. Nous avons donc fait le choix de travailler avec un maillage défini à une résolution de 100 m.

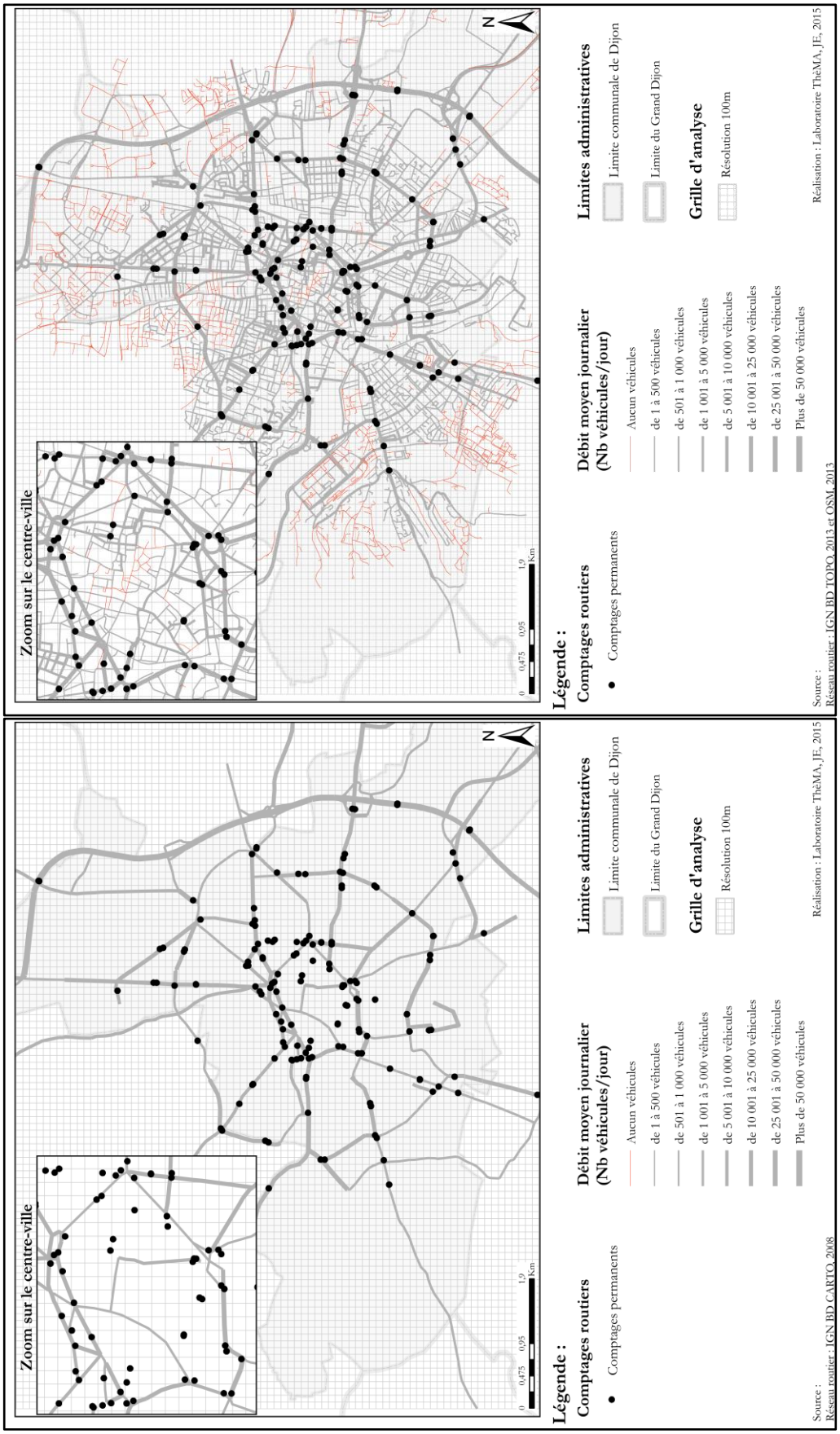



**Figure 3.23 : Les flux de circulation routières insérés en entrée du calcul des émissions de PAA (en haut : la BD TOPO exploitée avec le modèle SCAUP ; en bas : la BD CARTO exploitée par ATMOSF'AIR BOURGOGNE)**

### b. *Les caractéristiques relatives à la hiérarchisation du réseau routier*

Une donnée d'entrée fondamentale au calcul des émissions du trafic routier est à rattacher aux infrastructures de transports, qui en fonction de leurs caractéristiques, jouent un rôle important sur les niveaux d'émissions. Les informations relatives au réseau routier en entrée de CIRCUL'AIR ont été définies en tenant compte du réseau routier généré grâce au module rattaché à MOBISIM (*cf.* encart 5, p. 195). Cela nous permet d'apporter des informations permettant de caractériser les tronçons routiers de la ville de Dijon, informations qui seront ensuite nécessaires au calcul des émissions de PAA (PCIT, 2012 ; OMINEA, 2015). Au total, six informations relatives aux tronçons routiers ont été transmises à ATMOSF'AIR BOURGOGNE, dont nous proposons une description ci-dessous. Les caractéristiques du réseau routier sont différentes de celles présentées plus haut pour la génération du réseau routier pour SCAUP (*cf.* section 6.2.1.b, p. 193). C'est pourquoi nous présentons ici les informations nécessaires aux calculs des émissions de PAA pour CIRCUL'AIR. Ces informations visent à hiérarchiser le réseau routier, en s'appuyant dans un premier temps sur le référentiel de la BD TOPO :

- **Les hiérarchies** (*hiérarchie*) : elles sont issues de l'attribut « Importance » de la BD TOPO qui permet de décomposer le réseau routier de la ville en quatre classes (IGN, 2014) : les liaisons départementales, les liaisons interurbaines, les liaisons urbaines et les liaisons résidentielles.
- **Les vitesses de circulation** *(vitesse réelle)* : elles sont définies en tenant compte de la vitesse réglementaire pondérée en fonction de la pente, de la sinuosité du tronçon routier et de coefficients d'ajustement tenant compte de la densité de bâtiments situés à proximité des axes (Lunardi, 2013). À noter que les vitesses de circulation ont été testées et validées en s'appuyant sur une analyse du réseau routier via l'outil « *Network Analyst* » du logiciel SIG ArcGis[169] réalisé au cours d'un mémoire de Master (Tiphaine, 2014), et en s'appuyant sur les travaux de Manuel Appert et Laurent Chapelon (2003) sur la performance des réseaux routiers.
- **Les capacités** *(capacité)* : elles sont définies en tenant compte du nombre de voies, de leur largeur, de la vitesse de circulation et de la hiérarchie du tronçon. La capacité correspond au débit maximal de véhicules supporté par l'axe pour une heure, qui s'exprime en nombre de véhicules/heure.
- **Les longueurs** (*longueur*) : elles ont été calculées grâce aux SIG en mesurant la longueur effective de l'axe routier en mètres.
- **Les pentes** (*pente*) : elles sont calculées en fonction de la longueur mesurée précédemment et des attributs natifs de la BD TOPO (*Z min et Z max*). La pente est exprimée en un pourcentage de pente, qui croît proportionnellement à la pente. À noter qu'en fonction des pentes, les axes routiers sont ensuite décomposés en axes de montée et de descente (ASPA, 2005 ; PCIT, 2012).
- **Les TMJA simulés par SCAUP** : les valeurs de TMJA ont été obtenues par l'intermédiaire de SCAUP pour l'année 2008, dont nous avons présenté les différentes caractéristiques précédemment.

L'intégration de ces informations permet alors de hiérarchiser le réseau routier ainsi que les

[169] Présentation de l'outil Network Analyst de ArcGis : http://help.arcgis.com/fr/ARCGISDESKTOP/10.0/HELP/index.html#//004700000001000000.

données de simulation issues de SCAUP dans l'optique du calcul des émissions. Pour finir sur la hiérarchisation du réseau routier, on peut relever que depuis le départ nous avons exploité de nombreux référentiels pour hiérarchiser le réseau routier de la ville de Dijon, entre la BD TOPO pour la génération du réseau routier, OSM pour la simulation du trafic routier et la BD CARTO pour se rattacher au référentiel utilisé par ATMOSF'AIR BOURGOGNE. On remarque que malgré les fortes proximités existant entre chacun de ces référentiels, la hiérarchisation des axes routiers qu'ils proposent est différente. Ces éléments tendent à nous montrer la nécessité de s'accorder sur une même hiérarchie du réseau routier, en vue d'éviter la manipulation de nombreux référentiels comme c'est le cas dans ce travail. Certains travaux, comme ceux de Philippe Bonnin et Stéphane Douady (2013), et de Claire Lagesse *et al.* (2015), s'inscrivent dans ce sens en vue de caractériser la forme des réseaux routiers. Ces travaux s'appuient plus particulièrement sur la structure, la géométrie et la topologie des réseaux viaires pour en observer des structuralités sans faire appel à des référentiels préétablis (Lagesse *et al.*, 2015). Il apparaitrait intéressant d'exploiter ces travaux en vue de proposer une hiérarchisation du réseau routier généralisable à de nombreux espaces urbains.

### 8.2.2. **Les données d'entrées spécifiques au calcul des émissions de PAA**

Pour cette seconde étape, nous décrivons les données mises en place avec ATMOSF'AIR BOURGOGNE pour le calcul des émissions de PAA. Contrairement aux données précédentes qui sont à rattacher au modèle SCAUP, les données exploitées ici y sont extérieures, comme illustré en figure 3.22 (p. 271).

#### a. *Ventilation du parc automobile sur la ville de Dijon : une entrée primordiale pour l'évaluation de la PAA*

Dans un troisième temps, les débits de trafic routier simulés à partir de SCAUP sont redéfinis en une flotte de véhicules dont les caractéristiques s'appuient sur les données du *parc automobile* français. Le *parc automobile* correspond à « *l'ensemble des véhicules dont dispose à un moment donné un pays* »[170]. Plus précisément, c'est un inventaire du nombre de véhicules en circulation sur les routes nationales, ou régionales. Celui-ci se base sur les fichiers d'immatriculation nationaux et sur des données de projection établies à l'échelle du territoire français (Bourdeau, 1997 ; Hugrel et Joumard, 2004 ; OMINEA, 2015). À noter que le parc automobile « *ne tient pas compte du kilométrage parcouru par chacun d'entre eux* […] ; *il constitue une évaluation effective des véhicules. Même au sein d'une catégorie identique, les véhicules particuliers par exemple, l'usage des véhicules peut être très différent en fonction de certaines de leurs caractéristiques (carburants, cylindrée…). L'évaluation du parc automobile est nécessaire dès lors que l'on souhaite apprécier la pollution associée aux transports* » (Hugrel et Joumard, 2004, p. 6). Dans cette acception, le parc automobile est une donnée d'entrée primordiale pour le calcul des émissions de la pollution atmosphérique automobile (PAA).

C'est pourquoi il est ici nécessaire de ventiler les valeurs de débit simulé entre les différents types de véhicules que l'on rencontre sur les routes pour le calcul des émissions de la PAA. Pour cela, les AASQA s'appuient sur les données du parc automobile défini par le CITEPA, et dont les éléments méthodologiques se basent sur les travaux de doctorat initiés par Béatrice Bourdeau (1997), qui a établi les projections du parc automobile français jusqu'en 2020 (Hugrel et Joumard,

[170] Définition issues du Larousse en ligne, 2016 pour la notion de « **Parc** ».

2004). D'une manière générale, la ventilation des débits de trafic routier (*N*) vise à décomposer celui-ci en *K* sous-groupes de véhicules, en tenant compte du type de véhicules, entre les Véhicules Particuliers (VP), les Véhicules Utilitaires Légers (VUL), les Poids Lourds (PL), les Bus, les Autocars[171] et les Deux-Roues. La ventilation du parc automobile a été assurée par ATMOSF'AIR BOURGOGNE en suivant la relation suivante, en vue d'aboutir à un parc automobile représentatif de ce que nous pourrions retrouver sur la ville de Dijon en 2008 (*cf.* figure 3.24) :

1. DÉBIT DE SIMULATION → 2. TYPE DE VÉHICULES → 3. TYPE DE CARBURANT → 4. TYPE DE NORMES EURO → 5. PUISSANCE FISCALE DU VÉHICULE

Le processus de répartition du parc automobile ventilé sur la ville de Dijon est synthétisé au sein de la figure 3.24, et le tableau détaillé des clés de cette répartition est présenté en annexe 4 (p. vii) :

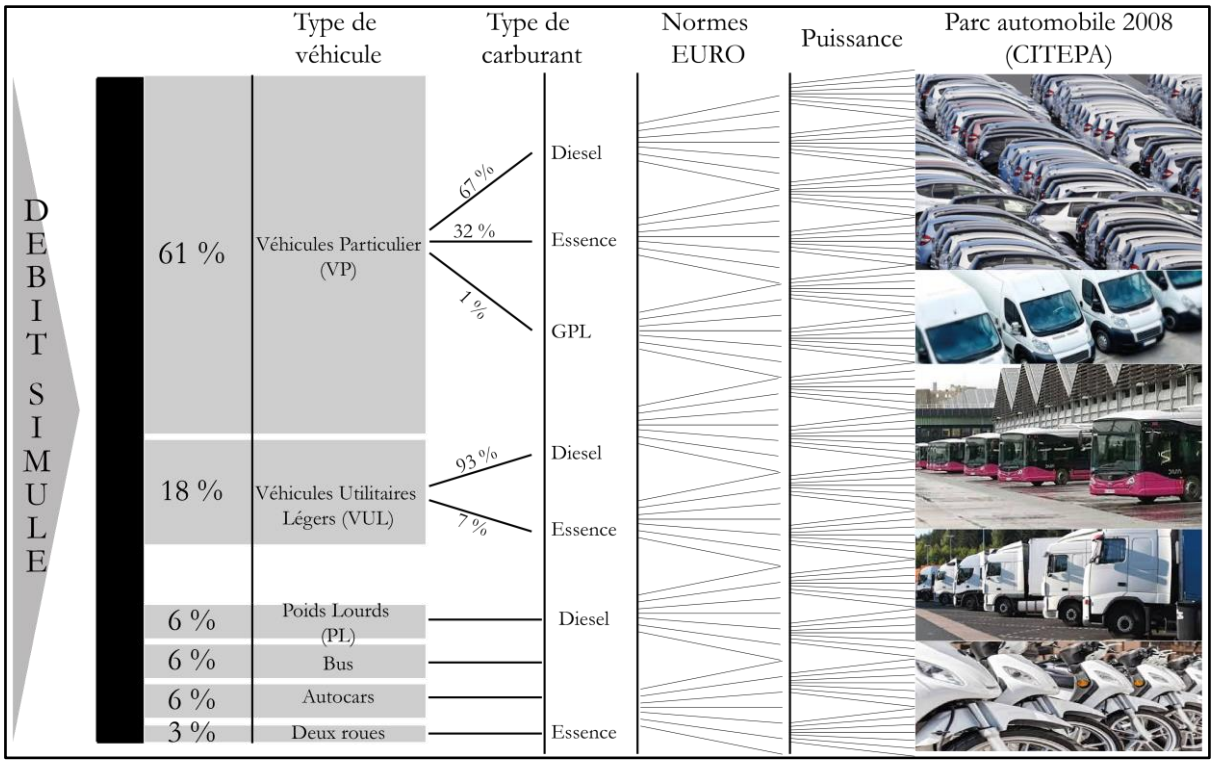



**Figure 3.24 : Processus de ventilation du parc automobile sur les données de trafic routier de SCAUP pour la ville de Dijon (d'après : ATMOSF'AIR BOURGOGNE, CITEPA, 2008)**

La ventilation du parc automobile sur la ville de Dijon se décompose en une majorité de Véhicules Particulier (VP) et de Véhicules Utilitaires Légers (VUL) qui à eux deux concentrent près de 60 % du parc automobile défini sur l'espace d'étude[172]. Ensuite, ces types de véhicules sont ventilés en fonction du carburant utilisé. À noter, que sur cette seconde ventilation, on retrouve la spécificité du parc automobile français, dont la forte diésélisation constitue son caractère atypique par rapport à d'autres pays (Dab et Roussel, 2001). Par exemple, en 2012, la part de VP diesel en France est de 73 % tandis qu'en Allemagne elle est de 48 %, et même de

[171] La dénomination des bus et des cars correspond à deux services différents : les Bus assurent les dessertes urbaines et les Autocars assurent les dessertes interurbaines : http://www.keolis.com/fr/activites/notre-expertise/nos-bus-et-cars.html.

[172] Les valeurs ont été définies par le CITEPA et constituent le référentiel exploité par ATMOSF'AIR BOURGOGNE.

28 % aux Pays Bas[173]. Cette tendance se retrouve pour les VP, pour lesquels les technologies diesel concentrent plus de 67 % des VP, contre 32 % de véhicules essence et 1 % de véhicules GPL pour l'année 2008 (*cf.* figure 3.24).

De surcroît, la prépondérance des technologies diesel en France présente aujourd'hui un véritable enjeu en matière de gestion de la qualité de l'air en milieu urbain. En effet, d'après le CITEPA (2015, p. 173) pour les émissions de $NO_x$, « *les trois postes contribuant aux émissions de $NO_x$ du transport routier en 2013 sont : les VP diesel catalysés (39 % des émissions), les PL*[174] *(37 %) et les VUL diesel catalysés (15 %)* ». Ces derniers éléments montrent d'autant plus que les véhicules routiers fonctionnant au Gazole présentent aujourd'hui un impact non-négligeable sur la qualité de l'air en milieu urbain. À titre indicatif, pour une même vitesse de déplacement les émissions unitaires de $NO_x$ rattachées aux véhicules essence sont cinq fois moins importantes que pour des véhicules diesel[175] (SETRA, 2009).

Pour ce travail sur la ville de Dijon, près de 230 types de véhicules ont été ventilés sur les données de débit (*cf.* annexe 4). Une fois le parc automobile défini, celui-ci est ensuite ajusté en tenant compte de trois types de voies de circulation rencontrés sur l'espace d'étude (Bourdeau, 1997 ; Hugrel et Joumard, 2004 ; OMINEA, 2015) : les voies urbaines, les routes et les autoroutes. En effet, « *compte tenu du fait que les émissions unitaires propres à un véhicule varient de manière sensible en fonction du type de voie* […] *on cherche ensuite à les répartir sur les trois types de voie de circulation* » (Hugrel et Joumard, 2004, p. 9). Par exemple, il apparait évident que le pourcentage de PL présents sur les autoroutes sera plus important que le taux de PL présents sur les voies urbaines ou les routes. C'est typiquement le cas des transports de marchandises les plus lourds, dont les restrictions de circulation en milieu urbain[176] limitent nécessairement le nombre en milieu urbain. Enfin, il faut noter que les pourcentages de types de véhicules sont renseignés pour chacun des axes du domaine d'étude, permettant, ensuite de bénéficier d'un trafic moyen pour chaque type de véhicules, qui sera intégré dans les calculs de CIRCUL'AIR.

### b. *Les profils temporels de fréquentation des axes routiers*

Pour finir, la dernière donnée d'entrée relative aux calculs des émissions de la PAA est à rattacher aux infrastructures de transport, en combinaison avec les comptages routiers exploités pour l'année 2008 (156 capteurs électromagnétiques). Ce traitement vise à associer à la hiérarchie des axes routiers des profils temporels de la fréquentation type observée par l'intermédiaire des stations.

Pour cela, un travail similaire à celui présenté au cours de la section 3.3.2 (p. 109), à travers l'analyse en ondelettes (*cf.* figure 1.40, p. 111) a été réalisé sur l'année 2008. Cette analyse vise à extraire les profils temporels de fréquentation des axes par une analyse de type *signal sur bruit*, en tenant compte des types de voies de la BD TOPO : les liaisons départementales, les liaisons interurbaines, les liaisons urbaines et les liaisons résidentielles. L'ensemble de ces profils type est

[173] Proposition de loi de la commission des finances du Senat relative à la nocivité du diesel pour la santé (11 Juin 2014) : http://www.senat.fr/rap/l13-600/l13-6001.html.
[174] À noter que tous les PL fonctionnent avec des technologies Diesel.
[175] Pour un kilomètre effectué à 50km/h, un véhicule diesel émet en moyenne 0,43g de $NO_x$ contre 0,086 g de $NO_x$ pour un véhicule essence (d'après les données issues de SETRA, 2009).
[176] Arrêté du 2 mars 2015 relatif à l'interdiction de circulation des véhicules de transport de marchandises à certaines périodes : https://www.legifrance.gouv.fr/eli/arrete/2015/3/2/DEVT1500238A/jo.

illustré en figure 3.25 (ci-dessous)[177] :

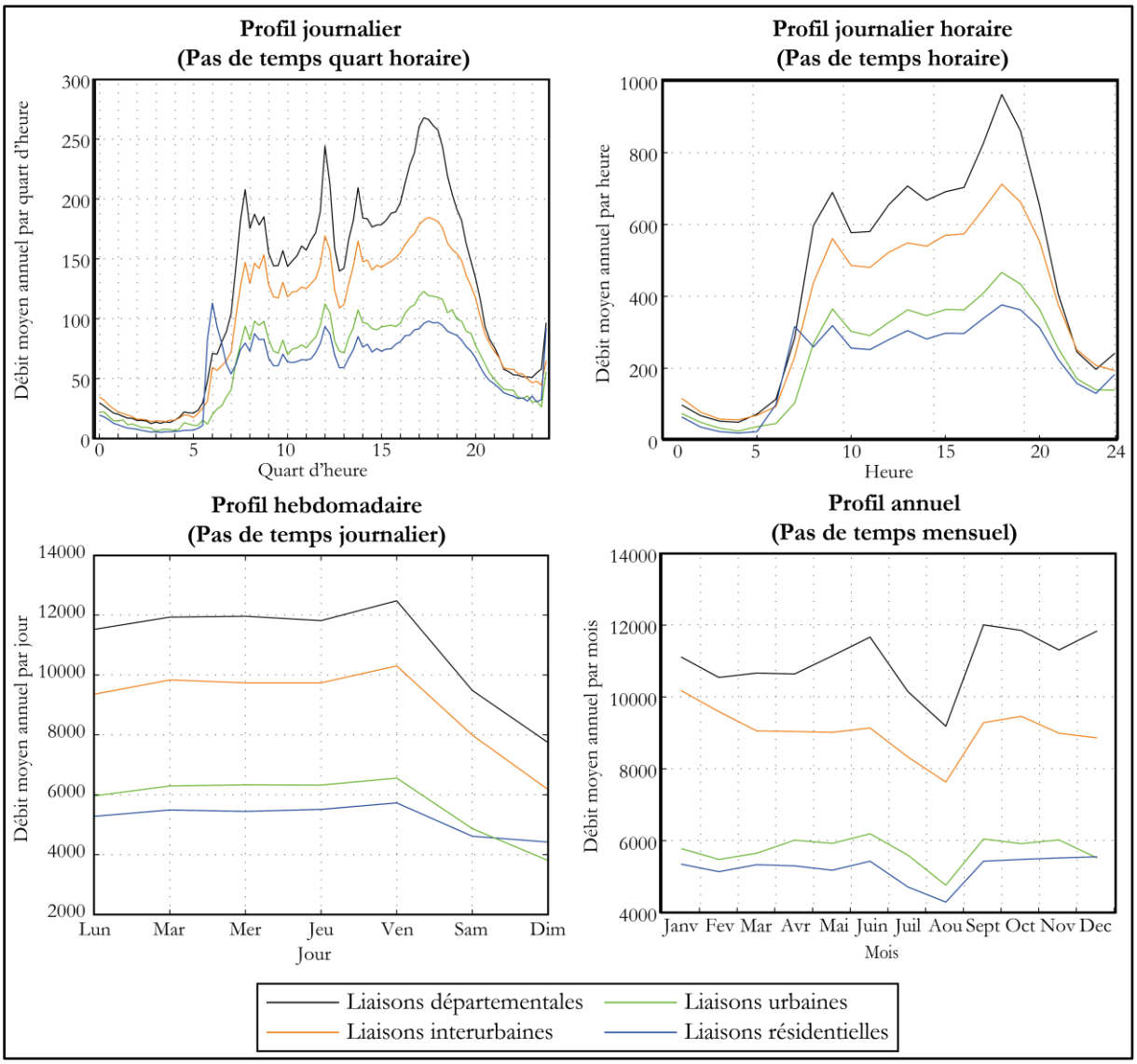



**Figure 3.25 : « Profil type » de la fréquentation des axes routiers observée à différentes temporalités selon la hiérarchie des voies issues de la BD TOPO pour l'année 2008**

En tenant compte de ces informations sur les *profils type* de fréquentation par type de voie (*cf.* figure 3.25), trois informations interdépendantes sont prises en compte pour l'estimation des volumes de circulation routière (ASPA, 2005 ; PCIT, 2012) :

- Un **profil annuel (au pas de temps mensuel)** correspondant à la répartition du trafic annuel entre chaque mois. Celui-ci se caractérise par une alternance saisonnière : on observe une baisse de la fréquentation routière lors de la période estivale et une période de plus forte fréquentation en hiver.
- Un **profil hebdomadaire (au pas de temps journalier)** correspondant à la répartition

[177] En outre, le travail d'extraction des *profils types* en fonction des données de comptages routiers est aujourd'hui utilisé par ATMOSF'AIR BOURGOGNE, lors de son exploitation du réseau de la BD TOPO sur la ville de Dijon, et plus particulièrement du le réseau routier résidentiel. Cette partie du travail sur les profils temporels a permis d'apporter une aide technique et méthodologique sur ces aspects de traitement de signal/bruit des capteurs électromagnétiques.

du trafic entre les jours ouvrables et les week-ends pour chaque mois (*profil mensuel*). À ce niveau, on observe une alternance entre les jours de la semaine où la fréquentation routière est plus importante et le week-end. On remarque par ailleurs que la fréquentation routière augmente progressivement la semaine pour atteindre son apogée le vendredi.

- Un **profil journalier (au pas de temps horaire)** correspondant à la répartition des volumes horaires de trafic pour chaque jour ouvrable et jour de week-end (*profil journalier*), et pour chaque mois (*profil mensuel*). Pour ce dernier profil, on observe la récurrence de la circulation routière journalière, où trois pics de fréquentation s'observent (matin, midi et soir).

À partir de ces clés temporelles, illustrées ci-dessus, CIRCUL'AIR calcule ensuite les volumes horaires de trafic routier pour chaque type de voie, par type de véhicule pour chaque jour de l'année, et de chaque mois de l'année (ASPA, 2005).

### 8.2.3. **Calcul des émissions annuelles avec CIRCUL'AIR : un calcul intégré**

Il est important de relever que, mis à part pour le réseau routier et pour les données de trafic routier qui lui sont associées, les données d'entrée au niveau de CIRCUL'AIR sont identiques aux données qui ont été mobilisées pour la réalisation du cadastre d'émissions de la PAA exploité par l'AASQA en 2008, comme c'est le cas du parc automobile. Cela permettra de comparer directement les sorties d'émissions établies à partir du modèle SCAUP et celles établies à partir des estimations réalisées par ATMOSF'AIR BOURGOGNE en 2008. En effet, le changement majeur en termes de données d'entrée est à rattacher à la géométrie du réseau routier (BD TOPO versus BD CARTO) ainsi qu'aux valeurs de circulation (Observation versus Simulation). Une fois les données d'entrée misent en place, le calcul des émissions est réalisé en s'appuyant sur le logiciel CIRCUL'AIR (version 3.0), dont le fonctionnement général est schématisé au sein de la figure 3.21 (p. 269).

C'est un logiciel développé par et pour les AASQA afin de mutualiser et de faciliter la réalisation des différentes étapes de calcul synthétisées en figure 3.21 (p. 269) en prenant « *en compte les éléments publiés dans le cadre du Pôle de Coordination national des Inventaires territoriaux routie*r *(PCIT, 2012)* » (ASPA, 2012, p. 24). Par ailleurs, il intègre dans le calcul des émissions l'ensemble des éléments permettant « *le calcul des émissions à chaud et à froid, mais également des émissions par évaporation et par usure des pneus, des freins et de la route ainsi que la remise en suspension calculées […]. Il intègre également des corrections des émissions en fonction de l'âge du véhicule, de l'utilisation de la climatisation, de l'évolution des carburants* » (ASPA, 2012, p. 24). Ces variables ont été définies en concordance avec les valeurs utilisées par ATMOSF'AIR BOURGOGNE, pour ne pas inclure trop de modification sur les données d'entrée et en vue d'assurer une cohérence des données calculées lors de la confrontation des résultats d'émissions de PAA.

D'une manière générale, en intégrant ces différents éléments qui sont difficilement exploitables par tous (émissions chaud/froid, usure, remise en suspension,…), le logiciel permet alors le calcul des émissions de PAA par la réalisation de trois grandes étapes de calculs, qui s'appuient sur les différentes données d'entrée précédemment exposées, dont le principe général est exposé en figure 3.21 (p. 269) :

- Le **calcul des volumes horaires et annuels** de trafic routier en s'appuyant sur les données d'entrée de débit (*a*), de réseau routier (*b*) et de profil temporel type (*d*)
- Le **calcul des coefficients de charge**, en s'appuyant sur les vitesses de circulation (*b*), la capacité des axes routiers (*b*), ainsi que sur les valeurs de débits (*a*)
- Le **calcul des équations COPERT**[178] (Ntziachristos et Zissis, 2009 ; AEE, 2011) pour les 230 types de véhicules (*c*), en s'appuyant sur les trafics horaires et annuels calculés précédemment, le parc automobile (*c*), la longueur et la pente des tronçons routiers (*b*).

Pour cette dernière étape, comme nous l'avons exposé précédemment (*cf.* section 5.3.3, p. 179), il est important de relever que COPERT est un modèle basé sur les vitesses de parcours horaire moyennes, permettant ensuite d'appliquer un facteur d'émissions moyen tenant compte du type de véhicule et de la vitesse moyenne d'un axe routier (ASPA, 2012 ; Ntziachristos et Zissis, 2009). Les estimations d'émissions qui en ressortent s'expriment sous la forme d'une quantité en g/km par unité de temps. Une fois l'ensemble des calculs mis en place, les résultats d'émissions sont ensuite agrégés au niveau des tronçons routiers pour tous les polluants majeurs que l'on est susceptible de respirer à savoir (ASPA, 2005, 2012): les principaux polluants toxiques (CO, $NO_x$, COV, PM, $NH_3$, $SO_2$, métaux lourds) et les principaux polluants radiatifs ($CO_2$, $N_2O$ et $CH_4$).

## 8.3. Spatialisation de la pollution atmosphérique automobile sur la ville de Dijon

Au cours de cette partie, nous nous focaliserons plus spécifiquement sur les émissions d'oxydes d'azote ($NO_x$) et à la présentation des résultats des calculs d'émissions de ces polluants établis en collaboration avec ATMOSF'AIR BOURGOGNE. En effet, à l'échelle locale, les polluants primaires comme les $NO_x$ sont ceux qui contribuent le plus à une mauvaise qualité de l'air en milieu urbain (OMS, 2006 ; Sportisse, 2007a). Si le trafic routier contribue en très grande partie aux émissions de ces trois polluants, il faut aussi relever que les $NO_x$ sont souvent privilégiés car ce sont de très bons polluants traceurs (*proxy*) du trafic routier (OMS, 2006 ; Keuken *et al.*, 2010 ; CITEPA, 2015). Une étude précédente, réalisée dans le cadre d'un mémoire (Emery, 2012), a permis de mettre en avant un lien entre les flux de circulation routière et les niveaux de $NO_x$ relevés à proximité des routes sur la ville de Dijon sur les temporalités journalières. Cette précédente étude a permis d'appréhender les relations complexes entre le flux de circulation routière, les conditions météorologiques (température, humidité) et les concentrations $NO_x$ sur la ville de Dijon. C'est pourquoi au cours de cette section les principaux résultats cartographiques et les analyses que nous apportons se focalisent plus spécifiquement sur les émissions de $NO_x$, dans le sens où l'étude de ce polluant apparait bien adaptée à la problématique de la pollution atmosphérique de proximité (*cf.* section 1.1.3, p. 31).

Différents questionnements seront soulevés : dans quelle mesure sommes-nous capables avec une approche de SMA telle que présentée dans le cadre de cette thèse d'améliorer les estimations de PAA à l'échelle d'un territoire comme la ville de Dijon ? Quels sont les potentiels opérationnels ? Quels sont les apports et les limites du modèle SCAUP pour l'estimation des émissions de PAA sur le territoire de la ville de Dijon en comparaison des estimations réalisées par les AASQA ? Dans la suite, nous présenterons toujours les résultats sur la base des

[178] La version de COPERT utilisée pour le calcul des émissions se base sur la version IV (4e génération) qui date de 2007 (AEE, 2011).

estimations réalisées par l'AASQA, qui constitue la seule base de données accessible et exploitable à l'échelle de la ville de Dijon.

### 8.3.1. Le bilan des émissions atmosphériques à l'échelle communale : un constat contrasté selon l'échelle d'étude

L'échelle communale constitue le premier niveau d'étude des cadastres d'émissions (PCIT, 2012). Pour cela, nous avons dans un premier temps confronté les valeurs cumulées d'émissions de polluants sur la ville de Dijon. Pour commencer il faut relever que sur la ville de Dijon, d'après les estimations d'ATMOSF'AIR BOURGOGNE établies dans le cadre du PPA en 2008, le secteur du trafic routier contribue à plus de 60 % des émissions atmosphériques sur la ville de Dijon (*cf.* tableau 3.3, ci-dessus). Ces premiers éléments montrent toute l'importance qu'il y a à appréhender les impacts atmosphériques du trafic routier sur notre terrain d'étude. De plus, ces premiers éléments sont en concordance avec les tendances observées à l'échelle nationale comme nous l'avons présenté au cours de la section 8.1.1 (p. 265) :

| | | Emissions de $NO_x$ (kg cumulé) | Différence* |
|---|---|---|---|
| **Atmosf'Air** | Trafic routier | 589 689,20 | |
| | Total des secteurs | 923 692,02 | |
| **SCAUP** | H10 | 441 175,57 | -25,19% |
| | H15 | 552 501,46 | -6,31% |
| | H20 | 661 672,94 | 12,21% |

* (Simulation - ATMOSF'AIR) / ATMOSF'AIR

**Tableau 3.5 : Synthèse des émissions de $NO_x$ (cumul en kg) estimées à l'échelle de la ville de Dijon pour l'année 2008**

Sur l'espace communal, la confrontation des valeurs d'émissions calculées à partir des données de simulation de SCAUP avec les émissions estimées par ATMOSF'AIR BOURGOGNE nous apporte quelques précisions sur les potentiels d'utilisation du modèle de SMA. En s'appuyant sur le tableau ci-dessus, on constate que le paramètre de simulation *H15* se rapproche le plus des estimations réalisées par l'AAQA, les paramètres *H10* et *H20* étant respectivement en dessous (- 25 %) et au-dessus (+ 12 %) des estimations réalisées par ATMOSF'AIR BOURGOGNE (*cf.* tableau 3.5). Cette première tendance se rapproche des éléments d'évaluation de SCAUP observés lors de la comparaison du trafic simulé avec les observations. En effet, en fonction des trois paramètres de simulation sélectionnés *H10, H15 et H20* (*cf.* section 7.3, pp. 246-254), on remarque des tendances relativement similaires aux constats précédents.

Ainsi, à notre sens, ces éléments sont à rattacher directement aux données de trafic routier issues de SCAUP. Ces résultats ne sont que peu surprenants du fait que les niveaux d'émissions en $NO_x$ tiennent compte, tout du moins en partie (Emery, 2012), de l'intensité des flux de circulation routière sur les tronçons routiers. Toutefois, ces premiers éléments soulèvent quelques questions concernant les estimations réalisées par l'AASQA. Effectivement, les réseaux routiers ayant des résolutions et une longueur totale foncièrement différentes (*cf.* figure 3.4, p. 232), on pourrait

intuitivement penser que le résultat du calcul des émissions issues des données de trafic routier simulé serait bien supérieur aux données d'émissions calculées par ATMOSF'AIR BOURGOGNE.

C'est pourquoi, dans un second temps, il est intéressant de se focaliser non plus sur une l'échelle communale, mais plutôt sur les réseaux routiers et leur hiérarchie pour mieux observer les différentes tendances calculées. Le tableau 3.6 (ci-dessous) illustre les différences d'émissions en $NO_x$ par type d'axe du territoire d'étude[179], et en tenant compte du référentiel issu de la BD CARTO (*cf.* section 5.3.2.a, p. 272) :

| Vocation Réseau | $NO_x$ (kg cumulé) | | | | | | |
|---|---|---|---|---|---|---|---|
| | ATMOSF'Air | SCAUP | | | Différence des estimations * | | |
| | | H10 | H15 | H20 | H10 | H15 | H20 |
| Ensemble | 382 245,82 | 480 601,64 | 605 132,79 | 725 506,14 | 26% | 58% | 90% |
| Liaisons locales | 27 345,03 | 28 136,10 | 39 485,82 | 51 298,94 | 3% | 44% | 88% |
| Liaison régionales | 280 852,14 | 130 975,93 | 170 267,29 | 216 192,84 | -53% | -39% | -23% |
| Liaisons principales | 158 462,62 | 250 058,18 | 302 545,20 | 343 177,89 | 58% | 91% | 117% |

* $\frac{\text{(Simulation - ATMOSF'AIR)}}{\text{ATMOSF'AIR}}$

**Tableau 3.6 : Synthèse des émissions de $NO_x$ (cumul en kg) estimées par type de voie selon la hiérarchie de la BD CARTO**

En s'appuyant sur le tableau 3.6, on remarque des résultats plus contrastés que précédemment, dans le sens où, sur l'ensemble des axes routiers pris en compte, les estimations réalisées à partir du modèle de SMA du trafic routier sont toujours supérieures au référentiel de l'AASQA. Ce constat montre que sur l'ensemble des axes exploités pour le calcul des émissions de $NO_x$, les différences d'estimation sont amplifiées de 25 % à 90 % par rapport aux estimations de l'AASQA. Cependant, on observe que si ces tendances se vérifient sur les liaisons locales et principales, au niveau des liaisons régionales la tendance est inverse. On remarque alors un effet du temps de stationnement sur ces différences, dans le sens où plus le temps de stationnement augmente plus la différence des estimations diminue.

Ces résultats montrent de nouveau les limites de la matrice de transition qui vise à répartir les véhicules en fonction de ces hiérarchies. En revanche, ces différents éléments nous montrent tout de même un premier apport de SCAUP à l'échelle des tronçons routiers, et plus spécifiquement sur les liaisons principales qui constituent les réseaux routiers et supportent les volumes de trafic routier les plus importants (*cf.* tableau 3.6 en bleu). Il faut toutefois nuancer nos propos, car les différences d'estimation sont aussi à rattacher à la dimension des réseaux routiers exploités, dont nous avons soulevé les caractéristiques précédemment (*cf.* tableau 3.4, p. 273). Dès lors, il est important d'affiner notre analyse en nous focalisant plutôt sur les tronçons routiers que sur la commune de Dijon, étant donné que ces premiers résultats ne permettent que de faire un constat général des estimations.

[179] On remarque que les émissions cumulées sont supérieures aux émissions présentées au sein du tableau 3.5. Ces différences s'explique par le zonage d'étude pris en compte : en effet certains axes (comme la rocade Georges Pompidou à l'Est de la ville) sont situés hors du zonage communal de Dijon.

### 8.3.2. **Bilan des émissions de $NO_x$ à l'échelle des tronçons routiers de la ville**

Les principaux enseignements à l'échelle communale montrent que les calculs d'émissions en cumul de $NO_x$ pour le paramètre *H15* sont relativement similaires aux estimations de l'AASQA locale (*cf.* tableau 3.5). Néanmoins, lorsque l'on se focalise sur la hiérarchie des tronçons routiers, on remarque que nos estimations leur sont plutôt supérieures sur l'ensemble des axes routiers. Cependant, en est-il de même pour les autres polluants ?

La figure 3.26 représente la répartition des émissions de polluants à l'échelle des tronçons routiers pour les quatre polluants que l'on rencontre le plus généralement en ville (CITEPA, 2015) : les $NO_x$, le CO, les $PM_{10}$ et les $PM_{2,5}$. À noter que le calcul des émissions de particules ($PM_{10}$ et $PM_{2,5}$) donne un résultat sous-estimé en raison de problèmes techniques qui sont inhérents au logiciel CIRCUL'AIR. Afin de pouvoir comparer les estimations sur une même base, les émissions de $NO_x$ à l'échelle des tronçons routiers sont représentées non plus en quantités de polluant totales, mais plutôt en quantités de polluants émis par mètre[180] comme illustré ci-dessous :

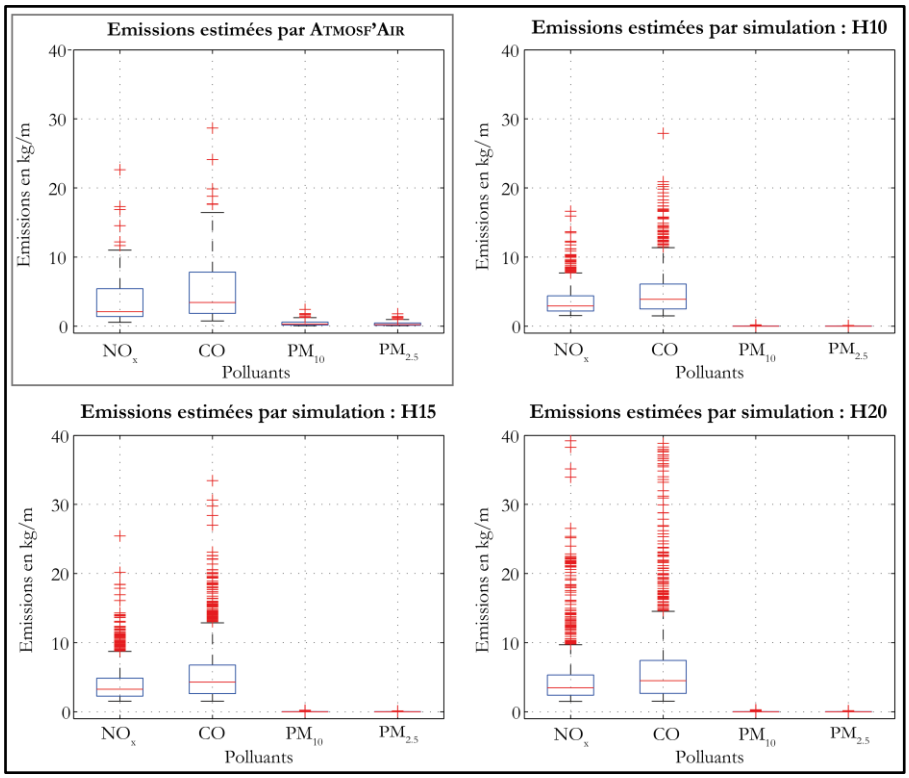



**Figure 3.26 : Répartition des émissions de polluants par axe routier (en kg/m) estimées à partir de CIRCUL'AIR**

À travers cette première illustration, on remarque que les estimations d'émissions de polluants réalisées par ATMOSF'AIR BOURGOGNE (A) et celles réalisées à partir des données de SCAUP en section 7.3.3 présentent une répartition relativement analogue (B, C et D). On constate un effet de la taille du réseau routier employé entre les estimations réalisées par l'AASQA et les estimations réalisées à partir du modèle de SMA du trafic routier. Ceci explique, en particulier pour les émissions de $NO_x$ et de CO, le positionnement des valeurs extrêmes (*outliers*) au-delà de

[180] Du fait que nous sommes au sein d'un espace urbain, nous avons privilégié le mètre, plutôt que le kilomètre, pour cartographier les estimations d'émissions en $NO_x$ issues des calculs de CIRCUL'AIR.

la frontière haute des boxplots au sein de la figure 3.26. En d'autres termes, pour les simulations, cette répartition montre une hétérogénéité des émissions de polluants entre un grand nombre d'axes routiers ayant des émissions relativement faibles (situées entre 1 et 10 kg de polluants par mètre) et un nombre d'axes moins important qui concentre les valeurs d'émissions les plus importantes (au-delà de 10 kg/m d'émissions de $NO_x$ et au-delà de 18 kg/m pour le CO). Si ces premières tendances sont relativement acceptables, en ce qui concerne la répartition des émissions de particules ($PM_{10}$ et les $PM_{2,5}$) les estimations montrent une répartition plus resserrée des concentrations par axe routier sur le territoire de la ville de Dijon.

On observe d'ailleurs que pour ces deux derniers polluants ($PM_{10}$ et les $PM_{2,5}$), les émissions sont bien plus faibles pour les émissions estimées à partir des données de SCAUP (0,0053 kg/m en moyenne) que pour celles estimées par ATMOSF'AIR BOURGOGNE (0,50 kg/m en moyenne), soit une différence de près de 100 fois moins de polluants cumulés calculés ; tandis que les émissions de $NO_x$ et de CO présentent des valeurs bien plus acceptables, avec pour ces deux polluants une différence très faible (pour un niveau de près de 200 g de polluants émis en moyenne). C'est pourquoi, dans un second temps, nous nous focaliserons plus spécifiquement sur la carte des émissions de $NO_x$ évaluées par ATMOSF'AIR BOURGOGNE, en figure 3.27 (ci-dessous), et de celles estimées à partir des données de simulation en moyenne horaire (en bas) et journalière (en haut), comme illustré en figure 3.28 pour *H10* (p. 286), en figure 3.29 pour *H15* (p. 287), et en figure 3.30 pour *H20* (p. 288), ci-après :

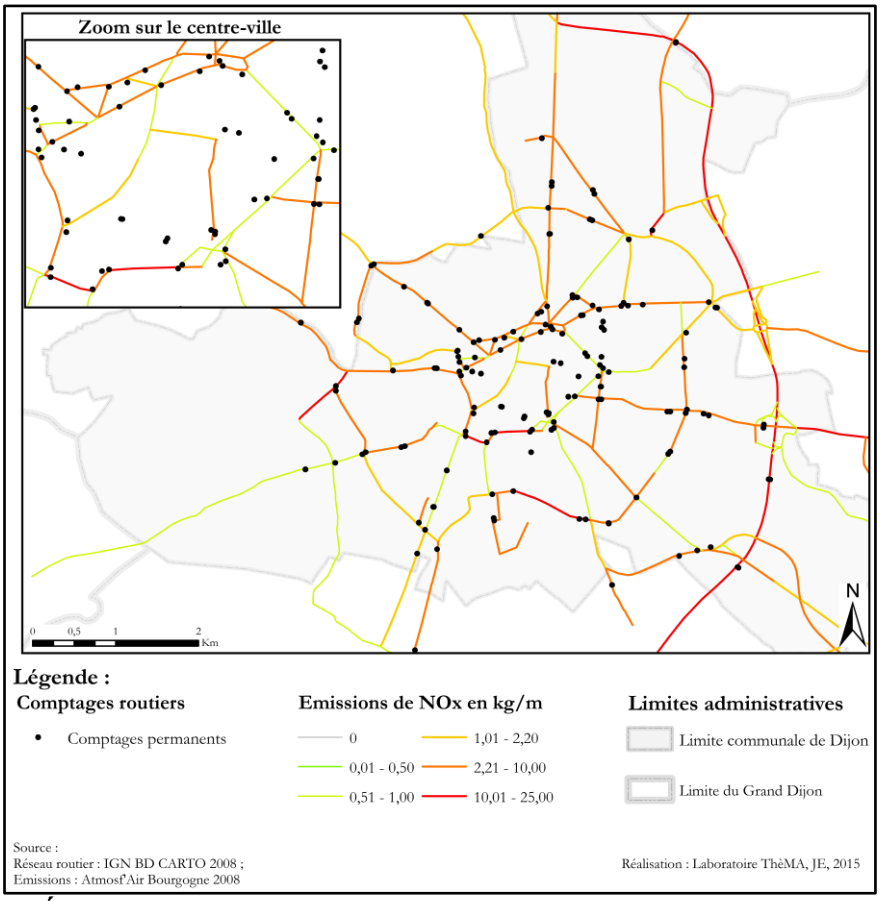



**Figure 3.27 : Émissions de $NO_x$ en kg/m obtenues pour l'année 2008 par ATMOSF'AIR BOURGOGNE (source : ATMOSF'AIR BOURGOGNE : PPA, 2008)**

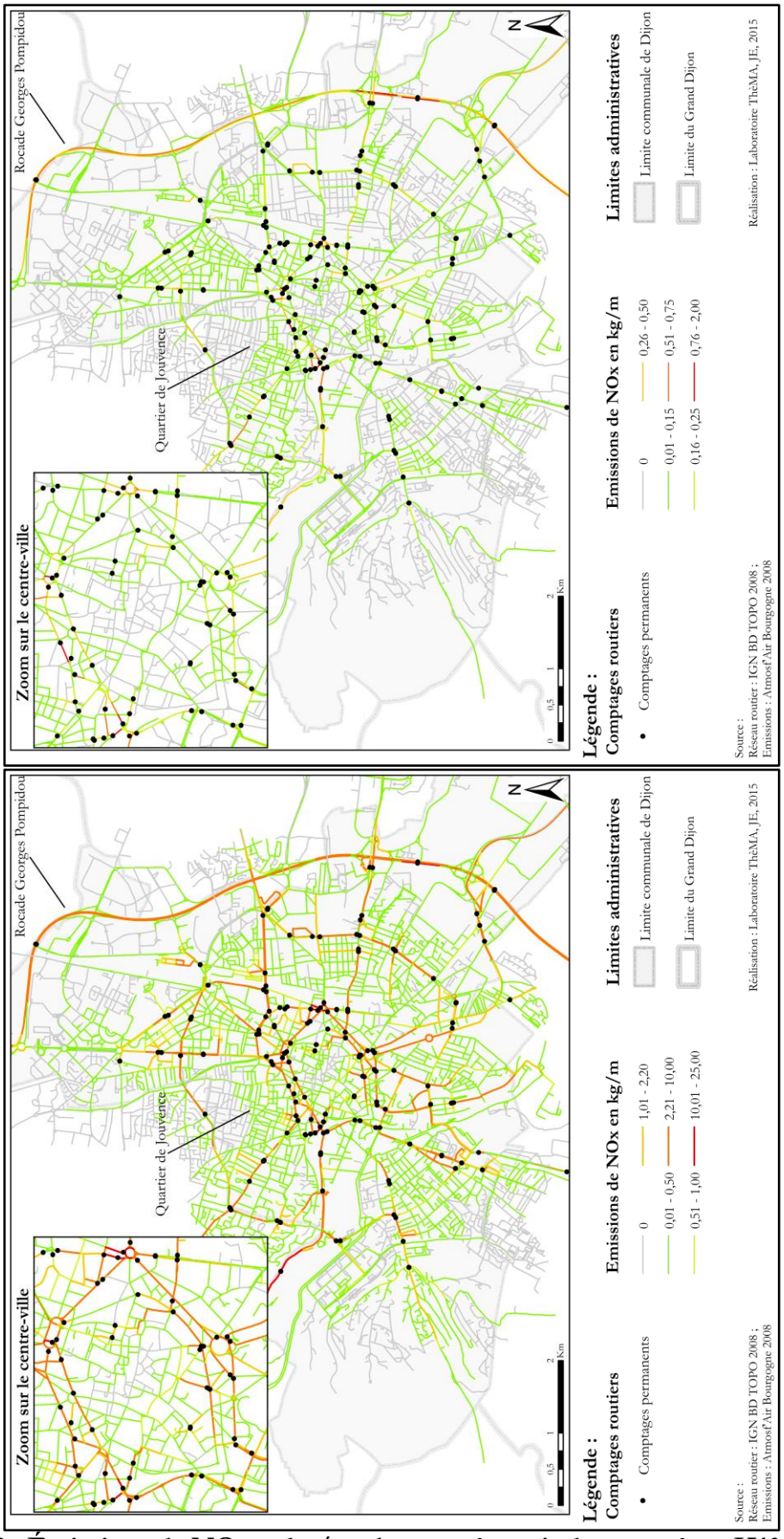



**Figure 3.28 : Émissions de $NO_x$ en kg/m obtenues à partir du paramètre *H10* de SCAUP en moyenne horaire (en haut) et en moyenne journalière (en bas)**

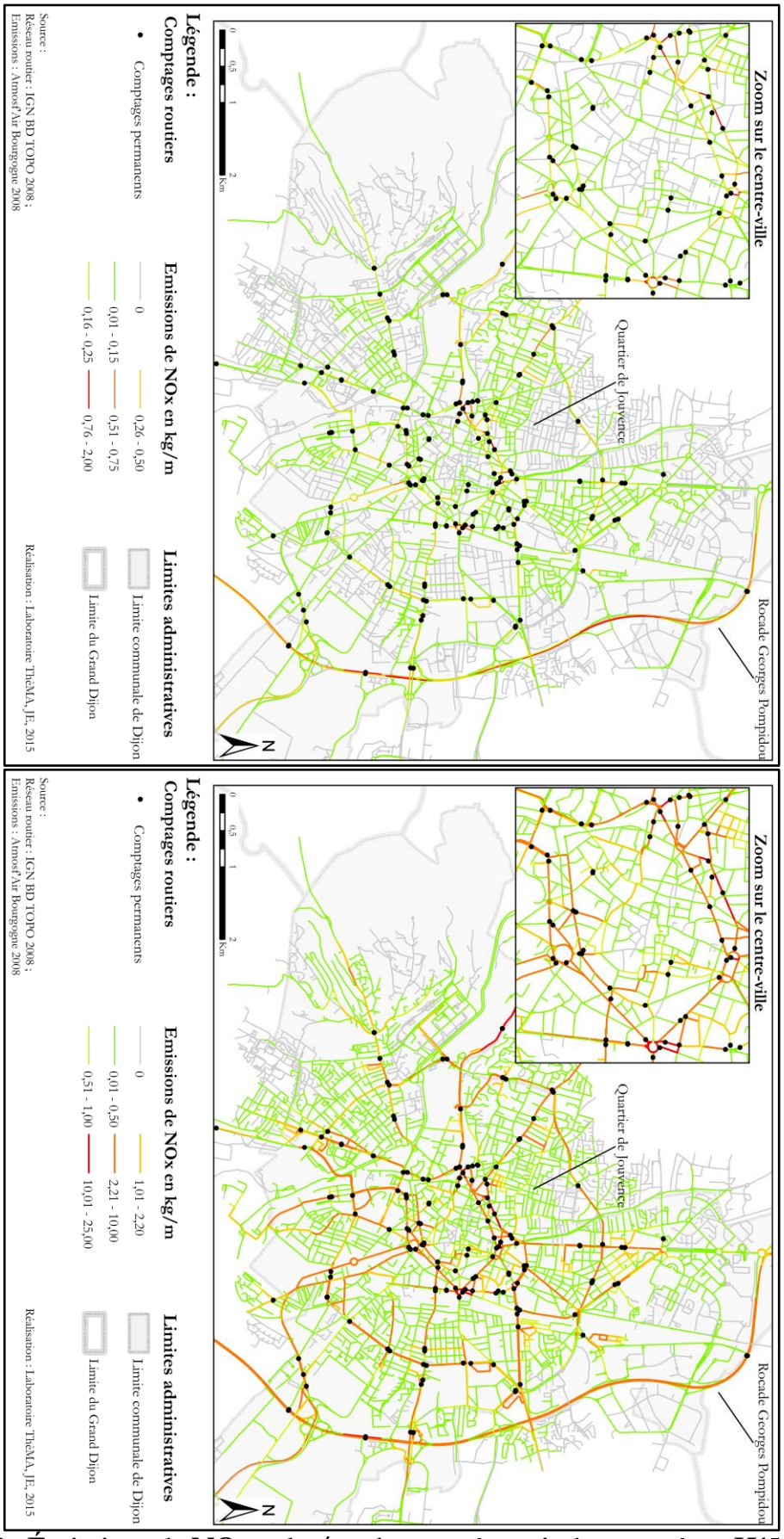



**Figure 3.29 : Émissions de $NO_x$ en kg/m obtenues à partir du paramètre *H15* de SCAUP en moyenne horaire (en haut) et en moyenne journalière (en bas)**




**Figure 3.30 : Émissions de $NO_x$ en kg/m obtenues à partir du paramètre *H20* de SCAUP en moyenne horaire (en haut) et en moyenne journalière (en bas)**

D'une manière générale, la répartition des émissions de $NO_x$ suit la morphologie du réseau routier dijonnais, où les boulevards périphériques ainsi que la rocade Georges Pompidou (située à l'Est du territoire) présentent les taux d'émission les plus importants de la ville. On observe alors que l'apport du réseau routier employé dans SCAUP (BD TOPO) met en exergue ces réseaux routiers, dont la forme générale rappelle la forme spatiale des estimations réalisées par ATMOSF'AIR BOURGOGNE (*cf.* figure 3.27*,* p. 285). En revanche, en intégrant les axes résidentiels ainsi qu'un plus grand nombre de pénétrantes au sein du territoire de la ville de Dijon, ces cartes illustrent plus finement les constats établis précédemment (*cf.* figure 3.26, p. 284). Ces premiers éléments expliquent aussi les différences d'estimation constatées pour les axes routiers au sein du tableau 3.6 (p. 283).

Dans le même ordre, on observe cette tendance au niveau des émissions du centre-ville (*cf.* encart du centre-ville au sein des illustrations cartographiques). Or on constate que la pénétrante Nord-Sud du centre-ville, qui apparait sur les données d'émissions d'ATMOSF'AIR BOURGOGNE (*cf.* figure 3.27), ne ressort pas aussi nettement sur les trois autres cartes (*cf.* figure 3.28*,* figure 3.29*, et* figure 3.30). Ici, on observe les différences d'émissions que nous avons constatées sur les liaisons régionales au sein du tableau 3.6 (p. 283), où le bilan global de nos estimations est inférieur aux données d'émissions estimées par l'AASQA sur ce type de voie. On remarque que cette forme pour les émissions journalières s'observe sur l'ensemble des cartes (en bas). En revanche, les estimations calculées en tenant compte du trafic moyen horaire ne permettent pas de faire ressortir nettement la hiérarchie du réseau routier. Ce constat s'explique en partie par les données de trafic routier, qui sont moins importantes que pour les données de trafic routier simulé à la journée. De plus, en se focalisant sur le positionnement des capteurs sur le réseau routier, on observe un effet de leur positionnement sur les émissions de $NO_x$ estimées à partir des débits moyens horaires. De ce fait, on remarque d'ailleurs qu'à cette granularité temporelle le positionnement des capteurs électromagnétiques apparait fondamental pour les estimations de PAA. En effet, on observe que lorsque la densité de capteurs sur le réseau routier est importante, les niveaux d'émissions le sont aussi. On retrouve ici les effets de notre approche guidée par les données dans le sens où les capteurs se situent en entrée du modèle SCAUP. *À contrario*, ces effets apparaissent moins nettement pour les estimations journalières que pour les estimations horaires.

Enfin, notre travail permet d'affiner spatialement les estimations de pollution atmosphérique tout en conservant la forme générale, où les réseaux routiers qui concentrent le plus de trafic routier sont ceux qui ont les taux d'émissions les plus importants. À ces « artères » urbaines se juxtaposent les réseaux résidentiels, qui présentent quant à eux un taux de $NO_x$ émis relativement faible, ce qui n'est que peu surprenant. On remarque d'ailleurs pour les estimations journalières (en bas des figures) un effet du temps de stationnement sur le calcul des émissions d'oxydes d'azote, dans le sens où plus le temps de stationnement augmente (de 10 à 20 minutes), plus la couverture du réseau routier résidentiel est importante. C'est particulièrement le cas de la partie Nord du quartier résidentiel de Jouvence (situé au Nord-Est du centre-ville), qui est de mieux en mieux couverte à mesure que le temps de stationnement augmente. Au niveau des estimations horaires, on remarque la tendance inverse, ce qui parait surprenant, car en effet, à ce niveau plus le temps de stationnement augmente plus les valeurs d'émissions horaires de $NO_x$ sont faibles, ce qui est particulièrement le cas de la rocade Georges Pompidou (à l'Est). Cette observation

apparait alors contre-intuitive par rapport aux émissions journalières, qui présentent une tendance inverse. On peut supposer que ce constat est probablement dû à l'extraction des débits moyens de trafic routier simulé à l'heure, qui a plutôt tendance à lisser les valeurs de forte fréquentation du trafic routier comme les pics de circulation routière (matin, midi, soir). Une grande partie de ces résultats est à associer aux valeurs de fréquentation simulées à partir du modèle SCAUP.

Si notre approche tend à apporter une première réponse pour affiner et pour répondre aux problématiques d'évaluation de la PAA de proximité en milieu urbain, ces éléments ne nous permettent pas réellement d'apprécier l'apport de notre travail pour l'estimation des émissions de PAA en comparaison des estimations réalisées par l'AASQA locale. Une grande partie de ces résultats est à associer aux valeurs de fréquentation simulées à partir de SCAUP. C'est pourquoi nous nous arrêterons, dans un dernier temps, plutôt sur une analyse conjointe des estimations réalisées par l'AASQA et des estimations réalisées à partir des données de trafic routier simulé à l'intérieur d'un même carroyage.

### 8.3.3. **Une réponse pour la spatialisation des émissions de PAA ?**

La confrontation des émissions de $NO_x$ a été réalisée en se basant sur un carroyage régulier d'une résolution de 100 mètres (*cf.* section 8.2.1.a, p. 272). À chacune des mailles du carroyage, les émissions de $NO_x$ estimées à partir des données fournies par ATMOSF'AIR BOURGOGNE et à partir de données simulées de SCAUP sont cumulées Dans ce sens, il est possible de confronter les données d'émissions en se reposant sur une analyse du taux d'évolution de nos estimations conformément au référentiel transmis par l'AASQA. La figure 3.31 synthétise la répartition du taux d'évolution des émissions de $NO_x$ cumulées par maille (en ordonnées), et pour chacun des paramètres de simulation sélectionné (en abscisses pour *H10*, *H15* et *H20*). Les représentations cartographiques des carroyages, ainsi que les relations linéaires entre d'une part les estimations d'ATMOSF'AIR BOURGOGNE et d'autre part les paramètres de simulation, sont représentées quant à elles en figure 3.32 pour *H10* (p. 291), en figure 3.33 pour *H15* (p. 292) et en figure 3.34 pour *H20* (p. 293). Avec la figure ci-dessous on constate, de prime abord, que l'ensemble des valeurs d'émissions est majoritairement inférieur aux émissions estimées par ATMOSF'AIR BOURGOGNE. Cependant, les cartes des différences d'estimations ont tendance à nuancer ce constat :

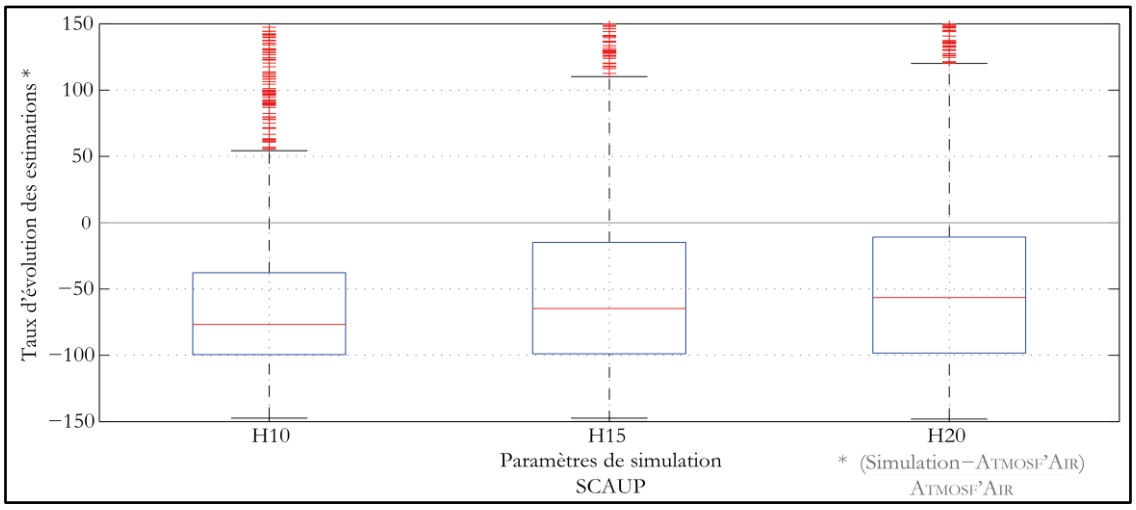



**Figure 3.31 : Taux d'évolution des émissions de $NO_x$ par paramètre de SCAUP en fonction des estimations réalisées par ATMOSF'AIR BOURGOGNE**

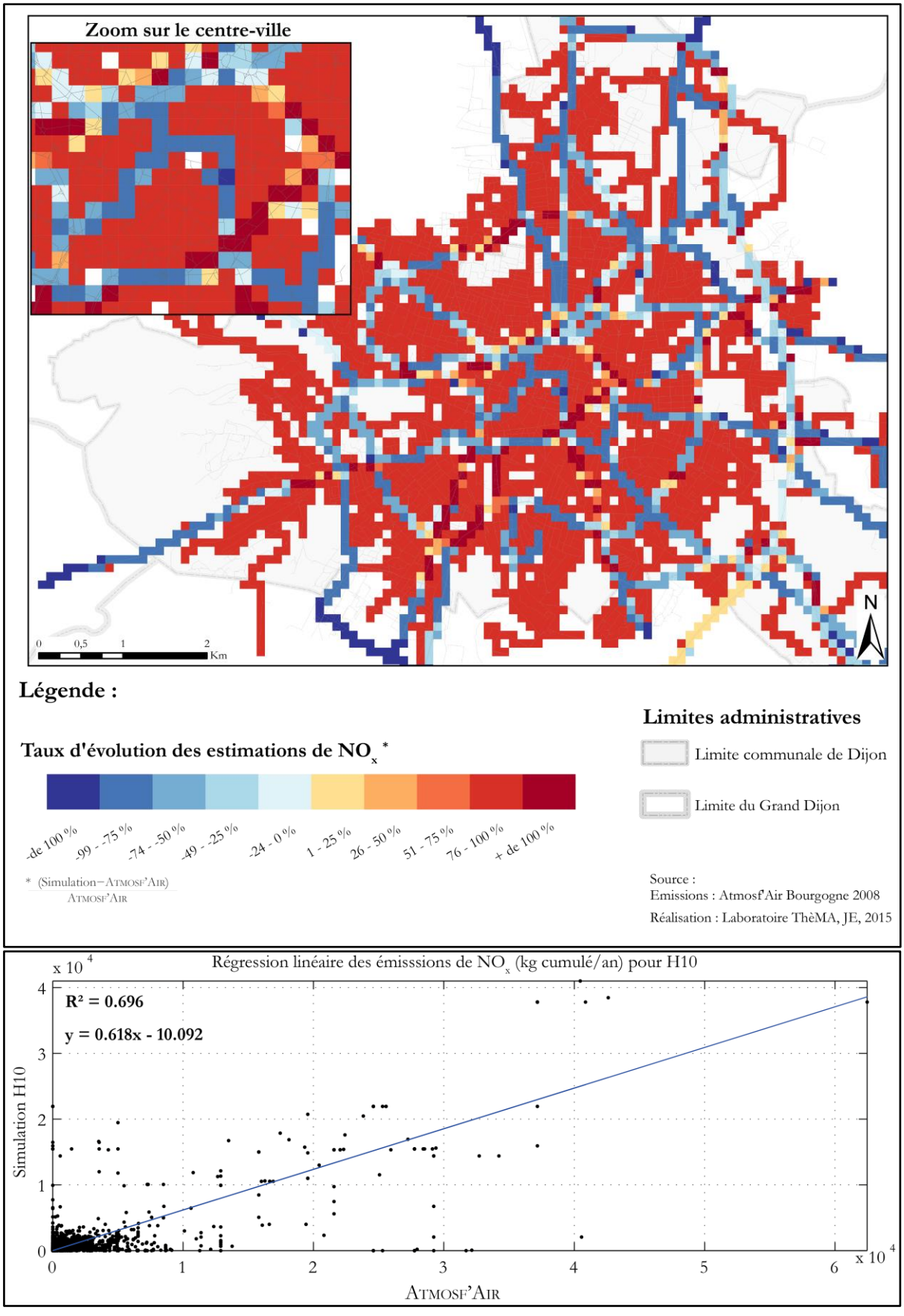



**Figure 3.32 : Analyse par carroyage des émissions de $NO_x$ pour le paramètre de SCAUP *H10* par taux d'évolution (en haut) et par régression linéaire (en bas)**

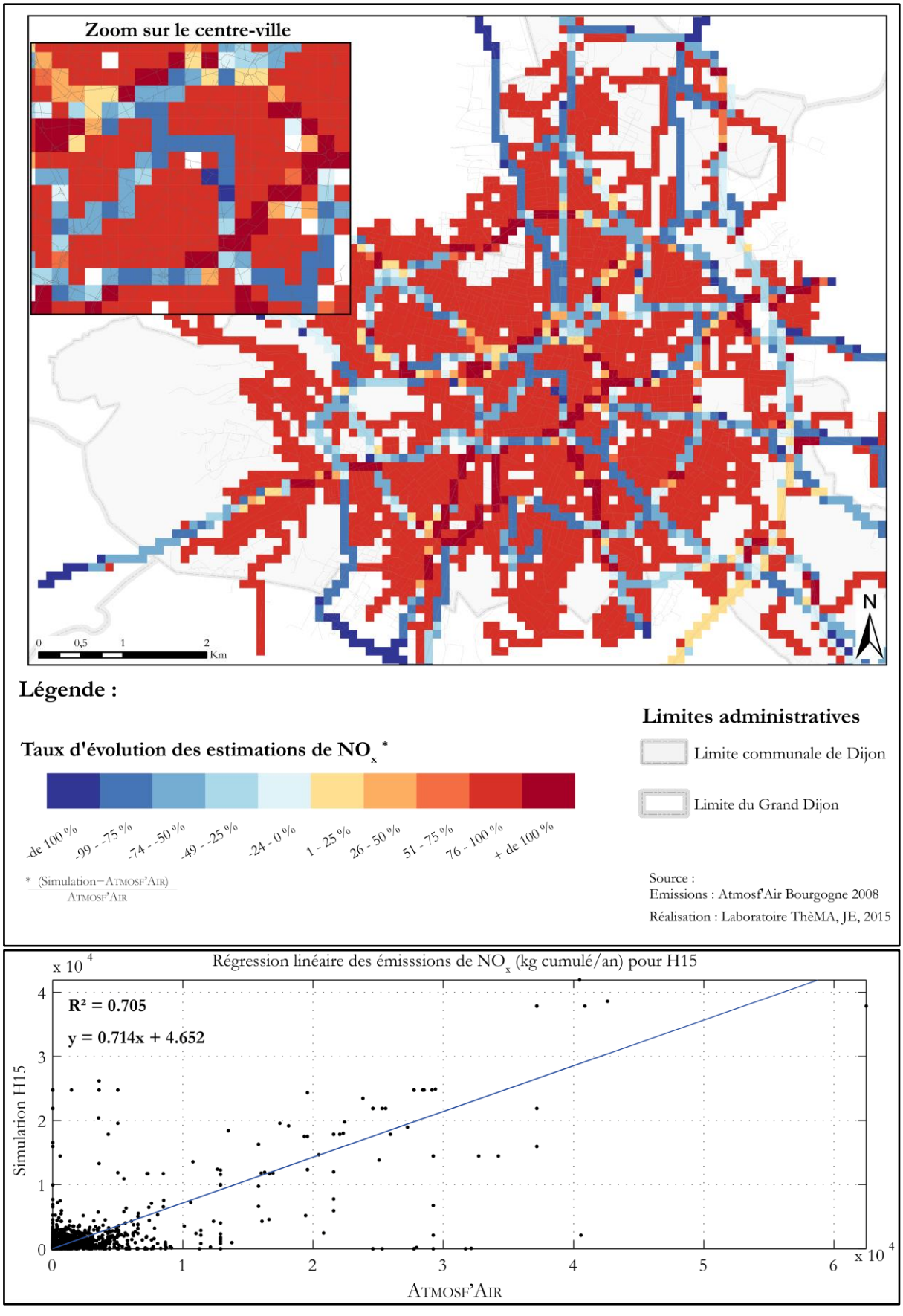



**Figure 3.33 : Analyse par carroyage des émissions de $NO_x$ pour le paramètre de SCAUP *H15* par taux d'évolution (en haut) et par régression linéaire (en bas)**

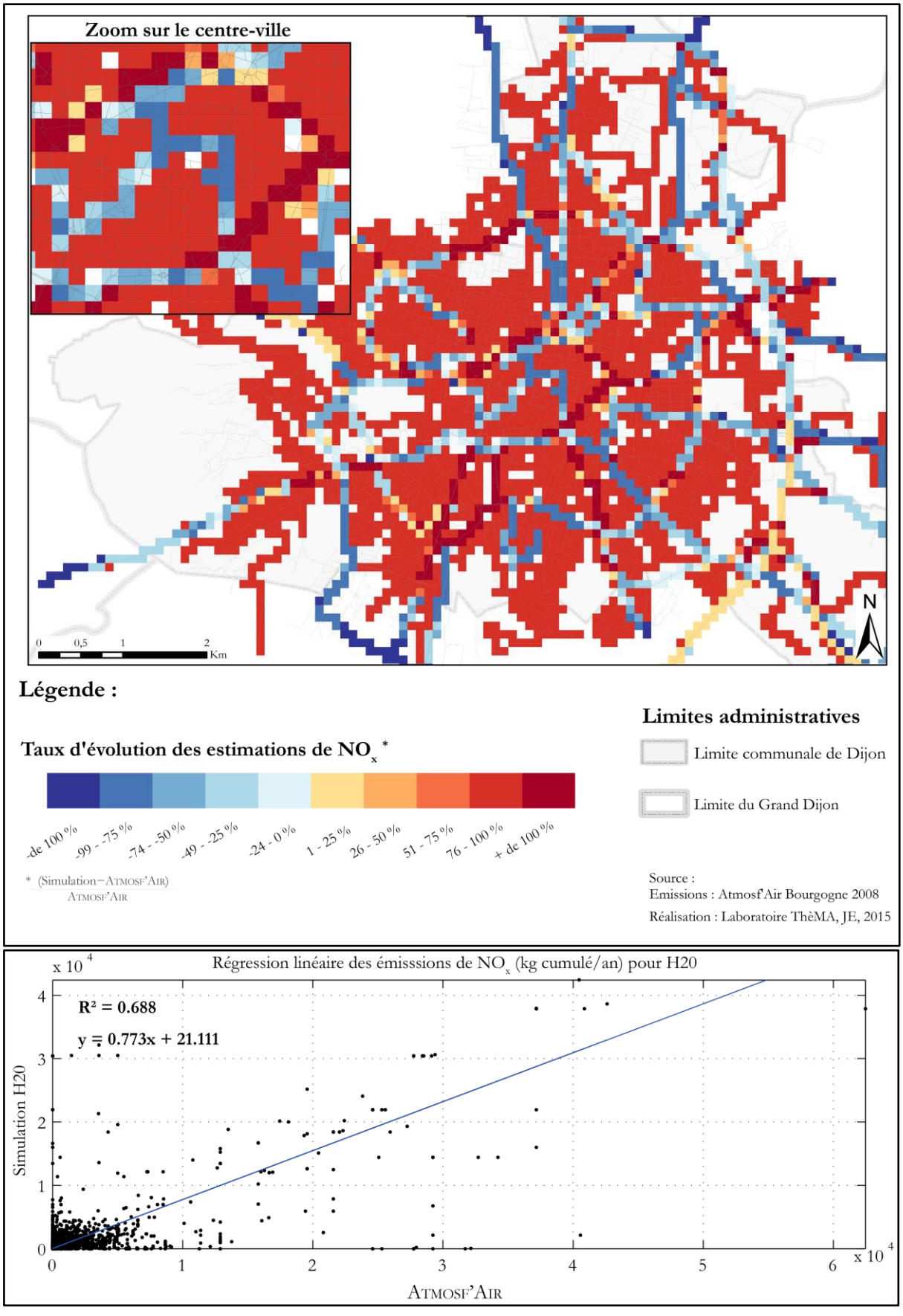



**Figure 3.34 : Analyse par carroyage des émissions de $NO_x$ pour le paramètre de SCAUP *H20* par taux d'évolution (en haut) et par régression linéaire (en bas)**

Il faut, avant tout, souligner que la représentation des résultats sous la forme d'un taux d'évolution permet une lecture plus intuitive des résultats. Néanmoins, cet indicateur présente aussi des limites pour notre analyse, car si les émissions estimées sont faibles, le taux d'évolution en sera bien plus important alors que les différences ne sont que peu importantes. C'est, par exemple, ce que nous constatons sur les régressions linéaires associées en bas des cartes (ci-dessus). Or, l'exploitation des résultats *via* cet indicateur fournit l'opportunité d'avoir un ordre de grandeur simple et intuitif des différences entre les différents jeux de données exploités. Des améliorations pour cette partie pourraient être envisagées, en s'inspirant par exemple des méthodes de validation des automates cellulaires.

Pour commencer, en nous appuyant sur les cartes précédentes, nos estimations sur les axes où des données d'émissions produites par ATMOSF'AIR BOURGOGNE sont disponibles sous-estiment globalement ces dernières. Ce premier résultat tend à nous montrer que les données d'émissions basées sur le modèle SCAUP n'apportent que peu de d'améliorations aux estimations réalisées par l'AASQA. D'une manière plus générale, on constate que les taux d'évolution de nos estimations sont constamment situés entre – 100 % et – 5 % de différence pour le 1er et le 3e quartiles (*cf.* figure 3.31, p. 290). On note d'ailleurs que, pour cette figure, la médiane est toujours inférieure à – 50 % de différence d'estimation. Cependant, en comparant les paramètres, on remarque qu'à mesure que le stationnement augmente, les erreurs d'estimation se rapprochent de la barre fatidique des 0 % en taux d'évolution.

Ces résultats sont à nuancer dans le sens où les données d'émissions à l'échelle des tronçons routiers sont confrontées au sein de chaque maille, mais ne sont pas toujours positionnées sur le même plan du fait de la géométrie des deux réseaux routiers exploités. Ce constat peut en partie expliquer ces tendances, avec des valeurs d'émissions qui peuvent varier du simple au double au sein de certaines mailles estimées. Par ailleurs, en s'appuyant sur les trois cartes du taux d'évolution, on observe plusieurs tendances intéressantes (*cf.* en haut de la figure 3.32, de la figure 3.33 et de la figure 3.34) :

- Une dominance visuelle du rouge sur l'ensemble des cartes[181] qui est à associer en très grande partie à la finesse du réseau routier employé lors de la simulation du trafic routier. C'est bien à l'intérieur de ces zonages (situés le plus généralement à l'intérieur des quartiers de la ville de Dijon) que les populations résidentes sont le plus présentes. C'est pourquoi affiner le calcul des émissions de PAA au sein de ces espaces est intéressant notamment dans le but de mieux caractériser les effets sanitaires et épidémiologiques des émissions atmosphériques d'origine automobile.
- La même tendance est observée sur quelques pénétrantes qui sont caractéristiques des liaisons locales de la ville de Dijon (*cf.* tableau 3.6, p. 283). Nos estimations apportent alors de nouvelles informations sur les émissions de $NO_x$, là où on peut observer des différences qui sont supérieures à plus de 100 % (en rouge foncé) par rapport aux estimations de l'AASQA.

[181] La carte des couleurs associées au taux d'évolution se lit comme suit : plus on est proche du rouge, plus la différence avec les estimations calculées à partir des données de simulation est supérieure aux estimations d'ATMOSF'AIR BOURGOGNE (+ 100 %), tandis que plus on se rapproche du bleu plus nos estimations sont inférieures aux estimations réalisées par ATMOSF'AIR BOURGOGNE.

- Une spécificité au niveau de la rocade Georges Pompidou est à relever. En effet, sur cette section courante, c'est-à-dire dont le tracé est rectiligne et dont les possibilités d'entrée et de sortie sont peu nombreuses, les estimations sont très proches voire meilleures que les estimations réalisées par l'AASQA. À ce niveau, on s'aperçoit qu'au positionnement d'un capteur est associé un taux d'évolution moyen qui se situe entre – 24 % et plus de + 100 %. En se reportant à la carte des $R^2$ à l'échelle des stations de comptage (*cf.* Annexe cartographique 1, p. vii), on s'aperçoit que sur ces six points de comptage les estimations de trafic routier simulé sont plutôt de bonne qualité (entre 0,9 et 0,95 de $R^2$). Dans ce sens, nos estimations apparaissent plus fidèles à ce qui a été observé sur le territoire. On observe d'ailleurs qu'ici plus le temps de stationnement augmente plus nos estimations se rapprochent et deviennent plus pertinentes que les estimations réalisées par ATMOSF'AIR BOURGOGNE[182].

Dès lors, il apparait nécessaire de se concentrer sur un indicateur statistique permettant de mettre en avant ces observations et de mieux distinguer les paramètres de simulation de SCAUP exploités au cours de ce chapitre. C'est pourquoi, à chacune des cartes réalisées, une régression linéaire entre les données estimées par ATMOSF'AIR BOURGOGNE (en abscisses) et les estimations calculées à partir des données de trafic simulées par SCAUP (en ordonnées) a été réalisée (*cf.* en bas de la figure 3.32, de la figure 3.33 et de la figure 3.34). Chaque point positionné dans l'espace géométrique correspond à une maille issue du carroyage mis en place. Cependant, ces éléments de validation sont à prendre avec des précautions, car on observe pour les trois paramètres de simulation exploités que les valeurs les plus faibles présentent une faible dispersion par rapport aux valeurs les plus fortes[183].

En se focalisant, dans un premier temps, sur les coefficients de détermination ($R^2$) calculés entre les deux échantillons de données[184], on constate que pour les trois paramètres de simulation (*H10*, *H15* et *H20*) l'indicateur calculé est de relativement bonne qualité, en se situant pour les trois simulations autour de 0,7 de coefficient de détermination (*cf.* en bas des figures ci-dessus) : respectivement 0,696 pour *H10*, 0,705 pour *H15* et 0,688 pour *H20*. Dans ce sens, on observe une tendance relativement similaire à la répartition des points de simulation projetés au sein du diagramme de Taylor (*cf.* figure 3.19, p. 256), où le paramètre *H15* constitue le paramètre de simulation dont les estimations en $NO_x$ sont les plus proches des estimations de l'AASQA bourguignonne. En revanche, lorsque l'on se focalise ensuite sur l'équation de la droite calculée, une autre information est à prendre en compte. En s'appuyant sur la droite de régression, on constate qu'à mesure que le temps de simulation augmente, la pente de la courbe (coefficient directeur) se rapproche de 1 : de 0,618 pour *H10* à 0,773 pour *H20*. Ce second constat conforte les observations établies au sein de la figure 3.31 (p. 290) sur la répartition du taux d'évolution observé sur chaque point de maille. En d'autres termes, plus la pente de la courbe sera proche de 1, plus le taux d'évolution calculé au sein de chaque maille sera proche des estimations réalisées

[182] Ce résultat met en avant les difficultés rencontrées par les AASQA pour estimer les volumes de trafic routier. Il permet aussi de mettre en avant les apports de notre travail du côté opérationnel, dans lequel s'inscrivent aussi les capacités opérationnelles de ce travail de thèse.

[183] Parallèlement, un traitement par régression linéaire en ajustant les valeurs les plus faibles par une fonction logarithmique, nous a permis de conforter nos résultats. En effet, en dispersant les valeurs par cette fonction, la même tendance linéaire a été observée. C'est pourquoi nous avons fait le choix de plutôt représenter les données brutes d'émissions que des données ajustées, dont l'ordre de grandeur est plus difficilement exploitable.

[184] Un échantillon estimé à partir des données de simulation et un échantillon estimé par ATMOSF'AIR BOURGOGNE.

par ATMOSF'AIR BOURGOGNE (0 %). Ce que l'on constate bien pour les émissions de $NO_x$ calculées à partir du paramètre *H20*. Enfin, la position des différentes données au sein du graphique de relation linéaire permet de montrer que les estimations réalisées à partir des données de simulation sont meilleures à mesure que les émissions de $NO_x$ sont plus importantes. Inversement, pour les valeurs les plus faibles, on constate plutôt que nos estimations sont plus faibles à mesure que la quantité de polluants émis est moins importante. Cette tendance apparait caractéristique du fonctionnement interne du modèle SCAUP, dans le sens où nous avons pu montrer que lorsque le débit routier simulé est important, la reproductibilité ainsi que les coefficients de détermination sont de meilleure qualité (*cf.* Chapitre 7, pp. 239-246).

Si nous devions retenir des éléments montrant l'apport de SCAUP par rapport aux estimations réalisées par ATMOSF'AIR BOURGOGNE, nous retiendrons les validations réalisées à l'échelle de la commune de Dijon, car elles permettent d'avoir une vision générale et synthétique des estimations de PAA réalisées à partir des données de simulation du trafic routier. Cependant, ce niveau apparait trop agrégé et on ne peut observer les apports spatiaux de l'approche proposée dans le cadre de ce travail. C'est pourquoi l'analyse des émissions de $NO_x$ à l'intérieur des mailles nous est apparue plus décisive pour apprécier les apports du modèle SCAUP. À ce niveau, notre bilan est plus nuancé et il est difficile de justifier avec précision si les estimations d'émissions de PAA sont représentatives de ce qui est émis sur le territoire. En définitive, en s'appuyant sur ces différents éléments d'*évaluation*, il apparait difficile de montrer « LE » paramètre de simulation qui répond le mieux à la problématique de la PAA. Cependant, les différents éléments abordés au cours de cette dernière section illustrent les capacités d'interopérabilité des données de simulation de SCAUP en vue de s'inscrire dans les référentiels utilisés par les AASQA.

Pour finir, il faut relever que le calcul des émissions de PAA présente des limites pour l'établissement des bilans d'émissions à l'échelle locale. C'est typiquement le cas du parc automobile, qui présentent ces limites selon que les estimations sont définies à partir du référentiel national, désagrégé ou local (Grassot *et al.*, 2012). En effet, en fonction du parc automobile, les auteurs montrent à travers l'aire urbaine Lyonnaise que « *l'utilisation du parc national plutôt que du parc local conduirait à une surestimation de 6 % des émissions de $CO_2$ et à une sous-estimation de 11 % de celles de $NO_x$ en 2006* » (Grassot *et al.*, 2012, p. 227). De surcroît, la composition du parc automobile à l'échelle locale est difficile à estimer avec précision sans faire appel à des observations in-situ pour relever les plaques minéralogiques des véhicules (André *et al.*, 2013 ; Carteret *et al.*, 2014). L'accès aux données d'immatriculation des véhicules auprès des préfectures est aussi difficile à obtenir. C'est pourquoi il est important de relever que « *la composition du parc automobile par catégorie de véhicules, motorisation, taille et réglementation de pollution est sujette à de grandes incertitudes* » (Carteret *et al.*, 2014, p. 240).

De plus, il faut relever aussi des incertitudes qui ne tiennent pas seulement aux méthodes employées, mais qui incombent plutôt aux informations transmises par les constructeurs automobiles qui communiquent les caractéristiques techniques de leurs véhicules, qui sont ensuite utilisées pour l'application de la méthodologie COPERT. En effet, les équations COPERT se fondent sur les caractéristiques techniques et technologiques qui ont été communiquées par les constructeurs automobiles (Ntziachristos et Zissis, 2009 ; AEE, 2011). Or, des événements

comme ceux de *l'affaire Volkswagen*[185] remettent en cause la fiabilité de ces informations transmises par les constructeurs automobiles. Le CITEPA s'est emparé de ces problèmes à travers un compte rendu pour évaluer les émissions unitaires des véhicules. Les premiers résultats de contrôle de émissions unitaires réalisé sur 52 véhicules ont permis de montrer que pour « *une large majorité des véhicules EURO 6 équipés de la technologie EGR + $NO_x$ trap*[186] *présente des anomalies, notamment en émissions réelles de $NO_x$ jusqu'à cinq fois la norme (80mg/km)* »[187].

En somme il y a un grand nombre d'incertitudes quant à la caractérisation du parc automobile à l'échelle locale tant sa variabilité est grande (Carteret *et al.*, 2014). Toutefois, l'utilisation de ces approches permet d'aborder les grandes tendances des émissions atmosphériques reliées au secteur du trafic routier au sein d'un territoire. Néanmoins, faute de mieux, la méthodologie d'estimation des PAA utilisée par les AASQA constitue aujourd'hui l'approche la plus habituellement utilisée pour la réalisation des cadastres d'émissions. S'inscrire dans ces référentiels nous a néanmoins permis de mettre en avant les atouts et les faiblesses de notre démarche expérimentale pour l'évaluation de la PAA en milieu urbain.

[185] Le Monde du 20 septembre 2015, « *Réduction frauduleuse des normes d'émissions en NOx et CO2* » : http://www.lemonde.fr/automobile/article/2015/09/20/accuse-de-tricherie-volkswagen-est-menace-de-sanctions-financieres-aux-etats-unis_4764397_1654940.html.

[186] La technologie *$NO_x$ trap ou piège à $NO_x$* correspond à un type de pot catalytique équipant les véhicules Diesel ; l'EGR (*Exhasut Gas Recirculation*) quant à lui correspond à un système de récupération des gaz d'échappement au niveau du moteur et non pas de l'échappement.

[187] Actualité du CITEPA du 7 Avril 2016 : http://www.citepa.org/fr/actualites/2314-7-avril-2016-controle-des-emissions-specifiques-des-voitures-premiere-reunion-de-la-commission-technique-independante

## Conclusion :

Pour conclure ce chapitre, nous souhaitons mettre en avant les atouts de l'approche SCAUP dans une optique opérationnelle. Effectivement, si plusieurs éléments sont discutables, la démarche de modélisation SCAUP que nous proposons s'inscrit, avant tout, dans une perspective expérimentale en vue d'apporter une réponse à une demande spécifique des AASQA lorsqu'aucune donnée de trafic routier issue des modèles classiques n'est accessible. Notre approche tend à proposer une démarche qui se situe à mi-chemin entre les comptages routiers et les modèles de trafic (*cf.* figure 3.21, p. 269). Elle est relativement simple à mettre en œuvre sur d'autres espaces d'étude, du moment que nous disposons de données relatives aux comptages routiers.

Notre travail s'insère aussi dans cette perspective de généralisation à d'autres espaces urbains que celui de la ville de Dijon. Les éléments que nous avons proposés tendent à nous montrer que les données simulées permettent de construire une *connaissance* sur les émissions de PAA, qui apparait relativement adéquate aux référentiels utilisés par les AASQA. Par ailleurs, inscrire notre travail en appui des acteurs du territoire constitue aussi un positionnement scientifique primordial afin de répondre aux enjeux socio-environnementaux que pose la pollution atmosphérique en milieu urbain. La recherche en soutien des acteurs du territoire est un élément important qui vise à faciliter et appuyer la mise en place d'actions à l'échelle locale et nationale. À travers cette application expérimentale du modèle SCAUP, s'inscrivant dans le cadre des inventaires territoriaux, notre travail permet d'apporter une réponse opérationnelle aux besoins des AASQA comme ATMOSF'AIR BOURGOGNE. Effectivement, lorsqu'aucune donnée n'est disponible, SCAUP permet de trouver un juste compromis entre les approches par comptage et les approches par modélisation, comme nous l'avons positionné au sein de la figure 3.21 (p. 269). De surcroît, l'approche SCAUP nécessite un temps de traitement moins important, permettant alors d'extraire toutes les granularités nécessaires aux calculs des émissions de PAA (TMJA, TMJH…).

C'est aussi à ce niveau que le modèle apporte une réponse opérationnelle à destination des AASQA, en facilitant alors le temps de collecte et de traitement des données de comptages routiers, et qu'il est à privilégier lorsque aucun modèle de trafic n'est disponible (PCIT, 2012 ; OMINEA, 2015). Cependant, on retrouve des éléments qui tendent à nous montrer que la répartition spatiale des véhicules sur les tronçons routiers reste un élément discutable, qui doit encore être amélioré et affiné au sein du modèle SCAUP, pour mieux apprécier les effets spatiaux des émissions de PAA. Mis à part pour ces aspects spatiaux, les objectifs visés par ce travail apparaissent majoritairement atteints, à notre sens :

- En **proposant une alternative au modèle de trafic routier classique** en vue d'estimer plus finement les débits de trafic observés sur un plus grand ensemble d'axes routiers. L'ensemble des données s'insère aussi dans un système de validation des estimations qui a été testé à travers le diagramme de Taylor
- En **générant des cartes de charge du trafic routier à différents moments** de la journée, ainsi qu'à une granularité relativement fine (le quart d'heure), sur un réseau routier urbain, tout en insérant un ensemble de données peu important pour alimenter le modèle SCAUP, dont le caractère génératif et dynamique constitue une force pour la modélisation du trafic routier à partir des comptages routiers

- **En se positionnant aussi dans une perspective opérationnelle** en support aux AASQA en charge de la réalisation des cadastres d'émissions de la pollution atmosphérique à l'échelle des espaces urbains.

Il faut relever un biais majeur de notre analyse ; en effet, on peut se demander si les estimations réalisées par les AASQA sont sous/surestimées par rapport aux émissions totales du territoire. Il apparait important de discuter de ces éléments : dans quelles conditions les estimations réalisées par les AASQA sont-elles robustes ? Or, il apparait difficile au niveau d'une étude sur les émissions atmosphériques de proposer un système de validation robuste des estimations réalisées à partir des données de trafic routier simulé par SCAUP et à partir des données estimées par ATMOSF'AIR BOURGOGNE. C'est pourquoi, dans une optique de validation de ce travail, il faudra ensuite s'insérer dans une optique d'immission plutôt que d'émission des polluants d'origine automobile. À ce niveau, nous pouvons disposer des quelques mesures des concentrations de polluants à proximité du trafic routier grâce aux *capteurs fixes* de surveillance de la qualité de l'air. Toutefois, le passage de l'émission à l'immission s'inscrit dans un contexte plus large, où, comme nous l'avons vu plus haut, à la pollution de proximité, se juxtapose et s'ajoute la pollution de fond urbain (*cf.* figure 1.2, p. 32). En effet, l'air que nous respirons est le résultat d'une *soupe atmosphérique*, dont l'origine des polluants en présence relève autant des secteurs résidentiel et industriel que routier. S'inscrire dans cette perspective nécessiterait alors la collecte d'un plus grand volume d'informations, plus difficilement accessibles et exploitables, ainsi que la mise en place d'une seconde chaîne de travail.

Il apparait alors intéressant de s'inscrire dans une perspective de couplage, plus que d'interopérabilité, du modèle SCAUP, avec l'intégration de la méthodologie COPERT au sein du processus de simulation. Elle constitue d'ailleurs une perspective de travail intéressante pour raffiner les estimations des émissions de PAA à l'échelle locale, ainsi que pour prolonger la démarche initiée ici. Dans ce sens, il est aussi important de relever les différentes perspectives de développement dans lesquelles ce travail s'inscrit, et dont nous aborderons les différents aspects et enjeux au cours du dernier chapitre de ce travail.

# Chapitre 9. LES PERSPECTIVES DE DÉVELOPPEMENT AUTOUR DE LA MATRICE TECHNIQUE : DE LA QUANTIFICATION À L'ÉVALUATION

Au cours du précédent chapitre, la démarche SCAUP a été envisagée dans une optique d'évaluation des émissions de PAA. Cependant cette dernière étape ne constitue pas une fin en soi, et de nombreuses portes restent ouvertes. Il s'agit alors de mentionner et de proposer non seulement un ensemble de développements et d'améliorations de la démarche de modélisation du trafic routier, mais aussi d'en élargir ses champs d'application pour proposer une démarche intégrée de l'étude de la PAA à l'échelle urbaine.

En effet, si l'approche de modélisation SCAUP permet d'observer un trafic routier émergent sous sa forme la plus élémentaire, le *véhicule*, celui-ci est encore loin d'être parfaitement représenté sur l'ensemble du réseau routier ; de même que les connaissances sur les niveaux de PAA qui en ressortent ne permettent pas d'apprécier avec finesse les caractères spatio-temporels du modèle SCAUP. Il est ainsi important d'inscrire ce travail dans une perspective plus large, en soulevant de nombreux éléments de développement de la démarche SCAUP. Ces différentes pistes de travail s'insèrent dans un contexte plus large, celui des impacts environnementaux du trafic routier en milieu urbain, en associant aux données d'émissions une troisième dimension, qui est l'atmosphère.

Ce chapitre vise alors à mettre en exergue les différents champs d'application et les développements de l'approche de modélisation SCAUP proposée dans ce travail doctoral. Pour cela, nous reviendrons sur les trois dispositifs de la matrice technique (*quantification, modélisation et évaluation*), pris non seulement à leur niveau niveau individuel, mais également en interaction. C'est pourquoi, pour ce dernier chapitre, nous structurerons nos propos autour de la matrice technique, en reprenant la figure 2.8 exposée précédemment (p. 146), mais aussi en nous focalisant autant sur les dispositifs que sur les relations qu'ils entretiennent entre eux, comme illustré en figure 3.35 (ci-dessous) :

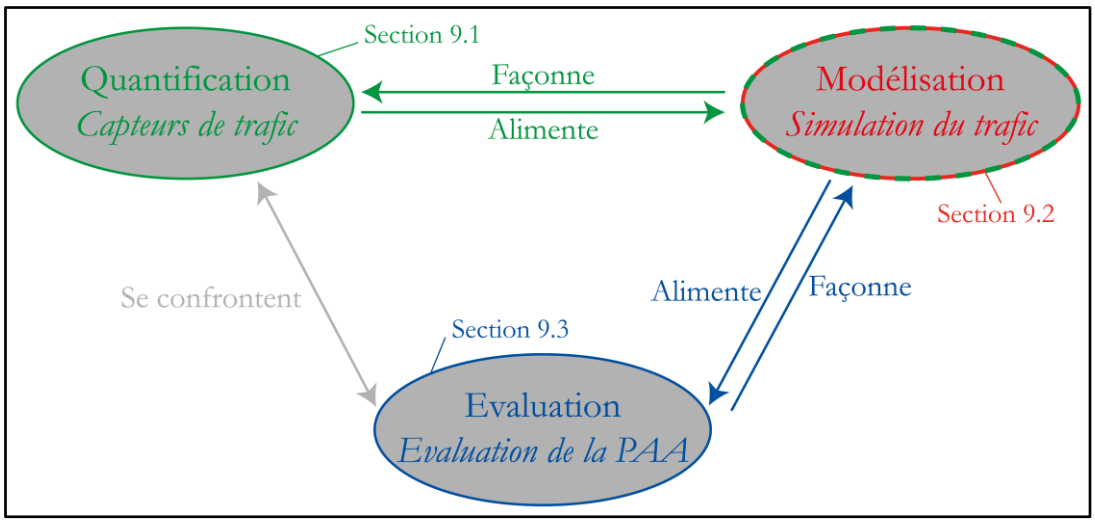



**Figure 3.35 : Structure du chapitre 9 : Une réflexion autour de la matrice technique et des relations entre les dispositifs de quantification, de modélisation et d'évaluation**

En effet, les principaux enseignements de notre *évaluation* tendent à soulever de nombreuses réflexions, portant autant sur le dispositif de quantification, que sur celui de modélisation et sur celui d'évaluation. Par ailleurs, en nous basant sur la matrice technique et sur les relations qu'entretiennent chacun des dispositifs, nous souhaitons terminer ce travail en revenant sur ces trois éléments qui en constituent la colonne vertébrale.

Ce dernier chapitre conclut la thèse en apportant, comme illustré ci-dessus, à chaque section un zoom sur un de ces dispositifs et sur ses relations avec les autres. Dans un premier temps, nous nous focaliserons sur le dispositif de quantification et ses relations avec la modélisation, en nous interrogeant aussi sur la manière avec laquelle le modèle SCAUP pourrait façonner le dispositif de quantification mis en place. Dans un deuxième temps, nous nous centrerons plus exclusivement sur le dispositif de modélisation, en proposant différentes améliorations du processus de simulation du modèle SCAUP. Enfin, le dispositif d'évaluation sera mobilisé en vue de réfléchir à une approche plus intégrée de l'évaluation des PAA en milieu urbain. En outre, à chacune des sections de ce chapitre, nous avons souhaité tester certaines de nos hypothèses en appui du modèle SCAUP, qui nous permettront de vérifier ses capacités heuristiques, d'adaptation et de généralisation à d'autres contextes urbains, comme nous l'illustrerons au cours ce dernier chapitre.

## 9.1. Les réflexions autour du dispositif de quantification pour alimenter et façonner le dispositif de modélisation

Comme nous l'avons vu précédemment, dans sa vocation première le dispositif de quantification alimente le dispositif de modélisation. Or, on remarque aussi que ce dernier façonne la manière avec laquelle le dispositif de quantification est construit (*cf.* figure 3.35, en vert). Pour cette première section, nous nous interrogerons sur les modalités de cette relation duale. Ainsi, dans quel cadre le dispositif de modélisation peut-il façonner le dispositif de quantification ? Comment améliorer la densité des comptages routiers pour la simulation du trafic routier ? Quels sont les potentiels d'utilisation d'autres capteurs urbains pour la simulation du trafic routier ?

### 9.1.1. Les questionnements liés à la répartition des capteurs : un cadre d'application heuristique du modèle SCAUP ?

Lors de la mise en place de l'approche de modélisation SCAUP, une des questions à laquelle nous avons été confrontés porte sur la couverture spatiale et la densité du réseau de comptage : la couverture spatiale du réseau routier est-elle significative ? Quelles sont la couverture et la répartition optimales du réseau de comptage routier pour qu'il soit représentatif de la circulation routière relevée en milieu urbain ?

Il est vite apparu difficile de répondre à ces questions, tant le positionnement des capteurs électromagnétiques dépend, non seulement de leur finalité, mais aussi, des pratiques des gestionnaires de trafic. Par exemple, dans le cas des stations SIREDO gérées par les départements, l'implantation des stations de mesure dépend en très grande partie du type de voie géré par le conseil départemental. Tandis que dans le cas d'une implantation pour la régulation du trafic routier en milieu urbain, le positionnement des capteurs est privilégié à proximité des carrefours équipés de feux tricolores. Pourtant, pour simuler le trafic routier avec SCAUP, il est naturel de dire que plus la densité du réseau de capteurs est importante, plus le trafic routier simulé sera de bonne qualité. Cependant, il est impensable d'équiper l'ensemble des tronçons routiers de capteurs électromagnétiques, autant pour des raisons financières que de gestion du réseau de capteurs.

Dans cette perspective, il apparait intéressant de présenter une application expérimentale de SCAUP liée à la couverture des réseaux de capteurs électromagnétiques. Cette expérience s'inscrit dans la deuxième relation du dispositif de quantification, dans le sens où la modélisation *façonne* elle-aussi la manière avec laquelle le dispositif de quantification alimente le dispositif de modélisation (Commenges, 2013). C'est pourquoi, il peut être intéressant d'employer le modèle SCAUP à des fins heuristiques en vue d'investiguer ces questionnements autour de l'implantation des comptages routiers. Pour cela, nous avons utilisé le réseau routier de simulation KISS (*cf.* section 6.3.1, p. 205), sur lequel nous avons testé différents taux de couverture du réseau de capteurs[188] (*cf.* section 6.3, p. 205). En somme, onze couvertures spatiales du réseau de capteurs fixes ont été testées : un test avec un seul axe équipé (1 %) et dix tests avec une densité de

[188] À noter que l'implantation des stations de comptage a été réalisée de manière aléatoire, en privilégiant toutefois une couverture optimale des axes possédant une hiérarchie plus importante. De plus, pour chacun des tests le même paramètre a été défini, à savoir un choix des axes routiers tenant compte de la hiérarchie du réseau routier, pour un temps de stationnement de 15 minutes.

capteurs couvrant de 10 % à 100 % du réseau routier employé. Notons ici que les tests ont été réalisés en adéquation avec les plans d'expérience présentés précédemment, en éprouvant le modèle de simulation par la mise en œuvre de 30 réplications de chacune des configurations testées. De plus, rappelons que, pour la ville de Dijon, le réseau de capteurs électromagnétiques couvre seulement 12,2 % des axes de l'espace (*cf.* section 3.2.2, p. 94).

La figure 3.36 (ci-dessous) illustre les débits moyens simulés pour chaque quart d'heure en fonction de la couverture spatiale du réseau de capteurs testée (par les courbes en couleur). D'une manière générale, on observe un résultat peu surprenant, dans la mesure où plus le réseau de mesure est dense, plus le trafic routier simulé devient important. Cette tendance s'observe aisément par une augmentation progressive du débit simulé par quart d'heure (gradient du bleu vers le rouge). Cependant, on remarque que le débit simulé augmente sensiblement jusqu'à une couverture spatiale de 40 % (en moyenne de 10 à 200 véhicules/quart d'heure), et qu'au-delà de ce seuil les courbes ont tendances à se resserrer à mesure que la densité des capteurs est importante. C'est-à-dire que le trafic routier simulé n'augmente plus aussi sensiblement que pour les autres configurations, en se concentrant autour d'un débit moyen simulé situé entre 300 et 400 véhicules/quart d'heure lors des périodes de forte fréquentation (7 h à 19 h). En s'appuyant sur ces premiers éléments, il apparait donc qu'une couverture de mesure de 40 % des axes d'un espace d'étude est suffisante, et qu'au-delà le trafic routier simulé n'en sera que faiblement amélioré.

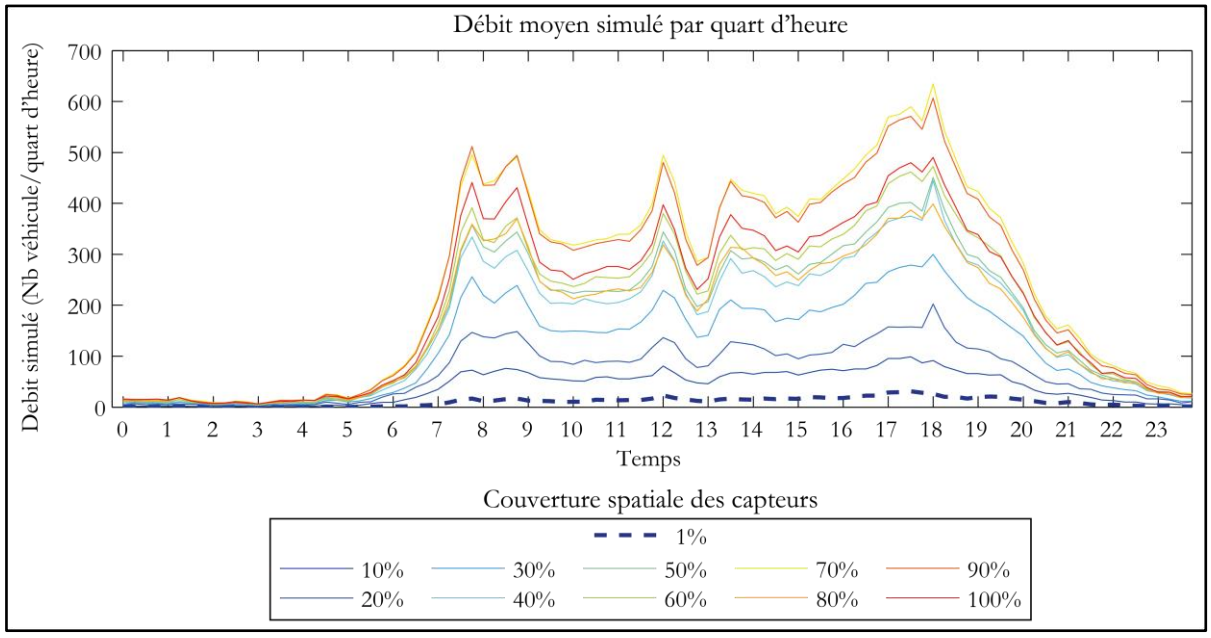



**Figure 3.36 : Débit moyen simulé par quart d'heure en fonction de la couverture des comptages routiers sur un réseau routier simple (KISS : 48 axes routiers)**

Ce premier constat est conforté lorsque l'on se concentre sur la reproductibilité des débits de véhicules simulés, comme illustrée en figure 3.37 (ci-dessous). Effectivement, au même titre que la figure précédente, on observe un resserrement des taux de reproductibilité à mesure que la densité des capteurs augmente. En définitive, une couverture de 40 % des axes routiers apparait suffisante pour être représentative du trafic routier de l'ensemble d'un espace d'étude quel qu'il soit. Il faut noter que dans cette figure seules les réplications sont analysées entre elles ; il n'est pas possible d'observer les relations entre chacune des configurations testées. En d'autres termes,

c'est la sensibilité du modèle SCAUP à simuler les mêmes conditions de trafic routier qui est représentée (emphase des simulations). Par conséquent cette analyse est complémentaire à la figure 3.36 (ci-dessus) et tend plus particulièrement à conforter nos constats et préconisations quant aux tests de couverture du réseau de capteurs sur le réseau routier (ci-dessous).

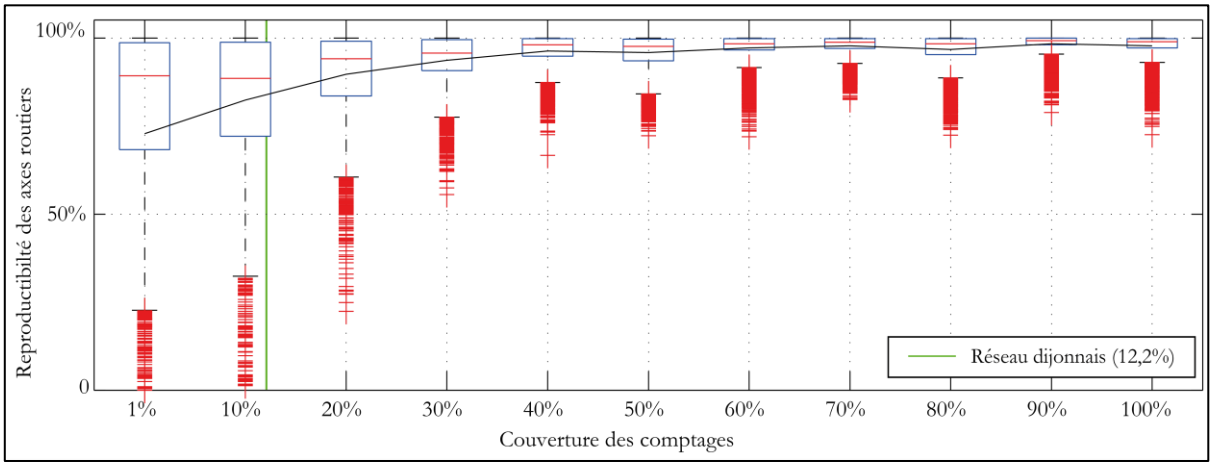



**Figure 3.37 : Répartition de la reproductibilité par axe routier en fonction de la densité du réseau de capteurs**

Cependant, ces constats sont à nuancer. En effet, les observations établies en figure 1.28 (p. 96) nous ont permis d'observer que l'implantation du réseau de capteurs s'opère en combinaison avec les volumes de trafic routier passant sur les infrastructures, c'est-à-dire en tenant compte de la dimension circulatoire des infrastructures de transports (Genre-Grandpierre, 2000, 2010). Comme nous l'avons vu, une faible couverture du réseau de mesure peut, malgré tout, être représentative de la majorité des flux de trafic routier d'un espace d'étude. Par exemple, dans notre cas, nous avons pu voir que près de 10 % des axes routiers de la ville de Dijon concentrent la plus grande part des flux de circulation routière (*cf.* figure 1.28, p. 96). C'est pourquoi, notre approche expérimentale nécessite d'être interprétée en toute relativité, et doit plutôt être appréhendée en tenant compte des volumes de trafic routier observés, ce qui n'a pas pu être réalisé à partir de l'environnement de simulation KISS testé ici.

Pour terminer, dans un contexte urbain *réel*, la mise en place d'une couverture de 40 % du réseau routier de la ville de Dijon nécessiterait l'implantation de près de 3 000 capteurs électromagnétiques. En outre, en excluant les réseaux résidentiels, il faudrait tout de même implanter près de 1 000 capteurs électromagnétiques sur l'espace d'étude. Dès lors il apparait peu envisageable de mettre cette solution en œuvre, que ce soit en termes de coûts structurels, de gestion ou d'exploitation du réseau de capteurs. Par contre, il est tout à fait envisageable d'optimiser et de densifier le réseau de mesure par le biais d'une combinaison d'autres dispositifs de quantification, comme les enquêtes O-D ainsi que d'autres capteurs (mobiles et/ou fixes) de plus en plus nombreux dans l'espace urbain.

### 9.1.2. **S'inscrire dans une approche de simulation du trafic routier mixte : la combinaison des comptages et des enquêtes O-D**

Les enquêtes O-D, évoquée dans la 1re partie, sont construites avant tout comme des dispositifs de quantification complémentaires aux comptages. C'est pourquoi, en vue de densifier le réseau de comptage et d'améliorer les aspects spatiaux du modèle SCAUP, l'une des premières pistes d'amélioration peut être investiguée du côté de ces données d'enquêtes.

Pour commencer, l'idée centrale de notre travail a été de postuler que les dynamiques intraurbaines ne sont que très peu approchées dans leurs caractéristiques spatio-temporelles. En effet, la richesse sémantique des données d'enquête constitue une force pour comprendre les comportements de mobilité, mais celles-ci restent imprécises lorsqu'il s'agit de saisir les mouvements dans l'espace et le temps (*cf.* section 2.1.2, p. 58). Ces limites sont en très grande partie à associer aux données d'entrée et au formalisme des modèles de trafic *classiques*, qui sont fondés sur l'utilisation systématique de données d'enquêtes O-D[189] (Bonnel, 2002 ; Commenges, 2013). De plus, à partir d'une approche *classique*, et spécifiquement lors de la mise en œuvre d'une approche désagrégée, le modélisateur se retrouve confronté à l'absence de données fines pour quantifier et évaluer les estimations réalisées. Dans cette perspective, l'utilisation de données issues des capteurs urbains fournit une alternative aux approches traditionnelles pour affiner la résolution temporelle des données d'enquêtes. Au cours des différents développements présentés *a posteriori*, nous avons pu relever deux éléments caractérisant les données de capteurs par rapport aux données d'enquêtes. En effet, si la dimension temporelle associée aux capteurs leur fournit un atout considérable par rapport aux données d'enquêtes, c'est plutôt sur leur dimension spatiale que l'usage des capteurs apparait encore discutable. Dans ce sens, on peut se demander si l'approche proposée ne peut pas se construire en complémentarité de données d'enquêtes O-D, en vue de mieux appréhender la modélisation du trafic routier en se reposant sur les atouts temporels des capteurs fixes et sur les atouts spatiaux des enquêtes O-D.

En effet, l'une des critiques récurrentes sur les modèles de trafic porte sur leur construction « hermétique », c'est-à-dire qu'ils ne comprennent que les déplacements mesurés à l'intérieur d'un espace excluant tout trafic extérieur (Masson, 2000). Dans une optique d'évaluation environnementale, à ce trafic « intérieur » il manque alors l'ensemble des flux de circulation routière qui se situent hors des zonages d'enquêtes, ce qui nécessite l'ajustement des trafics routiers estimés aux comptages routiers (ajustement, calibrage). Il apparait donc que les comptages routiers permettent de bien mieux appréhender ces flux qui se situent à l'extérieur, ainsi qu'à l'intérieur des zonages d'enquêtes. Effectivement, comme nous l'avons relevé ces capteurs se situent au cœur de l'évènement qu'ils mesurent, et ils peuvent dans ce cadre apparaitre bien plus exhaustifs que les données d'enquêtes. C'est pourquoi, pour affiner les données de trafic routier simulé avec SCAUP, il pourrait être intéressant de combiner ces deux dispositifs de quantification, en vue de proposer une démarche mixte s'inscrivant toujours dans une approche guidée par les données :

[189] Ici on soulève bien l'aspect relationnel montrant que le dispositif de modélisation est *façonné* par le dispositif de quantification (Commenges, 2013).

- en **s'appuyant sur les enquêtes O-D** afin de disposer d'informations permettant de répartir les flux de trafic routier sur l'ensemble d'un espace d'étude
- en **s'appuyant sur les comptages routiers** afin de disposer d'informations permettant de « rythmer » la circulation routière.

Différents constats autour du dispositif de quantification nous ont amené à proposer cette complémentarité. En effet, si le trafic routier simulé à partir de SCAUP fournit une première réponse aux limites intrinsèques des modèles à quatre étapes, il comporte aussi des limites quant à la manière avec laquelle les débits de véhicules sont répartis sur les axes routiers de l'agglomération. On peut alors se demander si le trafic routier simulé à partir des capteurs ne serait pas représentatif d'un trafic routier dit de « fond », c'est-à-dire représentatif des mouvements constants et sempiternels que l'on retrouve en milieu urbain. C'est un constat qui semble en adéquation avec les observations effectuées à partir du traitement de *signal sur bruit*, réalisé grâce à l'analyse en ondelettes (*cf.* figure 1.40, p. 111). Dans ce sens, les données de simulation s'apparenteraient à un *mouvement brownien* auquel il manquerait les déplacements intérieurs mesurés grâce aux enquêtes O-D. En d'autres termes, l'approche par enquête O-D permettrait d'avoir une répartition spatiale des flux de trafic routier moins discutable, alors que l'approche de simulation permettrait d'avoir une vision de la dynamique du trafic routier. C'est aussi dans cette optique, que les dispositifs de quantification du trafic routier que nous avons distingués plus haut s'avèrent être plutôt complémentaires qu'antagonistes (*cf.* section 2.1.2, p. 58).

Ainsi, exploiter les deux dispositifs de quantification (enquêtes et comptages) pour la même application de modélisation, puis de simulation, du trafic routier tendrait vers une situation idéale, dont la mise en œuvre se retrouverait confrontée à l'interopérabilité des deux dispositifs de quantification, au passage d'une vision statique à une vision dynamique, au couplage d'un modèle macroscopique à un modèle microscopique… L'ensemble de ces éléments comporte bien des limites et des verrous à lever, qui expliquent toutes les difficultés de mettre en œuvre cette approche mixte. Cependant, un premier travail intéressant serait de réaliser une confrontation entre ces deux types d'approches (enquête et comptages), pour juger de la qualité des calages, des calibrages et des différences de trafic routier qu'il en ressort.

Enfin, il faut aussi se préserver des éléments mentionnés plus haut (*cf.* section 4.2.1, p. 133). Effectivement, la mise en œuvre de ces deux types de mesures ne correspond pas aux mêmes applications : alors que les enquêtes O-D fournissent l'opportunité de prévoir et quantifier les flux de trafic sur le moyen et le long terme à travers la modélisation à quatre étapes (Bonnel, 2002 ; Commenges, 2013), le modèle SCAUP initié au cours de cette thèse ne permet que de rendre compte des flux de trafic routier observé à partir des comptages routiers (*cf.* section 9.2.1, p. 311). C'est pourquoi, en parallèle, il est d'autant plus important de prolonger les développements de l'approche de modélisation SCAUP, en s'éloignant voire en se dédouanant, des approches *classiques*, et en s'interrogeant plutôt sur une amélioration des dispositifs de quantification exploités tout au long de ce travail : les capteurs.

### 9.1.3. **Une réflexion sur l'usage des capteurs urbains pour renforcer la couverture des comptages routiers : les *smart cities* et les capteurs mobiles**

L'une des perspectives de travail pour optimiser la couverture des comptages routiers est de s'intéresser à d'autres systèmes de capteurs, qui peuvent être complémentaires aux capteurs électromagnétiques. Par ailleurs, il est intéressant de repositionner notre travail dans le contexte des *smart cities* (*cf.* section 2.1.3, p. 61), auquel le dispositif de quantification se rattache. En effet, dans le contexte actuel d'ouverture et d'accès aux données issues des différents capteurs, l'approche que nous proposons s'inscrit aussi dans une optique de généralisation de l'approche proposée à d'autres espaces urbains, voire à d'autres systèmes de capteurs. La collecte des données demeure, quant à elle, réalisée en un ou plusieurs points fixes du territoire (excluant par nature toutes les mesures mobiles, comme les mesures brutes issues des capteurs GPS ou GSM par exemple).

#### a. ***Les smart cities : un accès facilité aux données de comptages routiers***

Le contexte des *smart cities* (*cf.* figure 1.13, p. 61) s'inscrit en partie dans les problématiques d'instrumentation de la ville, qui visent à mutualiser un ensemble de données automatiques relevées par une multitude de capteurs intelligents (Batty, 2012 ; Chatzimilioudis *et al.*, 2012). Dans ce sens, l'accès à un grand nombre de sources de données ouvre non seulement la voie à une plus grande couverture des espaces urbains, mais aussi à un accès des données qui en sera d'autant plus facilité. En outre, cette tendance d'*instrumentation* de la ville fournit l'opportunité de densifier les mesures du trafic routier en exploitant un plus grand nombre d'informations mises à disposition, ainsi que l'opportunité d'ouvrir les champs d'applications de SCAUP à d'autres espaces urbains que celui de la ville de Dijon.

Dans ce sens, les *smart cities* offrent pour notre travail deux perspectives intéressantes, pour tester la capacité générique de notre approche de simulation, et pour couvrir des espaces d'étude plus vastes avec un plus grand ensemble de capteurs. En effet, comme nous l'avons soulevé, le domaine des transports constitue l'un des premiers domaines d'application des *smart cities* (Batty 2013). De surcroit, même si le contexte des *smart cities* n'en est qu'à son balbutiement, les données de comptage issues des capteurs électromagnétiques s'inscrivent dans cette tendance et constituent un bon exemple des potentiels futurs que vont proposer les *smart cities*. Tout d'abord les comptages routiers se situent à l'interface des données produites par l'usager (*prosumeur*) qui utilise la route (le service), et dont la collecte et la centralisation sont assurées par les collectivités territoriales. Ensuite la diffusion des informations liées à la circulation routière est actuellement de plus en plus facilitée, autant en France que dans d'autres pays, comme le montrent de nombreuses applications en ligne et en temps réel, telles que les applications *TomTom*, *Google Maps* ou *Waze*.

Même si le concept des *smart cities* n'est pas toujours mis en avant en tant que tel, pour le trafic routier en particulier, de nombreuses villes s'inscrivent aujourd'hui dans cette tendance d'informations en temps réel et en libre accès sur les conditions de circulation routière, issues des différents systèmes exploités par les collectivités territoriales. C'est par exemple le cas de villes

françaises comme Strasbourg[190] ou Bordeaux[191], qui communiquent en temps réel leurs données de circulation routière (CETE, 2013). Ces deux villes font office de *bons élèves* concernant l'accès aux données publiques, suite à l'application de la directive INSPIRE[192]. On notera d'ailleurs que les données strasbourgeoises sont d'autant plus intéressantes, par une mise à jour régulière des données de comptage toutes les 3 minutes. En outre, l'ensemble des informations liées aux comptages routiers sont accessibles sous différents formats adaptés à une approche géographique et informatique des comptages routiers (*.json, .xml, .gml*). Une perspective de travail est d'exploiter ces informations pour tester le modèle SCAUP sur ces deux espaces urbains. À l'échelle internationale, des villes nord-américaines comme San Francisco[193] ou Chicago[194] communiquent aussi leurs données de trafic à des résolutions spatio-temporelles intéressantes, voire en temps quasi-réel. Notons d'ailleurs que ces deux villes font partie des premières à initier et mettre en œuvre le concept de *smart cities* sur un grand nombre de thématiques, portant autant sur l'énergie et les transports que sur l'économie (la liste détaillée des données disponibles, sur le portail de la ville de Chicago en est une bonne illustration : https://data.cityofchicago.org/). De surcroit, d'autres espaces urbains, comme la ville de Wuxi en Chine[195], constituent des terrains d'étude potentiellement plus intéressants pour l'étude de la PAA. En effet ceux-ci se situent, aujourd'hui, dans des contextes de forte croissance de la circulation routière, et donc de croissance des impacts locaux du trafic routier, au premier rang desquels la pollution atmosphérique constitue une problématique sanitaire émergente (Guinot, 2008).

Enfin, il faut relever que nous avons utilisé la géométrie du réseau routier de la BD TOPO, car cette base de données présente peu de problèmes en termes de connexité et de topologie. Néanmoins, le modèle SCAUP est tout à fait adéquat à la simulation du trafic routier en exploitant le réseau routier issu d'OpenStreetMap (OSM), et ce d'autant plus que la matrice de transition tient compte du référentiel des voies issu d'OSM. Dans cette perspective, il est alors possible de formaliser le réseau routier d'un grand nombre d'espaces urbains, que ce soit au niveau des villes françaises, européennes ou mondiales. Nous reviendrons d'ailleurs au cours de la dernière section sur une application concrète à travers la ville de Marrakech (*cf.* section 9.3.2, p. 322).

#### b. *Les comptages routiers à partir des capteurs mobiles pour densifier la couverture des comptages routiers*

La densification du réseau de comptage routier peut aussi être envisagée par l'utilisation de comptages issus de systèmes de capteurs mobiles. En effet, la démocratisation des technologies de capteurs mobiles individuels, comme les *smartphones* ou les GPS, fournit l'opportunité de disposer d'une densité de comptages routiers bien plus importante que ce qui est permis avec les

[190] Données de trafic routier pour la CUS (Communauté Urbaine de Strasbourg) : http://www.strasbourg.eu/ma-situation/professionnel/open-data/donnees/mobilite-transport-open-data/trafic-routier-cus.
[191] Données de trafic routier pour la CUB (Communauté Urbaine de Bordeaux) : http://data.bordeaux-metropole.fr/dev/exemples/format_csv.php.
[192] Le rapport du CETE (2013) portant sur le trafic routier dresse un état des lieux des données ouvertes au niveau des collectivités territoriales françaises.
[193] Site de la régie des transports de San Francisco : http://www.sfcta.org/modeling-and-travel-forecasting/model-and-data-service-bureau.
[194] Portail des données de comptages routiers de la ville de Chicago : https://data.cityofchicago.org/Transportation/Average-Daily-Traffic-Counts/pfsx-4n4m.
[195] IEEE Smart Cities (Institute of Electrical and Electronics Engineers) « *Smart Environment Protection in Wuxi* » : http://smartcities.ieee.org/images/files/pdf/smart_enviromentprotection.pdf.

capteurs électromagnétiques. De plus, dans le contexte des *smart cities* et du *big data*, ils constituent un dispositif de comptage complémentaire d'autant plus intéressant à exploiter pour notre travail. Cependant, l'utilisation des données issues des capteurs mobiles soulève de nombreux écueils, liés autant à la protection des données individuelles qu'au système de collecte qui est par nature mobile : les comptages-trajectoires, pour lesquels nous avons estimé les contraintes plus haut (*cf.* figure 1.16, p. 68).

C'est pourquoi des systèmes de comptage « mixtes » ont vu le jour. Ces systèmes originaux se fondent sur l'utilisation des capteurs mobiles individuels, pour des comptages qui sont réalisés de manière collective en un point fixe. Dès lors, ces systèmes constituent une piste de travail intéressante en vue de bénéficier d'une meilleure couverture de l'espace urbain. Par ailleurs, ceux-ci fournissant des mesures *ponctuelles*, ils se rattachent aisément au système de comptage utilisé dans le cadre de ce travail. C'est typiquement le cas de la méthode des VTL (*Virtual Trip Line*), initiée dans le cadre du projet de recherche « *Mobile Millenium* [196]» (Herrera *et al.*, 2009, 2010). Cette méthode se fonde sur l'utilisation de téléphones mobiles, et la collecte des données individuelles est réalisée *via* des stations de comptage virtuelles et fixes (les VTL), placées au niveau des infrastructures de transport, comme illustré en figure 3.38 (Hoh *et al.*, 2008, 2012) :

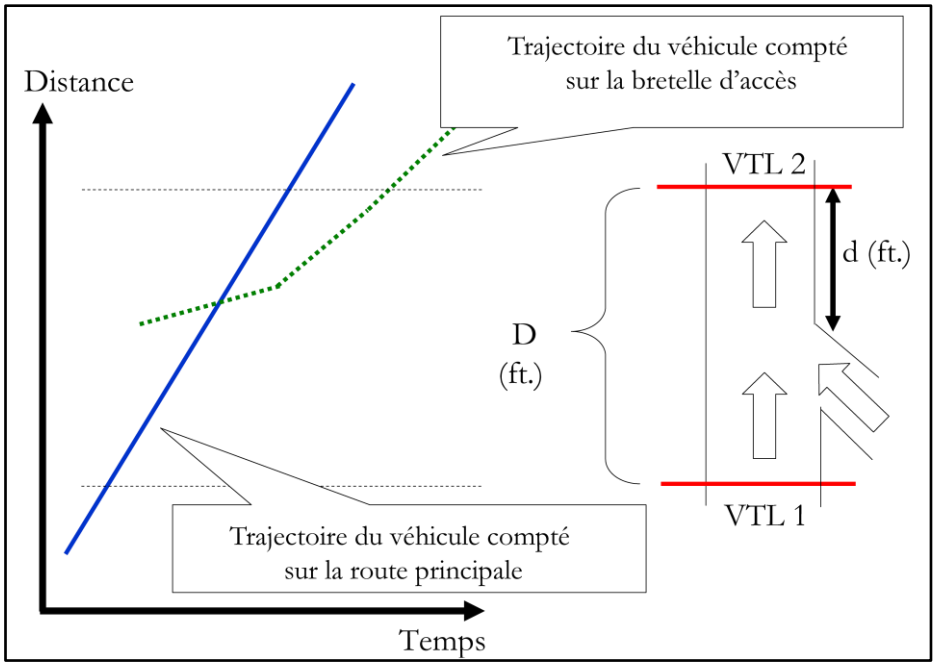



**Figure 3.38 : Principe de fonctionnement de mesure du trafic routier par VTL (*Virtual Trip Line*) sur une section autoroutière (source : Hoh *et al.*, 2008, p. 6)**

L'utilisation des VTL permet à la fois de recueillir une information fine sur le trafic routier, tout en ayant un objectif fort de protection de la vie privée (Hoh *et al.*, 2012). C'est pourquoi les données ne sont pas mesurées à l'échelle individuelle, mais plutôt à une échelle collective (l'infrastructure), par l'intermédiaire de stations virtuelles (VTL). En outre, on peut citer d'autres recherches qui s'inscrivent dans ce champ de la collecte d'informations sur le trafic routier à partir des systèmes de capteurs mobiles sans relever de traces personnelles [197] (That *et al.*, 2015).

[196] Site du projet Mobile Millenium : http://traffic.berkeley.edu/.
[197] SAGEO 2014, « *Atelier Système d'information en transport et mobilité à l'ère de la transition énergétique* : *Une approche de collecte participative d'informations sur le trafic sans traces personnelles* » Karine Zeitouni, Dai-Hai Ton That, Iulan Sandu Popa.

Le principe de mesure par VTL est illustré ci-dessus : lorsqu'un appareil mobile croise une VTL, l'appareil envoie alors un signal permettant de comptabiliser les véhicules passants, ainsi que leur direction. La vitesse, quant à elle, se mesure entre deux comptages virtuels positionnés en amont et en aval d'une infrastructure de transport (VTL1 & VTL2 en figure 3.38). La vitesse de déplacement est alors estimée en confrontant le temps de passage entre la VTL1 et la VTL2. Ainsi, avec ce type de comptage original, la mesure du trafic se fait donc au niveau de l'infrastructure, et se rapproche des comptages qui ont été exploités au long de ce travail (les capteurs électromagnétiques).

La mobilisation des capteurs fixes apparait aussi être une réponse aux enjeux liés à la protection de la vie privée, et elle permet de remettre au goût du jour la mobilisation des comptages électromagnétiques, pour lesquels les caractéristiques s'inscrivent dans les problématiques liées au *big data*. À notre sens, l'initiation de cette méthode vise en partie à compléter les lacunes des comptages par boucles électromagnétiques, car les emplacements de ces dernières ne permettent pas toujours de couvrir l'ensemble du réseau routier, en raison du coût important d'installation et d'entretien. De plus, l'implantation des VTL est flexible (Hoh *et al.*, 2008), ce qui leur confère un avantage supplémentaire par rapport à des capteurs électromagnétiques pour lesquels le déplacement est impossible sans détruire la chaussée où ils sont installés. La mobilisation des VTL constitue à notre sens un bon exemple d'un système de comptage mixte, combinant à la fois le capteur mobile pour la mesure et le capteur fixe pour la collecte. De surcroit, les VTL permettent d'offrir une couverture spatiale complémentaire à celle des comptages routiers bien plus importante et significative. Elles s'inscrivent aussi dans les tendances actuelles en vue d'exploiter les potentiels des capteurs mobiles pour la mesure des dynamiques urbaines (Herrera *et al.*, 2010 ; Hoh *et al.*, 2012).

## 9.2. Une approche de modélisation SCAUP à affiner

Après avoir abordé les aspects rattachés au dispositif de quantification et aux relations que celui-ci entretient avec le dispositif de modélisation, cette seconde section vise plutôt à se focaliser sur le dispositif de modélisation en lui-même (*cf.* figure 3.35, en rouge). Dans cette section, un premier test a été élaboré afin d'observer les réponses de notre modèle lors de l'insertion d'une perturbation liée à l'implantation du tramway sur la ville de Dijon. Ensuite, des pistes de développements seront proposées dans l'objectif d'améliorer le modèle SCAUP, à la fois sur le processus de simulation du trafic routier, mais également en portant un regard sur sa formalisation. Ici, ce n'est pas la démarche de modélisation (Chorley, 1964) que nous souhaitons interroger, mais bien le cœur du système de simulation multi-agents, en nous concentrant sur différents aspects du modèle SCAUP.

### 9.2.1. L'arrivée du tramway en 2013 : quelle(s) réponse(s) du modèle à une perturbation structurelle ?

Un second type de questionnements sur lequel porte notre travail vise aussi à appréhender les réponses du modèle à des changements structurels intervenus récemment sur le réseau routier dijonnais. Pour cela, nous avons introduit le réseau routier de la ville de Dijon à la suite de l'implantation du tramway en 2012. Ici, l'objectif est d'observer si SCAUP est capable de reproduire les mêmes conditions de circulation routière malgré l'insertion d'une perturbation liée

à l'implantation du tramway. Comme nous l'avons vu dans la première partie, la mise en place du tramway a profondément impacté le réseau de capteurs de la ville de Dijon. C'est pourquoi, l'implantation du tramway fournit l'opportunité d'étudier les réponses du modèle à ce changement d'état des réseaux (de comptage et routier), en observant si les mêmes conditions de circulation sont produites, bien que les données de comptages ne soient pas identiques. En somme, cette expérience vise à vérifier que le modèle SCAUP répond bien à ce à quoi il est destiné, afin de juger de ses aptitudes à produire un trafic routier observé en concordance avec les données d'observations collectées.

Pour cela, nous avons fait le choix de tester deux conditions initiales de simulation de SCAUP[198], en exploitant d'une part le réseau routier de la ville de Dijon de 2013, et d'autre part les données de comptage de 2008 et de 2013. Deux expériences ont été menées, en exploitant pour le premier test les stations de comptage disponibles en 2008 (156 stations), et pour le second test les stations de comptage disponibles en 2013 (110 stations). De surcroit, en vue de conserver des journées de comptage comparables, les journées de comptage sélectionnées pour les deux tests sont le mardi 9 septembre 2008 et le mardi 10 septembre 2013. Il est important de noter que mis à part pour les données de comptage, les mêmes paramètres en entrée du modèle ont été définis, en adéquation avec les résultats présentés plus haut (*cf.* section 7.3, p. 246, à savoir un déplacement tenant compte de la hiérarchie des axes routiers (paramètre de déplacement *H*) et un temps de stationnement défini à 15 minutes[199]. De même, l'idée étant de confronter les données de simulation des deux tests aux valeurs observées en 2013, les expériences ont été réalisées en adéquation avec le système de validation croisée mis en place en vue de bénéficier du maximum de comptages pour la simulation, la LOOCV (*cf.* section 7.1.3.b, p. 233).

Les résultats cartographiques en débit simulé par axes sont présentés en figure 3.39. Les coefficients de détermination au niveau de chaque station disponible en 2013 sont représentés parallèlement (*cf.* ci-dessous). En comparant les deux cartes, on remarque assez peu de différences significatives entre les deux tests de simulation en termes de débit simulé. Néanmoins, lorsque l'on se focalise sur les coefficients de détermination au niveau des stations, on remarque quelques différences, dans le sens où l'on observe une dominance du vert (coefficients situés entre 0,96 et 1) pour la simulation initialisée avec les stations de 2013, par rapport à celle initialisée avec les stations de 2008. À titre indicatif, près de 50 % des stations ont un indice inférieur à 0,9 pour la simulation initialisée avec les comptages 2008, alors que pour l'année 2013, 19 % des stations se situent sous ce seuil.

Cette observation n'est que peu surprenante. En effet, dans la mesure où notre approche de modélisation est guidée par les données (*data-driven*), le modèle SCAUP n'est que peu capable de produire un tout autre trafic routier que celui observé à une station donnée. En définitive, il est nécessaire d'avoir des données de comptage adéquates à la période que l'on souhaite simuler, sans quoi les débits simulés risquent d'être peu significatifs. En d'autres termes, les modifications structurelles du réseau routier induites par le tramway modifient sensiblement la structure du trafic routier sur la ville de Dijon.

[198] À noter qu'au même titre que le réseau routier de 2008, le réseau routier de 2013 a été conçu et réalisé par l'intermédiaire du module associé au modèle MOBISIM (p. 187).
[199] Dans le même ordre d'idée, la matrice de transition est identique à celle utilisée pour l'ensemble des simulations.

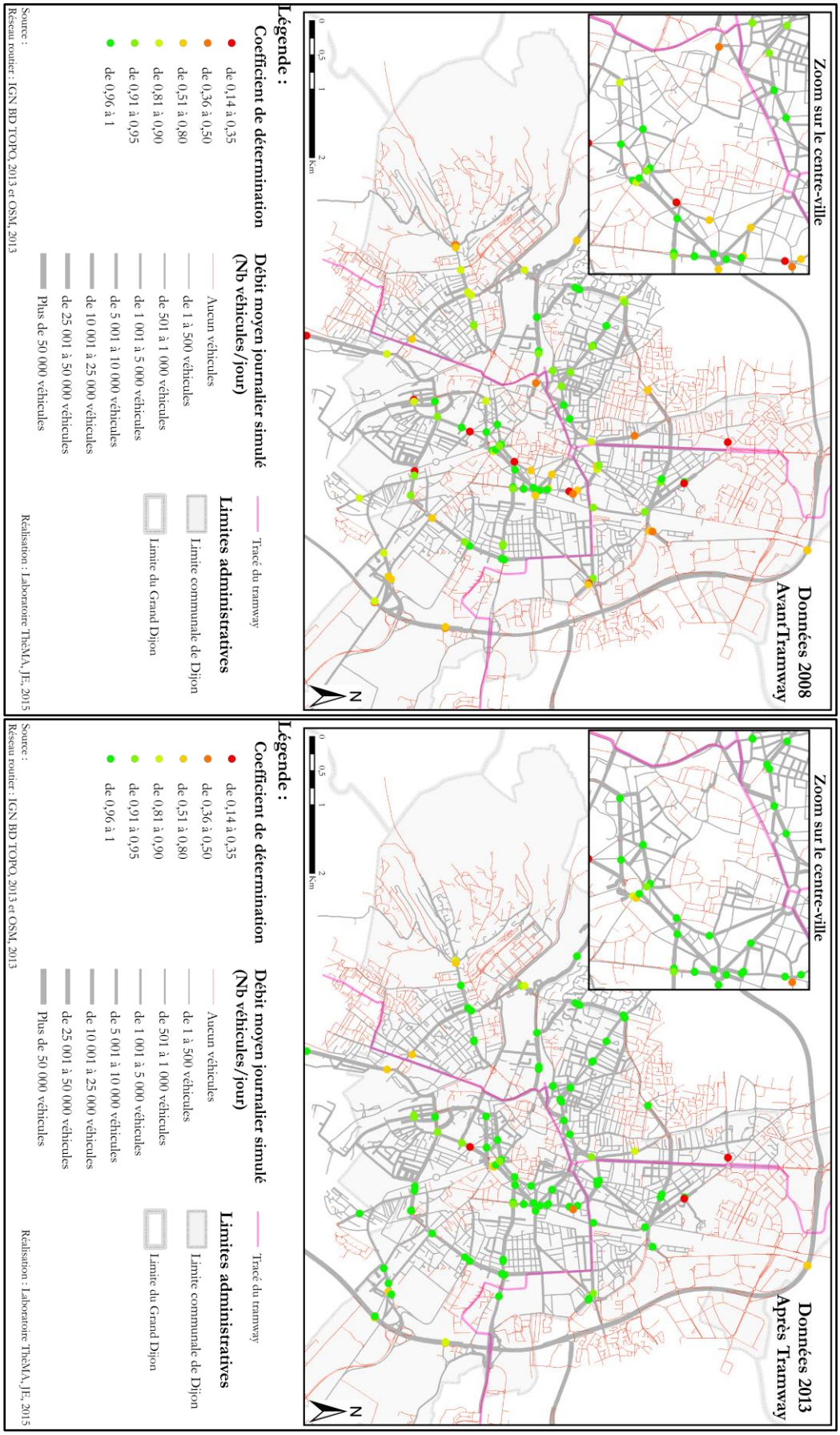



**Figure 3.39 : Trafic routier simulé après l'implantation du Tramway avec les données de comptage de 2008 (en haut) et de 2013 (en bas)**

Comme illustré sur le diagramme de Taylor (2001), présenté en figure 3.40, on observe le même résultat que sur les cartes précédentes. On remarque que les données simulées *via* les stations disponibles en 2013 (triangle bleu) fournissent des résultats de meilleure qualité que le trafic routier initialisé et simulé avec les stations disponibles en 2008 (étoile bleue). En d'autres termes, sur le diagramme, la distance aux observations de 2013 (en noir) est bien plus importante pour la simulation initialisée à partir des données de 2008 que pour celle initialisée à partir des données de 2013. De plus, sur le plan spatial (à droite) et sur le plan temporel (à gauche), on observe des tendances relativement proches de ce qui a été observé lors des expérimentations réalisées précédemment (*cf.* section 7.3.3, p. 254). C'est-à-dire que la dynamique temporelle du modèle SCAUP est de bien meilleure qualité que la répartition spatiale des véhicules simulés. Cependant, comparativement aux constats exposés plus haut (*cf.* figure 3.19, p. 256), sur le plan spatial (à gauche), le coefficient de corrélation est de bien meilleure qualité (0,41), alors que précédemment on dépassait beaucoup plus difficilement la ligne de significativité (illustrée en rouge). On observe alors que la matrice de transition permet de mieux répartir les véhicules sur le réseau routier de 2013 que sur le réseau routier de 2008 (p. 256).

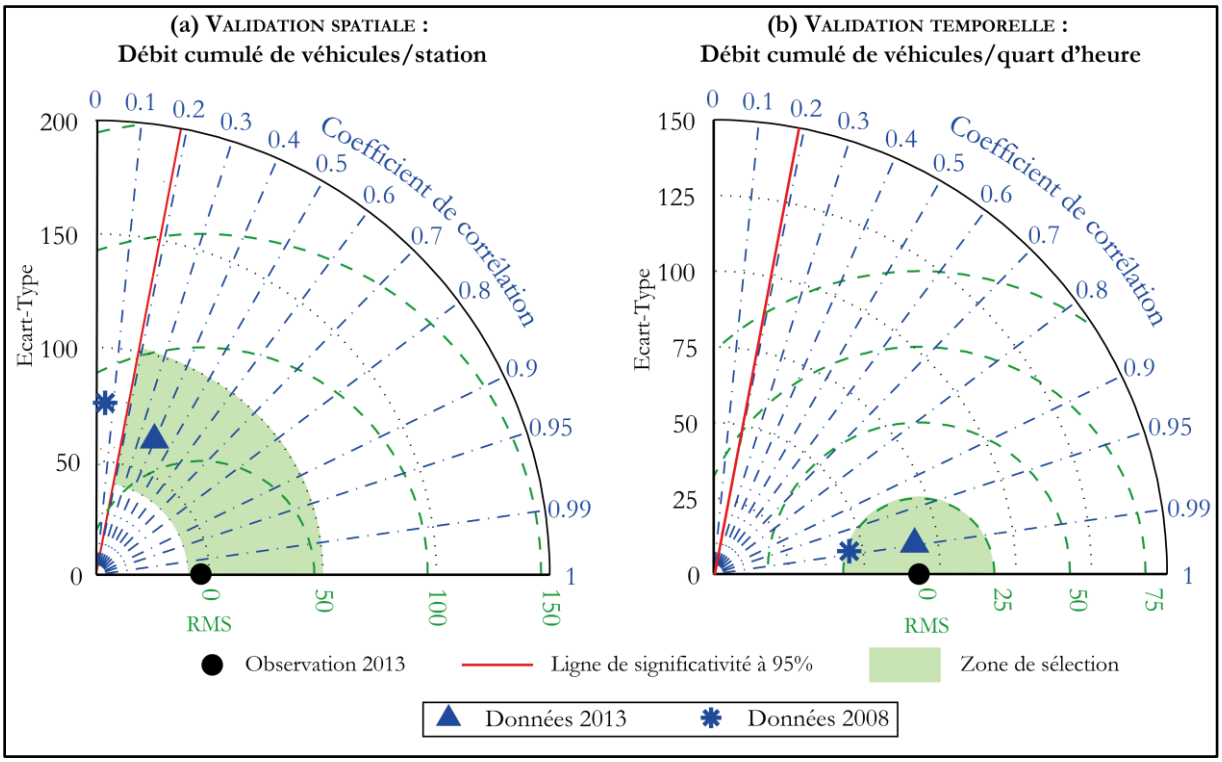



**Figure 3.40 : Évaluation des débits de trafic routier simulé avec SCAUP en fonction des observations du mardi 10 septembre 2013**

Ce constat s'explique probablement par deux aspects liés, autant aux données d'observations exploitées qu'à la modification du réseau routier, suite à la mise en service du tramway. D'une part, un nombre de stations plus restreint pour évaluer avec finesse la répartition des véhicules sur le réseau routier impacte nos validations. À titre indicatif pour l'évaluation, seulement 110 stations de référence sont disponibles pour l'année 2013. Dès lors, un nombre moins important de stations pour la validation spatiale ne nous permet pas d'apprécier avec finesse la répartition des véhicules. D'autre part, la nouvelle configuration du réseau routier (après 2013) apparait plus contraignante et plus hiérarchisée qu'en 2008, concentrant les déplacements sur non plus deux

mais une seule ceinture de boulevard périphérique. Ces deux hypothèses expliquent probablement en partie le positionnement des points de simulation sur la partie *spatiale* (à gauche) du diagramme de Taylor. C'est pourquoi il nous est apparu important de nuancer ces résultats.

### 9.2.2. **Comment améliorer le processus de simulation interne au modèle SCAUP ?**

Le processus de simulation interne à SCAUP porte sur la manière dont la simulation se déroule. Dans notre cas, au regard des résultats exposés précédemment, la période d'initialisation et de simulation du trafic routier apparait être une phase importante, pouvant être optimisée, afin de raffiner deux aspects du modèle de simulation. Le premier aspect vise à améliorer la reproductibilité du modèle SCAUP en s'inspirant des procédures employées en sciences du climat et en ingénierie du trafic routier. Le second aspect, quant à lui, est à rattacher à l'exploration de la matrice de transition, en vue d'optimiser la répartition des véhicules routiers sur l'espace d'étude. Du fait du nombre important de paramètres possibles de la matrice de transition (11^11 possibilités), son exploration ne peut être envisagée que par l'intermédiaire de la simulation, non seulement en s'appuyant sur les capacités de calcul (HPC), mais aussi, en s'aidant de procédures d'apprentissage se situant dans le prolongement des approches de Monte-Carlo (Millet, 2005).

#### a. ***Améliorer la reproductibilité de SCAUP : le processus de warm-up***

Pour commencer, l'un des constats que nous avons pu faire précédemment porte sur la manière avec laquelle la simulation est initialisée. Effectivement, au lancement de la simulation (habituellement à minuit) le réseau routier est vide, et ce n'est que par l'insertion progressive des valeurs de comptages routiers qu'il se charge (*au fil de l'eau*). Or, il apparait évident que lorsque la simulation est initialisée, un nombre de véhicules devrait être déjà présent sur les infrastructures de transport. Malgré tout, il faut souligner que ce nombre de véhicules est relativement faible. À titre indicatif, la moyenne des véhicules observés à minuit sur les stations de comptage exploitées est d'environ 20 véhicules (à Dijon). Cependant, il serait intéressant de bénéficier d'un trafic routier pré-chargé au démarrage de la simulation pour améliorer le trafic routier simulé.

En modélisation du climat, c'est une technique qui est bien connue lors de l'application des modèles numériques de climat, dont le bon fonctionnement nécessite un temps préalable d'apprentissage pour calculer les conditions dites d'initialisation des conditions atmosphériques de surface (Crétat *et al.*, 2011b). En d'autres termes, c'est un processus de *pré-simulation* du modèle qui est lancé pendant une période donnée, et qui ne sera pas pris en compte lors de l'analyse des résultats. Il permet alors de bénéficier d'une période de stabilisation du modèle. Ce processus est aussi intégré au sein de différents modèles de simulation du trafic routier comme SUMO (Behrisch *et al.*, 2011), VISSIM (Fellendorf et Vortisch, 2010) ou AIMSUN (Casas *et al.*, 2010). Pour ces modèles microscopiques, cette phase de *pré-simulation* est définie comme une période dite de « *warm-up* » (Antoniou *et al.*, 2014, p. 20) : « *en règle générale, la période de warm-up doit être suffisamment longue pour que le trafic routier soit suffisant sur la totalité ou la majorité du réseau routier avant le démarrage de la simulation principale* ». De surcroit, les auteurs relèvent que pour cette procédure, la plupart des manuels d'utilisation des modèles de trafic commerciaux soulignent que ce temps de pré-simulation est important (Barceló, 2010) en vue d'obtenir un « *état assez réaliste du trafic routier au début de la simulation* » (Antoniou *et al.*, 2014, p. 39) ; ce qui est particulièrement le cas des

modèles commercialisés par les entreprises privées comme TSS et PTV[200]. Pour réaliser cette procédure, l'utilisation des puissances de calcul devient encore plus indispensable car, dans le cas d'un modèle de climat la procédure de pré-simulation est généralement réalisée sur une année complète pour une même année de simulation (Crétat *et al.*, 2011b). La mise en œuvre de cette *pré-simulation* impacte alors le temps de calcul nécessaire à l'initialisation d'un modèle.

Dans le cas du modèle SCAUP, comme illustré en figure 3.41 (ci-dessous), on peut envisager cette procédure de *pré-simulation* sur une journée complète, ce qui permettrait par la suite de simuler le trafic routier d'une journée complète de simulation initialisée avec un réseau préalablement chargé. Dans ce sens, comme illustré ci-dessous, deux journées de comptage seraient simulées avec SCAUP, une première simulation dite *warm-up*, qui ne serait pas utilisée lors de l'*évaluation* des données simulées, et une seconde journée qui serait la période de simulation exploitée pour l'extraction et l'*évaluation* des résultats de simulation.

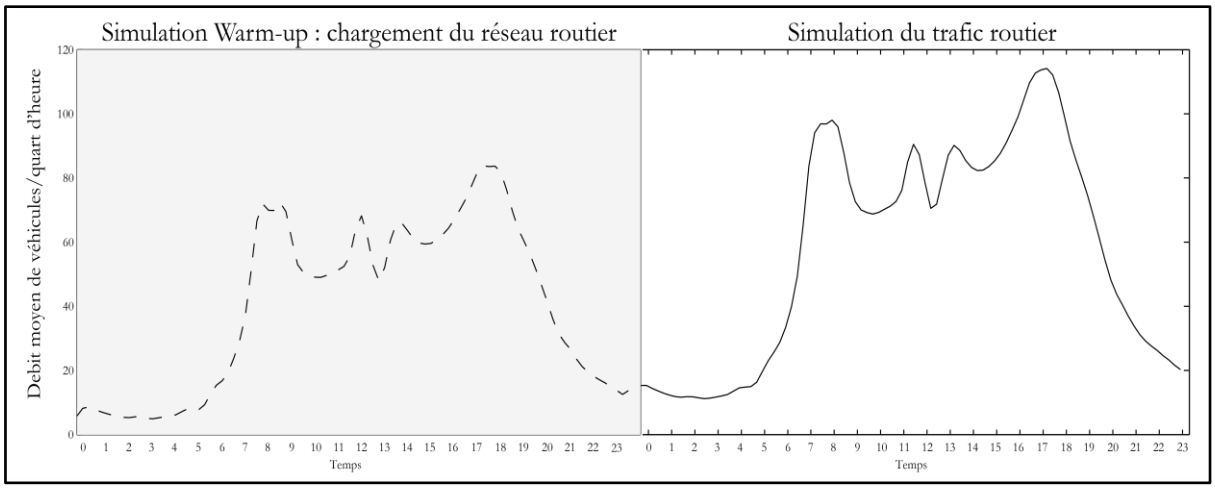



**Figure 3.41 : Charger le réseau routier par un processus de *warm-up* : un processus de *pré-simulation* pour la simulation du trafic routier**

Si le processus de *warm-up*, dans notre cas, permet de charger le réseau routier de l'espace d'étude, on peut aussi en attendre différentes améliorations lors de l'*évaluation* du modèle SCAUP. À notre sens, s'insérer dans cette perspective de développement du processus de simulation permettrait de raffiner deux aspects de la simulation, en améliorant :

- **la reproductibilité** du modèle, qui, comme nous l'avons observé (*cf.* figure 3.14, p. 245), est plus faible au démarrage de la simulation que par la suite. On peut alors en attendre une amélioration du pourcentage de reproductibilité pour la première période de simulation, qui, comme nous l'avons vu, est principalement impactée par un nombre de véhicules relativement faible (*cf.* section 7.2.3, p. 242).
- **la répartition** du nombre de véhicules sur les axes routiers, qui, comme nous l'avons observé en figure 3.16 (p. 248), est le plus généralement sous-estimé par rapport à ce qui est observé au niveau des stations de comptage. En simulant la période de comptage avec un réseau routier *pré-simulé*, on peut alors en attendre une relative augmentation du nombre de véhicules simulés sur les axes routiers, comme illustré en figure 3.41.

[200] TSS (Transport Simulation Systems) et PTV (Planung Transport Verkehr) sont respectivement les propriétaires des modèles de simulation du trafic routier AIMSUN et VISSIM-VISUM. À noter, que ce sont, aujourd'hui, les entreprises leaders pour la commercialisation des modèles de simulation du trafic routier.

Effectivement, la simulation du trafic routier s'effectue au fil de l'eau par l'insertion progressive des véhicules routiers. Naturellement, vu que le nombre de véhicules est implicitement augmenté, on pourrait alors attendre une relative amélioration en termes de véhicules passant sur les infrastructures de transport du domaine d'étude.

Cependant, la répartition des véhicules est plutôt définie par un paramètre essentiel, qui tient compte en très grande partie des paramètres de la matrice de transition, pour laquelle nous avons soulevé les limites à plusieurs reprises. C'est pourquoi il est aussi important d'en explorer plus finement le paramétrage, en vue de raffiner la répartition des véhicules sur les différents axes de l'espace d'étude.

### b. *Raffiner la matrice de transition par simulation et apprentissage*

Au cours de l'élaboration du plan d'expérience du modèle SCAUP, nous n'avons pas engagé d'exploration de la matrice de transition du déplacement des véhicules car, non seulement celle-ci a été définie en fonction des comptages permanents et ponctuels collectés, mais aussi car son exploration pas-à-pas n'est que peu envisageable du fait d'un nombre de possibilités trop important (*cf.* tableau 2.3 et tableau 2.4, p. 200). Par ailleurs, comme le mentionne André Berchtold (1998, p. 112) : « *l'utilisation des matrices de transition comporte généralement un grand nombre de paramètres à estimer* » ; ainsi il est souvent impossible d'en explorer l'ensemble des paramètres pas-à-pas. Étant dans ce contexte, nous rappelons que la matrice de transition associée au référentiel de OpenStreetMap nécessite l'exploration de 11^11 possibilités, ce qui rend peu envisageable l'exploration pas-à-pas de celle-ci. Or, nous avons constaté qu'il est important d'en raffiner les paramètres. Si son exploration détaillée n'est pas possible, il existe cependant différentes méthodes permettant de pallier ce problème comme les approches de Monte-Carlo (Millet, 2005) ou les chaînes de Markov (Berchtold, 1998). La chaîne de Markov, qui découle des approches de Monte-Carlo apparait intéressante à exploiter. En effet, c'est « *un processus dont les probabilités de transition sont des probabilités conditionnelles au passé* […] ; *il met en relation des observations successives d'une même variable en fonction d'un certain nombre d'observations passées de cette même variable* » (Berchtold, 1998, p. 16).

Cette méthode statistique est utilisée dans de nombreuses disciplines, comme la géographie, pour calibrer un modèle d'évolution d'occupation du sol (Antoni, 2006). Dans notre cas, elle pourrait fournir l'opportunité d'explorer le paramétrage de la matrice de transition, pour mieux se caler sur les données observées à l'échelle du réseau routier dijonnais. Il s'agirait par exemple de mettre en place d'un processus d'apprentissage de la matrice de transition tenant compte des valeurs estimées à l'échelle des tronçons routiers. Cependant, si le principe général de la chaîne de Markov apparait adapté à notre travail et au calcul des matrices de transition (Berchtold, 1998), il faut cependant relever que sa mise en œuvre nécessite des compétences avancées que nous ne pouvons pour le moment pas mettre en application. L'idée est ici de s'inscrire dans une perspective d'exploration des paramètres de la matrice de transition, qui vise à répartir chaque véhicule sur les tronçons routiers, en tenant compte des valeurs de fréquentation types observées à l'échelle du territoire.

Enfin, la mise en place de ce processus d'exploration de la matrice de transition joue sur la manière avec laquelle les véhicules se déplacent à l'échelle des tronçons routiers, en vue de mieux contraindre la répartition des véhicules (échelle microscopique), en adéquation avec les

observations réalisées sur le territoire (échelle macroscopique). C'est une approche que l'on peut assimiler au module d'affectation que nous avons réinterprété plus haut (*cf.* section 5.3.2, p. 172). Toutefois, dans une perspective de simulation du trafic routier, au module d'affectation se juxtapose un module d'écoulement qui vise à simuler le déplacement des véhicules tout au long d'une infrastructure de transport (Buisson, 1996). Une amélioration en entrainant une autre, nous nous intéresserons, dans un troisième temps, à la manière dont les véhicules parcourent les infrastructures de transport, en se focalisant sur le déplacement des véhicules pris individuellement, et en se centrant sur la relation que les véhicules entretiennent entre eux et avec l'infrastructure.

#### 9.2.3. D'un formalisme KISS vers un formalisme KIDS pour affiner le déplacement des véhicules sur les routes

L'approche de modélisation expérimentale SCAUP que nous avons décrite s'inscrit dans un formalisme des véhicules vus depuis l'extérieur, c'est-à-dire que leurs déplacements sont définis à un niveau général et relativement homogène. C'est pourquoi, la description de leur comportement de déplacement s'insère dans un formalisme KISS (*cf.* section 5.3.1.c, p. 169). Ce parti pris ne constitue cependant qu'une première pierre pour une approche de modélisation SCAUP plus complète. En effet, dans une perspective d'évaluation de la PAA, il est fondamental d'affiner le comportement de déplacement des agents « véhicules » en s'inscrivant dans un formalisme KIDS du mouvement des véhicules routiers. En d'autres termes, s'inscrire dans le formalisme KIDS vise à plutôt se centrer sur l'intérieur des véhicules, en y apportant un niveau de détails bien plus élevé qu'auparavant, afin de reproduire des comportements de déplacement non plus homogènes mais bien hétérogènes et propres à chaque véhicule.

Fondamentalement, affiner le déplacement des véhicules sur le réseau routier ne nécessite aucun changement sur la structure du modèle conceptuel, mais c'est plutôt au niveau du formalisme du modèle de simulation (*étape 2* de la démarche de modélisation comme illustré en figure 2.1, p. 128) que ces modifications entrainent des changements. On peut identifier deux points en vue de raffiner le formalisme du modèle SCAUP. Le premier point porte sur la description du déplacement du véhicule (*agent*), tandis que le second porte sur les interactions que les véhicules ont entre eux (*agent-agent*). C'est-à-dire qu'il est nécessaire de se focaliser sur les différentes règles de conduite qui régissent le mouvement et la friction des véhicules tout au long de leurs parcours. En s'inscrivant dans cette optique, deux niveaux peuvent alors être identifiés :

- **Pour les agents** : en intégrant aux véhicules une taille, ainsi que des règles de conduite sur les tronçons. Par exemple, en formalisant un code de la route intégrant les priorités à droite et une signalétique au niveau des carrefours routiers, tandis que l'intégration d'une taille pour chaque véhicule permettra de leurs attribuer une meilleure emprise sur les tronçons routiers.
- **Entre les agents** : en intégrant des règles d'interaction entre les agents véhicules, comprenant non seulement les distances de sécurité réglementaires définies par le code de la route, mais aussi, des règles de décélération et d'accélération qui dépendent de la présence des autres véhicules.

Ces deux éléments visent à définir, au niveau des véhicules, une *friction* entre les agents, afin que ces mobiles perçoivent, interprètent et prennent des décisions relatives à leur positionnement sur la route à chaque étape du processus de simulation. Cependant on peut se demander quelle(s) amélioration(s) attendre de l'insertion de règles de conduite au sein du modèle SCAUP ?

Tout d'abord, l'intégration d'un comportement de déplacement plus détaillé devrait assez naturellement nous permettre de reproduire la relation entre le débit et la vitesse (courbe débit/vitesse) grâce à la nature générative de la SMA. Si ce cas de figure se vérifie, cela nous permettrait de simuler différentes conditions d'écoulement du trafic routier : un trafic fluide et un trafic congestionné (*cf.* encart 1, p. 75). De plus, les règles d'accélération ainsi que de décélération des véhicules sur les tronçons routiers devraient aussi permettre de produire un trafic routier « en accordéon ». C'est à dire que la flotte de véhicules simulés se déplace non plus dans des conditions optimales (fluides), mais dans des conditions de déplacement variables à *l'intérieur* même d'une voie de circulation où les véhicules sont en interaction (les intersections), et donc proches, et les moments où les véhicules sont éloignés (la section courante). Enfin, les vitesses ainsi que les comportements de déplacement des véhicules n'étant pas non plus homogènes mais bien hétérogènes, on approche d'une conceptualisation de l'écoulement du trafic routier plus proche de ce qui est observé sur le territoire. De surcroit, dans une optique d'évaluation de la PAA, le raffinement des vitesses de circulation par la reproduction des flux de circulation en accordéon, ainsi que par des comportements de déplacement des véhicules plus fins, est une perspective de développement d'autant plus importante dans le cadre d'une évaluation environnementale des impacts du trafic routier en milieu urbain.

Ces différents postulats autour de la relation débit-vitesse et des variations de vitesse pour la modélisation des conditions de circulation, ont déjà été montrés à plusieurs reprises, que ce soit à l'échelle macroscopique ou microscopique (Buisson, 1996 ; Bonnel, 2002 ; Leclercq, 2002 ; Champion, 2003 ; Behrisch *et al.*, 2011 ; Ortúzar et Willumsen, 2011). C'est pourquoi nous supposons que ces phénomènes peuvent affiner la manière avec laquelle les véhicules se déplacent sur les infrastructures de transport. Par ailleurs, en disposant de données issues des capteurs électromagnétiques, il est aussi possible d'appréhender ces dynamiques de circulation en associant aux débits de trafic routier les comptages relevés sous la forme d'un taux d'occupation[201] (*cf.* tableau 1.5*,* p. 85). Dans ce sens, avec les données issues des capteurs électromagnétiques, nous disposons alors de données permettant de qualifier les conditions de circulation routière, et donc de valider les conditions de trafic routier simulées avec SCAUP. Dans cette optique, on retrouve ici les aspects empiriques et inductifs de la démarche de modélisation SCAUP initiée au cours de ce travail.

Pour finir sur les développements associés au dispositif de modélisation, il est important de souligner que ces raffinements s'inscrivent dans l'optique de mieux comprendre et observer les phénomènes de Pollution Atmosphérique Automobile (PAA). L'amélioration du déplacement des véhicules fournit l'opportunité d'intégrer plus directement la méthodologie COPERT au sein même du modèle de trafic routier. En effet, comme nous l'avons observé au cours du Chapitre 8 (pp. 263- 299), l'inscription de notre travail de simulation au sein des référentiels nationaux

[201] À noter que les données de taux d'occupation (TO) n'ont pas été collectées auprès de la ville de Dijon, car leur récupération demandait un temps de recueil trop important pour le gestionnaire de trafic, par rapport aux données de débit qui elles étaient déjà disponibles en l'état.

exploités par les AASQA n'est pas assez fin pour apprécier la dynamique portée par le SCAUP et donc *a fortiori* des émissions de PAA qui en seraient estimées sur des pas-de-temps très courts (minutes ou quart d'heure). Or, ces dynamiques sont caractéristiques de la pollution de proximité, qui pour rappel s'inscrivent dans des temporalités très courtes, allant de la seconde à la minute ainsi qu'à l'intérieur d'un même tronçon routier sous la forme de *hot spots* (*cf.* section 1.1.3, p. 31). Nous verrons d'ailleurs qu'en parallèle de la méthodologie COPERT, d'autres méthodes de calcul des émissions de PAA présentent des potentiels intéressants pour mieux prendre en compte les conditions de circulation routière (*cf.* section 9.3, ci-dessous).

Dans ce sens, raffiner la simulation du déplacement des véhicules, ainsi que les conditions de circulation routière, tout au long du processus de modélisation SCAUP, s'inscrit aussi dans les relations qu'entretiennent dispositif de modélisation et dispositif d'évaluation au sein de la matrice technique (*cf.* figure 3.35). C'est pourquoi dans un troisième temps nous nous intéresserons aux perspectives de développement associées au dispositif d'évaluation mis en place, en s'inscrivant, non plus dans une optique d'interopérabilité avec les référentiels existants, mais plutôt dans une optique de couplage entre simulation du trafic routier et émissions de PAA.

## 9.3. S'inscrire dans une démarche intégrée d'*évaluation* de la PAA en milieu urbain : de l'émission à l'imission

Parallèlement aux développements qui se rattachent au modèle SCAUP, cette dernière section vise à mettre en avant ceux qui se rattachent plutôt au dispositif d'évaluation, afin d'offrir une vision complète des perspectives de travail de l'approche proposée (*cf.* figure 3.35). En effet, lors du calcul des émissions de PAA avec CIRCUL'AIR, la démarche que nous avons présentée ne permet pas de mettre en exergue les aspects dynamiques du modèle SCAUP, ce qui joue en partie sur les estimations de PAA à un niveau de détail bien trop agrégé pour apprécier les dynamiques de la PAA dans toutes ses dimensions. C'est pourquoi il s'avère nécessaire de s'inscrire dans une optique de couplage entre simulation du trafic routier et modèle d'émission du trafic routier au sein même de la démarche de modélisation. Cette dernière section s'inscrit ainsi dans une vision plus large de l'évaluation en vue d'offrir un regard complet des perspectives de développement de l'approche SCAUP.

### 9.3.1. **Un couplage nécessaire pour affiner les estimations de PAA ?**

Nous avons soulevé une des limites de l'approche présentée se rattachant aux référentiels d'évaluation des AASQA : les émissions de polluants sont trop agrégées, ce qui ne met que peu en avant le caractère spatio-temporel de la PAA. Dès lors, en combinaison d'un comportement de déplacement des véhicules plus fin, une perspective de développement primordiale de la démarche de modélisation SCAUP vise à intégrer et à appliquer la méthodologie COPERT (intégrée aussi à CIRCUL'AIR) au sein même du modèle.

Effectivement, en s'appuyant sur les vitesses de déplacement simulées, il est possible d'appliquer les règles de calculs d'émissions unitaires des véhicules inclus au sein de la méthodologie COPERT. Dans ce sens, en utilisant les vitesses de déplacement simulées et la définition d'un parc automobile, il apparait envisageable d'avoir une approche du calcul des émissions de PAA intégrée à SCAUP, en se dédouanant des référentiels utilisés par les AASQA, qui s'avèrent trop agrégés et généraux. Des recherches ont été initiées dans ce sens en s'appuyant sur le modèle de

simulation du trafic routier AIMSUN (Chanut et Chevallier, 2012). D'autres modèles de simulation du trafic routier s'inscrivent aussi dans cette optique. C'est particulièrement le cas de SYMUVIA[202], développé par l'IFSTTAR et le LICIT, qui intègre des modules permettant d'évaluer les émissions du trafic routier, ainsi que des calculs de bruits. Cependant, la méthodologie COPERT se base sur les vitesses moyennes de circulation par tronçon, et ne permet pas d'estimer avec précision la dynamique du trafic routier à l'intérieur d'un tronçon routier. C'est pourquoi, en parallèle de la méthodologie COPERT, d'autres couplages peuvent être envisagés, comme avec l'outil HBEFA[203], qui permet de mieux considérer les effets locaux du trafic routier, par la prise en compte de la voie et de la « situation de trafic » (André *et al.*, 2013).

De plus, il faut relever qu'en s'inscrivant dans ce couplage intégré, avec des outils comme COPERT ou HBEFA, l'une des difficultés à laquelle nous sommes confrontés porte sur la définition d'un parc automobile représentatif de l'espace local simulé. De même que le calcul des surémissions « à froid » s'avère aussi plus complexe que dans une approche agrégée. Pour pallier ces limites, le plus généralement, l'ensemble de ces éléments est défini grossièrement (Chanut et Chevallier, 2012). Par conséquent, comme le relève le guide du PCIT (2012, p. 177) pour le calcul des surémissions à froid au sein de la méthodologie COPERT, on utilise « *une valeur du paramètre β (% de véhicules circulant à froid) adaptée à la situation locale. Des hypothèses locales peuvent être fournies par les CETE* [aujourd'hui CEREMA] *et certains modèles de trafic fournissent cette information pour chaque brin modélisé* ». C'est-à-dire qu'en s'appuyant sur des pourcentages estimés à partir des données nationales ou des données locales, le taux de véhicule « à froid » est défini sur la base de descriptions générales qui tentent d'approcher des grandes tendances observées.

Malgré tout, la force de se placer dans une démarche intégrée plus complète s'inscrit dans l'optique de proposer deux types de sorties à partir du modèle SCAUP :

- En termes de **trafic routier simulé**, pour alimenter les référentiels des AASQA, ce qui a été présenté au cours de ce travail
- En termes **d'émissions de PAA**, pour mieux appréhender les caractéristiques spatio-temporelles de la pollution de proximité automobile, qui se situent dans le prolongement de nos travaux.

Un premier développement a été initié dans ce sens à travers une application sur la ville de Marrakech, intégrant un calcul des émissions du trafic routier au modèle SCAUP sur la base de la méthodologie COPERT. Cette première application permet d'illustrer non seulement les perspectives de développement d'une approche intégrée de calcul de la PAA, mais aussi le caractère généraliste de l'approche de modélisation SCAUP à d'autres espaces que celui de la ville de Dijon.

[202] Vidéo de présentation du modèle SYMUVIA : https://www.youtube.com/watch?v=LqBD8nKiHN8.
[203] Handbook Emission Factors for Road Transport.

### 9.3.2. **L'application du modèle SCAUP sur la ville de Marrakech : Une première application de couplage avec la méthodologie COPERT**

Ce travail, lancé en Mai 2016, s'inscrit dans l'optique de réaliser une animation en trois dimensions des émissions de PAA en vue de la COP 22 qui se tiendra à Marrakech du 7 au 18 Novembre 2016[204]. L'ensemble de ce projet pour la COP 22 est coordonné par l'IRD (Institut de Recherche en Développement), visant à « *sensibiliser la société civile, en particulier les jeunes mais également les ONG et les universitaires, aux enjeux de la lutte contre le changement climatique* »[205]. Dans ce cadre, le modèle SCAUP a été exploité pour la simulation du trafic routier sur la ville de Marrakech, et il sera présenté et utilisé pour une animation sur les impacts de la PAA en milieu urbain. Cette application s'inscrit dans l'objectif de vulgariser, mais aussi de sensibiliser le grand public aux effets de la PAA à l'échelle d'un territoire, sous la forme d'un modèle interactif. En effet, par l'intermédiaire d'une interface interactive (tablette numérique), les utilisateurs pourront interagir avec le modèle SCAUP en modifiant différents paramètres d'entrée, par exemple la forme du réseau routier ou les caractéristiques du parc automobile, en vue d'apprécier les effets de ces modifications sur les émissions de PAA.

Au-delà de ces aspects, cette application fournit aussi l'opportunité de tester les capacités du modèle SCAUP sur un espace urbain autre que celui de la ville Dijon, en modifiant seulement l'environnement de simulation généré à partir d'OpenStreetMap et les comptages réalisés à partir d'un système de comptage original. De plus, cela permet de s'inscrire dans une perspective de couplage de la méthodologie COPERT au sein même du modèle SCAUP. L'initialisation du modèle de simulation de trafic routier nécessite peu de manipulation. En effet, seul le réseau routier Marrakéchois issu d'OpenStreetMap et les comptages routiers sont modifiés en entrée du modèle. Cependant, la ville de Marrakech ne disposant que de très peu de comptages permanents, une application de comptage a été développée pour collecter des données de trafic routier sur le territoire d'étude. Effectués via une application mobile[206], les comptages sont réalisés par des enquêteurs qui, par l'intermédiaire de l'application, recensent le passage des véhicules à différents endroits. C'est un dispositif de comptage original, qui allie nouvelles technologies et comptage manuel, tout en permettant d'automatiser, centraliser et contrôler l'ensemble des comptages effectués une fois ceux-ci synchronisés sur un serveur. Si ces comptages ne sont pas issus du même système de comptage que les capteurs électromagnétiques, ils demeurent cependant adéquats pour initialiser et alimenter le modèle SCAUP en données de comptages ponctuels positionnés sur le réseau routier. C'est pourquoi nous pouvons estimer qu'en combinaison de données de comptages sur le trafic routier, la démarche SCAUP défendue au cours de ce travail peut techniquement être facilement applicable à d'autres espaces urbains que celui de la ville de Dijon.

Le couplage avec la méthodologie COPERT s'appuie sur le référentiel communiqué par SETRA (2009), qui permet de disposer de facteurs d'émissions unitaires tenant compte de la vitesse moyenne de circulation pour plusieurs catégories de véhicules. Il est important de relever que les

[204] Page d'accueil de la COP 22 : http://www.cop22-morocco.com/.
[205] Page d'accueil des activités de l'IRD pour la COP 22 : http://maroc.ird.fr/cop-22.
[206] L'application a été développée par Saad Thoubi dans le cadre de sa thèse portant sur l'*Élaboration d'un modèle multi-agents pour la mobilité urbaine : optimisation des itinéraires* sur la ville de Marrakech : présentation et descriptif du sujet de thèse : https://edite-de-paris.fr/public/phd/html/10342762.

facteurs unitaires sont définis pour un véhicule moyen du parc automobile en circulation, et de plus, les facteurs d'émissions pour une vitesse donnée sont des valeurs moyennes, c'est-à-dire que « *les plus basses vitesses doivent être considérées comme représentatives d'une circulation urbaine caractérisée par de nombreux arrêts, alors que les cycles à haute vitesse traduisent une circulation plus fluide* » (SETRA, 2009, p. 5). Ces courbes d'émissions, dont un exemple pour les $NO_x$ est donné en figure 3.42, apparaissent adaptées pour tester une première approche de couplage à partir de SCAUP pour le calcul intégré des émissions de PAA.

Notons que nous avons fait le choix de nous inscrire dans une approche plus simpliste (KISS) en choisissant de simuler un parc automobile moyen, et en distinguant seulement deux types de véhicules particuliers (VP) en fonction de leur carburant : les VP essence et les VP diesel. À partir de cette distinction, comme illustré en figure 3.42, deux types de parc automobile ont été définis : un parc automobile moyen de 2007 et un parc automobile moyen de 2020.

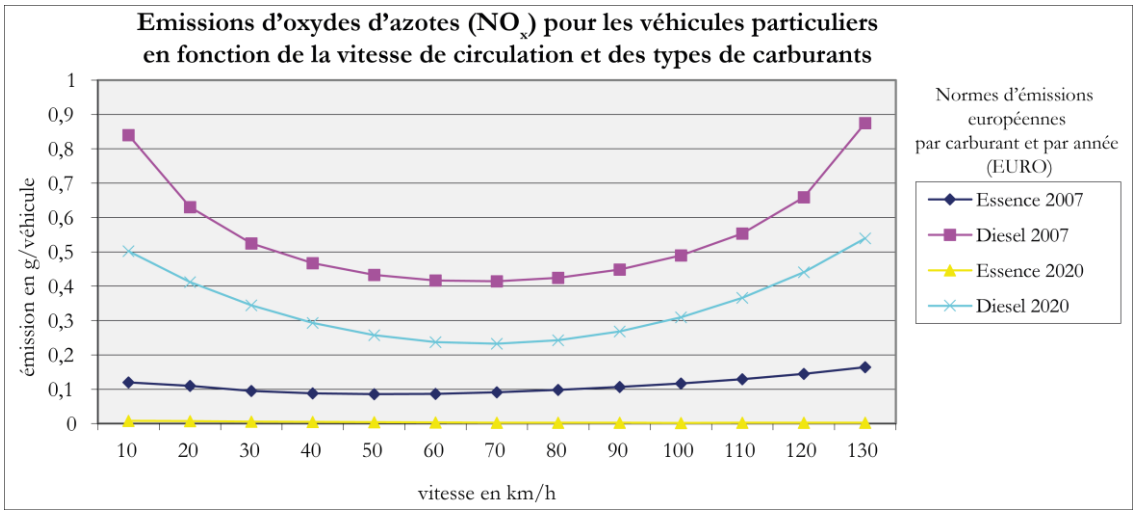



**Figure 3.42 : Facteurs d'émissions unitaires d'oxydes d'azote ($NO_x$) intégrés à SCAUP en fonction des types de carburants pour deux périodes : 2007 et 2020 (source : SETRA, 2009)**

Le calcul des émissions unitaires est spatialisé sous la forme d'une grille permettant de calculer et de visualiser l'évolution des niveaux de PAA pendant et après le processus de simulation du trafic routier, comme représenté en figure 3.43 (ci-dessous). S'inscrire dans cette optique est intéressant car les véhicules diesel présentent un facteur d'émissions près de 5 fois supérieur aux émissions unitaires des véhicules essence (*cf.* figure 3.42). Par conséquent, il est intéressant de tester différents cas d'école en jouant sur les parts de VP diesel et essence afin d'illustrer et de visualiser, par exemple, l'impact d'une trop forte diésélisation du parc automobile comparativement à un parc automobile sans véhicules diesel. L'application du parc automobile de 2020 permettra aussi de montrer les impacts de l'insertion des nouvelles technologies et des nouvelles normes environnementales, les facteurs d'émissions pour les véhicules diesel étant, à cette date, près de deux fois inférieurs aux normes du parc automobile de 2007 (*cf.* figure 3.42).

Cette approche constitue une première phase de développement et demeure relativement réductrice, car cette application s'inscrit, comme nous l'avons vu, dans l'optique d'une animation de vulgarisation destinée au public. Notons ici que l'application complète de la méthodologie COPERT dans SCAUP implique un temps de développement, mais aussi de test, important. En

outre, cette application nous a permis de réaliser un test opérationnel de simulation du trafic routier sur la ville de Marrakech, qui pour l'instant, faute de comptages de terrain, n'a pas encore été évalué et testé. Cependant, nous pouvons voir par l'intermédiaire de la figure 3.43 (ci-dessous) que l'ensemble de la démarche SCAUP construite au cours de cette thèse fonctionne sur un autre espace d'étude, autant en termes de trafic routier simulé que d'émissions de polluants atmosphériques comme ici pour les $NO_x$ :

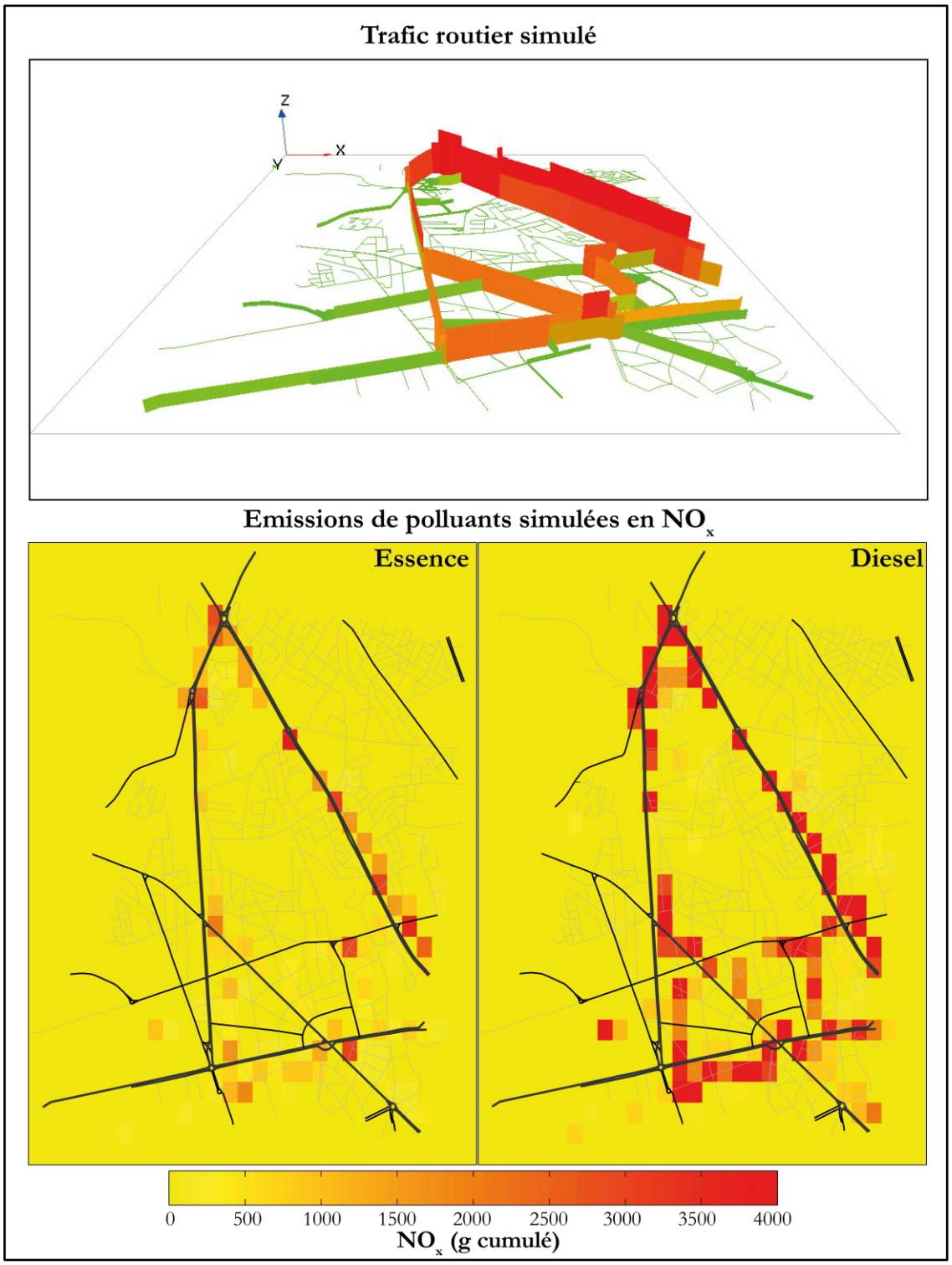



**Figure 3.43 : Illustration de deux *scénarios types* d'émissions de $NO_x$ simulés par SCAUP, en fonction des véhicules : essence (à gauche) et diesel (à droite)**

### 9.3.3. **Le passage aux émissions : un enjeu de recherche encore d'actualité**

Pour finir, nous souhaitons ici mentionner différentes perspectives de développements qui se situent dans le prolongement du dispositif d'évaluation de la PAA proposé dans ce travail. Au cours du Chapitre 8 (p. 263), nous avons distingué les immissions des émissions. L'une des perspectives de travail, en vue de mieux apprécier les phénomènes de PAA, serait de passer des valeurs d'émissions atmosphériques en quantité (kg) à des valeurs d'immissions sous la forme de concentration ($\mu g/m^3$). Ceci peut être envisagé grâce à l'utilisation de données de circulation atmosphérique issues des modèles numériques de climat. De plus, avoir accès aux émissions fournirait l'opportunité de confronter les niveaux de polluants modélisés aux données relevées et observées au niveau des stations de qualité de l'air de la ville de Dijon. S'inscrire dans cette perspective nous permettrait alors de mieux juger de la qualité des données de trafic routier simulées à partir du modèle SCAUP.

Le passage aux émissions permettrait, de surcroit, de valider les données de qualité de l'air modélisées avec les stations d'observation de la qualité de l'air, ce qui n'a pu être réalisé lors de l'analyse des données d'émissions présentées au cours du Chapitre 8 (p. 263). Pour modéliser les immissions, et comme illustré en figure 3.44 (ci-dessous), deux itinéraires peuvent être identifiés afin d'obtenir un bilan assez complet des immissions en milieu urbain et à une échelle relativement fine :

- en **utilisant la simulation multi-agents** (illustré par l'itinéraire en noir) pour avoir accès aux cycles de l'azote et de l'ozone
- en **utilisant les modèles numériques du climat à l'échelle régionale**, en combinaison d'un modèle de diffusion atmosphérique (CTM : Chimie Transport Model), ainsi que d'un modèle de canopée urbaine (Rouïl, 2012 ; Michelot *et al.*, 2015), pour avoir accès aux concentrations des polluants atmosphériques comme les oxydes d'azote ($NO_x$) ou les particules (PM).

L'idée que nous proposons, au sein de la figure 3.44, vise à faire l'interface entre d'une part, le modèle SCAUP (à gauche), et d'autre part la modélisation numérique de la qualité de l'air (à droite). Le premier itinéraire vise à exploiter les données météorologiques d'un modèle de climat pour améliorer l'estimation des émissions du modèle SCAUP. Tandis que le second itinéraire vise plutôt à alimenter en données d'émissions un modèle de la qualité de l'air pour améliorer l'estimation des concentrations de PAA à l'échelle urbaine.

Le premier itinéraire identifié peut être envisagé en s'appuyant sur la simulation multi-agents (SMA), en vue de reproduire les interactions entre trafic routier, émissions de polluants et météorologie (Emery, 2012). En s'appuyant sur les données météorologiques, il est envisageable de formaliser un cycle atmosphérique simple (KISS), en reproduisant par exemple les cycles de l'azote et de l'ozone[207]. Cependant, il faut relever que les phénomènes de pollution atmosphérique sont des phénomènes prenant en compte un grand nombre de relations physico-chimiques, et ce d'autant plus à des échelles locales (Fenger, 2009). En effet, à ce niveau, les polluants sont très réactifs et présentent des dynamiques allant de la minute à quelques heures (Fenger, 1999 ;

[207] Des modèles *types* en fournissent les bases conceptuelles et méthodologiques, en s'inspirant de ces cycles. C'est, par exemple, le cas du modèle Netlogo « *SMOG* » développé par Madison Fitzpatrick en 2013 : http://ccl.northwestern.edu/netlogo/models/community/Smog.

Sportisse, 2007a ; Fenger, 2009). Par exemple, le dioxyde d'azote ($NO_2$) est très réactif, et il entre en interaction non seulement avec d'autres polluants comme l'oxyde d'azote (NO) et l'ozone ($O_3$), mais aussi avec les conditions météorologiques (rayonnement solaire, vent, humidité…).

C'est pourquoi il est également intéressant de s'inscrire dans un positionnement de modélisation KIDS, qu'il ne serait pas possible de mettre en œuvre *via* la simulation multi-agents. Par conséquent, la simulation de la PAA par l'intermédiaire de la simulation multi-agents nécessiterait un développement complet des relations physico-chimiques des polluants, tandis que la modélisation numérique du climat permet plus aisément de modéliser la pollution atmosphérique dans ses caractéristiques physico-chimiques, en faisant appel à des modèles dit de chimie-transport[208] et de canopée urbaine, et ce indépendamment du trafic routier. Comme illustré ci-dessous, nous proposons dans cette perspective d'utiliser le modèle SCAUP pour alimenter en données de trafic, et/ou d'émissions de PAA, la simulation numérique de la qualité de l'air.

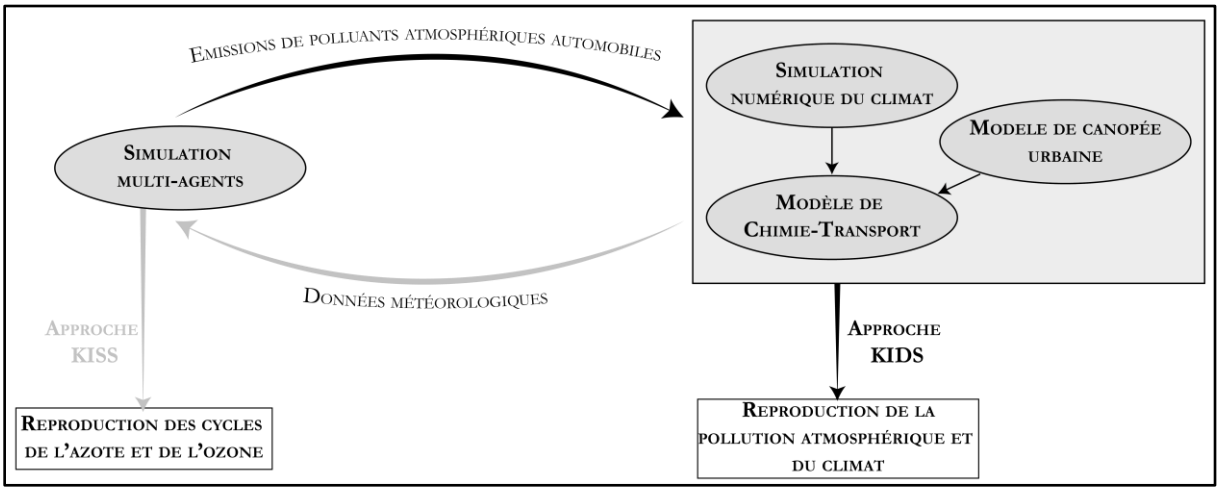



**Figure 3.44 : Deux itinéraires de couplage pour passer de l'émission à l'immission de la pollution atmosphérique automobile : quelle intégration des données de simulation de SCAUP pour le calcul des concentrations de PAA ?**

C'est pourquoi, ensuite, il pourrait être plus intéressant d'exploiter plus directement un modèle numérique de climat comme WRF[209] ou MESO-NH[210]. Il est à noter qu'une première application de modélisation numérique du climat avec WRF a été réalisée à l'échelle de la ville de Dijon, permettant alors de bénéficier de données météorologiques (température, humidité, vent, rayonnement solaire…) à une haute résolution (150 m). Cette application a été initiée dans le cadre du projet MUSTARDijon[211] porté par l'ADEME Bourgogne, l'agglomération du Grand Dijon et le Centre de Recherche en Climatologie (CRC) de l'Université de Bourgogne (Pohl *et al.*, 2015), et auquel nous avons activement participé. Néanmoins, pour la modélisation de la pollution atmosphérique, le modèle numérique de climat n'est pas suffisant, et il est nécessaire de faire appel à d'autres modèles, en complément de celui-ci, comme illustré en figure 3.44 :

[208] Modèle de chimie-transport ou modèle CTM (Chimie Transport Model).
[209] Weather Research and Forecasting model (WRF) : http://www.wrf-model.org/index.php.
[210] MESOscale Non-Hydrostatic model (MESO-NH) : http://mesonh.aero.obs-mip.fr/mesonh52.
[211] Le projet MUSTARDijon porte sur la caractérisation de l'Ilôt de Chaleur Urbain (ICU), dont l'objectif vise à observer et mieux comprendre les différences de températures entre la ville et la campagne, par l'intermédiaire de stations d'observations météorologiques et d'un modèle numérique de climat régional (Pohl *et al.*, 2015).

- Un **modèle de canopée urbaine** permettant de prendre en compte les caractéristiques morphologiques de la ville. En effet, par rapport à un espace rural, la ville modifie sensiblement la circulation atmosphérique, ainsi que la dispersion des polluants, dans l'atmosphère. C'est pourquoi, pour l'échelle urbaine, il faut ajouter un modèle de canopée urbaine, permettant alors de prendre en compte la morphologie de la ville sous différentes formes, portant sur l'occupation du sol, la morphologie des bâtiments (hauteur et largeur) ou l'agencement des rues (orientation, hauteur, largeur…). Il est à noter que dans le cadre du projet MUSTARDijon, les développement réalisés ont permis de définir et d'intégrer les éléments de morphologie urbaine à l'aide du modèle de canopée urbaine BEP-BEM[212], couplé au modèle de climat régional WRF (Pohl *et al.*, 2015).

- Un **modèle de chimie-transport** qui permet de simuler la dispersion des polluants atmosphérique (Arteta, 2005 ; Michelot *et al.*, 2015). Ces modèles de chimie-transport ne produisent pas de données météorologiques, c'est pourquoi ils sont associés aux modèles de simulation numérique de climat[213]. On parle alors de modèle de qualité de l'air *offline*, car « *la météorologie est une donnée d'entrée qui n'est pas influencée par le calcul des concentrations* » (Michelot *et al.*, 2015, p. 92). Au même titre, pour que le modèle puisse estimer les niveaux de concentration, il est nécessaire d'intégrer au modèle de chimie-transport des données relatives aux émissions de polluants. Dans ce sens, le modèle de SMA du trafic routier fournit l'opportunité d'y intégrer les données d'émissions calculées préalablement. La modélisation de la qualité de l'air nécessite alors d'identifier les échelles spatiale et temporelle pertinentes aux objectifs visés. En effet, les modèles de chimie-transport présentent une formalisation physico-chimique très détaillée, et les gains en termes de résolution spatio-temporelle coûtent extrêmement cher en temps de simulation. C'est pourquoi, pour la modélisation de la qualité de l'air urbain, Nicolas Michelot *et al.* (2015) identifient comme pertinentes les échelles régionale et locale, et comme ils le mentionnent (p.94), « *à fine échelle, c'est le temps court qui est ciblé, c'est-à-dire de l'heure à quelques jours* ». À cette échelle, deux modèles de chimie-transport peuvent être identifiés, qui sont le modèle CHIMERE[214] et le modèle ADMS-Urban[215]. Ces deux modèles sont d'ailleurs utilisés par les AASQA pour la cartographie des concentrations de polluants (Rouïl, 2012). Alors que le premier permet une bonne représentation des échelles régionales (Martiny *et al.*, 2016), il est pour le moment limité pour la modélisation de la qualité de l'air à l'échelle intraurbaine du fait des données disponibles ainsi que de sa résolution spatio-temporelle (Sportisse, 2007b). À l'inverse, ADMS-Urban est un modèle dit « canyon », permettant alors une modélisation relativement fine de la pollution atmosphérique de proximité à l'échelle urbaine (Michelot, 2014). Contrairement à CHIMERE, ce dernier fournit l'opportunité d'intégrer directement les données de trafic routier simulé sans avoir à y

[212] Ces deux modèles complémentaires permettent pour BEP d'intégrer les perturbations associées aux bâtiments (murs et matière), ainsi que les turbulences associées à l'orientation des rues, tandis que BEM permet de prendre en compte les flux de chaleur latente émis par les logements (Kohler, 2015).

[213] Dans la plupart des applications, ce sont non pas des données de simulation, mais plutôt des données météorologiques observées et réanalysées, qui sont intégrées en entrée du modèle de chimie-transport, comme c'est le cas avec les données issues de ERA-Interim (http://www.ecmwf.int/en/research/climate-reanalysis/era-interim). Ainsi, le cheminement proposé ici se distingue de la démarche classique.

[214] Page de présentation de CHIMERE : http://www.lmd.polytechnique.fr/chimere/.

[215] Page de présentation de ADMS-Urban : http://www.cerc.co.uk/environmental-software/ADMS-Urban-RMI.html.

calculer les émissions de PAA. De surcroit, comme on peut l'observer sur la présentation des données en entrée de ADMS-Urban[216], celui-ci présente d'ailleurs des données d'entrée proches de celles nécessaires au fonctionnement de CIRCUL'AIR : profil temporel, trafic routier (*cf.* section 8.2, p. 271). Enfin, on peut aussi relever que ces deux modèles présentent l'opportunité d'être couplés avec le modèle numérique de climat WRF employé dans le cadre du projet MUSTARDijon (Pohl *et al.*, 2015).

Dans ce sens, les modèles numériques de climat, en combinaison d'un modèle de chimie-transport et d'un modèle de canopée urbaine, fournissent l'opportunité d'intégrer des données d'émissions de polluants calculées à partir du modèle SCAUP en entrée du modèle de chimie-transport. Cette perspective permettrait alors de qualifier l'ambiance atmosphérique à des échelles relativement fines, et dans une vision bien plus large que celle présentée dans cette thèse. Cependant, on comprend que le nombre de données pour la modélisation de la qualité de l'air *via* la simulation numérique du climat est très important et constitue déjà un enjeu non négligeable. Effectivement, comme le souligne Nicolas Michelot (2014, p. 306) : « *un logiciel de dispersion est habituellement composé de plusieurs modèles ou modules, plus ou moins performants, parfois optionnels, dont les performances peuvent dépendre de celles des autres modules. Ainsi, les performances dépendent sensiblement de la qualité des données d'entrée et de la configuration réalisée par l'utilisateur* ». Dans le même ordre d'idée, il souligne que plus la recherche d'une échelle spatio-temporelle est fine (locale), plus la pollution atmosphérique simulée aura tendance à être approximative.

Dès lors, le passage des émissions aux immissions demeure une thématique complexe, autant au niveau des données à insérer en entrée des différents modèles (climat, canopée urbaine et chimie-transport), qu'au niveau des échelles spatiales fines. L'ensemble de ces propos tend à nous montrer que la modélisation de la pollution atmosphérique à l'échelle urbaine demeure un enjeu de recherche d'actualité, comme nous le montre, par exemple, le workshop sur « *Le climat urbain : modélisation des échanges de chaleur et de polluants* »[217]. Un autre enjeu semble important, c'est celui de l'évaluation de la température en ville, comme l'Ilôt de Chaleur Urbain (ICU), car les enjeux liés à la mesure des concentrations de polluants à l'échelle urbaine y sont étroitement liés (Dab et Roussel, 2001 ; Cantat, 2004). Ainsi, les travaux sur l'ICU menés dans le cadre du projet MUSTARDijon (Pohl *et al.*, 2015) s'inscrivent aussi dans ces perspectives de recherches sur le climat urbain.

Pour finir, les données météorologiques modélisées dans le cadre du projet MUSTARDijon pourraient alors être exploitées dans deux itinéraires, comme nous le proposons au sein de la figure 3.44. Ce travail de recherche fournit effectivement l'opportunité de tester les apports du modèle de climat régional, non seulement pour la modélisation de la PAA par simulation multi-agents (*cheminement KISS*), mais aussi pour alimenter un module de chimie-transport (*cheminement KIDS*). Ces deux itinéraires sont complémentaires, car si la première démarche KISS est potentiellement applicable, pour la seconde de nombreux verrous restent à lever, car les modèles de chimie-transport ont une résolution spatiale qui demeure encore faible. En effet, plus l'échelle spatiale visée est fine plus le temps de paramétrisation, ainsi que le recours aux calculs haute

[216] Description des données en entrée et en sortie de ADMS-Urban : http://www.cerc.co.uk/environmental-software/ADMS-Urban-model/data.html.
[217] Présentations du workshop sur « Le climat urbain : modélisation des échanges de chaleur et de polluants » (Avril, 2015) : https://www6.inra.fr/basc/Recherche/Autres-projets-finances/MOTUS/WS-Climat-urbain-avril-15.

performance, deviennent croissants (Sportisse, 2007b). Dès lors, la modélisation de la qualité de l'air à l'échelle urbaine demeure un domaine de recherche d'actualité, qui n'en est qu'à ses débuts et dont la mise en place implique un temps de développement important, voire « *de remettre à plat l'ensemble de la chaine de modélisation* [des modèles de chimie-transport], *puis ensuite les procédures d'optimisation* » (Sportisse, 2007b, p. 66).

### 9.3.4. **Informer et sensibiliser le grand public par l'intermédiaire du *Géoweb***

Pour terminer, l'ensemble de ce travail vise aussi à s'inscrire non seulement dans le domaine de l'épidémiologie, mais aussi dans le domaine des *smart cities*, par la diffusion et la communication des valeurs estimées et simulées au grand public. Au-delà de la simple sensibilisation du grand public, comme nous avons pu le présenter pour la ville de Marrakech, on peut aussi s'insérer dans une perspective de diffusion (*monitoring*) des informations collectées et simulées en temps quasi-réel par SCAUP. En effet, en s'appuyant sur le caractère automatique des données de comptage, celles-ci peuvent alors être utilisées pour alimenter un outil de *Géoweb* en vue de diffuser une information en temps réel, sur les conditions de circulation routière, mais également sur les niveaux de pollution atmosphérique et sonore associés au trafic routier (*monitoring environnemental*).

Le terme de *Géoweb* désigne les outils en ligne combinant à la fois information géographique et données en ligne (Mericskay, 2011). Le développement des applications comme Google Maps, Google Earth ou Street View en sont une bonne illustration. Les données spatiales produites par Google, mais surtout par les utilisateurs via l'utilisation des nouvelles technologies, s'insèrent dans ce champ : « *Sensors everywhere : votre téléphone sait beaucoup de choses sur le monde qui vous entoure. Si vous prenez ces informations et que vous les combinez avec celles de tous les autres téléphones, nous avons une photographie instantanée de ce qui se passe dans le monde. Les bulletins météorologiques peuvent se baser non plus sur une centaine de capteurs, mais bien sur des centaines de millions de capteurs mobiles. Les informations sur la circulation routière peuvent être basées non plus sur des capteurs de trafic ou des observations, mais plutôt en s'appuyant sur la densité, la vitesse des téléphones (des individus) coincés dans les embouteillages* »[218]. Par ailleurs, les informations en temps réel sont aussi mobilisées comme des données historiques permettant de proposer aux utilisateurs des profils types de trafic routier, comme il est possible de le faire avec les données issues des capteurs électromagnétiques.

C'est pourquoi, à notre sens, les données issues de ces capteurs peuvent, elles aussi, être exploitées à cette fin de *Géoweb*. En effet, ces données automatiques sont mises à jour toutes les 15 minutes, permettant alors de disposer d'informations en temps réel et différé, qui sont collectées et centralisées par les gestionnaires du trafic routier. Elles peuvent alors être exploitées, comme illustré en figure 3.45, en proposant une application du même type, *via* un monitoring de ces données vers une interface *géoweb*. En effet, ces systèmes de capteurs fixes exploités par les CIGT[219] apparaissent bien plus fiables que les données mobiles issues des dispositifs comme Google Maps ou TomTom, dont on ne connait pas les données qui sont exploitées pour relever le trafic routier. À travers le modèle SCAUP, l'usage des comptages fournit alors l'opportunité d'apporter une alternative à ces services de *géoweb*, le modèle SCAUP permettant, quant à lui, d'avoir un processus intermédiaire, voire tampon, en vue de fournir une information simulée sur la circulation routière, ainsi que sur les niveaux de pollution atmosphérique, en temps quasi-réel.

[218] Andy Rubin, Google Official Blog : http://googleblog.blogspot.fr/2008/09/future-of-mobile.html.
[219] CIGT : Centres d'Ingénierie et de Gestion du Trafic (*cf.* Chapitre 3, p. 81)

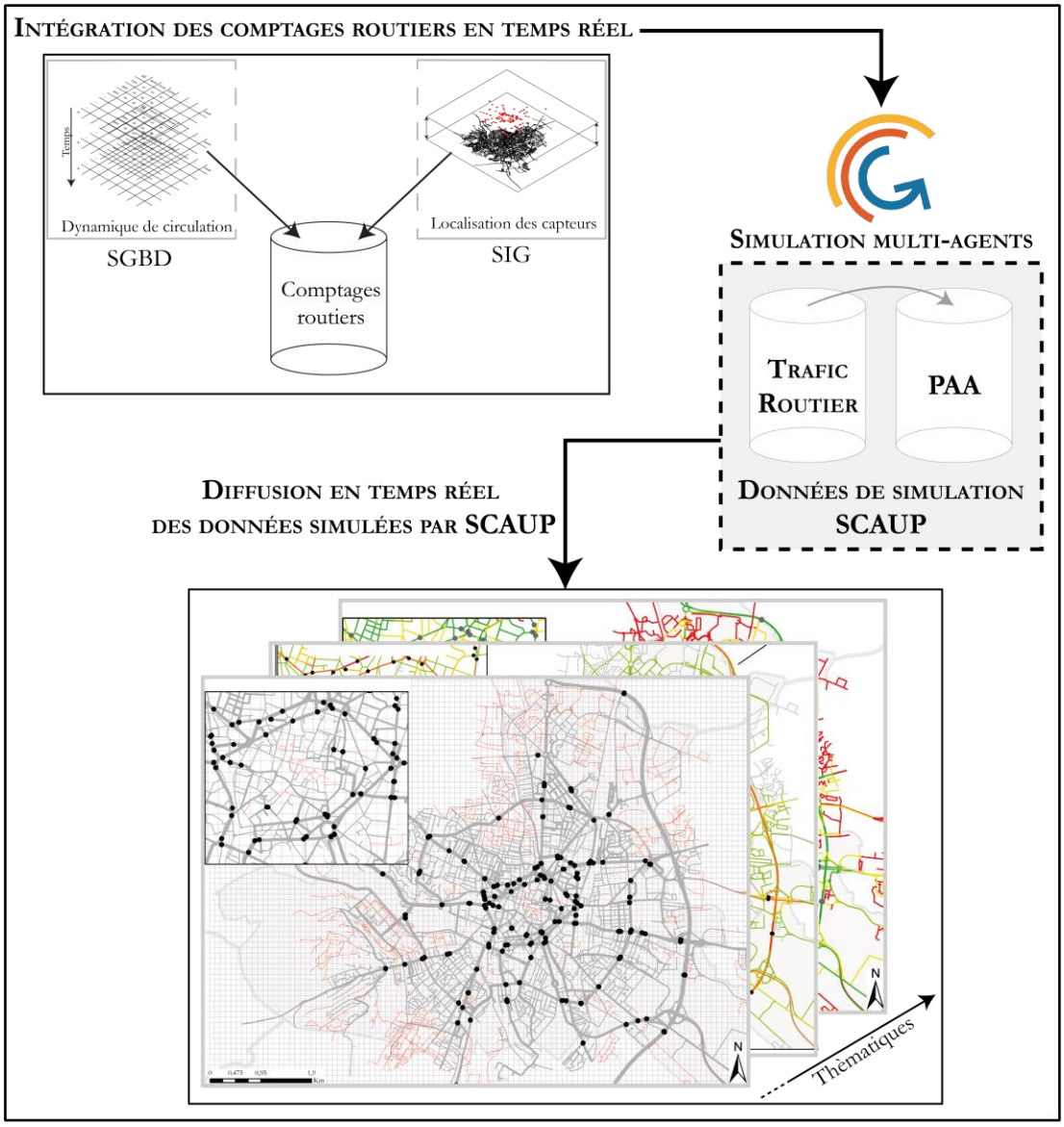



**Figure 3.45 : Un schéma d'intégration et de monitoring des données de comptages et de simulation sous une interface de Géoweb**

Outre ces aspects, c'est bien l'accessibilité aux données spatiales issues des systèmes de capteurs urbains (fixes ou mobiles) qui se démocratise, autant pour les experts que pour les usagers. Le monitoring en temps réel des données de comptages routiers, par l'intermédiaire du modèle SCAUP ainsi que d'une interface de *géoweb*, s'agence d'autant plus au sein des *smart cities*, en vue de produire et de diffuser une information à destination des usagers, mais aussi des collectivités territoriales, voire des AASQA, afin de fournir une meilleure connaissance sur la circulation routière, mais également sur les impacts environnementaux du trafic routier en milieu urbain.

Pour finir, ces différents développements, rattachés à la démarche de modélisation SCAUP, sont encore loin d'être mis en œuvre. Avant cela, il faut déjà aborder et tester une à une chacune des améliorations proposées. Au cours de ce chapitre, il est apparu important de revenir sur des éléments de discussions et de perspectives, portant autant sur le dispositif de quantification que sur celui d'évaluation en passant aussi par le dispositif de modélisation. Cette mise en perspective de la démarche de modélisation SCAUP initiée à travers ce travail ouvre de nombreuses portes.

Nos propos se sont aussi insérés dans une approche plus générale de l'observation et de la simulation des impacts environnementaux du trafic routier en milieu urbain. Ce travail de thèse ne présente pas une fin en soi, il n'est qu'une première pierre à un édifice en faveur de la ville durable et intelligente (*smart cities*). En effet, l'ensemble des perspectives abordées au cours de ce dernier chapitre s'inscrit dans ce sens, et ce d'autant plus que les données tiennent une place centrale dans la démarche mise en œuvre au cours de cette thèse.

Les enjeux autour de l'usage des capteurs urbains, de leur exploitation, mais également de leur diffusion, constituent des enjeux de recherche toujours aussi fertiles dans l'optique d'une évaluation environnementale toujours plus précise des impacts du trafic routier à l'échelle locale. Bien que certains des développements proposés au cours de ce dernier chapitre puissent apparaître ambitieux, il est avant tout important d'adopter une vision plus large de l'approche de modélisation SCAUP initiée au cours de cette thèse, en se rattachant aux différents enjeux actuels et futurs de la ville durable et intelligente, autant dans sa dimension environnementale que sociétale.

## CONCLUSION DE LA TROISIÈME PARTIE

Cette troisième partie nous a permis d'observer et de mettre en avant les choix méthodologiques ainsi que les enjeux empiriques de la démarche de modélisation SCAUP (*Simulation multi-agents à partir de CApteurs Urbains pour la Pollution atmosphérique automobile*) mise en œuvre à travers un cadre expérimental testé sur la ville de Dijon. Comme nous l'avons souligné à plusieurs reprises, c'est bien un processus de construction de la *connaissance sur les émissions de polluants atmosphérique automobile* qui a été mis en œuvre (*cf.* figure 1.5, p. 37). Par conséquent, au cours de cette dernière partie, le dispositif de quantification a été placé au centre de la démarche d'évaluation, que ce soit lors du passage de la *donnée* de comptage à une *information* sur la circulation routière de la ville de Dijon, ou lors du passage de cette *information* simulée à une *connaissance* sur les niveaux d'émissions de polluants atmosphériques automobiles à l'échelle de la ville de Dijon.

En parallèle, il faut aussi noter que l'outil informatique et les données géographiques tiennent une place prépondérante pour réaliser l'ensemble de cette démarche de modélisation, ces éléments caractérisant d'autant plus notre positionnement conceptuel, mais également méthodologique, au sein des outils de la *géocomputation* (*cf.* figure 1.10, p. 49). En effet, à travers la démarche de validation croisée (LOOCV) et l'utilisation du diagramme de Taylor (2001), nous avons tout d'abord tenté d'offrir un système d'évaluation et d'interprétation du modèle SCAUP qui soit en adéquation avec les défis qu'imposent la mise en œuvre d'une démarche guidée par les données : l'approche *deepening* (Hassan *et al.*, 2010a, 2010b). C'est-à-dire que les données interviennent à chaque étape du processus, tout en s'assurant de l'entière indépendance des échantillons utilisés à des fins de validation et de ceux utilisés à des fins de simulation.

Ensuite, le passage aux émissions de PAA, qui s'est appuyé sur cette démarche de *construction de la preuve* (Fouillé *et al.*, 2012), nous a permis de mettre en avant les atouts opérationnels de notre approche en vue de répondre aux besoins des AASQA locales lors de la mise en place des inventaires et des cadastres d'émissions de la pollution atmosphérique à l'échelle urbaine. Enfin, en nous appuyant sur l'expérience réalisée sur la ville de Dijon, nous avons pu relever différents éléments de discussion qui induisent différentes perspectives de développement de l'approche SCAUP. C'est pourquoi nous sommes revenu dans un dernier temps sur l'ensemble des perspectives de l'approche initiée dans la thèse, en nous arrêtant sur chacun des dispositifs de la matrice technique, ainsi que sur leurs relations (quantification, modélisation, évaluation). En effet, l'expérience réalisée sur la ville de Dijon ouvre la porte à un travail de recherche d'autant plus important et intéressant qu'il n'en est encore qu'à sa première phase.

De notre point de vue, la mise en œuvre d'une approche plus intégrée de la démarche SCAUP s'insère dans un contexte interdisciplinaire en vue d'offrir une vision plus complète des impacts environnementaux du trafic routier en milieu urbain, en y intégrant la pollution atmosphérique en combinaison des conditions météorologiques. Parallèlement, les potentiels d'exploitation des capteurs urbains, et plus particulièrement des capteurs électromagnétiques, dans le contexte de la directive INSPIRE et des *smart cities*, présentent l'opportunité de s'inscrire dans un système de communication et de diffusion via un monitoring web des données de comptage et des données qui sont ensuite simulées. On peut alors relever deux aspects de ces perspectives de développement, qui s'inscrivent dans un premier temps dans une recherche plutôt fondamentale

sur la place des données dans la construction d'une démarche scientifique ; puis dans un second temps dans une recherche appliquée, autant à destination des décideurs locaux et des AASQA pour l'élaboration des outils de planification urbaine, qu'à destination du grand public avec un caractère plus informatif sur les impacts environnementaux du trafic routier.

C'est pourquoi améliorer la connaissance (*cf.* figure 1.5, p. 37) sur les impacts du trafic routier, autant dans le temps que dans l'espace, s'avance également comme un enjeu politique, environnemental et social. En effet, l'évaluation des impacts environnementaux du trafic routier en milieu urbain se situe à l'interface de l'urbanisme, du transport, de la climatologie et de l'épidémiologie. Elle s'insère nécessairement dans un contexte pluridisciplinaire, pour une meilleure prise en compte des impacts environnementaux et sanitaires du trafic routier. Par ailleurs, les populations résidentes se saisissent de plus en plus de ces problématiques. En effet, l'enquête nationale sur les préoccupations environnementales des français en 2015 montre que « *la question du manque de transports en commun et les nuisances sonores sont les problèmes de proximité les plus fréquemment évoqués en premier. Même si la pollution de l'air ne semble pas être le principal souci rencontré par les enquêtés, elle s'impose toutefois comme un réel enjeu de préoccupation* » (CGDD, 2016, p. 2).

# CONCLUSION GÉNÉRALE

L'objectif central de cette thèse étant dédié à la proposition d'une démarche de modélisation du trafic routier en milieu urbain, nous souhaitons tout d'abord mettre en avant les éléments caractéristiques du modèle SCAUP (*Simulation multi-agents à partir de CApteurs Urbains pour la Pollution atmosphérique automobile*) proposée au cours de cette recherche doctorale :

- C'est une **démarche complète de construction de la connaissance**, allant du comptage routier aux émissions de polluants atmosphériques automobiles, en passant par la circulation routière simulée à l'échelle d'un territoire.
- Le modèle de simulation du trafic SCAUP est **capable de produire des tendances observées** au niveau des capteurs électromagnétique. Cependant, si la dynamique du trafic routier est très bien conservée et portée par le modèle, c'est plutôt au niveau de la répartition spatiale des véhicules simulés que le modèle SCAUP éprouve des difficultés.
- La **démarche de modélisation SCAUP est complémentaire** des démarches existantes pour le calcul des émissions de PAA (Pollution Atmosphérique Automobile). En effet, comme nous l'avons illustré, le modèle SCAUP a été conçu en partie comme une alternative pour l'estimation des flux de circulation routière, en se situant à l'interface des modèles de trafic routier *classiques* et des approche par comptages routiers.
- En nous rattachant au référentiel national des AASQA, comme nous l'avons illustré avec le logiciel de calcul des émissions CIRCUL'AIR, nous sommes en mesure d'alimenter des outils d'évaluation de la pollution atmosphérique à l'échelle locale. C'est donc une **réponse opérationnelle et adaptée** aux besoins des AASQA locales pour alimenter les inventaires et les cadastres d'émissions de la pollution atmosphérique.
- C'est une **démarche généralisable** et applicable à d'autres espaces urbains, comme nous avons pu le voir à travers l'exemple de la ville de Marrakech. De plus, l'intérêt de notre approche réside dans le fait qu'elle nécessite, par rapport à une approche à quatre étapes, très peu de données en entrée, ce qui rend d'autant plus facile son utilisation et son application à d'autres contextes urbains. Le peu de données à insérer en entrée du modèle SCAUP (comptages, capteurs et réseau routier) permet, en outre, d'éprouver et de tester la démarche sur un nombre important d'espaces urbains.

Ces quelques points caractérisent la démarche de modélisation SCAUP proposée au cours de cette thèse. À partir de ces points, la mise en œuvre de la démarche constitue un premier aboutissement, qui nous a permis de tester les potentiels d'exploitation des capteurs urbains pour la mesure d'autres phénomènes. C'est une première forme d'aboutissement de l'approche de modélisation SCAUP qui a été proposée dans cette thèse. Ce travail nous permet aujourd'hui d'ouvrir ses champs d'application vers différentes pistes de développement, au premier rang desquelles le couplage du modèle SCAUP à d'autres modèles comme COPERT et les modèles de climat nous apparait primordial.

Il apparait maintenant important de revenir sur le cheminement qui a conduit à l'initiation et à la réalisation de l'approche et du travail défendus ici. C'est pourquoi, nous proposons plutôt une discussion autour de plusieurs éléments clés qui ont permis de conduire ce travail.

## Une approche à l'intersection de deux géographies

Si la géographie est notre point de départ, on peut aussi relever au cours de ce travail notre goût pour une géographie se situant plus particulièrement à l'interface de la *géographie humaine* et de la *géographie physique*, tant par la thématique traitée ici que par les méthodes que nous avons utilisées. C'est aussi un reflet de notre formation, qui s'est inscrite depuis plusieurs années autant dans le champ de la géographie des transports que dans le champ de la climatologie. En confrontant ces deux composantes de la géographie (*physique et humaine*), qui, au premier abord, peuvent paraître foncièrement distinctes par les *objets* qui y sont étudiés, nous avons pu cependant remarquer que les termes de « *capteurs* », « *données* », « *analyses statistiques* », « *modèles* », « *simulations* » ou « *évaluations* » sont communs aux deux disciplines géographiques. Ainsi, l'instrumentation et les questions rattachées aux méthodes employées permettent de faire le pont entre ces deux géographies. Dans ce sens, pourquoi ne pas s'entourer des compétences des uns pour enrichir celles des autres ? Et ce d'autant plus que les deux disciplines partagent des *objets* en commun ; alors pourquoi ne pas s'entourer d'outils et de méthodes exploités depuis longtemps par les climatologues pour la réalisation d'un modèle de simulation multi-agents des dynamiques du trafic routier, et inversement ?

Pourtant, s'il demeure une distinction entre les deux disciplines, pour nous c'est bien au sein de la *géocomputation* que les deux disciplines se réunissent et se mélangent, atténuant alors la séparation entre ces deux géographies. C'est aussi dans ce sens que notre positionnement scientifique au sein de la *géocomputation* s'est progressivement inscrit et affirmé pendant ces dernières années. Effectivement, le climatologue s'inscrit aussi dans la *géocomputation*, par exemple, lors de la mise en œuvre des modèles numériques de climat, où il y manipule autant la géographie, que la statistique et l'informatique pour mettre en œuvre sa démarche. De plus, lors de l'application des modèles, la place de l'informatique est tout autant centrale en climatologie que pour la modélisation des dynamiques urbaines, et ce autant pour la simulation du modèle que pour l'analyse et *l'évaluation* des données produites.

La *géocomputation* est aussi un positionnement fédérateur, tant les deux disciplines géographiques peuvent s'enrichir l'une et l'autre. C'est typiquement le cas de l'exploitation que nous avons faite du diagramme de Taylor (2001), qui est à l'origine un outil dédié au modèle numérique de climat, et que nous avons considéré ici pour l'*évaluation* du modèle SCAUP. En outre, les questionnements entre *quantification*, *modèle* et *évaluation* constituent une colonne vertébrale commune aux deux disciplines. Par conséquent, nous avons pu remarquer au fil de notre expérience que la matrice technique, remise au goût du jour par Hadrien Commenges (2013), fournit un cadre de réflexion immuable à la démarche de modélisation (Chorley, 1964). Elle vise en effet à construire une connaissance sur un phénomène quelconque, qu'il soit géographique ou non. Si nous avons tenté de mettre en avant ces positionnements à chaque étape de notre réflexion, il semble que nous soyons parvenus ici à trouver un juste compromis pour nous positionner à l'interface de la géographie physique et humaine.

## La matrice technique : une colonne vertébrale pour la construction d'une connaissance scientifique ?

Comme nous l'avons vu, la connaissance scientifique quelle qu'elle soit s'inscrit dans une certaine subjectivité, où le modélisateur et le thématicien apportent à chaque étape du processus un regard de plus en plus personnel. Ce regard subjectif joue alors sur la manière avec laquelle une connaissance en est ensuite extraite par l'intermédiaire d'un modèle. Pour nous, cette subjectivité s'est foncièrement inscrite dans une vision du trafic routier sous un angle géographique, et non pas issue de la socio-économie des transports, qui aurait sans nul doute construit la démarche de simulation du trafic routier d'une manière bien différente de celle proposée ici. Or la vision de chacun joue forcément sur la manière avec laquelle un modèle est mis en œuvre. Ainsi, la notion de *matrice technique*, et plus spécifiquement les dispositifs mis en relation, montrent que notre démarche ne s'éloigne pas de la conception *classique* de la démarche de modélisation, car comme le relève Hadrien Commenges (2013, p. 82), « *c'est tout l'intérêt de considérer qu'il existe une matrice technique constituée de trois dispositifs et que c'est au sein de cette matrice, et non au sein de chaque dispositif pris isolément, que peuvent être compris les principaux objets et méthodes de la socio-économie des transports. En effet, pour que ces relations existent il faut des objets communs qui puissent circuler d'un dispositif à l'autre* ».

Dès lors, la modification de l'un des dispositifs techniques, que ce soit celui de quantification, de modélisation ou d'évaluation, implique de repenser chacun d'entre eux sous une toute autre forme, modifiant alors l'ensemble de la matrice technique. C'est ce que nous avons tenté de faire, lors de la mise en œuvre de la démarche de modélisation SCAUP. En effet, deux dispositifs ont été modifiés : le premier tient plutôt compte de la finalité de notre approche, qui s'inscrit dans l'optique d'une *évaluation* non plus économique, mais plutôt environnementale du trafic routier à travers la Pollution Atmosphérique Automobile (PAA).

La seconde modification se rattache à la *quantification*, qui est construite non plus en s'appuyant sur des enquêtes Origine-Destination (O-D), mais en exploitant des données de comptage issues de systèmes de capteurs fixes. En outre, à travers la littérature exploitée, il n'existe à notre connaissance aucun modèle de trafic routier qui se construise en dehors de la mobilisation des matrices O-D. Par conséquent, les modèles de trafic routier commerciaux, qu'ils soient statiques, dynamiques, microscopiques ou macroscopiques, ne répondent pas à nos besoins. Cet état des lieux nous a alors amené à proposer une toute autre approche de simulation du trafic routier, visant autant à rapprocher la quantification de ce qui est mesuré, qu'à s'inscrire dans la continuité des développements actuels, en nous reposant sur la simulation multi-agents (SMA), et en nous inscrivant dans l'optique d'une approche guidée par les données : l'approche *deepening* (Hassan *et al.*, 2010a, 2010b).

## Le géographe, l'informaticien et le statisticien : un ancrage dans l'interdisciplinarité

Au même titre que le biologiste réalise ses expériences par l'intermédiaire des éprouvettes, aujourd'hui, à travers la modélisation, et plus particulièrement les plates-formes de SMA, le géographe se dote de son propre laboratoire, au sein duquel il peut tester ses hypothèses et en observer le fonctionnement en toute autonomie. Ces tendances sont aussi le reflet de la démocratisation des outils informatiques depuis les années 1990 dans notre discipline. On

remarque par ailleurs que la modélisation de l'espace tient aujourd'hui une place prépondérante, que ce soit pour l'étude des dynamiques urbaines, l'analyse spatiale ou plus généralement pour la géographie (Batty *et al.*, 1998 ; Banos, 2013 ; Sanders, 2001b ; Daudé et Langlois, 2006). Malgré tout, force est de constater que la modélisation en géographie ne peut être réalisée de manière hermétique à notre discipline, et cette démarche s'inscrit d'autant plus dans une vision transversale de l'environnement urbain, et de ses *objets* où de nombreuses disciplines interagissent (géographie, sociologie, économie, climatologie, biologie, physique, mathématique…), peuvent discuter, partager et échanger sur leurs visions et leurs interprétations des *objets* qui constituent la ville. De notre point de vue, l'étude des impacts environnementaux du trafic routier en milieu urbain doit s'inscrire dans cet ancrage interdisciplinaire. Comme le relève Hélène Mathian et Lena Sanders (2014, p. 63) : « *si certains se spécialisent dans un domaine ou l'autre, c'est par la connaissance des enjeux associés à chacun que l'on peut mettre en place une chaîne de traitement cohérente* ». En effet, dans le contexte actuel, la construction de connaissance quelle qu'elle soit doit s'inscrire dans un contexte interdisciplinaire où les échanges, par exemple, entre le thématicien et l'informaticien permettent d'enrichir la manière avec laquelle nous construisons une *connaissance*.

L'inscription de ce travail dans la *géocomputation* ainsi que dans les méthodes axées sur les données (*data-driven*) a nécessité un élargissement des compétences du géographe, en allant emprunter des outils qui se rattachent à d'autres disciplines que la nôtre, comme les plates-formes de SMA. C'est pourquoi ce travail s'inscrit avant tout dans l'interdisciplinarité, ce qui nous a permis d'entrevoir que la recherche géographique ne peut se construire par une seule personne pour aboutir à une approche complète de l'espace urbain. Cet ancrage a été enrichissant, et il guide aujourd'hui la vision que nous avons de la géographie à travers ces méthodes, ces outils, ainsi que les expériences mises en place au cours de cette thèse. Si aujourd'hui la place de l'informatique prend de plus en plus importance dans notre discipline, cette tendance ne constitue pas non plus une fin en soi. Pourtant, au cours de ce travail, la délimitation entre la géographie et l'informatique est parfois floue, tant nous avons placé (au même titre que la donnée) l'outil au centre de notre démarche et de nos propos. Il faut alors bien noter que le géographe n'est pas seul, même si « *les pratiques de modélisation et de simulation en géographie et en sciences humaines et sociales se libèrent* » (Banos, 2013). Pour nous, la géographie se construit et s'affirme d'autant plus dans un contexte (*pluri*)interdisciplinaire[220], et sa place tient plutôt à la manière avec laquelle nous en construisons une connaissance sur le fonctionnement de l'espace urbain et de ses dynamiques, qui rend l'intérêt pour notre discipline d'autant plus important : comment formaliser les dynamiques urbaines sous la forme d'un système complexe à travers le trafic routier ? Comment interpréter et analyser les résultats issus de la simulation du trafic routier ? Comment qualifier la simulation du trafic routier ?

À notre sens, le géographe prend de plus en plus part à ces débats, et il permet aujourd'hui de positionner sa discipline dans un contexte de recherche fédérateur, où chacun peut prendre sa place. C'est ce que nous avons tenté de faire transparaitre, à notre niveau et tout au long de cette thèse, en nous inscrivant dans une démarche de modélisation fédérative associant géographie, informatique, climatologie et statistique. C'est aussi un travail de thèse qui se positionne autant

[220] Pour nous, des événements comme l'école d'été de l'Institut des Systèmes Complexes d'Île-de-France (ISC-PIF), ou l'école thématique MAPS (Modélisation multi-agents Appliquée aux Phénomènes Spatialisés) ont permis de comprendre l'importance et les enjeux de recherche de cette interdisciplinarité.

dans une recherche fondamentale, lorsque nous avons abordé les thématiques des méthodes axées sur les données, que dans une recherche appliquée à destination des collectivités territoriales et des Associations Agréées de Surveillance de la Qualité de l'Air (AASQA).

Enfin dans un contexte d'interdisciplinarité, la démarche traditionnelle de constitution des modèles sous leur forme *input-output* n'est plus suffisante. Il faut alors être en mesure de décrire l'ensemble des *boîtes noires* pour éviter une certaine défiance envers le modèle qui a été construit. C'est pourquoi nous avons dans la mesure du possible mis en avant et explicité l'ensemble du fonctionnement, autant interne qu'externe, du simulateur du trafic routier. L'intégration d'un volet explicatif des modèles en géographie est aujourd'hui incontournable, et des langages ou des formalismes permettent à notre sens d'éclairer l'ensemble des hypothèses et des paramètres d'un modèle sans faire nécessairement appel au formalisme mathématique. C'est, par exemple, le cas du formalisme Unified Modelling Language (UML) qui permet non seulement de formaliser le modèle par l'intermédiaire des diagrammes de classes et d'actions, mais qui permet aussi de faire le pont entre les différents acteurs participant à la constitution d'un modèle, en vue de comprendre, mais aussi de faire comprendre, la démarche initiée. Il nous apparait important aujourd'hui de nous inscrire dans cette optique d'ouverture des *boîtes noires* et de lisibilité des modèles. C'est ce que nous avons tenté de faire en nous situant dans une description complète de notre approche, ainsi qu'en tentant de rapprocher l'observation de ce qui est modélisé et des transformations qui s'ensuivent à travers le cadre de réflexion : données → informations → connaissances.

Enfin, à travers les comptages routiers, le rapprochement que nous avons tenté de faire entre ce qui est mesuré et ce qui est ensuite simulé nécessite un formalisme d'autant plus important pour comprendre et faire comprendre les atouts de cet ancrage empirique des modèles de simulation. En effet, les questionnements autour des données issues des capteurs et de leur exploitation constituent aujourd'hui des problématiques d'autant plus importantes pour la manière avec laquelle cette *réalité* est ensuite communiquée, interprétée et exploitée par les acteurs du territoire, que ce soit à l'échelle locale ou à l'échelle nationale.

## De l'*homme moyen* à l'*homme capteur* : vers une "gouvernance" du capteur ?

*Capteurs, big data, smart cities* ont été des termes très souvent évoqués dans le fond de nos propos. Ce sont des concepts qui demeurent émergents et intéressants, autant pour les chercheurs et les décideurs publics que pour les usagers. À notre sens, ces termes s'insèrent dans un questionnement de fond plus large qui vise à rapprocher la mesure de la décision publique. Pour nous, la construction d'une approche ancrée dans les données a permis de mettre en avant les atouts d'un rapprochement des *objets* mesurés, pour en simuler ensuite ces mêmes *objets*. Dans le fond, ce sont des questionnements qui portent sur la place que tiennent et tiendront les données issues des capteurs dans la quantification et la planification de l'espace urbain. Nous avons défendu l'idée que ces données peuvent mesurer des *choses* qui jusqu'alors ne sont que peu mesurables à travers les enquêtes classiques. Par ailleurs, nous avons réussi à montrer qu'à travers l'utilisation d'un système de capteurs fixes, il est possible de produire une connaissance sur les impacts environnementaux du trafic routier (PAA) en utilisant exclusivement ce système de mesure. C'est pourquoi, il est important de s'arrêter sur les modifications que l'avènement des

capteurs entrainera sur les outils de planification urbaine : quelle place le capteur tiendra-t-il dans la décision publique et dans les mesures (au sens politique) qui en ressortiront ? En quoi l'utilisation des capteurs va modifier l'instrumentalisation de l'espace urbain ?

Ces derniers questionnements plus généraux sur la place que tient la donnée dans la gouvernance actuelle des territoires nous paraissent importants. En effet, comme nous l'avons vu, l'observation des territoires, qui a été avant tout institutionnelle, apparait se situer aujourd'hui dans une phase de désinstitutionalisation (Manovich, 2011). En effet, les observations sont aujourd'hui de plus en plus souvent collectées et entre les mains de grandes entreprises, comme Google par exemple, qui par l'utilisation des *smartphones Android* dispose d'un potentiel énorme de capteurs sous la forme d'usagers (*sensors everywhere*). Pourtant, à travers la *smart city*, on a tendance à percevoir une volonté de (ré)institutionnalisation des mesures issues des systèmes de capteurs. Effectivement, si la « *politique des grands nombres* » (Desrosières, 2000) est une tendance relativement bien acceptée, avec la *smart city* nous nous dirigeons plutôt vers une *institutionnalisation et une gouvernance du capteur* dans la planification urbaine. Ces questionnements rejoignent les querelles autour du réalisme et de la nature des techniques permettant de quantifier et de mesurer les comportements humains, qui ont eu tendance à cristalliser les débats pendant plusieurs années. Dans son ouvrage, Alain Desrosières souligne ces points, ainsi que la prépondérance que prend la mesure, ou l'*objet* en lui-même, dans le champ de l'action publique, pour décrire le monde et agir sur lui à travers des êtres crées sous la forme d'un « *homme moyen* »[221] (Desrosières, 2000, p. 88). S'il montre qu'avec des chiffres, on peut *tout leur faires dire* selon la manière avec laquelle on les agence, structure et transforme pour en produire une connaissance quelconque, il montre aussi l'impact que cette instrumentation a sur l'ensemble de la décision publique locale : « *les questions d'hygiène publique, d'épidémiologie, de délinquance, impliquent des mesures (au sens des décisions) administratives et politiques, dont la discussion et la justification peuvent prendre appui sur des mesures (au sens de quantifications) des fins à atteindre et des moyens qui y sont consacrés. L'alchimie qui transforme des actes individuels libres et aléatoires en agrégats déterminés et stables fournit au débat des points de référence, des objets transmissibles, parce qu'extérieurs aux personnes. Elle constitue le cœur de l'instrumentation statistique d'un espace public* ». (Desrosières, 2000, p. 401).

Pour actualiser les propos d'Alain Desrosières, la généralisation des capteurs tend à produire de nouvelles mesures qui ne décrivent plus un *homme moyen*, mais plutôt un *homme capteur*, branché sous électrodes 24 h/24. Dès lors, la question de l'*exhaustivité* ne constitue plus tellement une contrainte, comme cela peut être le cas avec les enquêtes. Et les questionnements glissent non seulement vers ces problèmes d'*exhaustivité* des données, mais aussi vers une institutionnalisation de ces capteurs (protection privée, accessibilité, diffusion…). Ainsi, la recherche sur les dynamiques urbaines s'enrichit de ces systèmes de mesures, qui permettent d'observer des tendances qui jusqu'alors n'étaient que peu quantifiables par l'intermédiaire des enquêtes. Derrière cette quête de l'exhaustivité à tout crin, les pouvoirs publics sauront-ils garder la main sur ces données ?

[221] À l'époque où Alain Desrosières écrit ces propos dans sa première édition en 1993 (la deuxième édition date de 2000), les modèles et leurs résultats sont déjà fortement ancrés et institutionnalisés dans la décision publique, comme le montrent les ouvrages sur l'évolution de la modélisation des transports en France de Konstantinos Chatzis (2009) ou de Gilles Debizet (2011).

Pour terminer, les capteurs vont également modifier profondément l'instrumentation de la décision publique, comme par exemple, la mise en place des *smart cities*, qui visent à placer la mesure au cœur de l'information publique en milieu urbain. Cependant la *smart city* demeure encore une tendance marginale et émergente, qui n'en est encore qu'à ses balbutiements, et on constate qu'aucune définition n'en est encore bien délimitée (Languillon-Aussel *et al.*, 2016). Pourtant, l'ensemble des services urbains et de leur instrumentation sera concerné par l'évolution des *smart cities*. L'inscription de la donnée de capteur dans la décision publique fournit une toute autre réponse à l'instrumentation de l'espace public, qui, en réponse aux propos d'Alain Desrosières, fournit l'opportunité de modifier l'instrumentation statistique de l'espace public, autant pour la mesure du trafic routier que pour son évaluation environnementale. La place du capteur, l'exploitation de ses données dans le contexte du *big data* et de la *smart city*, sont autant de portes ouvertes qui permettront de continuer et d'affirmer le travail présenté dans cette thèse, et auquel le géographe devra sans nul doute prendre part…

# BIBLIOGRAPHIE


AEE, 2014. Air quality in Europe - 2014 report (Rapport de l'Agence Européenne pour l'environnement No. 5/2014). Copenhague.

AEE, 2011. COPERT 4 : Estimating emissions from road transport (Publication). European Environment Agency, Copenhague.

Aguiton, C., Cardon, D., Smoreda, Z., 2009. Living Maps: New data, new uses, new problems, in: Engaging Data.First International Forum on the Application and Management of Personal Electronic Information. Cambridge, Massachussets.

Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E., 2002. A survey on sensor networks. IEEE Commun. Mag. 40, 102–114. doi:10.1109/MCOM.2002.1024422

Amblard, F., Daudé, E., Gaudou, B., Grignard, A., Hutzler, G., Lang, C., Nicod, J.-M., Sheeren, D.D., Taillandier, P., 2015. Introduction à NetLogo, in: Simulation Spatiale À Base D'agents Avec NetLogo, Partie 1, Systèmes D'information, Web et Informatique Ubiquitaire. ISTE, pp. 73–112.

Amblard, F., Phan, D., 2006. Modélisation et Simulation Multi-agents Applications pour les Sciences de l'Homme et de la Société. Hermes Science Publications, Paris;

Amblard, F., Rouchier, J., Bommel, P., Varenne, F., Phan, D., 2006. Evaluation et validation de modèles multi-agents, in: Modélisation et Simulation Multi-Agents : Applications Pour Les Sciences de l'Homme et de La Société, Lavoisier. Paris, pp. 103–140.

Anderson, C., 2008. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. WIRED.

André, M., Roche, A.-L., Bourcier, L., 2013. Statistiques de parcs et trafic pour le calcul des émissions de polluants des transports routiers en France (Rapport IFSTTAR-LTE). IFSTTAR-LTE.

Andrienko, G., Andrienko, N., 1999. Interactive maps for visual data exploration. Int. J. Geogr. Inf. Sci. 13, 355–374. doi:10.1080/136588199241247

Andrienko, G., Andrienko, N., Rinzivillo, S., Nanni, M., Pedreschi, D., Giannotti, F., 2009. Interactive visual clustering of large collections of trajectories, in: IEEE Symposium on Visual Analytics Science and Technology, 2009. VAST 2009. Presented at the IEEE Symposium on Visual Analytics Science and Technology, 2009. VAST 2009, pp. 3–10. doi:10.1109/VAST.2009.5332584

Anselin, L., 1989. What is special about spatial data?: Alternative perspectives on spatial data analysis. Presented at the Spatial Statistics, Past, Present and Future, National Center for Geographic Information and Analysis, Syracuse University.

Antoni, J.-P., 2010. Modéliser la ville - Formes urbaines et politiques de transport. Economica, Paris.

Antoni, J.-P., 2006. Calibrer un modèle d'évolution de l'occupation du sol urbain. L'exemple de Belfort. Cybergeo Eur. J. Geogr. doi:10.4000/cybergeo.2436

Antoni, J.-P., Vuidel, G., 2010. MobiSim : un modèle multi-agents et multi-scalaire pour simuler les mobilités urbaines, in: Antoni, J.-P. (Ed.), Modéliser La Ville. Forme Urbaine et Politiques de Transport. Economica, Méthodes et approches, pp. 50–77.

Antoniou, C., Barcelo, J., Brackstone, M., Celikoglu, H., Ciuffo, B., Punzo, V., Sykes, P., Toledo, T., Vortisch, P., Wagner, P., 2014. Traffic Simulation: Case for guidelines [WWW Document]. URL http://publications.jrc.ec.europa.eu/repository/handle/111111111/30680 (accessed 6.16.16).

Appert, M., Chapelon, L., 2003. Variabilité de la performance des résaeux routiers : application à

la région urbaine de Londres, in: Graphes et réseaux. Modélisation multiniveau. Hermes Science Publications, Paris, pp. 51–76.

Appert, M., Chapelon, L., 2002. Planification des transports régionaux en Languedoc-Roussillon et Nord-Pas-de-Calais : Analyse comparée des chaînes de transport à dominante routière et ferroviaire (Rapport d'études), INRETS - Groupement Régional Nord-Pas-de-Calais pour la recherche dans les tranports. UMR 6012 ESPACE, Montpellier.

Arrêté du 24 août 2011 relatif au SNIEBA, 2011. , Journal officiel de la république française.

Arrêté interministériel du 26 Mars 2014 relatif au déclenchement des procédures préfectorales en cas d'épisodes de pollution de l'air ambiant, 2014. , Journal officiel de la république française.

Arteta, J., 2005. Etude de l'impact du mécanisme chimique et des taux de photolyse "online" sur les simulations 3D d'épisodes de pollution régionale (Physique de l'Atmosphère). Clermont-Ferrand 2, Clermont-Ferrand.

Article L411-6 Instruction interministérielle sur la signalisation routière, 1991. , Code de la route.

ASPA, 2012. Réalisation de l'inventaire des émissions sur l'Eurodistrict Strasbourg-Ortenau (Rapport d'études). ASPA : Association pour la Surveillance et l'Etude de la Pollution Atmosphérique en Alsace, Strasbourg.

ASPA, 2005. Méthodologie de calcul des émissions routières : Application à la Lorraine (Rapport Technique). ASPA : Association pour la Surveillance et l'Etude de la Pollution Atmosphérique en Alsace, Starsbourg.

Audard, F., Carpentier, S., Oliveau, S., 2014. Les « big data » sont-elles l'avenir de la géographie [théorique et quantitative] ? Presented at the 20e Biennale de géographie d'Avignon - Géopoint 2014 - « Controverses et géographies », pp. 1–4.

Aw, T., 2010. La ville nouvelle de Marne-la-Vallée et son insertion dans la dynamique francilienne : évaluation des enjeux du renforcement de la structure polycentrique sur les systèmes de déplacements. Paris Est.

Badariotti, D., Weber, C., 2002. La mobilité résidentielle en ville. Modélisation par automates cellulaires et système multi-agents à Bogota. L'Espace Géographique tome 31, 97–108.

Bailly, A., Ferras, R., 2010. Eléments d'épistémologie de la géographie, 2e édition. ed. Armand Colin, Paris.

Bakis, H., Valentin, J., 2010. Amateurisme cartographique et géographique à l'heure du web 2.0. Netcom Réseaux Commun. Territ. 109–132. doi:10.4000/netcom.556

Balmer, M., 2007. Travel demand modeling for multi-agent traffic simulations: Algorithms and systems. Université de Zurich, Zurich.

Balmer, M., Nagel, K., Raney, B., 2004. Large-Scale Multi-Agent Simulations for Transportation Applications. J. Intell. Transp. Syst. 8, 205–221. doi:10.1080/15472450490523892

Banos, A., 2013. Pour des pratiques de modèlisation et de simulation libérées en géographie et SHS (thesis). Université Paris 1 Panthéon Sorbonne.

Banos, A., 2006. Des trajectoires individuelles à la ville en mouvement : reconstruire les mobilités urbaines quotidiennes, in: Actes Des 13ièmes Journées de Rochebrune. pp. 289–300.

Banos, A., 2005. La voie de l'étonnement : favoriser l'abduction dans les Systèmes d'Information Géographique, in: Apport Des SIG À La Recherche. Orléans, pp. 237–254.

Banos, A., Boffet-Mas, A., Chardonnel, S., Lang, C., Marilleau, N., Thévenin, T., 2010. Simuler la mobilité urbaine quotidienne : le projet MIRO XX.

Banos, A., Caillou, P., Gaudou, B., Marilleau, N., 2015. Exploration de modèles agent, in: Banos, A., Lang, C., Marilleau, N. (Eds.), Simulation spatiale à base d'agents avec NetLogo 1 : introduction et bases. ISTE, Londres, pp. 113–162.

Banos, A., Sanders, L., 2012. Vers une cartographie sémantique des modèles urbains : des individus aux systèmes de villes, in: La Modélisation de La Ville : Du Modèle Au Projet Urbain, RéférenceS de La DRI. Paris, pp. 52–64.

Banos, A., Thévenin, T., 2008. Création de champs de potentiel et simulation d'itinéraires à partir de l'enquête ménages-déplacement, in: Information géographique et dynamiques urbaines 1 - Analyse et simulation de la mobilité des personnes. Hermes science : Lavoisier (Traité IGAT), pp. 119–138.

Banos, A., Thévenin, T., 2005. Révéler les rythmes urbains quotidiens par la carte animée. Revue Internationale de Géomatique 15, pp.11-31. doi:10.3166/rig.15.11-31

Barceló, J., 2010. Models, Traffic Models, Simulation, and Traffic Simulation, in: Barceló, J. (Ed.), Fundamentals of Traffic Simulation, International Series in Operations Research & Management Science. Springer New York, pp. 1–62.

Barles, S., 2008. Comprendre et maîtriser le métabolisme urbain et l'empreinte environnementale des villes. Ann. Mines - Responsab. Environ. 52, 21. doi:10.3917/re.052.0021

Batty, M., 2013. Big data, smart cities and city planning. Dialogues Hum. Geogr. 3, 274–279. doi:10.1177/2043820613513390

Batty, M., 2012. Smart cities, big data. Environ. Plan. B Plan. Des. 39, 191 – 193. doi:10.1068/b3902ed

Batty, M., 2007. Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals. The MIT Press.

Batty, M., Dodge, M., Doyle, S., Hudson-Smith, A., 1998. Modelling virtual urban environments [WWW Document]. CASA Work. Pap. Cent. Adv. Spat. Anal. UCL Lond. UK 1998. URL http://discovery.ucl.ac.uk/219/ (accessed 7.1.15).

Bauzer-Medeiros, C., Carles, O., Vuyst, F.D., Hugueney, B., Joliveau, M., Jomier, G., Manouvrier, M., Naija, Y., Scémama, G., Steffan, L., 2006. Vers un entrepôt de données pour le trafic routier. Presented at the Entrepôt de données et Analyse en Ligne EDA'06.

Bavoux, J.-J., Beaucire, F., Chapelon, L., Zembri, P., 2005. Géographie des transports. Armand Colin, Paris.

Beckman, R.J., Baggerly, K.A., McKay, M.D., 1996. Creating synthetic baseline populations. Transp. Res. Part Policy Pract. 30, 415–429. doi:10.1016/0965-8564(96)00004-3

Behrisch, M., Bieker, L., Erdmann, J., Krajzewicz, D., 2011. SUMO - Simulation of Urban MObility - an Overview. Presented at the SIMUL 2011, The Third International Conference on Advances in System Simulation, pp. 55–60.

Benenson, I., Torrens, P.M., 2004. Geosimulation: object-based modeling of urban phenomena. Comput. Environ. Urban Syst. 28, 1–8.

Berchtold, A., 1998. Chaine de Markov et modèles de transition : Applications aux sciences sociales. Hermes Science Publications.

Besnier, J.-M., 2011. Anatomie des thèories, in: Les thèories de la connaissance, Que sais-je ? Presses Universitaires de France, p. 128.

Biernat, E., Lutz, M., LeCun, Y., 2015. Data science : fondamentaux et études de cas : Machine learning avec Python et R. Eyrolles, Paris.

Blanchard, M., Borelli, D., Chardon, B., Declercq, C., Fabre, P., Host, S., Jusot, J.-F., Larrieu, S., Lefranc, A., Medina, S., Pascal, L., Wagner, V., 2008. Évaluation de l'impact sanitaire de la pollution atmosphérique urbaine : Concepts et méthodes. Extrapol Pollut. Atmos. Santé, Institut de veille sanitaire.

Bommel, P., 2009. Definition d'un cadre methodologique pour la conception de modeles multi-agents adaptee a la gestion des ressources renouvelables (These). These de doctorat -- Informatique.

Bonnafous, A., 1990. Le siècle des ténèbres de l'économie. Economica, Paris.

Bonnel, P., 2010. Estimer la demande de transport, in: Système de Transport Urbain : Caractérisation de L'offre et Estimation de La Demande, Traité de l'IGAT. Hermes science : Lavoisier (Traité IGAT), Paris, pp. 109–144.

Bonnel, P., 2002. Prévision de la demande de transport (HDR). Université Lumière - Lyon II, Lyon II.

Bonnin, P., Douady, S., 2013. Analyse et modélisation de la morphogenèse du réseau des rues, in: Les réseaux dans le temps et dans l'espace. Presented at the ctes de la deuxième journée d'étude du groupe fmr, Paris, pp. 47–73.

Boquet, Y., Desse, R.-P., Collectif, 2010. Commerces et mobilités. Editions Universitaires de Dijon, Paris.

Borsu, M., 2004. Les apports des nouvelles technologies pour le recueil des données de trafic (Rapport de fin d'étude pour le CERTU). ENTPE, Paris.

Bouju, A., Bertrand, F., Malle-Noyon, V., Servigne, S., Devogele, T., Ray, C., Martin, H., Gensel, J., 2009. Gestion de données spatio-temporelles au sein de base de données capteurs. Presented at the COGIST 09, Première Conférence Francophone sur les Technologies de l'Information, de la Communication et de la Géolocalisation dans les Systèmes de Transports, St Quay Portrieux (Côtes d'Armor).

Bouquet, F., Chipeaux, S., Lang, C., Marilleau, N., Nicod, J.-M., Taillandier, P., 2015. Chapitre 1 : Introduction à l'approche agent, in: Simulation Spatiale À Base D'agents Avec NetLogo 1. ISTE, London, pp. 17–38.

Bourdeau, B., 1997. Evolution du parc automobile francais entre 1970 et 2020. Chambéry.

Bourrel, E., 2003. Modélisation dynamique de l'écoulement du trafic routier: du macroscopique au microscopique (Thèse doctorat). Institut national des sciences appliquées, Lyon, France.

Bréheret, L., Schettini, F., Bernauer, E., Barbier, M., Centre d'études sur les réseaux, les transports, SODIT, 2000. Traitements de données de trafic : besoins, état de l'art, exemple de mise en oeuvre [WWW Document]. URL http://lara.inist.fr/handle/2332/925 (accessed 5.28.14).

Briassoulis, H., 2000. Analysis of Land Use Change: Theoretical and Modeling Approaches. Regional Research Institute, West Virginia University.

Brunet, R., Ferras, R., Théry, H., 2005. Les mots de la géographie : Dictionnaire critique, Édition : 3e édition revue et augmentée. ed. La Documentation Française, Montpellier; Paris.

Brunsdon, C., 2014. Spatial science – Looking outward. Dialogues Hum. Geogr. 4, 45–49. doi:10.1177/2043820614525709

Brunsdon, C., Singleton, A.D. (Eds.), 2015. Geocomputation: A Practical Primer. SAGE Publications Ltd, Thousand Oaks, CA.

Buard, É., 2013. Dynamiques des interactions espèces - espace : mise en relation des pratiques de déplacement des populations d'herbivores et de l'évolution de l'occupation du sol dans le parc de Hwange (Zimbabwe). Paris 1.

Buguellou, J.-B., 2012. Micro-simulation des déplacements par système multi-agents : exploration multi-niveaux. Tours.

Buhler, T., 2012. Eléments pour la prise en compte de l'habitude dans les pratiques de déplacements urbains : Le cas des résistances aux injonctions au changement de mode de déplacement sur l'agglomération lyonnaise (phdthesis). INSA de Lyon.

Buisson, C., 1996. Analyse bibliographique et première classification des algorithmes de simulation de l'écoulement du trafic routier. Rech. Transp. Sécurité 35–48.

Buisson, C., Lesort, J.-B., 2010. Comprendre le trafic routier : Méthodes et calculs. CERTU, Lyon.

Cabrera Delgado, J., Bonnel, P., 2012. Quelle dynamique temporelle pour le modèle à 4 étapes ? (Rapport Final). Laboratoire d'Economie des Tranports (LET), Lyon.

Cambien, A., 2010. La modélisation urbaine : une approche historique, in: Modéliser la ville - Formes urbaines et politiques de transport. Economica, Paris, pp. 19–49.

Cantat, O., 2004. L'îlot de chaleur urbain parisien selon les types de temps. Norois Environ. Aménage. Société 75–102. doi:10.4000/norois.1373

Carteret, M., André, M., Pasquier, A., 2014. Évaluation de la composition du parc automobile en Ile-de-France pour le calcul des émissions de polluants liés au trafic routier. Pollut.

Atmos. doi:10.4267/pollution-atmospherique.4342
Casas, J., Ferrer, J.L., Garcia, D., Perarnau, J., Torday, A., 2010. Traffic Simulation with Aimsun, in: Barceló, J. (Ed.), Fundamentals of Traffic Simulation, International Series in Operations Research & Management Science. Springer New York, pp. 173–232.
Cazala, A., Crespy, G., De Monchy, G., Deterne, J., Garnier, P., Rimatteu, P., 2006. Rapport sur la comparaison au niveau européen des coûts de construction, d'entretien et l'exploitation des routes (Mission d'audit de modernisation). Conseil général des ponts et chaussée, Paris.
CEREMA, 2014. Impact du cycle de feux sur la capacité, l'acceptabilité de l'attente aux feux et la sécurité : rapport d'étape phase 1 (Rapport d'études). Direction Territoriale de Méditerranée.
CEREMA, 2008. Agir contre l'effet de serre, la pollution de l'air et le bruit dans les plans de déplacements urbains: Approches et méthodes. CEREMA (Centre d'etudes et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement), Paris.
CERTU, 2003. Modélisation des déplacements urbains de voyageurs : Guide pratique (Guide pratique). CERTU, Paris.
CERTU, 1999. Les logiciels de planification des déplacements urbains (Catalogue de logiciel). CERTU Centre d'Etudes sur les Réseaux, les Transports, et l'Urbanisme), Paris.
CETE, 2013. Open data : Données de circulation routière (Rapport Final). CETE Lyon, Lyon.
CGDD, 2016. Opinions et pratiques environnementales des Français en 2015 (Opinion No. 750), Chiffres et statistiques. Commissariat Général au Développement Durable (CGDD), Paris.
CGDD, 2012. Modèlisation urbaine : de la représentation au projet, RéférenceS » de la Direction de la recherche et de l'innovation (DRI). Commissariat Général au Développement Durable (CGDD), Paris.
CGDD, 2010. La mobilité des Français - Panorama issu de l'enquête nationale transports et déplacements 2008, La Revue. Service de l'Observation et des Statistiques ( SOeS) du Commissariat Général au Développement Durable (CGDD ).
Champion, A., 2003. Mécanisme de coordination multi-agent fondé sur des jeux : application à la simulation comportementale de trafic routier en situation de carrefour. Valenciennes.
Chanut, S., Chevallier, E., 2012. Estimating the atmospheric impacts of traffic management projects: application of theoretical models to concrete cases. Rech. Transp. Sécurité 28, 1–14. doi:10.1007/s13547-011-0018-4
Chatzimilioudis, G., Konstantinidis, A., Laoudias, C., Zeinalipour-Yazti, D., 2012. Crowdsourcing with Smartphones. IEEE Internet Comput. 16, 36–44. doi:10.1109/MIC.2012.70
Chatzis, K., 2013. Une histoire comparée de la modélisation des déplacements urbains en France et en Amérique du Nord (Etats-Unis et Canada) sur la longue durée (1950 à nos jours) (report).
Chatzis, K., 2009. Etat et privé dans la modélisation des déplacements urbains en France, 1960-2005 : quel processus de « production » ? (report).
Cheng, T., Haworth, J., Manley, E., 2012. Advances in geocomputation (1996–2011). Comput. Environ. Urban Syst., Special Issue: Advances in Geocomputation 36, 481–487. doi:10.1016/j.compenvurbsys.2012.10.002
Cheylan, J.-P., 2007. Les processus spatio-temporels: quelques notions et concepts préalables à leur représentation. Doss. Mappemonde « Cartogr. Animée » 3, 21.
Chorley, R.J., 1964. Geography and Analogue Theory. Ann. Assoc. Am. Geogr. 54, 127–137. doi:10.1111/j.1467-8306.1964.tb00478.x
CITEPA, 2015. Inventaire des émissions de polluants atmosphériques et de gaz à effet de serre en France – Séries sectorielles et analyses étendues. CITEPA (Centre Interprofessionnel Technique d'Etude de la Pollution Atmosphérique), Paris.

Citton, Y., 2012. Traiter les données : entre économie de l'attention et mycélium de la signification. Multitudes 49, 143–149. doi:10.3917/mult.049.0143

Cohen, G., Salomon, I., Nijkamp, P., 2002. Information–communications technologies (ICT) and transport: does knowledge underpin policy? Telecommun. Policy 26, 31–52. doi:10.1016/S0308-5961(01)00052-0

Commenges, H., 2013. L'invention de la mobilité quotidienne. Aspects performatifs des instruments de la socio-économie des transports. Université Paris-Diderot - Paris VII.

Coria, J., Bonilla, J., Grundström, M., Pleijel, H., 2015. Air pollution dynamics and the need for temporally differentiated road pricing. Transp. Res. Part Policy Pract. 75, 178–195. doi:10.1016/j.tra.2015.03.004

Corti, P., Kraft, T.J., Mather, S.V., Park, B., 2014. PostGIS Cookbook. Packt Publishing.

Costeseque, G., 2011. Analyse et modélisation du trafic routier: Passage du microscopique au macroscopique. Mém. Maitrise À L'Université Lyon–2011.

Crétat, J., Macron, C., Pohl, B., Richard, Y., 2011a. Quantifying internal variability in a regional climate model: a case study for Southern Africa. Clim. Dyn. 37, 1335–1356. doi:10.1007/s00382-011-1021-5

Crétat, J., Pohl, B., Richard, Y., 2011b. Les modèles climatiques régionaux : outils de décomposition des échelles spatio-temporelles., in: Dixièmes Rencontres de ThéoQuant. Besançon, France, pp. 1–11.

Crooks, A., 2015. Agents-based modeling and geographical information systems, in: Geocomputation: A Practical Primer. SAGE Publications Ltd, Thousand Oaks, CA, pp. 63–77.

Cura, R., Tannier, C., 2015. Construction et évaluation d'un modèle de simulation dans un contexte interdisciplinaire, in: Douzièmes Rencontres de Théo Quant. Besançon, France.

Cuyala, S., 2014. Analyse spatio-temporelle d'un mouvement scientifique. L'exemple de la géographie théorique et quantitative européenne francophone. (phdthesis). Université Paris 1 Panthéon-Sorbonne.

Czura, G., Taillandier, P., Tranouez, P., Daudé, É., 2015. MOSAIIC: City-Level Agent-Based Traffic Simulation Adapted to Emergency Situations, in: Takayasu, H., Ito, N., Noda, I., Takayasu, M. (Eds.), Proceedings of the International Conference on Social Modeling and Simulation, plus Econophysics Colloquium 2014. Springer International Publishing, Cham, pp. 265–274.

Dab, W., Roussel, I., 2001. L'Air et la ville. Hachette, Paris.

Daudé, E., 2005. Systèmes multi-agents pour la simulation en géographie : vers une Géographie Artificielle 353–380.

Daudé, E., 2004. Apports de la simulation multi-agents à l'étude des processus de diffusion. Cybergeo Eur. J. Geogr. doi:10.4000/cybergeo.3835

Daudé, E., Langlois, P., 2006. Introduction à la modélisation multi-agent des systèmes complexes en géographie, in: Modélisation et Simulation Multi-agents Applications pour les Sciences de l'Homme et de la Société. Hermes Science Publications, Paris;, pp. 23–48.

De Coninck, F., 2015. L'invention des routines. Espace Popul. Sociétés Space Popul. Soc.

De Palma, A., 2013. Simulation dynamique du trafic urbain (L'évaluation socio-économique en période de transition No. Tome 2). Commissariat général à la stratégie et à la prospective, Paris.

De Runz, C., 2008. Imperfection, temps et espace : modélisation, analyse et visualisation dans un SIG archéologique (phdthesis). Université de Reims - Champagne Ardenne.

Debizet, G., 2011. L'évolution de la modélisation des déplacements urbains en France 1960-2005. Flux 8–21.

Debizet, G., 2004. Déplacements urbains de personnes : de la planification des transports à la gestion durable de la mobilité. Mutations d'une expertise. Université Panthéon-Sorbonne - Paris I.

Delay, E., 2015. Réflexions géographiques sur l'usage des systèmes multi agents dans la compréhension des processus d'évolution des territoires viticoles de fortes pentes : le cas de la Côte Vermeille et du Val di Cembra. Limoges.

Demazeau, Y., 2003. Créativité émergente centrée utilisateur. Tech. Sci. Inform. 22, 31–36.

Desrosières, A., 2000. La politique des grands nombres : histoire de la raison statistique. LA DECOUVERTE, Paris.

Devys, E., Gesquière, G., 2012. Interopérabilité et intégration des données et modèles urbains : standards, normes et tendances pour les SIG, in: La Modélisation de La Ville : Du Modèle Au Projet Urbain, RéférenceS de La DRI. Paris, pp. 92–103.

Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. Numer. Math. 1, 269–271. doi:10.1007/BF01386390

Dobruszkes, F., Lanneaux, M.-A., 2010. Conclusions : comme un reflet de la géographie des transports à la française. Belg. Rev. Belge Géographie 245–246. doi:10.4000/belgeo.6819

Drevon, G., Jambon, F., Chardonnel, S., Christophe, S., André-Poyaud, I., Davoine, P.-A., Lutoff, C., 2014. Évaluation comparée de l'apport de l'assistance GPS aux enquêtes de mobilité. Netcom Réseaux Commun. Territ. 13–34. doi:10.4000/netcom.1527

Ducruet, C., 2010a. Les mesures locales d'un réseau (Rapport de synthèse). Groupe FMR (Flux, Matrices , Réseaux), Paris.

Ducruet, C., 2010b. Les mesures globales d'un réseau (Rapport de synthèse). Groupe FMR (Flux, Matrices , Réseaux), Paris.

Dupuy, 1992. L'urbanisme des réseaux : Théories et méthodes. Armand Colin, Paris.

Dupuy, G., 1995. Les territoires de l'automobile. Economica, Paris.

Dupuy, G., 1993. 6. De l'informatique municipale à la "ville intelligente" : tendances de l'informatisation urbaine. Annu. Collectiv. Locales Tome 13, 93–103.

Dupuy, G., 1987. Vers une théorie territoriale des réseaux : une application au transport urbain. Ann. Géographie 96, 658–679. doi:10.3406/geo.1987.20639

Dupuy, G., 1984. Villes, systèmes et réseaux : le rôle historique des techniques urbaines. Réseaux 2, 3–23. doi:10.3406/reso.1984.1100

Dupuy, G., 1975. Une technique de planification au service de l'autombile : Les modèles de trafic urbain. Ministère de l'équipement, Paris.

Dupuy, G., Benguigui, L.G., 2015. Sciences urbaines : interdisciplinarités passive, naïve, transitive, offensive. Métropoles.

Edmonds, B., Hales, D., 2003. Replication, Replication and Replication: Some Hard Lessons from Model Alignment. J. Artif. Soc. Soc. Simul. 6(4).

Edmonds, B., Moss, S., 2004. From KISS to KIDS – An "Anti-simplistic" Modelling Approach, in: Davidsson, P., Logan, B., Takadama, K. (Eds.), Multi-Agent and Multi-Agent-Based Simulation, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 130–144.

Elichegaray, C., Henry, Y., Morchéoine, A., 2008. La pollution de l'air : Sources, effets, prévention. Dunod, Paris.

Elissalde, B., 2000. Géographie, temps et changement spatial. Espace Géographique 29, 224–236. doi:10.3406/spgeo.2000.2008

Emery, J., 2012. La qualité de l'air liée au transport routier en milieu urbain : Analyse des concentrations en oxydes d'azote sur l'agglomération dijonnaise (Mémoire de Recherche Master 2). Université de Bourgogne, Dijon.

Epstein, J.M., 1999. Agent-based computational models and generative social science. Complexity 4, 41–60. doi:10.1002/(SICI)1099-0526(199905/06)4:5<41::AID-CPLX9>3.0.CO;2-F

Faye, S., Chaudet, C., Demeure, I., 2012. Contrôle du trafic routier urbain par un réseau fixe de capteurs sans fil (Rapport Technique). Telecom ParisTech Groupe RMS : Réseaux, Mobilité et Services, Paris.

Fellendorf, M., Vortisch, P., 2010. Microscopic Traffic Flow Simulator VISSIM, in: Barceló, J.

(Ed.), Fundamentals of Traffic Simulation, International Series in Operations Research & Management Science. Springer New York, pp. 63–93.

Fen-Chong, J., 2012. Organisation spatio--temporelle des mobilités révélées par la téléphonie mobile en Ile--de--France (phdthesis). Université Panthéon-Sorbonne - Paris I.

Fenger, J., 2009. Air pollution in the last 50 years – From local to global. Atmos. Environ., Atmospheric Environment - Fifty Years of Endeavour 43, 13–22. doi:10.1016/j.atmosenv.2008.09.061

Fenger, J., 1999. Urban air quality. Atmos. Environ. 33, 4877–4900. doi:10.1016/S1352-2310(99)00290-3

Ferber, J., 2006. Concepts et méthodologie multi-agents, in: Modélisation et Simulation Multi-agents Applications pour les Sciences de l'Homme et de la Société. Hermes Science Publications, Paris;, pp. 23–48.

Ferber, J., 1995. Les Systèmes multi-agents: vers une intelligence collective. InterEditions.

Foltête, J.-C., Genre-Grandpierre, C., Josselin, D., 2008. Impacts des réseaux viaires sur les mobilités urbaines : quelques illustrations, in: Thériault, M., Rosiers, F.D. (Eds.), Information Géographique et Dynamiques Urbaines 1 - Analyse et Simulation de La Mobilité Des Personnes, Traité IGAT. Hermès-Lavoisier, pp. 139–165.

Fotheringham, A.S., 1998. Trends in quantitative methods II: stressing the computational. Prog. Hum. Geogr. 22, 283–292. doi:10.1191/030913298670827726

Fouillé, L., Broc, J.-S., Bourges, B., Bougnol, J., Mestayer, P., 2012. La place des modèles de trafic dans les récentes modélisations des impacts environnementaux des transports. Importance de l'explicitation des méthodes et hypothèses. Rech. Transp. Sécurité 28, 190–200. doi:10.1007/s13547-012-0037-5

Gahegan, M., 1999. Guest Editorial: What is Geocomputation? Trans. GIS 3, 203–206. doi:10.1111/1467-9671.00017

Gallez, C., Kaufmann, V., 2009. Aux racines de la mobilité en sciences sociales. 41–55.

Genre-Grandpierre, C., 2010. Caractériser la forme et le fonctionnement des réseaux de transports, in: Système de Transport Urbain : Caractérisation de L'offre et Estimation de La Demande. Hermes science : Lavoisier (Traité IGAT), Paris, pp. 83–106.

Genre-Grandpierre, C., 2000. Forme et fonctionnement des réseaux de transport : approche fractale et réflexions sur l'aménagement des villes. Besançon.

Giamo, G., Anderson, R., Wargelin, L., Stopher, P., 2010. Will it work? Pilot results from the first large-scale GPS-based household travel survey in the United States. Presented at the Transportation Research Board Annual Meeting, 89th, 2010, Washington, DC, USA.

Giannopoulos, G.A., 2004. The application of information and communication technologies in transport. Eur. J. Oper. Res., New Technologies in Transportation Systems 152, 302–320. doi:10.1016/S0377-2217(03)00026-2

Gilbert, N., Troitzsch, K.G., 2005. Simulation for the Social Scientist, 2nd ed. Open University Press, Maidenhead, England ; New York, NY.

Gleyze, J.-F., 2005. La vulnérabilité structurelle des réseaux de transport dans un contexte de risques (phdthesis). Université Paris-Diderot - Paris VII.

Goodchild, M., 2009. NeoGeography and the nature of geographic expertise. J. Locat. Based Serv. 3, 82–96. doi:10.1080/17489720902950374

Goodchild, M.F., 2013. Prospects for a Space–Time GIS. Ann. Assoc. Am. Geogr. 103, 1072–1077. doi:10.1080/00045608.2013.792175

Goodchild, M.F., 2010. Twenty years of progress: GIScience in 2010. J. Spat. Inf. Sci. 0, 3–20. doi:10.5311/josis.v0i1.32

Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography. GeoJournal 69, 211–221. doi:10.1007/s10708-007-9111-y

Goodchild, M.F., 2000. GIS and Transportation: Status and Challenges. GeoInformatica 4, 127–139. doi:10.1023/A:1009867905167

Goodchild, M.F., 1992. Geographical information science. Int. J. Geogr. Inf. Syst. 6, 31–45. doi:10.1080/02693799208901893

Goodchild, M.F., Longley, P.P.A., 2014. The Practice of Geographic Information Science, in: Fischer, M.M., Nijkamp, P. (Eds.), Handbook of Regional Science. Springer Berlin Heidelberg, pp. 1107–1122.

Grassot, L., Nicolas, J.-P., Pluvinet, P., 2012. De l'intérêt de contrôler l'impact des hypothèses de composition du parc automobile sur l'estimation des émissions liées au trafic routier. Rech. Transp. Sécurité 28, 215–227. doi:10.1007/s13547-012-0040-x

Grataloup, C., 2015. Introduction à la géohistoir, Armand Colin. ed, Cursus. Dunod, Paris.

Guinot, B., 2008. Pollution atmosphérique et développement urbain en Chine : quelles perspectives ? Perspect. Chin. 105, 67–76. doi:10.3406/perch.2008.3674

Gulliver, J., Briggs, D.J., 2005. Time-space modeling of journey-time exposure to traffic-related air pollution using GIS. Environ. Res. 97, 10–25. doi:10.1016/j.envres.2004.05.002

Gwiazdzinski, L., Drevon, G., 2014. Des méthodes et des outils au service d'une nouvelle intelligence des mobilités. Netcom Réseaux Commun. Territ. 5–12.

Gwiazdzinski, L., Klein, O., 2014. Du suivi GPS des individus à une approche chronotopique. Netcom Réseaux Commun. Territ. 77–106. doi:10.4000/netcom.1604

Hadachi, A., 2013. Travel Time Estimation Using Sparsely Sampled Probe GPS Data in Urban Road Networks Context. Rouen, INSA.

Hägerstrand, T., 1970. What about people in Regional Science? Pap. Reg. Sci. Assoc. 24, 6–21. doi:10.1007/BF01936872

Haggett, P., 1965. Locational analysis in modern geography. Arnold, Londres.

Hassan, S., Antunes, L., Arroyo, M., 2009. Deepening the Demographic Mechanisms in a Data-Driven Social Simulation of Moral Values Evolution, in: David, N., Sichman, J.S. (Eds.), Multi-Agent-Based Simulation IX, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 167–182.

Hassan, S., Antunes, L., Pavón, J., Gilbert, N., 2010a. Stepping on Earth: A Roadmap for Data-driven Agent-Based Modelling., in: Proceedings of the 5th Conference of the European Social Simulation Association (ESSA08). Presented at the The European Social Simulation Association Annual Conference, Brescia, Italy.

Hassan, S., Pavón, J., Antunes, L., Gilbert, N., 2010b. Injecting Data into Agent-Based Simulation, in: Takadama, K., Cioffi-Revilla, C., Deffuant, G. (Eds.), Simulating Interacting Agents and Social Phenomena, Agent-Based Social Systems. Springer Japan, pp. 177–191.

Henn, V., 2001. Information routière et affectation du trafic : vers une modélisation floue. Saint-Etienne.

Héran, F., 2002. Le reflux des rues à sens unique 48–49, 83–93.

Herrera, J.C., Work, D.B., Herring, R., Ban, X. (Jeff), Jacobson, Q., Bayen, A.M., 2010. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. Transp. Res. Part C Emerg. Technol. 18, 568–583. doi:10.1016/j.trc.2009.10.006

Herrera, J.C., Work, D.B., Herring, R., Ban, X.J., Bayen, A.M., 2009. Evaluation of Traffic Data Obtained via GPS-Enabled Mobile Phones: the Mobile Century Field Experiment. UC Berkeley Cent. Future Urban Transp. Volvo Cent. Excell.

Hey, T., Tansley, S., Tolle, K. (Eds.), 2009. The Fourth Paradigm: Data-Intensive Scientific Discovery, 1 edition. ed. Microsoft Research, Redmond , Washington.

Hirtzel, J., 2015. Exploration prospective des mobilités résidentielles dans une agglomération urbaine au moyen d'un modèle de simulation multi-agents (Mobisim) (phdthesis). Université de Franche-Comté.

Hoh, B., Gruteser, M., Herring, R., Ban, J., Work, D., Herrera, J.-C., Bayen, A.M., Annavaram, M., Jacobson, Q., 2008. Virtual Trip Lines for Distributed Privacy-preserving Traffic

Monitoring, in: Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services, MobiSys '08. ACM, New York, NY, USA, pp. 15–28. doi:10.1145/1378600.1378604

Hoh, B., Iwuchukwu, T., Jacobson, Q., Work, D., Bayen, A.M., Herring, R., Herrera, J.C., Gruteser, M., Annavaram, M., Ban, J., 2012. Enhancing Privacy and Accuracy in Probe Vehicle-Based Traffic Monitoring via Virtual Trip Lines. IEEE Trans. Mob. Comput. 11, 849–864. doi:10.1109/TMC.2011.116

Hugrel, C., Joumard, R., 2004. Transport routier-Parc, usage et émissions des véhicules en France de 1970 à 2025 (Rapport de recherche). INRETS : Institution national de recherche sur les transports et leur sécurité, Bron.

IGN, 2014. Descriptif de la BD TOPO (Descriptif de contenu No. Révision du document de 2011). Saint Mandé.

IGN, 2010. Descriptif de la BD CARTO (Descriptif de contenu No. Révision du document de 2015). Saint Mandé.

Joliveau, T., 2004. Géomatique et gestion environnementale du territoire. Recherches sur un usage géographique des SIG (Habilitation à diriger des recherches). Université de Rouen, Rouen.

Juan, S., 2015. Le concept de routine dans la socio-anthropologie de la vie quotidienne. Espace Popul. Sociétés Space Popul. Soc.

Keuken, M.P., Jonkers, S., Wilmink, I.R., Wesseling, J., 2010. Reduced NOx and PM10 emissions on urban motorways in The Netherlands by 80 km/h speed management. Sci. Total Environ. 408, 2517–2526. doi:10.1016/j.scitotenv.2010.03.008

Kitchin, R., 2014. Big Data, new epistemologies and paradigm shifts. Big Data Soc. 1, 2053951714528481. doi:10.1177/2053951714528481

Kitchin, R., 2013. Big data and human geography Opportunities, challenges and risks. Dialogues Hum. Geogr. 3, 262–267. doi:10.1177/2043820613513388

Klein, O., 2010. Visualiser les mobilités quotidiennes : vers d'autres modes de représentations, in: Système de Transport Urbain : Caractérisation de L'offre et Estimation de La Demande, Traité de l'IGAT. Hermes science : Lavoisier (Traité IGAT), Paris, pp. 145–186.

Klein, O., 2007. Modélisation et représentations spatio-temporelles des déplacements quotidiens urbains : Application à l'aire urbaine Belfort-Montbéliard. Strasbourg 1.

Kohler, M., 2015. Assessement of the building energy requirements : added value of the use of the urban climate modeling. Strasbourg, Strasbourg.

Krajzewicz, D., Herkorn, G., Wagner, P., Rössel, C., 2002. SUMO (simulation of urban mobility) : An open-source traffic simulation. Proc 4th Middle East Symp. Simul. Model. 183.187.

Krzyżanowski, M., Kuna-Dibbert, B., Schneider, B., 2005. Health effects of transport-related air pollution, WHO Regional Office for Europe. World Health Organisation.

Ksontini, F., 2013. Modèle d'agent fondé sur les affordances : application à la simulation de trafic routier. Valenciennes.

Kuhn, T.-S., 1970. La structure des révolutions scientifiques. Flammarion, Paris.

Kwan, M.-P., 2013. Beyond Space (As We Knew It): Toward Temporally Integrated Geographies of Segregation, Health, and Accessibility. Ann. Assoc. Am. Geogr. 103, 1078–1086. doi:10.1080/00045608.2013.792177

Kwan, M.-P., 2012. The Uncertain Geographic Context Problem. Ann. Assoc. Am. Geogr. 102, 958–968. doi:10.1080/00045608.2012.687349

Lagesse, C., Bordin, P., Bonnin, P., Douady, S., 2015. Méthodologie de modélisation et de caractérisation des réseaux spatiaux. Application au réseau viaire de Paris. Flux Patrim. Commun. Entrep. Réseau.

Langlois, P., 2006. Approche conceptuelle de l'espace : structuration de l'espace, du temps et des objets dans un contexte de modélisation multi-agent, in: Modélisation et Simulation

Multi-agents Applications pour les Sciences de l'Homme et de la Société. Hermes Science Publications, Paris;, pp. 363–384.

Langlois, P., Blanpain, B., Daudé, E., 2015. MAGéo, une plateforme de modélisation et de simulation multi-agent pour les sciences humaines. Cybergeo Eur. J. Geogr. doi:10.4000/cybergeo.27236

Languillon-Aussel, R., Leprêtre, N., Granier, B., 2016. La stratégie de la « smart city » au Japon : expérimentations nationales et circulations globales. EchoGéo. doi:10.4000/echogeo.14598

Lannoy, P., 2003. L'automobile comme objet de recherche, Chicago, 1915-1940. Rev. Fr. Sociol. 44, 497. doi:10.3917/rfs.443.0497

Leclercq, L., 2002. Modélisation dynamique du trafic et applications à l'estimation du bruit routier (Thèse de doctorat). École Doctorale Mécanique, Energétique, Génie Civil, Acoustique, MEGA, France.

Lévy, J., Lussault, M., 2003. Dictionnaire de la géographie. Belin, Paris.

Lhomme, S., 2012. Les réseaux techniques comme vecteur de propagation des risques en milieu urbain - Une contribution théorique et pratique à l'analyse de la résilience urbaine (phdthesis). Université Paris-Diderot - Paris VII.

Livet, P., 2006. Essai d'épistémologie de la simulation multi-agents en sciences sociales, in: Modélisation et Simulation Multi-agents Applications pour les Sciences de l'Homme et de la Société. Hermes Science Publications, Paris;, pp. 193–218.

Lunardi, N., 2013. Mobisim, un outil d'aide à l'évaluation de la durabilité des politiques de transports en milieu urbain. Mém. Master À L'Université Franche-Comté.

MacEachren, A.M., Gahegan, M., Pike, W., Brewer, I., Cai, G., Lengerich, E., Hardisty, F., 2004. Geovisualization for Knowledge Construction and Decision Support. IEEE Comput. Graph. Appl. 24, 13–17.

MacEachren, A.M., Kraak, M.-J., 2001. Research Challenges in Geovisualization. Cartogr. Geogr. Inf. Sci. 28, 3–12. doi:10.1559/152304001782173970

Makri, A., Stilianakis, N.I., 2008. Vulnerability to air pollution health effects. Int. J. Hyg. Environ. Health 211, 326–336. doi:10.1016/j.ijheh.2007.06.005

Maneerat, S., Daudé, E., 2013. Geosimulation of dengue vector dynamics in urban areas.

Manovich, L., 2011. Trending: The promises and challenges of Big Social Data, in: Debates in the Digital Humanities. The University of Minnesota Press, Minneapolis.

Marilleau, N., 2006. Méthodologie, formalismes et outils de modélisation-simulation pour l'étude des systèmes complexes : application à la mobilité géographique (Thèse d'informatique). Université de Franche-Comté, Besançon.

Martiny, N., Roucou, P., Menut, L., Chiapello, I., Broutin, H., 2016. Mineral dust, a key indicator of the meningistis risk at the scale of the belt. Submitt. Atmospheric Environ.

Masson, S., 2000. Les interactions entre système de transport et système de localisation en milieu urbain et leur modélisation (Economie des transports). Lyon 2.

Massot, M.-H., Orfeuil, J., 2008. Mobilité résultante et mobilité organisatrice : les paradigmes de la mobilité au service de la compréhension de la ville, in: CHALAS, A. (Ed.), La Mobilité Qui Fait La Ville - Actes Des 3èmes Rencontres Internationales En Urbanisme de l'Institut d'Urbanisme de Grenoble. Université Pierre Mendès France-Grenoble) CERTU, p. 326.

Mathian, H., Sanders, L., 2014. Objets géographiques et processus de changement, Collection Systèmes d'information géographique et intelligence territoriale. ed. ISTE Editions, London.

Mathis, P., 2003. Graphes et réseaux. Modélisation multiniveau. Hermes Science Publications, Paris.

McFadden, D., 2000. Disaggregate Behavioral Travel Demand's RUM Side - A 30-Year Retrospective, in: In Travel Behavior Research: The Leading. Elsevier, pp. 17–64.

McFadden, D., Reid, F., 1975. Aggregate travel demand forecasting from disaggregated behavioral models. Transp. Res. Rec. 534, 24–37.

Mérenne, E., 2008. Géographie des transports : Contraintes et enjeux, édition revue et corrigée. ed. PU Rennes, Rennes.

Mericskay, B., 2011. Les Sig et la cartographie à l'ère du géoweb. L'Espace Géographique Tome 40, 142–153.

Meunier, D., 2013. Les modèles de trafic et l'évaluation économique (L'évaluation socio-économique des investissements publics No. Tome 2). Paris.

Michelot, N., 2014. L'influence des topoclimats sur la pollution de l'air aux particules dans le sud-ouest des Alpes-Maritimes (Thèse de géographie). Nice, Nice.

Michelot, N., Carrega, P., Rouïl, L., 2015. Panorama de la modélisation de la dispersion atmosphérique. Pollut. Atmos. Clim. Santé Société 92–100.

Mille, M., 2000. Des densités habitantes aux densités mouvantes l'exemple de la métropole lilloise. Cybergeo Eur. J. Geogr. doi:10.4000/cybergeo.1866

Miller, H.E., 2013. Big-data in cloud computing: A taxonomy of risks. InformationResearch 18(1) paper 571.

Miller, H.J., 2015. Spatio-temporal knowledge discovery, in: Brunsdon, C., Singleton, A.D. (Eds.), Geocomputation: A Practical Primer. SAGE Publications Ltd, Thousand Oaks, CA, pp. 97–109.

Miller, H.J., Goodchild, M.F., 2014. Data-driven geography. GeoJournal 1–13. doi:10.1007/s10708-014-9602-6

Miller, H.J., Han, J., 2009. Geographic Data Mining and knowledge discovery : an overview, in: Geographic Data Mining and Knowledge Discovery, CRC Press. Boca Raton.

Millet, A., 2005. Méthodes de Monte-Carlo (Mémoire de Recherche Master). Université Paris 7 et Paris 1, Paris.

Minar, N., Burkhart, R., Langton, C., al, et, 1996. The Swarm Simulation System: A Toolkit for Building Multi-Agent Simulations.

Moulin, B., Ali, W., 2008. Une méthode de conception de géosimulation pour la modèlisation multi-agents des déplacements individuels, in: Thériault, M., Rosiers, F.D. (Eds.), Information Géographique et Dynamiques Urbaines 1 - Analyse et Simulation de La Mobilité Des Personnes, Traité IGAT. Hermès-Lavoisier, Paris, pp. 253–276.

Nguyen-Luong, D., Coindet, J.-P., Fradin, J.-R., Courel, J., Messenet, A., 2012. Faisabilité d'une enquête globale transports (EGT) intégrale par association d'un GPS, d'un SIG et d'un Système expert en Île-de-France (Rapport Final). Institut d'aménagement et d'urbanisme IDF, Paris.

North, M.J., Collier, N.T., Ozik, J., Tatara, E.R., Macal, C.M., Bragen, M., Sydelko, P., 2013. Complex adaptive systems modeling with Repast Simphony. Complex Adapt. Syst. Model. 1, 1–26. doi:10.1186/2194-3206-1-3

North, M.J., Macal, C.M., 2007. Managing Business Complexity: Discovering Strategic Solutions with Agent-Based Modeling and Simulation. Oxford University Press.

Noucher, M., 2009. La donnée géographique aux frontières des organisations : approche socio-cognitive et systémique de son appropriation. (phdthesis). Ecole Polytechnique Fédérale de Lausanne (EPFL).

Ntziachristos, L., Zissis, S., 2009. Exhaust emissions from road transport. EMEP/EEA emission inventory guidebook - COPERT, in: EMEP/EEA Air Pollutant Emission Inventory Guidebook 2013 : Technical Guidance to Prepare National Emission Inventories. AEE : European Environment Agency, Copenhague, p. 1.A.3.b.i-iv (Chapitre technique).

OMINEA, 2015. Organisation et Méthodes des Inventaires Nationaux des Emissions Atmosphériques en France : OMINEA (Rapport d'études). CITEPA (Centre Interprofessionnel Technique d'Etude de la Pollution Atmosphérique), Paris.

OMS, 2006. Air Quality Guidelines Global Update 2005 : particulate matter, ozone, nitrogen

dioxide, and sulfur dioxide. Regional Office for Europe, & World Health Organization.
Openshaw, S., 2000. Chapitre 1 :Geocomputation, in: Geocomputation. CRC Press, Londres, pp. 1–30.
Openshaw, S., Abrahart, R.J., 2000. Geocomputation, 2nde Edition Taylor et Francis. ed. CRC Press, Londres.
Openshaw, S., Charlton, M., Wymer, C., Craft, A., 1987. A Mark 1 Geographical Analysis Machine for the automated analysis of point data sets. Int. J. Geogr. Inf. Syst. 1, 335–358. doi:10.1080/02693798708927821
Orfeuil, J.-P., 2004. Accessibilité, mobilité, inégalités : regards sur la question en France aujourd'hui, in: Transports, pauvretés, exclusions. Paris, pp. 22–47.
Orfeuil, J.-P., 2000. L'évolution de la mobilité quotidienne. INRETS, Arcueil.
Ortúzar, J. de D., Willumsen, L.G., 2011. Modelling Transport, 4 edition. ed. Wiley, Chichester, West Sussex, United Kingdom.
O'Sullivan, D., 2005. Geographical information science: time changes everything. Prog. Hum. Geogr. 29, 749–756.
Oxley, T., Valiantis, M., Elshkaki, A., ApSimon, H.M., 2009. Background, Road and Urban Transport modelling of Air quality Limit values (The BRUTAL model). Environ. Model. Softw. 24, 1036–1050. doi:10.1016/j.envsoft.2009.02.011
PCIT, 2012. Méthode d'élaboration des inventaires territoriaux des émissions atmosphériques (Guide méthodologique). Paris.
Peuquet, D.J., 2002. Representations of Space and Time. The Guilford Press, New York.
Pfoser, D., 2008. Floating Car Data, in: Encyclopedia of GIS. Springer US, pp. 321–321.
Planchon, V., 2005. Traitement des valeurs aberrantes: concepts actuels et tendances générales. Biotechnol. Agron. Société Environ. 9, 19–34.
Plassard, F., 1992. Les réseaux de transport et de communication, in: Encyclopédie de géographie. Economica, Paris, pp. 533–556.
Pohl, B., 2015. Variabilité Climatique Régionale en Afrique Australe : Observations, Modélisation (HDR). Université de Bourgogne, Dijon.
Pohl, B., Douville, H., 2011. Diagnosing GCM errors over West Africa using relaxation experiments. Part II: intraseasonal variability and African easterly waves. Clim. Dyn. 37, 1313–1334. doi:10.1007/s00382-011-1106-1
Pohl, Richard, Kohler, Emery, Castel, De Lapparent, Roux, Bientz, Codet-Hache, Dumaitre, Toussaint, Tissot, Thevenin, Thevenin, Pergaud, 2015. The Urban Heat Island of a middle-size French city as seen by high-resolution numerical experiments and in situ measurements the case of Dijon, Burgundy., in: ResearchGate. Presented at the ICUC9.
Primequal-Predit, 2012. Pollution de l'air et transports terrestres. Dix ans de recherche. Paris.
Pumain, D., 2014a. Observation, observation, observation. Cybergeo Eur. J. Geogr.
Pumain, D., 2014b. Une théorie géographique des villes. BSGLg.
Pumain, D., Saint-Julien, T., 2010a. Analyse spatiale : les localisations, 2e édition. ed. Armand Colin, Paris.
Pumain, D., Saint-Julien, T., 2010b. Analyse spatiale : les interactions, 2e édition. ed. Armand Colin, Paris.
Quinet, E., 1998. Principes de l'économie des transports. Economica, Paris.
Ramat, E., 2006. Introduction à la modèlisation et à la simulation à événements discrets, in: Modélisation et Simulation Multi-agents Applications pour les Sciences de l'Homme et de la Société. Hermes Science Publications, Paris;, pp. 49–73.
Reix, R., 2011. Systèmes d'information et management des organisations, Édition : 6e édition. ed. VUIBERT, Paris.
Resnick, M., 1996. StarLogo: An Environment for Decentralized Modeling and Decentralized Thinking, in: Conference Companion on Human Factors in Computing Systems, CHI '96. ACM, New York, NY, USA, pp. 11–12. doi:10.1145/257089.257095

Reuillon, R., Leclaire, M., Rey-Coyrehourcq, S., 2013. OpenMOLE, a workflow engine specifically tailored for the distributed exploration of simulation models. Future Gener. Comput. Syst., Including Special sections: Advanced Cloud Monitoring Systems & The fourth IEEE International Conference on e-Science 2011 — e-Science Applications and Tools & Cluster, Grid, and Cloud Computing 29, 1981–1990. doi:10.1016/j.future.2013.05.003

Reungoat, P., Chiron, M., Momas, I., 2004. Évaluation de l'exposition à la pollution atmosphérique liée au trafic routier dans les études épidémiologiques : une revue de littérature. Rev. DÉpidémiologie Santé Publique 52, 271–296. doi:10.1016/S0398-7620(04)99052-9

Robert, D., 2001. Pertinence des zonages géographiques pour l'analyse des transports. PREDIT-METL, Paris.

Roncayolo, M., 1990. La Ville et ses territoires. Folio, Paris.

Rouïl, L., 2012. Les modèles régionaux et les modèles urbains : nouveaux outils d'aide à la décision. Pollut. Atmos. Clim. Santé Société 119–121.

Roussel, I., 2010. Chapitre 23 : Pollutions atmosphériques locales et développement durable, in: Développement durable et territoire. Presses Universitaires du Septentrion, Villeneuve d'Ascq, France.

Roussel, I., 2006. Climatologie géographiques et pollutions atmosphériques : quelles synergies ?, in: Climat et Société : L'apport Des Géographes-Climatologues. Presented at the Journées de la climatologie, Nice, p. 10.

Sanders, L., 2007. Objets géographiques et simulation agent, entre thématique et méthodologie. Rev. Int. GéomatiqueInternational J. Geomat. Spat. Anal. 17, p135-160.

Sanders, L., 2001a. Modèles en analye spatiale: introduction 17–29.

Sanders, L., 2001b. Modèles en analye spatiale: introduction 17–29.

Servigne, S., Ray, C., Bouju, A., Devogele, T., Bertrand, F., Gutiérrez, C., Guillaume, N., 2009. Gestion de masses de données en temps réel au sein de bases de données capteurs. Rev. Int. Géomat. 19, 133–150. doi:10.3166/geo.19.133-150

SETRA, 2012. Panorama des systèmes de recueil de données de trafic routier.

SETRA, 2009. Emissions routières de polluants atmosphériques : Courbes et facteurs d'influence (Note d'information No. 92), Série Economie Environnement Conception. SETRA (Service d'Etudes sur les Transports, les Routes et leurs Aménagements), Bagneux.

Sportisse, B., 2007a. Pollution atmospherique des processus à la modélisation. Springer, Paris.

Sportisse, B., 2007b. Modélisation et simulation de la pollution atmosphérique (Habilitation à diriger des recherches). Joseph Fourier, Grenoble.

Stone, M., 1974. Cross-Validatory Choice and Assessment of Statistical Predictions. J. R. Stat. Soc. Ser. B Methodol. 36, 111–147. doi:10.2307/2984809

Sykes, P., 2010. Traffic Simulation with Paramics, in: Barceló, J. (Ed.), Fundamentals of Traffic Simulation, International Series in Operations Research & Management Science. Springer New York, pp. 131–171.

Taillandier, P., Grignard, A., Gaudou, B., Drogoul, A., 2014. Des données géographiques à la simulation à base d'agents : application de la plate-forme GAMA. Cybergeo Eur. J. Geogr. doi:10.4000/cybergeo.26263

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmospheres 106, 7183–7192. doi:10.1029/2000JD900719

That, D.H.T., Popa, I.S., Zeitouni, K., 2015. PPTM: Privacy-aware Participatory Traffic Monitoring Using Mobile Secure Probes. Presented at the 31èmes journées Bases de Données Avancées (BDA '15). Demo paper, p. 4 pages.

Thériault, M., Rosiers, F.D., 2008. Information géographique et dynamiques urbaines : 2 volumes, Traité IGAT. ed. Hermès-Lavoisier, Paris.

Thévenin, T., 2011. Transport, espace, temps: regard croisé entre time geography et géo-histoire

(Habilitation à diriger des recherches). Université de Bourgogne, Dijon.
Thévenin, T., Chardonnel, S., Cochey, É., 2007. Explorer les temporalités urbaines de l'agglomération de Dijon. Espace Popul. Sociétés Space Popul. Soc. 179–190. doi:10.4000/eps.2069
Tiphaine, K., 2014. L'évolution de l'accessibilité du réseau routeir dijonnais, un an après la mise en place du tramway (Mémoire de Recherche Master 1). Université de Bourgogne, Dijon.
Tisue, S., Wilensky, U., 2004. NetLogo: A simple environment for modeling complexity, in: In International Conference on Complex Systems. pp. 16–21.
Torrence, C., Compo, G.P., 1998. A Practical Guide to Wavelet Analysis. Bull. Am. Meteorol. Soc. 79, 61–78. doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2
Torrens, P.M., 2015. Geographical agents in three dimensions, in: Brunsdon, C., Singleton, A.D. (Eds.), Geocomputation: A Practical Primer. SAGE Publications Ltd, Thousand Oaks, CA.
Torrens, P.M., 2010. Geography and computational social science. GeoJournal 75, 133–148. doi:10.1007/s10708-010-9361-y
Torrens, P.M., 2000. How land-use-transportation models work (Working / discussion paper), CASA.
Tukey, J.W., 1977. Exploratory Data Analysis, 1st ed. Pearson, Reading, Mass.
Tukey, J.W., 1962. The Future of Data Analysis. Ann. Math. Stat. 33, 1–67. doi:10.1214/aoms/1177704711
Uherek, E., Halenka, T., Borken-Kleefeld, J., Balkanski, Y., Berntsen, T., Borrego, C., Gauss, M., Hoor, P., Juda-Rezler, K., Lelieveld, J., Melas, D., Rypdal, K., Schmid, S., 2010. Transport impacts on atmosphere and climate: Land transport. Atmos. Environ., Transport Impacts on Atmosphere and Climate: The ATTICA Assessment Report 44, 4772–4816. doi:10.1016/j.atmosenv.2010.01.002
Varenne, F., 2013. Introduction : Modèles et simulations dans l'enquête scientifique: variétés traditionnelles et mutations contemporaines, in: Epistémologies et Pratiques de La Modélisation et de La Simulation, Sciences et Philosophie. Matériologiques, Paris, pp. 10–98.
Varenne, F., 2008. Epistémologie des modèles et des simulations, in: Epistémologie Des Modèles et Des Simulations : Tour D'horizon et Tendances. Paris, France.
Wegener, M., 2004. Overview of land use transport models, in: Handbook in Transport, Vol. 5. Transport, Geography and Spatial Systems. Oxford: Pergamon/Elsevier. pp. 127–146.
Wegener, M., 1994. Operational Urban Models State of the Art. J. Am. Plann. Assoc. 60, 17–29. doi:10.1080/01944369408975547
White, J., Quick, J., Philippou, P., 2004. The use of mobile phone location data for traffic information, in: 12th IEE International Conference on Road Transport Information and Control, 2004. RTIC 2004. Presented at the 12th IEE International Conference on Road Transport Information and Control, 2004. RTIC 2004, pp. 321–325. doi:10.1049/cp:20040048
Wiel, M., 2005. Ville et mobilité : un couple infernal ? Editions de l'Aube, La Tour d'Aigues.
Wilensky, U., Rand, W., 2007. Making Models Match: Replicating an Agent-Based Model. J. Artif. Soc. Soc. Simul. 10(4), 2.

# TABLE DES MATIÈRES

**INTRODUCTION GÉNÉRALE ..... 9**

**PARTIE 1 - LES MESURES DU TRAFIC ROUTIER EN MILIEU URBAIN : L'APPORT DES DISPOSITIFS DE COMPTAGE PAR CAPTEUR ..... 17**

INTRODUCTION DE LA PREMIÈRE PARTIE ..... 19
CHAPITRE 1. SIMULER L'ESPACE URBAIN EN MOUVEMENT ..... 23
*1.1. Quelles échelles spatio-temporelles pour étudier les dynamiques urbaines ?* ..... *24*
1.1.1. Dynamiques urbaines et système urbain ..... 24
a. Observer l'espace urbain à des niveaux plus fins : vers une approche objet-centrée ? ..... 24
b. La mise en mouvement de l'espace urbain : une recherche dans la complexité ? ..... 26
1.1.2. Étudier les dynamiques urbaines par la géographie des transports ..... 29
1.1.3. Une approche spatio-temporelle pour l'évaluation de la pollution atmosphérique automobile (PAA) en milieu urbain ..... 31
*1.2. Entre rationalisme et empirisme* ..... *35*
1.2.1. De la donnée à la connaissance d'un phénomène : quel positionnement épistémologique ? ..... 36
1.2.2. Le big data : une question de paradigme scientifique ? ..... 40
1.2.3. Allier temps et espace en géographie : des défis d'exploration à relever ..... 42
*1.3. La Géocomputation : un arsenal d'outils pour la constitution et l'étude de l'information spatio-temporelle* ..... *45*
1.3.1. Les approches guidées par les données en géographie : un travail s'inscrivant dans l'interdisciplinarité ? ..... 45
1.3.2. Qu'est-ce que la *Géocomputation* ? ..... 47
1.3.3. Les outils pour la *géocomputation* : des méthodes axées sur les données ..... 49
*Conclusion* ..... *53*
CHAPITRE 2. L'ÉTUDE DU TRAFIC ROUTIER EN MILIEU URBAIN, ENTRE DISPOSITIF DE MESURES ET DISPOSITIF DE CONSTRUCTION ..... 55
*2.1. Capter le trafic routier* ..... *56*
2.1.1. Trafic versus Mobilité ..... 56
2.1.2. Enquêtes versus Comptages : quelle(s) mesure(s), pour quel(s) usage(s) ? ..... 58
2.1.3. Les notions de *big data* et *les smart cities* pour replacer les capteurs au service de l'information urbaine ..... 61
*2.2. Les systèmes de capteurs mobiles pour capter la trajectoire des véhicules* ..... *64*
2.2.1. Les caractéristiques des capteurs mobiles ..... 65
2.2.2. Les comptages-trajectoires : une mesure individuelle à appréhender ..... 68
2.2.3. Les limites des capteurs mobiles pour la mesure du trafic routier ..... 69
*2.3. Les données issues des capteurs fixes : une mesure collective* ..... *71*
2.3.1. Les caractéristiques des différents capteurs fixes : ..... 71
2.3.2. Les comptages ponctuels : une mesure collective pour le trafic routier ..... 74
Encart 1 La relation entre Débit, Vitesse et Concentration pour caractériser les conditions générales de la circulation routière par section ..... 75
2.3.3. Les avantages des capteurs fixes pour la mesure du trafic ..... 75
*Conclusion :* ..... *78*
CHAPITRE 3. DU COMPTAGE ROUTIER AU DISPOSITIF DE QUANTIFICATION DU TRAFIC ROUTIER : LE CAS DE LA VILLE DE DIJON ..... 81
*3.1. Un rôle pivot des données de comptages routier pour la régulation de la circulation routière en milieu urbain* ..... *83*
3.1.1. Les capteurs électromagnétiques : fonctionnement et nature des mesures pour le suivi du trafic routier ..... 83
3.1.2. La gestion et la régulation de la circulation routière en milieu urbain : le rôle central des comptages routiers ..... 85
Encart 2 L'exemple des ondes vertes : une régulation globale de la circulation urbaine ..... 88
*3.2. Présentation du réseau de mesure de la ville de Dijon* ..... *89*
3.2.1. La gestion du trafic routier de la ville de Dijon ..... 89
a. La cellule régulation de la circulation de la ville de Dijon ..... 89
b. Le réseau de comptage routier de la ville de Dijon ..... 91
3.2.2. Les données de comptages recueillies : un réseau de comptage hiérarchisé à l'image du réseau routier ..... 94
3.2.3. Du comptage au dispositif de quantification du trafic routier ..... 97
a. Homogénéisation et stockage des données de comptages routier ..... 99

b. Spatialisation et caractéristiques des stations de mesures au sein du SIG ..... 99
Encart 3 Un traitement spécifique pour les données de la DIR Centre-Est ..... 100
c. Une gestion spatio-temporelle des données de comptages par l'intermédiaire d'un SGBD et d'un SIG intégrés ..... 101
*3.3. Que nous révèlent les capteurs de la ville de Dijon pour la visualisation du trafic routier en milieu urbain ?* ..... *102*
3.3.1. État des lieux des capteurs électromagnétiques : un réseau de mesures tributaire de l'évolution du réseau urbain ..... 102
a. Qualification des données de comptage : une analyse temporelle de la fiabilité du réseau de capteurs ..... 103
b. État des lieux du réseau de mesures sur la période d'étude : disponibilité et évolution du réseau de capteurs dans le temps ..... 106
3.3.2. Se focaliser sur la temporalité des données pour révéler les rythmes de l'espace urbain ..... 109
3.3.3. Les défis à relever pour l'étude spatio-temporelle du trafic routier ..... 114
CONCLUSION DE LA PREMIÈRE PARTIE ..... 117

**PARTIE 2 - DE LA MODÉLISATION À LA SIMULATION MULTI-AGENTS DU TRAFIC ROUTIER POUR LA PAA : UNE APPROCHE ORIENTÉE SUR LES DONNÉES ..... 119**

INTRODUCTION DE LA DEUXIÈME PARTIE ..... 121
CHAPITRE 4. MODÉLISER ET SIMULER LE TRAFIC ROUTIER EN MILIEU URBAIN POUR LA QUALITÉ DE L'AIR : DÉFIS ET ENJEUX ..... 125
*4.1. Un retour sur les concepts de modèle(s) et de simulation en sciences humaines* ..... *127*
4.1.1. Un « *modèle de modèles* » pour construire une démarche de modélisation ..... 127
4.1.2. La simulation pour animer et explorer les modèles : le cas des *modèles de simulation* et de *géosimulation* ..... 130
*4.2. Quel modèle de trafic retenir pour l'étude de la PAA ?* ..... *133*
4.2.1. Quels modèles de trafic, pour quels usages ? ..... 133
Encart 4 Le formalisme de la modélisation à quatre étapes ..... 136
4.2.2. Du macroscopique au microscopique : quelle classification retenir ? ..... 137
4.2.3. Quelle approche de modélisation du trafic routier pour l'étude de la pollution atmosphérique automobile ? ..... 139
*4.3. Les enjeux liés à une approche de simulation du trafic routier à partir des comptages routiers* ..... *141*
4.3.1. Les enjeux liés à la matrice origine-destination : un élément clé pour la simulation du trafic routier ? ..... 142
4.3.2. Les enjeux liés à la matrice technique : de la quantification à l'évaluation de la pollution atmosphérique automobile ..... 145
4.3.3. Les enjeux liés à la formalisation du modèle : entre approche mathématique à approche comportementale du trafic routier ..... 148
*Conclusion :* ..... *150*
CHAPITRE 5. LES APPORTS DE LA SIMULATION MULTI-AGENTS POUR L'ÉTUDE DU TRAFIC ROUTIER ..... 151
*5.1. Le renouveau de la modélisation du trafic routier par la simulation-multi-agents* ..... *152*
5.1.1. Genèse de la modélisation *classique* du trafic routier ..... 152
5.1.2. L'émergence des critiques de la modélisation *classique* du trafic routier : « *zone don't travel ; people travel !* » ..... 154
5.1.3. L'apport des outils de géosimulation dans les modèles de trafic : un bouleversement méthodologique ..... 156
*5.2. Les approches de géosimulation pour la représentation du trafic routier : quelle méthode privilégier ?* ..... *159*
5.2.1. La géosimulation : des approches de simulation de l'espace géographique ..... 159
5.2.2. Automate Cellulaire (AC) ou Simulation Multi-Agents (SMA) pour le trafic routier ? ..... 161
5.2.3. Un pas nécessaire vers la simulation multi-agents pour la mise en mouvement des dynamiques intraurbaines ? ..... 163
*5.3. La SMA pour la modélisation du trafic routier à partir des capteurs de trafic routier : le modèle SCAUP* ..... *165*
5.3.1. Une conceptualisation spatio-temporelle du trafic routier : quelles entrées ? ..... 165
a. Quelle conceptualisation du trafic routier ? ..... 165
b. Inscrire les SMA dans la donnée : l'approche « deepening » ..... 167
c. L'approche expérimentale en SMA : du KISS au KIDS ..... 169
5.3.2. Le fonctionnement du modèle SCAUP : le rôle structurant des infrastructures de transports pour le déplacement des véhicules ..... 172
a. Une réinterprétation de l'affectation et l'écoulement des véhicules : une problématique spatio-temporelle ..... 172
b. Comment caractériser la morphologie du réseau routier ? ..... 173
c. La représentation du réseau routier : un élément fondamental du fonctionnement spatio-temporel de SCAUP ..... 174
5.3.3. Les sorties pour l'évaluation de la PAA : les enjeux d'une approche intégrée ..... 179
*Conclusion :* ..... *182*
CHAPITRE 6. EXPÉRIMENTATION ET MISE EN MOUVEMENT DU TRAFIC ROUTIER PAR LA SIMULATION MULTI-AGENTS ..... 183
*6.1. Les plates-formes de simulation multi-agents :* ..... *184*

6.1.1. Quelle classification des plates-formes de simulation multi-agents ? ..... 184
6.1.2. La plate-forme de simulation multi-agents GAMA (*GIS & Agent-based Modelling Architecture*) ..... 187
*6.2. Formalisation du modèle de simulation multi-agents : du projet à l'expérimentation de SCAUP ..... 190*
6.2.1. Agents et environnement : spatialisation et initialisation du système multi-agents SCAUP ..... 192
a. Du capteur au véhicule sur la route : formalisation des agents ..... 192
b. La génération du réseau routier : l'environnement des agents ..... 193
Encart 5 Le module de génération du réseau routier de MOBISIM : ..... 195
6.2.2. Mise en mouvement du modèle de trafic routier : formalisation des interactions au sein du modèle SCAUP ..... 196
a. Interaction agent-agent : initialisation des véhicules routiers ..... 196
b. Interaction agent-environnement : appréhender le déplacement des véhicules sur le réseau routier ..... 198
c. Disparition et stationnement des véhicules lors de la simulation ..... 201
6.2.3. Le suivi du processus de simulation par l'utilisateur ..... 203
*6.3. Expérimentation et plan d'expérience du modèle SCAUP : ..... 205*
6.3.1. Une démarche expérimentale de simulation abductive ..... 205
6.3.2. « Répliquer, répliquer et répliquer » ..... 207
6.3.3. Calibrage et ajustements des axes équipés de capteurs : contrôler les flux de trafic entrants et sortants du réseau routier ..... 210
6.3.4. L'impact de l'insertion des choix routiers : hiérarchiser les flux de trafic routier ..... 213
6.3.5. Exploration du temps de stationnement : comment gérer le nombre de véhicules présents lors du processus de simulation ? ..... 214
CONCLUSION DE LA DEUXIÈME PARTIE ..... 217

**PARTIE 3 - DU TRAFIC ROUTIER À LA POLLUTION ATMOSPHÉRIQUE AUTOMOBILE : VALIDATION, ÉVALUATION ET PERSPECTIVES ..... 219**

INTRODUCTION DE LA TROISIÈME PARTIE ..... 221
CHAPITRE 7. SIMULATION DU TRAFIC ROUTIER SUR LA VILLE DE DIJON : DE LA SIMULATION À L'ÉVALUATION ..... 223
*7.1. Exploration du modèle SCAUP à travers le cas de la ville de Dijon ..... 225*
7.1.1. Plan d'expérience de SCAUP sur la ville de Dijon : l'approche KIDS ..... 225
7.1.2. Expérimentation de SCAUP sur la ville de Dijon : données d'entrée, disponibilités et initialisation de la simulation multi-agents ..... 228
7.1.3. Les enjeux liés à une approche guidée par les données : les questionnements liés à la validation et à l'évaluation du modèle SCAUP ..... 231
a. Les principes généraux de la validation des modèles guidée par les données ..... 231
b. La méthode LOOCV : tester le modèle de simulation avec l'ensemble du dispositif de quantification à disposition ..... 233
*7.2. Analyse préliminaire de l'approche expérimentale de simulation du trafic routier ..... 235*
7.2.1. Un trafic routier généralisé et simulé sur la ville de Dijon ..... 235
7.2.2. La reproductibilité spatiale de SCAUP : comment le trafic routier est-il simulé ? ..... 239
7.2.3. La fréquentation intra-journalière : analyse du signal temporel ..... 242
*7.3. Comment choisir la meilleure simulation : de la simulation à la réalité ..... 246*
7.3.1. L'analyse statistique classique pour mesurer les relations entre données de simulation et données d'observation ..... 246
7.3.2. Un indicateur synthétique issu de l'ingénierie du trafic routier pour l'évaluation des modèles de trafic : l'indicateur GEH ..... 250
7.3.3. Un indicateur synthétique pour l'évaluation des modèles : le diagramme de Taylor (2001) ..... 254
*Conclusion : ..... 260*
CHAPITRE 8. UN DISPOSITIF D'ÉVALUATION ENVIRONNEMENTALE : DU TRAFIC ROUTIER AUX ÉMISSIONS DE POLLUANTS ATMOSPHÉRIQUES AUTOMOBILES ..... 263
*8.1. Les enjeux de la spatialisation des émissions de polluants atmosphériques automobiles ..... 265*
8.1.1. La spécificité du trafic routier : un secteur à part entière ? ..... 265
8.1.2. Présentation de la démarche d'évaluation environnementale de la PAA ..... 268
*8.2. Données d'entrée et méthodologie de calcul des émissions avec CIRCUL'AIR ..... 271*
8.2.1. La quantification du trafic routier : intégration des données issues de SCAUP ..... 271
a. Les débits de trafic routier : une donnée d'entrée spatio-temporelle ..... 272
b. Les caractéristiques relatives à la hiérarchisation du réseau routier ..... 275
8.2.2. Les données d'entrées spécifiques au calcul des émissions de PAA ..... 276
a. Ventilation du parc automobile sur la ville de Dijon : une entrée primordiale pour l'évaluation de la PAA ..... 276

b. Les profils temporels de fréquentation des axes routiers ..... 278
8.2.3. Calcul des émissions annuelles avec CIRCUL'AIR : un calcul intégré ..... 280
*8.3. Spatialisation de la pollution atmosphérique automobile sur la ville de Dijon* ..... *281*
8.3.1. Le bilan des émissions atmosphériques à l'échelle communale : un constat contrasté selon l'échelle d'étude ..... 282
8.3.2. Bilan des émissions de $NO_x$ à l'échelle des tronçons routiers de la ville ..... 284
8.3.3. Une réponse pour la spatialisation des émissions de PAA ? ..... 290
*Conclusion :* ..... *298*
CHAPITRE 9. LES PERSPECTIVES DE DÉVELOPPEMENT AUTOUR DE LA MATRICE TECHNIQUE : DE LA QUANTIFICATION À L'ÉVALUATION ..... 301
*9.1. Les réflexions autour du dispositif de quantification pour alimenter et façonner le dispositif de modélisation* ..... *303*
9.1.1. Les questionnements liés à la répartition des capteurs : un cadre d'application heuristique du modèle SCAUP ? 303
9.1.2. S'inscrire dans une approche de simulation du trafic routier mixte : la combinaison des comptages et des enquêtes O-D ..... 306
9.1.3. Une réflexion sur l'usage des capteurs urbains pour renforcer la couverture des comptages routiers : les *smart cities* et les capteurs mobiles ..... 308
a. Les smart cities : un accès facilité aux données de comptages routiers ..... 308
b. Les comptages routiers à partir des capteurs mobiles pour densifier la couverture des comptages routiers ..... 309
*9.2. Une approche de modélisation SCAUP à affiner* ..... *311*
9.2.1. L'arrivée du tramway en 2013 : quelle(s) réponse(s) du modèle à une perturbation structurelle ? ..... 311
9.2.2. Comment améliorer le processus de simulation interne au modèle SCAUP ? ..... 315
a. Améliorer la reproductibilité de SCAUP : le processus de warm-up ..... 315
b. Raffiner la matrice de transition par simulation et apprentissage ..... 317
9.2.3. D'un formalisme KISS vers un formalisme KIDS pour affiner le déplacement des véhicules sur les routes ..... 318
*9.3. S'inscrire dans une démarche intégrée d'évaluation de la PAA en milieu urbain : de l'émission à l'imission* ..... *320*
9.3.1. Un couplage nécessaire pour affiner les estimations de PAA ? ..... 320
9.3.2. L'application du modèle SCAUP sur la ville de Marrakech : Une première application de couplage avec la méthodologie COPERT ..... 322
9.3.3. Le passage aux émissions : un enjeu de recherche encore d'actualité ..... 325
9.3.4. Informer et sensibiliser le grand public par l'intermédiaire du *Géoweb* ..... 329
CONCLUSION DE LA TROISIÈME PARTIE ..... 333

**CONCLUSION GÉNÉRALE ..... 335**

**BIBLIOGRAPHIE ..... 345**

**TABLE DES MATIÈRES ..... 361**

**TABLE DES FIGURES ..... 365**

**TABLE DES TABLEAUX ..... 371**

**ANNEXES ..... I**

***Annexe 1. Synthèse des informations sur les réseaux de capteurs en France ..... ii***
***Annexe 2. Macro VBA de mise en forme des fichiers de comptages issus du SRT ..... iii***
***Annexe 3. Extraction automatique des données de comptage ..... v***
***Annexe 4. Clés de répartition du parc automobile définies sur la ville de Dijon ..... vii***
***Annexe 5. Régressions linéaires des paramètres explorés avec SCAUP ..... ix***

**ANNEXES CARTOGRAPHIQUES ..... VI**

***Annexe cartographique 1. Simulation du trafic moyen journalier à partir de SCAUP ..... vii***
***Annexe cartographique 2. Indice de reproductibilité de SCAUP ..... xv***
***Annexe cartographique 3. Cartographie du GEH à partir de SCAUP ..... xxiii***

# TABLE DES FIGURES

## PARTIE 1 : LES MESURES DU TRAFIC ROUTIER EN MILIEU URBAIN : L'APPORT DES DISPOSITIFS DE COMPTAGE PAR CAPTEUR


Figure 1.1 : L'approche traditionnelle en trois sous-systèmes urbains pour la modélisation du trafic routier (d'après Masson, 2000) ..... 27

Figure 1.2 : L'intégration des échelles dans l'étude de la pollution atmosphérique (source : Roussel, 2006, p. 9) ..... 32

Figure 1.3 : Les variations horaires de la pollution de l'air pour les villes de Stockholm, Londres en dioxyde d'azotes ($NO_2$) et en particules ($PM_{10}$) ..... 34

Figure 1.4 : Les contextes spatio-temporels pour l'étude des pollutions atmosphériques automobiles en milieu urbain ..... 35

Figure 1.5 : Le processus de construction de l'information et de la connaissance ..... 37

Figure 1.6 : Schéma et démarche de constitution de la connaissance en épistémologie des sciences (d'après Besnier, 2011) ..... 38

Figure 1.7 : L'abduction : un double processus de réflexion entre empirisme et rationalisme ..... 39

Figure 1.8 : La conceptualisation des temporalités en géographie (d'après Klein, 2007) ..... 43

Figure 1.9 : La *Géocomputation* : un positionnement interdisciplinaire pour l'analyse et la modélisation des données spatiales (source : Alex Singleton, Big data in the real world : *http ://www.alex-singleton.com/talk/2014/12/02/Big-Data-in-the-real-world/)* ..... 47

Figure 1.10 : La *Géocomputation* vue par les outils pour l'analyse des données spatio-temporelles : des approches interdisciplinaires ..... 49

Figure 1.11 : Les caractéristiques spatiales de la circulation ..... 57

Figure 1.12 : Typologie des dispositifs de quantification du trafic routier ..... 59

Figure 1.13 : Le concept de *smart city* au service de l'information publique pour et par les acteurs de l'espace urbain (les usagers et les collectivités) ..... 61

Figure 1.14 : Les différents types de capteurs en milieu urbain pour le trafic routier ..... 65

Figure 1.15 : Extrait d'une carte du trafic routier sur la ville de Dijon depuis le site Waze. La localisation des utilisateurs est individualisée ; les événements de trafic routier sont décrits par le motif de l'évènement, la vitesse du trajet, le temps de trajet impacté ..... 67

Figure 1.16 : L'exemple de comptage-trajectoire du trafic automobile par relevés GPS ..... 68

Figure 1.17 : Un système NFC placé sur le mobilier urbain (source : http://www.orange-business.com/fr/clear-channel) ..... 72

Figure 1.18 : Illustration d'une boucle électromagnétique située sur le Boulevard Paul Doumer, à l'est de la commune de Dijon ..... 73

Figure 1.19 : Les relations entre débit Q, vitesse V et concentration D ..... 75

Figure 1.20 : Principe de fonctionnement d'une boucle électromagnétique pour le comptage du trafic routier (d'après Borsu, 2004) ..... 84

Figure 1.21 : Exemple type d'un plan de feux en un cycle de deux phases sur un carrefour en milieu urbain (d'après Faye *et al.*, 2012) ..... 86

Figure 1.22 : Principe de fonctionnement et de synchronisation d'une onde verte sur un tronçon routier à sens unique ..... 88

Figure 1.23 : Illustration du PC de la circulation routière de la ville de Dijon (au centre, l'écran de contrôle à diodes des carrefours à feux ; de chaque côté de l'écran de contrôle des écrans numériques permettent de visualiser l'état du trafic routier en direct par l'intermédiaire de caméras) ..... 90

Figure 1.24 : Illustration des dispositifs de détection autonomes pour les feux tricolores sur un même carrefour entre le Boulevard Gabriel et l'Avenue Alain Savary (à gauche : un radar de détecteur de présence ; à droite : un capteur électromagnétique) ..... 90

Figure 1.25 : Le réseau de comptage routier permanent et ponctuel géré par la cellule de régulation de la circulation de la ville de Dijon ..... 92

Figure 1.26 : Répartition des villes en fonction de la densité des réseaux de comptage ..... 93

Figure 1.27 : Localisation des comptages automatiques effectués sur le territoire de l'agglomération dijonnaise ..... 95

Figure 1.28 : Classification en 9 classes selon la méthode des K-means des axes routiers en fonction du TMJA

observé pour 828 axes routiers mesurés sur la période 2003-2013 (deux sens confondus) ..... 96
Figure 1.29 : Du dispositif de comptage au dispositif de quantification : traitement des comptages routiers permanents collectés au sein des CIGT ..... 98
Figure 1.30 : Architecture des données de comptages routiers au sein du SRT (à gauche) et pour le SGBD (à droite) ..... 99
Figure 1.31 : Diagramme de classe UML formalisant le réseau de comptage routier sous PostGIS : à gauche la structuration dite « temporelle » (la relation entre les capteurs et les données de comptages) ; à droite, la structuration dite « spatiale » (la localisation des capteurs sur le réseau routier) ..... 100
Figure 1.32 : Codage du sens de comptage des stations de mesures sur le réseau routier (1 : correspond au sens de comptage identique au sens de digitalisation du réseau routier et inversement pour un sens de comptage inverse au sens de digitalisation) ..... 101
Figure 1.33 : Processus pour le traitement statistique des données de comptages routiers ..... 102
Figure 1.34 : Illustration du traitement des VA d'une station de comptage pour le mois de Septembre 2008 (en abscisse les valeurs de débit ; en ordonnée une journée de comptage correspondant à 96 comptages) ..... 104
Figure 1.35 : Traitement des VM par station de comptage (à gauche, une station disponible pour l'ensemble de la période 2001 à 2012 ; à droite, une station disponible pour 7 années : 2003 à 2009) ..... 105
Figure 1.36 : Périodicité des valeurs manquantes sur l'ensemble de la base de données trafic (les pourcentages sont définis sur l'ensemble des valeurs de débit au quart d'heure de la période 2001-2012) ..... 106
Figure 1.37 : Représentation des valeurs manquantes sur la période 2001-2012 (à droite, traitement sans distinction de la disponibilité des stations ; à gauche, traitement avec filtrage des stations disponibles) ..... 107
Figure 1.38 : Disponibilité des 208 stations de comptage sur la ville de Dijon pour la période 2001-2012 : une illustration des impacts de l'implantation du tramway de Dijon sur le réseau routier ..... 108
Figure 1.39 : Pourcentage de disponibilité des stations par mois et année sur l'ensemble de la période de recueil (en abscisse les mois ; en ordonnée les années ; le gradient de couleur illustre le pourcentage de disponibilité des stations pour 208 stations) ..... 109
Figure 1.40 : Caractéristiques de la fréquentation routière de la ville de Dijon sous la forme de cycles temporels pour la période 2002 -2012 en quarts horaires pour les 208 PM permanents: ..... 111

## PARTIE 2 : DE LA MODÉLISATION À LA SIMULATION MULTI-AGENTS DU TRAFIC ROUTIER POUR LA PAA : UNE APPROCHE ORIENTÉE SUR LES DONNÉES

Figure 2.1 : La démarche de modélisation : un processus de formalisation en trois étapes ..... 128
Figure 2.2 : Chronologie des approches de simulation en sciences humaines : les apports des outils informatiques (d'après Gilbert et Troitzsch, 2005, p. 7) ..... 131
Figure 2.3 : Enjeux, données et temporalités des études sur la circulation routière : une place prépondérante de la modélisation ..... 134
**Figure 2.4 : La structure séquentielle du modèle à quatre étapes (d'après Ortúzar et Willumsen (2011), p. 21)** ..... 136
Figure 2.5 : Le lien entre échelle de temps et représentation des véhicules : une descente des échelles spatio-temporelles à prendre en compte ..... 138
Figure 2.6 : Le rôle pivot de la matrice origine-destination pour la simulation du trafic routier (en rouge, le cheminement classique ; en vert, le cheminement de la thèse) ..... 142
Figure 2.7 : Revaloriser les comptages routiers pour la simulation du trafic routier : quel positionnement de la donnée dans le processus de modélisation ? ..... 145
Figure 2.8 : Une grille de lecture de la modélisation classique du trafic routier à travers le concept de la matrice technique (source : Commenges, 2013, p. 76) ..... 146
Figure 2.9 : L'imbrication des dispositifs de la matrice technique pour l'évaluation de la PAA : de l'observation à la connaissance ..... 147
Figure 2.10 : Un bouleversement méthodologique avec les approches désagrégées : l'approche *top-down* ..... 158
Figure 2.11 : Processus d'organisation et d'observation de deux approches de géosimulation : le système multi-agents (à droite) et l'automate cellulaire (à gauche) ..... 161
Figure 2.12 : La représentation d'une infrastructure routière sous la forme d'une espace discontinu (en haut) et d'un espace continu (en bas) pour le déplacement d'un véhicule ..... 162
Figure 2.13 : La conceptualisation du trafic routier sous le modèle SCAUP ..... 166
Figure 2.14 : Schéma d'intégration des données au sein du modèle SCAUP ..... 167

Figure 2.15 : Les différentes approches de modélisation en simulation multi-agents : de l'approche déductive à l'approche inductive (source : Hassan *et al.*, 2010a, p. 7) ..... 168
Figure 2.16 : Le « *fer à cheval* » : une grille de lecture à deux entrées pour positionner les modèles de simulation multi-agents (d'après Banos et Sanders, 2012, p. 54) ..... 170
Figure 2.17 : Du modèle KISS au modèle KIDS : application du cadre expérimental de SCAUP à la ville de Dijon ..... 171
Figure 2.18 : Hiérarchies du réseau routier issues de la BD TOPO pour la ville de Dijon ..... 176
Figure 2.19 : Hiérarchies du réseau routier retenues sur la ville de Dijon (en haut : les vitesses réglementaires ; en bas : la hiérarchie issue de OpenStreetMap) ..... 177
Figure 2.20 : Illustration du facteur d'émission depuis la méthodologie COPERT pour l'estimation des oxydes d'azote (source : SETRA, 2009) ..... 180
Figure 2.21 : Une classification des plates-formes de simulation multi-agents ..... 184
Figure 2.22 : Trois manières de définir un modèle SMA avec RePast : l'exemple du modèle « Zombie » (Source des images : http://repast.sourceforge.net/screenshots.php) ..... 186
Figure 2.23 : Interface graphique et environnement de travail sous la plate-forme GAMA ..... 188
Figure 2.24 : Structuration UML du modèle d'affectation des comptages routier au sein de SCAUP : du capteur au trafic routier ..... 191
Figure 2.25 : Génération de l'environnement, des agents capteurs (en noir) et des agents véhicules (en rouge) sur le réseau de la ville de Dijon ..... 193
Figure 2.26 : Formalisation UML de la génération des véhicules à partir des capteurs ..... 196
Figure 2.27 : Illustration de la génération et de l'action des agents « miroir » sur un tronçon routier pour la régulation du trafic routier ..... 197
Figure 2.28 : Formalisation UML de la dynamique de déplacement des véhicules sur le réseau routier ..... 198
Figure 2.29 : Synthèse des débits moyens journaliers comptabilisés par toutes les stations de comptages routiers (870 stations) pour la période 2001-2014 ..... 199
Figure 2.30 : Paramétrage du stationnement des véhicules au cours de la simulation selon une loi gaussienne (temps de déplacement de 15 minutes pour un écart de 5 minutes) ..... 202
Figure 2.31 : Illustration des interfaces utilisateur lors du processus de simulation à trois horaires différentes (7 h 30, 12 h 00 et 00 h 00) sur la ville de Dijon ..... 204
Figure 2.32 : Plan d'expérience du modèle SCAUP : du réseau théorique à la ville de Dijon ..... 205
Figure 2.33 : Signature de la réplication de SCAUP pour un paramètre de l'approche KISS (1000 réplications pas-à-pas pour un débit moyen de véhicules simulés) ..... 208
Figure 2.34 : Analyse de la réplication de SCAUP par le signal de variance moyen de débit de véhicules simulés (1000 réplications) ..... 209
Figure 2.35 : Indice de reproductibilité moyen (30 réplications) spatial (en haut) et temporel (en bas) du modèle de simulation sans paramétrage, ni contraintes de déplacement ..... 211
Figure 2.36 : Calibrage des axes possédant des stations de mesure (moyenne des 15 axes équipés de stations) ..... 212
Figure 2.37 : Simulation du trafic routier sans (en haut) et avec (en bas) contraintes des choix routiers (moyenne des débits simulés pour les 30 réplications) : hiérarchiser la fréquentation des axes routiers ..... 213
Figure 2.38 : Variation du temps de stationnement (en abscisse) pour trois conditions initiales de simulations (en ordonnées, à droite, est présenté l'écart-type moyen, ainsi que la moyenne des débits simulés, à gauche, pour les 30 réplications) ..... 215

## PARTIE 3 : DU TRAFIC ROUTIER À LA POLLUTION ATMOSPHÉRIQUE AUTOMOBILE : VALIDATION, ÉVALUATION ET PERSPECTIVES

Figure 3.1 : Plan d'expérience du modèle SCAUP sur la ville de Dijon ..... 225
Figure 3.2 : Les stations de mesure opérationnelles (156 postes) pour la simulation du trafic routier sur une journée représentative : le mardi 9 septembre 2008 ..... 228
Figure 3.3 : Observation du trafic routier pour les 156 stations de mesure opérationnelles : le mardi 9 septembre 2008 ..... 230
Figure 3.4 : Principe général de la validation croisée : simulation et observation du modèle de simulation multi-agents du trafic routier ..... 232
Figure 3.5 : Principe de la validation croisée et d'échantillonnage d'apprentissage pour la simulation du trafic routier à partir des stations de mesure : l'approche *Leave One Out Cross Validation* (LOOCV) ..... 234

Figure 3.6 : Débit moyen journalier pour le paramètre *R15* (moyenne journalière des 156 simulations en débit/jour) ..... 236
Figure 3.7 : Débit moyen journalier pour le paramètre *H15* (en haut) et le paramètre *S15* (en bas) : moyenne journalière des 156 simulations en débit/jour ..... 237
Figure 3.8 : Pourcentage des axes sans véhicules en fonction du temps de stationnement ..... 238
Figure 3.9 : Pourcentage de fréquentation des axes en fonction de la typologie OpenStreetMap pour les simulations établies à 15 minutes de stationnement (moyenne des débits simulés pour les 156 simulations des paramètres *R15*, *H15*, *S15*) ..... 238
Figure 3.10 : Répartition des indices de reproductibilité spatiale pour les 7 506 axes de la ville de Dijon sur l'ensemble des paramètres d'exploration ..... 239
Figure 3.11 : Reproductibilité spatiale des 156 simulations pour les paramètres de stationnement définis à 15 minutes, pour les choix routiers aléatoire (R : en haut) et hiérarchique (H : en bas) ..... 241
Figure 3.12 : Relation entre le débit simulé et la reproductibilité pour chaque axe du domaine d'étude selon le paramètre *H15* ..... 242
Figure 3.13 : Moyenne de la fréquentation intra-journalière du trafic routier pour les différents paramètres de SCAUP explorés sur la ville de Dijon (hors axes possédant une station de comptage routier : 7 350 axes routiers) ..... 243
Figure 3.14 : Reproductibilité temporelle moyenne des différents paramètres de simulation (en ordonnées) au pas de temps quart d'heure (en abscisses) ..... 245
Figure 3.15 : Coefficients de détermination ($R^2$) pour les 156 stations de comptage (en ordonnées) et pour l'ensemble des paramètres explorés (en abscisses) : le tableau associé à l'abscisse correspond à la moyenne des 156 coefficients de détermination calculés ..... 247
Figure 3.16 : Relation entre le nombre de véhicules simulés et observés par station de comptage pour les 156 stations d'observation : paramètre *H15* (les points correspondent aux stations) ..... 248
Figure 3.17 : Régression linéaire (en haut) et analyse des résidus (en bas) en moyenne temporelle des 156 stations d'observation (Nb de véhicules moyen/Quart d'heure) : paramètre *H15* (les points correspondent au quart d'heure de la journée) ..... 249
Figure 3.18 : Spatialisation du GEH pour l'ensemble des stations de comptage ..... 253
Figure 3.19 : Diagramme de Taylor (2001) de la validation spatiale (a) et temporelle (b) du plan d'exploration du modèle SCAUP ..... 256
Figure 3.20 : La contribution du secteur des transports (à gauche) et des différents modes de transport (à droite) aux émissions des principaux polluants atmosphériques en France (source : CITEPA, 2015, p. 172 : Émissions pour l'année 2013) ..... 266
Figure 3.21 : Diagramme des étapes de calcul des émissions du trafic routier (d'après PCIT, 2012, p. 168, figure 28) ..... 269
Figure 3.22 : Données nécessaires et *workflow* pour le calcul des émissions des polluants atmosphériques automobiles ..... 271
Figure 3.23 : Les flux de circulation routières insérés en entrée du calcul des émissions de PAA (en haut : la BD TOPO exploitée avec le modèle SCAUP ; en bas : la BD CARTO exploitée par ATMOSF'AIR BOURGOGNE) ..... 274
Figure 3.24 : Processus de ventilation du parc automobile sur les données de trafic routier de SCAUP pour la ville de Dijon (d'après : ATMOSF'AIR BOURGOGNE, CITEPA, 2008) ..... 277
Figure 3.25 : « Profil type » de la fréquentation des axes routiers observée à différentes temporalités selon la hiérarchie des voies issues de la BD TOPO pour l'année 2008 ..... 279
Figure 3.26 : Répartition des émissions de polluants par axe routier (en kg/m) estimées à partir de CIRCUL'AIR ..... 284
Figure 3.27 : Émissions de $NO_x$ en kg/m obtenues pour l'année 2008 par ATMOSF'AIR BOURGOGNE (source : ATMOSF'AIR BOURGOGNE : PPA, 2008) ..... 285
Figure 3.28 : Émissions de $NO_x$ en kg/m obtenues à partir du paramètre *H10* de SCAUP en moyenne horaire (en haut) et en moyenne journalière (en bas) ..... 286
Figure 3.29 : Émissions de $NO_x$ en kg/m obtenues à partir du paramètre *H15* de SCAUP en moyenne horaire (en haut) et en moyenne journalière (en bas) ..... 287
Figure 3.30 : Émissions de $NO_x$ en kg/m obtenues à partir du paramètre *H20* de SCAUP en moyenne horaire (en haut) et en moyenne journalière (en bas) ..... 288
Figure 3.31 : Taux d'évolution des émissions de $NO_x$ par paramètre de SCAUP en fonction des estimations réalisées par ATMOSF'AIR BOURGOGNE ..... 290
Figure 3.32 : Analyse par carroyage des émissions de $NO_x$ pour le paramètre de SCAUP *H10* par taux d'évolution

(en haut) et par régression linéaire (en bas) ..... 291
Figure 3.33 : Analyse par carroyage des émissions de $NO_x$ pour le paramètre de SCAUP *H15* par taux d'évolution (en haut) et par régression linéaire (en bas) ..... 292
Figure 3.34 : Analyse par carroyage des émissions de $NO_x$ pour le paramètre de SCAUP *H20* par taux d'évolution (en haut) et par régression linéaire (en bas) ..... 293
Figure 3.35 : Structure du chapitre 9 : Une réflexion autour de la matrice technique et des relations entre les dispositifs de quantification, de modélisation et d'évaluation ..... 301
Figure 3.36 : Débit moyen simulé par quart d'heure en fonction de la couverture des comptages routiers sur un réseau routier simple (KISS : 48 axes routiers) ..... 304
Figure 3.37 : Répartition de la reproductibilité par axe routier en fonction de la densité du réseau de capteurs ..305
Figure 3.38 : Principe de fonctionnement de mesure du trafic routier par VTL (*Virtual Trip Line*) sur une section autoroutière (source : Hoh *et al.*, 2008, p. 6) ..... 310
Figure 3.39 : Trafic routier simulé après l'implantation du Tramway avec les données de comptage de 2008 (en haut) et de 2013 (en bas) ..... 313
Figure 3.40 : Évaluation des débits de trafic routier simulé avec SCAUP en fonction des observations du mardi 10 septembre 2013 ..... 314
Figure 3.41 : Charger le réseau routier par un processus de *warm-up* : un processus de *pré-simulation* pour la simulation du trafic routier ..... 316
Figure 3.42 : Facteurs d'émissions unitaires d'oxydes d'azote ($NO_x$) intégrés à SCAUP en fonction des types de carburants pour deux périodes : 2007 et 2020 (source : SETRA, 2009) ..... 323
Figure 3.43 : Illustration de deux *scénarios types* d'émissions de $NO_x$ simulés par SCAUP, en fonction des véhicules : essence (à gauche) et diesel (à droite) ..... 324
Figure 3.44 : Deux itinéraires de couplage pour passer de l'émission à l'immission de la pollution atmosphérique automobile : quelle intégration des données de simulation de SCAUP pour le calcul des concentrations de PAA ? ..... 326
Figure 3.45 : Un schéma d'intégration et de monitoring des données de comptages et de simulation sous une interface de Géoweb ..... 330

# TABLE DES TABLEAUX

## PARTIE 1 : LES MESURES DU TRAFIC ROUTIER EN MILIEU URBAIN : L'APPORT DES DISPOSITIFS DE COMPTAGE PAR CAPTEUR


Tableau 1.1 : Les temporalités du système urbain (d'après Wegener, 1994, 2004) ..... 26
Tableau 1.2 : Les quatre paradigmes scientifiques ..... 41
Tableau 1.3 : Les différentes technologies de capteurs mobiles pour la mesure du trafic routier en milieu urbain (source : Borsu, 2004 ; SETRA, 2012) ..... 70
Tableau 1.4 : Les caractéristiques des capteurs fixes et mobiles pour la mesure du trafic routier en milieu urbain 78
Tableau 1.5 : La nature des mesures par type de capteurs électromagnétiques ..... 85
Tableau 1.6 : Pourcentage des comptages routiers réalisés par la ville de Dijon en fonction de la nature du réseau routier ..... 94


## PARTIE 2 : DE LA MODÉLISATION À LA SIMULATION MULTI-AGENTS DU TRAFIC ROUTIER POUR LA PAA : UNE APPROCHE ORIENTÉE SUR LES DONNÉES


Tableau 2.1 : Relation entre le type de voie issu de la BD TOPO et l'estimation des vitesses réglementaires intégrée à MOBISIM (Antoni et Vuidel, 2010) ..... 175
Tableau 2.2 : Caractéristiques générales des attributs du réseau routier (source : Lunardi, 2013) et http://thema.univ-fcomte.fr/mobisim/) ..... 195
Tableau 2.3 : Synthèse des comptages routiers en fonction de la vitesse réglementaire ..... 200
Tableau 2.4 : Synthèse des comptages routiers par type de route (hiérarchie OpenStreetMap) ..... 200


## PARTIE 3 : DU TRAFIC ROUTIER À LA POLLUTION ATMOSPHÉRIQUE AUTOMOBILE : VALIDATION, ÉVALUATION ET PERSPECTIVES


Tableau 3.1 : Nomenclature des acronymes et des codes couleur des expériences réalisées lors de l'expérimentation du modèle SCAUP sur la ville de Dijon ..... 227
Tableau 3.2 : Pourcentage de stations où le trafic routier est sous-estimé par SCAUP (nombre de stations où le trafic routier simulé est inférieur au trafic routier observé) ..... 248
Tableau 3.3 : Pourcentage des indicateurs du GEH compris entre 0 et 5 (excellents), entre 5 et 10 (bons) et supérieurs à 10 (moyens) pour l'ensemble des paramètres de simulation ..... 251
Tableau 3.4 : Comparaison de la longueur cumulée des réseaux routiers utilisés sur la ville de Dijon pour le calcul des émissions de la PAA ..... 273
Tableau 3.5 : Synthèse des émissions de $NO_x$ (cumul en kg) estimées à l'échelle de la ville de Dijon pour l'année 2008 ..... 282
Tableau 3.6 : Synthèse des émissions de $NO_x$ (cumul en kg) estimées par type de voie selon la hiérarchie de la BD CARTO ..... 283

# ANNEXES


ANNEXE 1. SYNTHÈSE DES INFORMATIONS SUR LES RÉSEAUX DE CAPTEURS EN FRANCE ..... II
ANNEXE 2. MACRO VBA DE MISE EN FORME DES FICHIERS DE COMPTAGES ISSUS DU SRT ..... III
ANNEXE 3. EXTRACTION AUTOMATIQUE DES DONNÉES DE COMPTAGE ..... V
ANNEXE 4. CLÉS DE RÉPARTITION DU PARC AUTOMOBILE DÉFINIES SUR LA VILLE DE DIJON ..... VII
ANNEXE 5. RÉGRESSIONS LINÉAIRES DES PARAMÈTRES EXPLORÉS AVEC SCAUP ..... IX

## Annexe 1. Synthèse des informations sur les réseaux de capteurs en France

Le tableau ci-dessous détaille les informations recensées sur les réseaux de capteurs dédiés à la régulation du trafic routier en France :

| Commune | Année | Nombre de capteur | Superficie estimée (sourc : INSEE, 2015) | Densité du réseau de capteur (nb de capteur/km²) |
|---|---|---|---|---|
| DIJON | 2015 | 215 | 40,41 | **5,32** |
| BELFORT | 2015 | 103 | 17,10 | **6,02** |
| BESANCON | 2015 | 220 | 65,05 | **3,38** |
| BORDEAUX | 2014 | 200 | 49,36 | **4,05** |
| MONTPELLIER | 2013 | 201 | 56,88 | **3,53** |
| NANTES | 2014 | 344 | 65,19 | **5,28** |
| REIMS | 2012 | 200 | 46,9 | **4,26** |
| ROUEN | 2015 | 118 | 21,38 | **5,52** |
| STRASBOURG | 2015 | 345 | 78,26 | **4,41** |
| TOULOUSE | 2000 | 400 | 118,3 | **3,38** |
| | | | **Moyenne** | **4,43** |
| | | | **Ecart-type** | **0,91** |

**Figure A : Synthèse des dispositifs de comptages routiers pour la régulation du trafic routier pour 10 communes Françaises**

L'ensemble des informations ont été glanées à plusieurs endroits d'ici ou de la, en effet peu de communes mettent à disposition ces informations. C'est pourquoi, nous avons mobilisé différentes canaux d'informations dont la liste est explicitée ci-dessous :

- Certaines communes mettent à disposition leurs données sur les réseaux de capteurs mobilisés comme les communes (CETE, 2013) :
  - **BORDEAUX** : http://data.bordeaux-metropole.fr/dev/exemples/format_csv.php
  - **MONTPELLIER** : http://opendata.montpelliernumerique.fr/Comptage-vehicules-particuliers,
  - **STRASBOURG** : http://www.strasbourg.eu/fr/ma-situation/professionnel/open-data/donnees/mobilite-transport-open-data/trafic-routier-cus
- D'autres informations ont été directement transmises par les services en charge de la gestion de la voirie et des faux tricolores que nous avons contactés :
  - **BELFORT** : Service de gestion de la voirie de Belfort (Bertrand Delavelle)
  - **BESANÇON** : Service de gestion de la voirie de Besançon (Clément Billet)
  - **DIJON** : Service de gestion de la voirie de Dijon (Florent Bouvier)
  - **REIMS** : http://dl-reims.fr/infos/reims/services_divers/remi/remi.htm#feux
- Enfin, pour d'autres communes, nous sommes appuyés sur des projets de recherche s'appuyant sur les données de comptages :
  - **NANTES** : PROJET EVAL-PDU : *Fouillé L. et al, 2012, « Eval-PDU: urban traffic and its environmental impacts modelling to assess Urban Mobility Master Plan : Conception of a methodology based on the Nantes case »development " , 12. p. 9*
  - **ROUEN** : PROJET MOSAIIC : *Czura G.et al, 2015, « MOSAIIC: City-Level Agent-Based Traffic Simulation Adapted to Emergency Situations ». In : Takayasu H, Ito N, Noda I, Takayasu M (éd.). Proceedings of the International Conference on Social Modeling and*

*Simulation, plus Econophysics Colloquium 2014 ; Cham : Springer International Publishing, p. 265-274.*

- **TOULOUSE** : PROJET CADDY : *Bauzer-Medeiros C. et al, 2012, « Vers un entrepôt de données pour le trafic routier ». In : Entrepôt de données et Analyse en Ligne EDA'06. , 2006, cf. p.9*

## Annexe 2. Macro VBA de mise en forme des fichiers de comptages issus du SRT

La mise en forme de l'ensemble des fichiers Excel par station de comptage a été réalisée pour une année donnée. Le code VBA étant trop long seul l'étape de mise en forme du premier jour de comptages est reporté ci-dessous (attention, le temps de calcul peut être long car le travail est réalisé sous des fenêtres actives) :

```
'----------------------------------------------------------------------
Sub MEF_FICHIER_Excel_PostGre()
'----------------------------------------------------------------------
Monfichier = Dir("*.xls")
'----------------------------------------------------------------------
'Boucle sur l'ensemble des fichiers par station de mesure'
'----------------------------------------------------------------------

    While Monfichier <> ""
        Workbooks.Open Monfichier
        Monfichier = Dir()
        classeur_actif = ActiveWorkbook.Name
        '----------------------------------------------------------
        'Suppression des feuilles inutiles'
        '----------------------------------------------------------

        Sheets(Array("Saisie", "valeur hexa", "Feuil1")).Select
        Sheets("Feuil1").Activate
        Application.DisplayAlerts = False
        ActiveWindow.SelectedSheets.Delete
            '----------------------------------------------------
            'Mise en forme des feuilles par mois = Sheets.Select'
            '----------------------------------------------------

        While Sheets <> ""

            Sheets.Select
            Rows("1:1").Select
            Selection.Delete Shift:=xlUp
            '----------------------------------------------------
            'Mise en forme des dates à partir d'un fichier
            '   source -- 'MEF_Date.xls'
            '----------------------------------------------------

            Windows("MEF Date.xls").Activate
            Sheets("janvier").Select
            Range("A1").Select
            Range(Selection, Selection.End(xlToRight)).Select
            Selection.Copy
            Windows(classeur_actif).Activate
            Sheets("janvier").Select
            Range("A1").Select
            ActiveSheet.Paste
            '----------------------------------------------------
            ' Affecter l'identifiant des stations de mesures 'nomfichier'
            '----------------------------------------------------

            Range("A2").Select
            Application.CutCopyMode = False
                ActiveCell.FormulaR1C1
                =_"=MID(CELL(""nomfichier""),FIND(""[",CELL(""nomfichier""))+1,SUM(FIND({""[""
                ;""]""},CELL(""nomfichier""))*{-1 ;1})-1)"
            Range("A2").Select
            Selection.AutoFill Destination:=Range("A2:A97")
            Range("A2:A97").Select
            '----------------------------------------------------
            ' Mise en forme de l'année de comptage
            '----------------------------------------------------

            Range("B2").Select
            'ActiveCell.FormulaR1C1 = "2001"  ' A modifier en fonction
        l'année à considérer
```

```
                ActiveWindow.WindowState = xlMaximized
                Selection.AutoFill Destination:=Range("B2:B97")
                Range("B2:B97").Select
                '----------------------------------------------------------
                ' Mise en forme de quart d'heure 'Format'
                '----------------------------------------------------------

                    Range("C1").Select
                    Range(Selection, Selection.End(xlToRight)).Select
                    Selection.NumberFormat = "[$-F400]h:mm:ss AM/PM"
                '----------------------------------------------------------
                ' Format d'une seule colonne de date (jour/mois/année heure:minute)
                '----------------------------------------------------------

                    Columns("B:B").Select
                    Selection.NumberFormat = "d/m/yy h:mm ;@"
                    Range("B3").Select
                    ActiveCell.FormulaR1C1 = "=R[-1]C+R1C4"
                    Range("B3").Select
                    Selection.AutoFill Destination:=Range("B3:B2977"), Type:=xlFillDefault
                    Range("B3:B2977").Select

                    Columns("B:B").Select
                    Selection.Copy
                    Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
                        :=False, Transpose:=False
                '----------------------------------------------------------
                'Copie sur toutes les feuilles mensuelles
                '----------------------------------------------------------

                Columns("A:C").Select
                Range("C1").Activate
                Application.CutCopyMode = False
                Selection.Copy
                '----------------------------------------------------------
                'Import des données en format colonne
                '----------------------------------------------------------

                Columns("C:C").Select
                Application.CutCopyMode = False
                Selection.Insert Shift:=xlToRight
                Range("C1").Select
                ActiveCell.FormulaR1C1 = "Comptage"
                '----------------------------------------------------------
                ' Agencement des données colonne en ligne pour une journée
                ' Fonction collage spécial transposé
                '----------------------------------------------------------

                Range("D2").Select
                Range(Selection, Selection.End(xlToRight)).Select
                Selection.Copy
                Range("C2").Select
                Selection.PasteSpecial Paste:=xlPasteAll, Operation:=xlNone, SkipBlanks:=
                            _False, Transpose:=True
                ... répétition en fonction du nombre de jour de comptage
                ...
                ...

            Wend
        '----------------------------------------------------------
        'Sauvegarde du fichier Excel
        '----------------------------------------------------------

        ActiveWorkbook.Save
        '----------------------------------------------------------
        'Fermeture de la fenêtre active
        '----------------------------------------------------------

        ActiveWindow.Close

    Wend
'----------------------------------------------------------------------------
End Sub
'----------------------------------------------------------------------------
```

Illustration du traitement sur un classeur Excel, à droite les données issues du SRT de la ville de Dijon (chaque colonne correspond à un jour et chaque ligne à un quart d'heure), à gauche, les données mis en forme par la macro ci-dessus sous trois colonnes (une ligne correspond à un quart d'heure de comptage) :

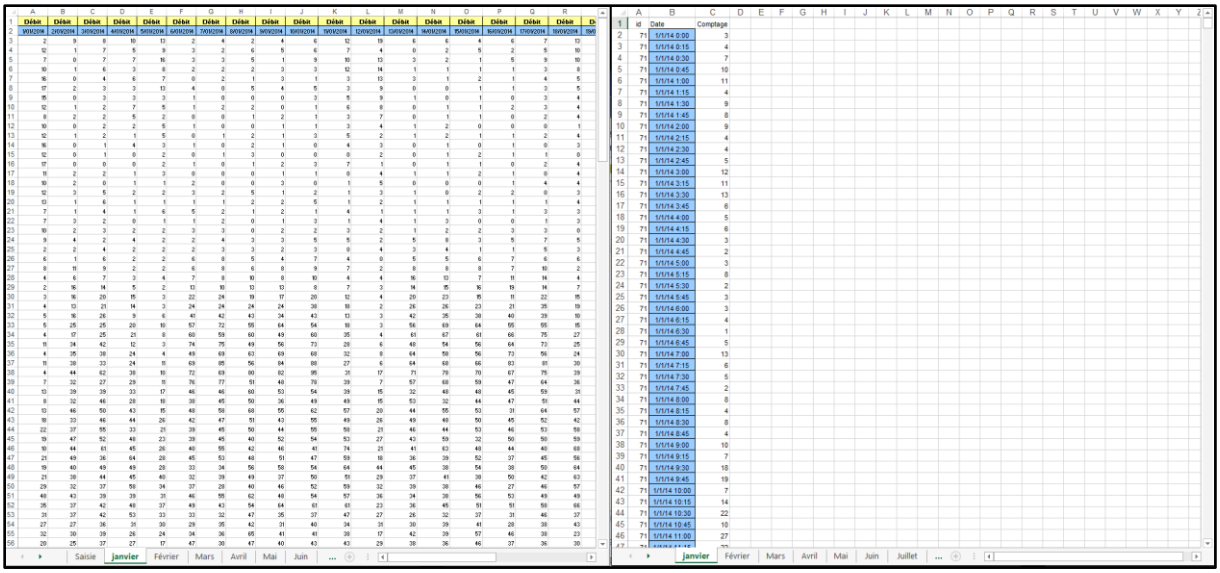


**Figure B : Mise en forme de la base de données pour l'insertion dans la base de données trafic (à gauche, structure des données brutes ; à droite, structure des données après traitement)**

## Annexe 3. Extraction automatique des données de comptage

Au cours du mois de Mai 2014, le service déplacement du Grand Dijon nous a contactés pour obtenir des données de trafic routier sur la période 2001 à 2012 sous la forme de TMJA. Cette demande s'inscrit dans les études d'impact du tramway sur la circulation routière. Les illustrations suivantes illustrent le processus d'extraction depuis la base de données trafic centralisée sous PostGIS-PostgreSQL. Le temps total de la requête pour l'extraction des données de TMJA et l'export cartographique prend 7 minutes et 6 secondes. Habituellement, le temps d'extraction depuis le SRT de la ville de Dijon varie de 1 à 2 heures pour une année de comptage.

L'ensemble de la procédure d'extraction est illustrée ci-dessous :

1- **<u>Extraction des données de TMJA sous la forme d'une requête SQL ci-dessous pour l'année 2012 (les données sont centralisée dans la table 'comptages')</u>** :

**EXTRACTION DES VALEURS PAR JOUR**

```
----------------------------------------
CREATE TABLE FREQUENTATION_2012 AS
SELECT
   id_pm,
   EXTRACT(year from date) AS annee,
   EXTRACT(month from date) AS mois,
   EXTRACT(day from date) AS jour,
   SUM(comptage) AS TMJA
FROM comptages
WHERE EXTRACT(year from date) = 2012
GROUP BY  id_pm, annee, mois, jour
ORDER BY  id_pm, annee, mois, jour ;
----------------------------------------
```

**EXTRACTION DES TABLES GEOMETRIQUES**

```
----------------------------------------
CREATE TABLE PM_TMJA_2012 AS
SELECT ROW_NUMBER
() over () AS gid,
a.id_pm, AVG(a.TMJA), b.geom
FROM   FREQUENTATION_2012 a,
          pm_cg_bus_rgf93 b
WHERE a.id_pm = b.id
GROUP BY a.id_pm, b.geom ;
----------------------------------------
```

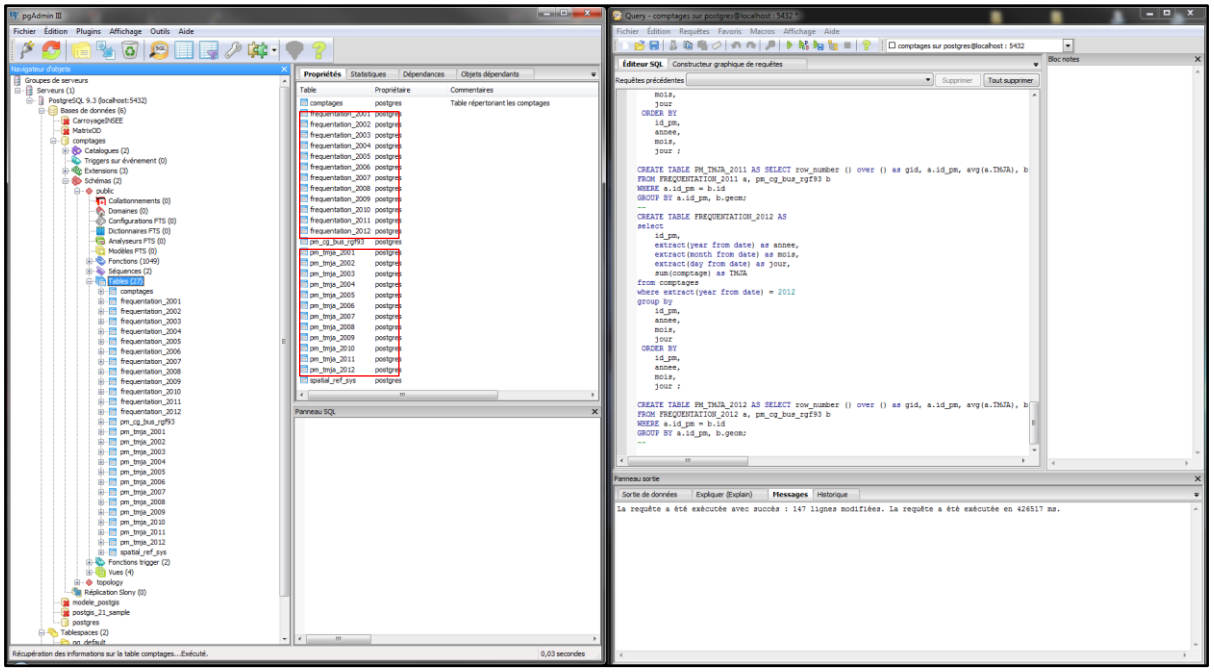


**Figure C : Extraction des valeurs de TMJA depuis PostgreSQL-PostGIS (à gauche, illustration des tables de données ; à droite, commandes SQL pour le traitement des données de TMJA par année)**

## Extraction et visualisation des fichiers géographiques sous un SIG :

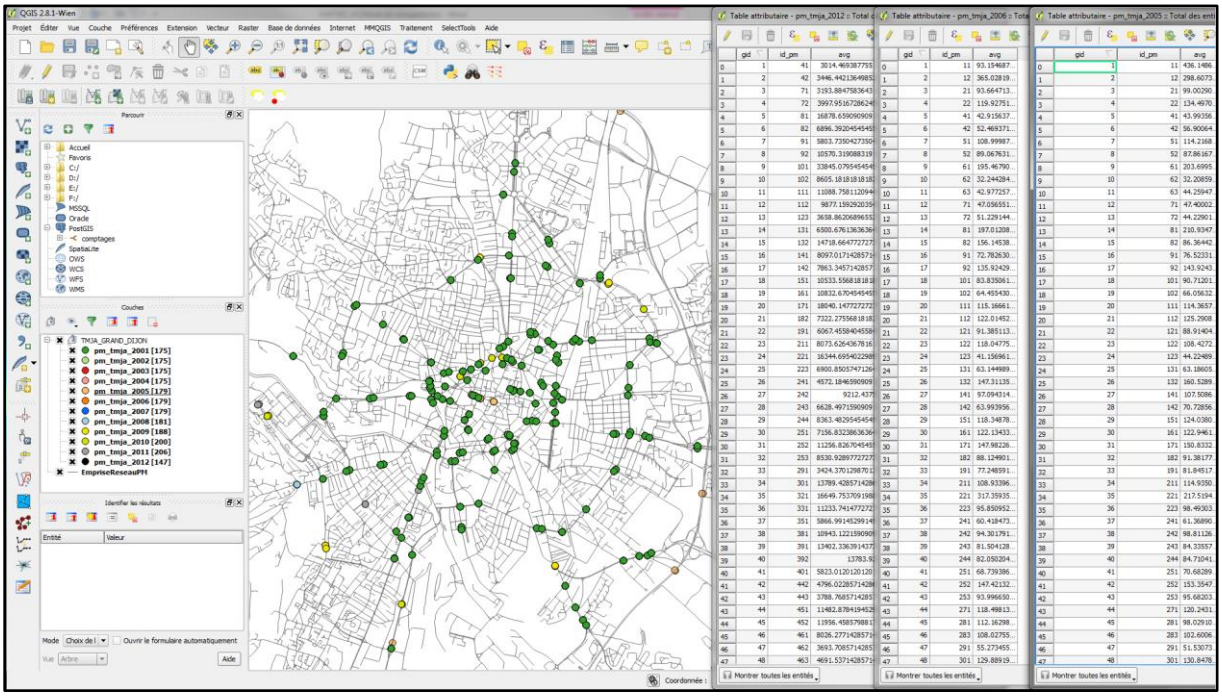


**Figure D : Illustration des données géométriques de TMJA pour les années 2001 à 2012 sous un SIG (à droite, importation des shapes ; à gauche, structure des tables attributaires avec les valeurs de TMJA par station de mesure)**

# Annexe 4. Clés de répartition du parc automobile définies sur la ville de Dijon

| Type | Sous-type | Norme | Puissance | Global | Urbain | Route | Autoroute |
|---|---|---|---|---|---|---|---|
| Véhicules Particulier (VP) | VP Essence | ECE 15/04 (1987-1992) | CC<1.4 | 0,055 | 0,057 | 0,059 | 0,048 |
| | VP Essence | ECE 15/04 (1987-1992) | 1.4<CC<2.0 | 0,040 | 0,039 | 0,037 | 0,047 |
| | VP Essence | ECE 15/04 (1987-1992) | CC>2.0 | 0,002 | 0,002 | 0,002 | 0,003 |
| | VP Essence | Euro 1 (1993-1996) | CC<1.4 | 0,108 | 0,111 | 0,114 | 0,092 |
| | VP Essence | Euro 1 (1993-1996) | 1.4<CC<2.0 | 0,056 | 0,054 | 0,052 | 0,065 |
| | VP Essence | Euro 1 (1993-1996) | CC>2.0 | 0,012 | 0,010 | 0,010 | 0,017 |
| | VP Essence | Euro 2 (1997-2000) | CC<1.4 | 0,158 | 0,163 | 0,168 | 0,136 |
| | VP Essence | Euro 2 (1997-2000) | 1.4<CC<2.0 | 0,101 | 0,098 | 0,094 | 0,118 |
| | VP Essence | Euro 2 (1997-2000) | CC>2.0 | 0,011 | 0,010 | 0,009 | 0,016 |
| | VP Essence | Euro 3 (2001-2005) | CC<1.4 | 0,144 | 0,148 | 0,153 | 0,124 |
| | VP Essence | Euro 3 (2001-2005) | 1.4<CC<2.0 | 0,102 | 0,099 | 0,095 | 0,119 |
| | VP Essence | Euro 3 (2001-2005) | CC>2.0 | 0,013 | 0,011 | 0,010 | 0,018 |
| | VP Essence | Euro 4 (2006-2010) | CC<1.4 | 0,122 | 0,126 | 0,130 | 0,105 |
| | VP Essence | Euro 4 (2006-2010) | 1.4<CC<2.0 | 0,067 | 0,065 | 0,062 | 0,078 |
| | VP Essence | Euro 4 (2006-2010) | CC>2.0 | 0,009 | 0,008 | 0,008 | 0,014 |
| | VP Diesel | Pre-Euro (...-1992) | CC<2.0 | 0,034 | 0,034 | 0,035 | 0,033 |
| | VP Diesel | Pre-Euro (...-1992) | CC>2.0 | 0,004 | 0,004 | 0,004 | 0,005 |
| | VP Diesel | Euro 1 (1993-1996) | CC<2.0 | 0,082 | 0,082 | 0,083 | 0,079 |
| | VP Diesel | Euro 1 (1993-1996) | CC>2.0 | 0,029 | 0,028 | 0,026 | 0,034 |
| | VP Diesel | Euro 2 (1997-2000) | CC<2.0 | 0,150 | 0,151 | 0,153 | 0,146 |
| | VP Diesel | Euro 2 (1997-2000) | CC>2.0 | 0,041 | 0,040 | 0,037 | 0,048 |
| | VP Diesel | Euro 3 (2001-2005) | CC<2.0 | 0,269 | 0,270 | 0,274 | 0,260 |
| | VP Diesel | Euro 3 (2001-2005) | CC>2.0 | 0,053 | 0,051 | 0,048 | 0,061 |
| | VP Diesel | Euro 4 (2006-2010) | CC<2.0 | 0,297 | 0,299 | 0,303 | 0,288 |
| | VP Diesel | Euro 4 (2006-2010) | CC>2.0 | 0,040 | 0,039 | 0,037 | 0,047 |
| | VP GPL | Pre-Euro (...-1992) | | 0,030 | 0,030 | 0,030 | 0,030 |
| | VP GPL | Euro 1 (1993-1996) | | 0,085 | 0,085 | 0,085 | 0,085 |
| | VP GPL | Euro 2 (1997-2000) | | 0,172 | 0,172 | 0,172 | 0,172 |
| | VP GPL | Euro 3 (2001-2005) | | 0,329 | 0,329 | 0,329 | 0,329 |
| | VP GPL | Euro 4 (2006-2010) | | 0,383 | 0,383 | 0,383 | 0,383 |
| Véhicules Utilitaires Légers (VUL) | VUL Essence | Pre-Euro (...-1994) | | 0,228 | 0,246 | 0,225 | 0,210 |
| | VUL Essence | Euro 1 (1995-1998) | | 0,172 | 0,169 | 0,173 | 0,174 |
| | VUL Essence | Euro 2 (1999-2000) | | 0,280 | 0,279 | 0,282 | 0,281 |
| | VUL Essence | Euro 3 (2001-2005) | | 0,244 | 0,232 | 0,246 | 0,258 |
| | VUL Essence | Euro 4 (2006-2011) | | 0,076 | 0,075 | 0,075 | 0,077 |
| | VUL Diesel | Pre-Euro (...-1994) | | 0,068 | 0,069 | 0,071 | 0,063 |
| | VUL Diesel | Euro 1 (1995-1998) | | 0,104 | 0,105 | 0,106 | 0,101 |
| | VUL Diesel | Euro 2 (1999-2000) | | 0,245 | 0,246 | 0,246 | 0,244 |
| | VUL Diesel | Euro 3 (2001-2005) | | 0,418 | 0,417 | 0,415 | 0,424 |
| | VUL Diesel | Euro 4 (2006-2011) | | 0,164 | 0,163 | 0,162 | 0,168 |
| Poids lourds | PL Rigid Diesel < 7.5 t | Pre-Euro (...-1993) | | 0,001 | 0,002 | 0,001 | 0,001 |
| | PL Rigid Diesel 7.5 t - 12 t | Pre-Euro (...-1993) | | 0,007 | 0,008 | 0,007 | 0,007 |
| | PL Rigid Diesel 12 t - 14 t | Pre-Euro (...-1993) | | 0,004 | 0,004 | 0,004 | 0,003 |
| | PL Rigid Diesel 14 t - 20 t | Pre-Euro (...-1993) | | 0,011 | 0,012 | 0,011 | 0,010 |
| | PL Rigid Diesel 20 t - 26 t | Pre-Euro (...-1993) | | 0,003 | 0,003 | 0,003 | 0,003 |
| | PL Rigid Diesel 26 t - 28 t | Pre-Euro (...-1993) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Rigid Diesel 28 t - 32 t | Pre-Euro (...-1993) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 14 t - 20 t | Pre-Euro (...-1993) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 20 t - 28 t | Pre-Euro (...-1993) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 28 t - 34 t | Pre-Euro (...-1993) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 34 t - 40 t | Pre-Euro (...-1993) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 40 t - 50 t | Pre-Euro (...-1993) | | 0,001 | 0,001 | 0,001 | 0,001 |
| | PL Rigid Diesel < 7.5 t | EURO I (1994-1996) | | 0,003 | 0,003 | 0,003 | 0,002 |

**Figure E : Clés de répartition du parc automobile définies sur la ville de Dijon : Partie 1 (Source : ATMOSF'AIR BOURGOGNE 2008)**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Poids lourds | PL Rigid Diesel 7.5 t - 12 t | EURO I (1994-1996) | | 0,014 | 0,015 | 0,014 | 0,013 |
| | PL Rigid Diesel 12 t - 14 t | EURO I (1994-1996) | | 0,005 | 0,006 | 0,005 | 0,005 |
| | PL Rigid Diesel 14 t - 20 t | EURO I (1994-1996) | | 0,022 | 0,024 | 0,022 | 0,020 |
| | PL Rigid Diesel 20 t - 26 t | EURO I (1994-1996) | | 0,007 | 0,007 | 0,007 | 0,007 |
| | PL Rigid Diesel 26 t - 28 t | EURO I (1994-1996) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Rigid Diesel 28 t - 32 t | EURO I (1994-1996) | | 0,001 | 0,001 | 0,001 | 0,001 |
| | PL Articulated Diesel 14 t - 20 t | EURO I (1994-1996) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 20 t - 28 t | EURO I (1994-1996) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 28 t - 34 t | EURO I (1994-1996) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 34 t - 40 t | EURO I (1994-1996) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 40 t - 50 t | EURO I (1994-1996) | | 0,009 | 0,008 | 0,009 | 0,010 |
| | PL Rigid Diesel < 7.5 t | EURO II (1997-2001) | | 0,009 | 0,011 | 0,010 | 0,008 |
| | PL Rigid Diesel 7.5 t - 12 t | EURO II (1997-2001) | | 0,044 | 0,048 | 0,045 | 0,041 |
| | PL Rigid Diesel 12 t - 14 t | EURO II (1997-2001) | | 0,011 | 0,013 | 0,012 | 0,011 |
| | PL Rigid Diesel 14 t - 20 t | EURO II (1997-2001) | | 0,076 | 0,083 | 0,078 | 0,071 |
| | PL Rigid Diesel 20 t - 26 t | EURO II (1997-2001) | | 0,038 | 0,038 | 0,038 | 0,038 |
| | PL Rigid Diesel 26 t - 28 t | EURO II (1997-2001) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Rigid Diesel 28 t - 32 t | EURO II (1997-2001) | | 0,007 | 0,007 | 0,007 | 0,007 |
| | PL Rigid Diesel > 32 t | EURO II (1997-2001) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 14 t - 20 t | EURO II (1997-2001) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 20 t - 28 t | EURO II (1997-2001) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 28 t - 34 t | EURO II (1997-2001) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 34 t - 40 t | EURO II (1997-2001) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 40 t - 50 t | EURO II (1997-2001) | | 0,080 | 0,071 | 0,077 | 0,085 |
| | PL Rigid Diesel < 7.5 t | EURO III (2002-2006) | | 0,013 | 0,016 | 0,013 | 0,011 |
| | PL Rigid Diesel 7.5 t - 12 t | EURO III (2002-2006) | | 0,051 | 0,056 | 0,052 | 0,047 |
| | PL Rigid Diesel 12 t - 14 t | EURO III (2002-2006) | | 0,015 | 0,017 | 0,016 | 0,014 |
| | PL Rigid Diesel 14 t - 20 t | EURO III (2002-2006) | | 0,088 | 0,097 | 0,090 | 0,083 |
| | PL Rigid Diesel 20 t - 26 t | EURO III (2002-2006) | | 0,059 | 0,058 | 0,059 | 0,059 |
| | PL Rigid Diesel 26 t - 28 t | EURO III (2002-2006) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Rigid Diesel 28 t - 32 t | EURO III (2002-2006) | | 0,020 | 0,020 | 0,020 | 0,020 |
| | PL Rigid Diesel > 32 t | EURO III (2002-2006) | | 0,001 | 0,001 | 0,001 | 0,001 |
| | PL Articulated Diesel 20 t - 28 t | EURO III (2002-2006) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 28 t - 34 t | EURO III (2002-2006) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 34 t - 40 t | EURO III (2002-2006) | | 0,001 | 0,001 | 0,001 | 0,001 |
| | PL Articulated Diesel 40 t - 50 t | EURO III (2002-2006) | | 0,209 | 0,186 | 0,203 | 0,223 |
| | PL Articulated Diesel 50 t - 60 t | EURO III (2002-2006) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Rigid Diesel < 7.5 t | EURO IV (2007-2009) | | 0,004 | 0,005 | 0,004 | 0,004 |
| | PL Rigid Diesel 7.5 t - 12 t | EURO IV (2007-2009) | | 0,013 | 0,014 | 0,013 | 0,012 |
| | PL Rigid Diesel 12 t - 14 t | EURO IV (2007-2009) | | 0,006 | 0,006 | 0,006 | 0,005 |
| | PL Rigid Diesel 14 t - 20 t | EURO IV (2007-2009) | | 0,030 | 0,033 | 0,031 | 0,028 |
| | PL Rigid Diesel 20 t - 26 t | EURO IV (2007-2009) | | 0,024 | 0,024 | 0,024 | 0,024 |
| | PL Rigid Diesel 26 t - 28 t | EURO IV (2007-2009) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Rigid Diesel 28 t - 32 t | EURO IV (2007-2009) | | 0,012 | 0,012 | 0,012 | 0,012 |
| | PL Rigid Diesel > 32 t | EURO IV (2007-2009) | | 0,001 | 0,000 | 0,001 | 0,001 |
| | PL Articulated Diesel 14 t - 20 t | EURO IV (2007-2009) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 20 t - 28 t | EURO IV (2007-2009) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 28 t - 34 t | EURO IV (2007-2009) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 34 t - 40 t | EURO IV (2007-2009) | | 0,000 | 0,000 | 0,000 | 0,000 |
| | PL Articulated Diesel 40 t - 50 t | EURO IV (2007-2009) | | 0,099 | 0,088 | 0,097 | 0,106 |
| Bus | Bus Urbain Midi < 15 t | Pre-Euro (...-1993) | | 0,018 | 0,018 | 0,018 | x |
| | Bus Urbain Midi < 15 t | EURO I (1994-1996) | | 0,011 | 0,011 | 0,011 | x |
| | Bus Urbain Midi < 15 t | EURO II (1997-2001) | | 0,034 | 0,034 | 0,034 | x |
| | Bus Urbain Midi < 15 t | EURO III (2002-2006) | | 0,056 | 0,056 | 0,056 | x |
| | Bus Urbain Midi < 15 t | EURO IV (2007-2009) | | 0,026 | 0,026 | 0,026 | x |
| | Bus Urbain Std 15 t - 18 t | Pre-Euro (...-1993) | | 0,018 | 0,018 | 0,018 | x |
| | Bus Urbain Std 15 t - 18 t | EURO I (1994-1996) | | 0,011 | 0,011 | 0,011 | x |
| | Bus Urbain Std 15 t - 18 t | EURO II (1997-2001) | | 0,032 | 0,032 | 0,032 | x |
| | Bus Urbain Std 15 t - 18 t | EURO III (2002-2006) | | 0,030 | 0,030 | 0,030 | x |
| | Bus Urbain Std 15 t - 18 t | EURO IV (2007-2009) | | 0,006 | 0,006 | 0,006 | x |
| | Bus Urbain Artic > 18 t | Pre-Euro (...-1993) | | 0,110 | 0,110 | 0,110 | x |
| | Bus Urbain Artic > 18 t | EURO I (1994-1996) | | 0,068 | 0,068 | 0,068 | x |
| | Bus Urbain Artic > 18 t | EURO II (1997-2001) | | 0,203 | 0,203 | 0,203 | x |
| | Bus Urbain Artic > 18 t | EURO III (2002-2006) | | 0,266 | 0,266 | 0,266 | x |
| | Bus Urbain Artic > 18 t | EURO IV (2007-2009) | | 0,112 | 0,112 | 0,112 | x |

**Figure F : Clés de répartition du parc automobile définies sur la ville de Dijon: Partie 1 (Source : ATMOSF'AIR BOURGOGNE 2008)**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Car | Car (Coaches) Std < 18 t | Pre-Euro (...-1993) | | 0,067 | 0,067 | 0,067 | 0,067 |
| | Car (Coaches) Std < 18 t | EURO I (1994-1996) | | 0,041 | 0,041 | 0,041 | 0,041 |
| | Car (Coaches) Std < 18 t | EURO II (1997-2001) | | 0,123 | 0,123 | 0,123 | 0,123 |
| | Car (Coaches) Std < 18 t | EURO III (2002-2006) | | 0,156 | 0,156 | 0,156 | 0,156 |
| | Car (Coaches) Std < 18 t | EURO IV (2007-2009) | | 0,058 | 0,058 | 0,058 | 0,058 |
| | Car (Coaches) 3 Axes > 18 t | Pre-Euro (...-1993) | | 0,068 | 0,068 | 0,068 | 0,068 |
| | Car (Coaches) 3 Axes > 18 t | EURO I (1994-1996) | | 0,043 | 0,043 | 0,043 | 0,043 |
| | Car (Coaches) 3 Axes > 18 t | EURO II (1997-2001) | | 0,128 | 0,128 | 0,128 | 0,128 |
| | Car (Coaches) 3 Axes > 18 t | EURO III (2002-2006) | | 0,217 | 0,217 | 0,217 | 0,217 |
| | Car (Coaches) 3 Axes > 18 t | EURO IV (2007-2009) | | 0,100 | 0,100 | 0,100 | 0,100 |
| Deux roues | Deux roues | Pre-Euro (...-1999) | 2T < 50 cc | 0,374 | 0,374 | 0,374 | x |
| | Deux roues | Euro 1 (2000-2004) | 2T < 50 cc | 0,048 | 0,048 | 0,048 | x |
| | Deux roues | Euro 2 (2005-2006) | 2T < 50 cc | 0,578 | 0,578 | 0,578 | x |
| | Deux roues | Pre-Euro (...-1999) | 2T > 50 cc | 0,015 | 0,015 | 0,016 | 0,013 |
| | Deux roues | Euro 1 (2000-2004) | 2T > 50 cc | 0,069 | 0,069 | 0,072 | 0,062 |
| | Deux roues | Euro 2 (2005-2006) | 2T > 50 cc | 0,047 | 0,047 | 0,050 | 0,043 |
| | Deux roues | Euro 3 (2007-...) | 2T > 50 cc | 0,033 | 0,033 | 0,035 | 0,030 |
| | Deux roues | Pre-Euro (...-1999) | 4T < 250 cc | 0,012 | 0,012 | 0,012 | 0,012 |
| | Deux roues | Euro 1 (2000-2004) | 4T < 250 cc | 0,102 | 0,102 | 0,100 | 0,103 |
| | Deux roues | Euro 2 (2005-2006) | 4T < 250 cc | 0,095 | 0,095 | 0,094 | 0,097 |
| | Deux roues | Euro 3 (2007-...) | 4T < 250 cc | 0,083 | 0,083 | 0,082 | 0,084 |
| | Deux roues | Pre-Euro (...-1999) | 4T - 250 - 750 cc | 0,020 | 0,020 | 0,020 | 0,021 |
| | Deux roues | Euro 1 (2000-2004) | 4T - 250 - 750 cc | 0,134 | 0,134 | 0,132 | 0,136 |
| | Deux roues | Euro 2 (2005-2006) | 4T - 250 - 750 cc | 0,093 | 0,093 | 0,092 | 0,095 |
| | Deux roues | Euro 3 (2007-...) | 4T - 250 - 750 cc | 0,065 | 0,065 | 0,064 | 0,066 |
| | Deux roues | Pre-Euro (...-1999) | 4T > 750 cc | 0,013 | 0,013 | 0,013 | 0,014 |
| | Deux roues | Euro 1 (2000-2004) | 4T > 750 cc | 0,106 | 0,106 | 0,105 | 0,108 |
| | Deux roues | Euro 2 (2005-2006) | 4T > 750 cc | 0,066 | 0,066 | 0,065 | 0,067 |
| | Deux roues | Euro 3 (2007-...) | 4T > 750 cc | 0,049 | 0,049 | 0,048 | 0,050 |

**Figure G : Clés de répartition du parc automobile définies sur la ville de Dijon : Partie 3 (Source : ATMOSF'AIR BOURGOGNE 2008)**

## Annexe 5. Régressions linéaires des paramètres explorés avec SCAUP

Le tableau ci-dessous sert d'appui en se référant aux codex couleurs ainsi qu'aux les acronymes utilisé au sein des figures suivantes :

| | Temps de stationnement | 10 min | 15 min | 20 min | 30 min | 45 min |
|---|---|---|---|---|---|---|
| Choix routiers | Aléatoire (R) | R10 | R15 | R20 | R30 | R45 |
| | Hierarchie (H)* | H10 | H15 | H20 | H30 | H45 |
| | Vitesse (S) | S10 | S15 | S20 | S30 | S45 |
| * Nomendature de la hiérarchie des routes issue de OSM | | | | | | |

**Figure H : Acronymes des tests d'exploration établis lors de l'expérimentation du modèle de simulation multi-agents du trafic routier**

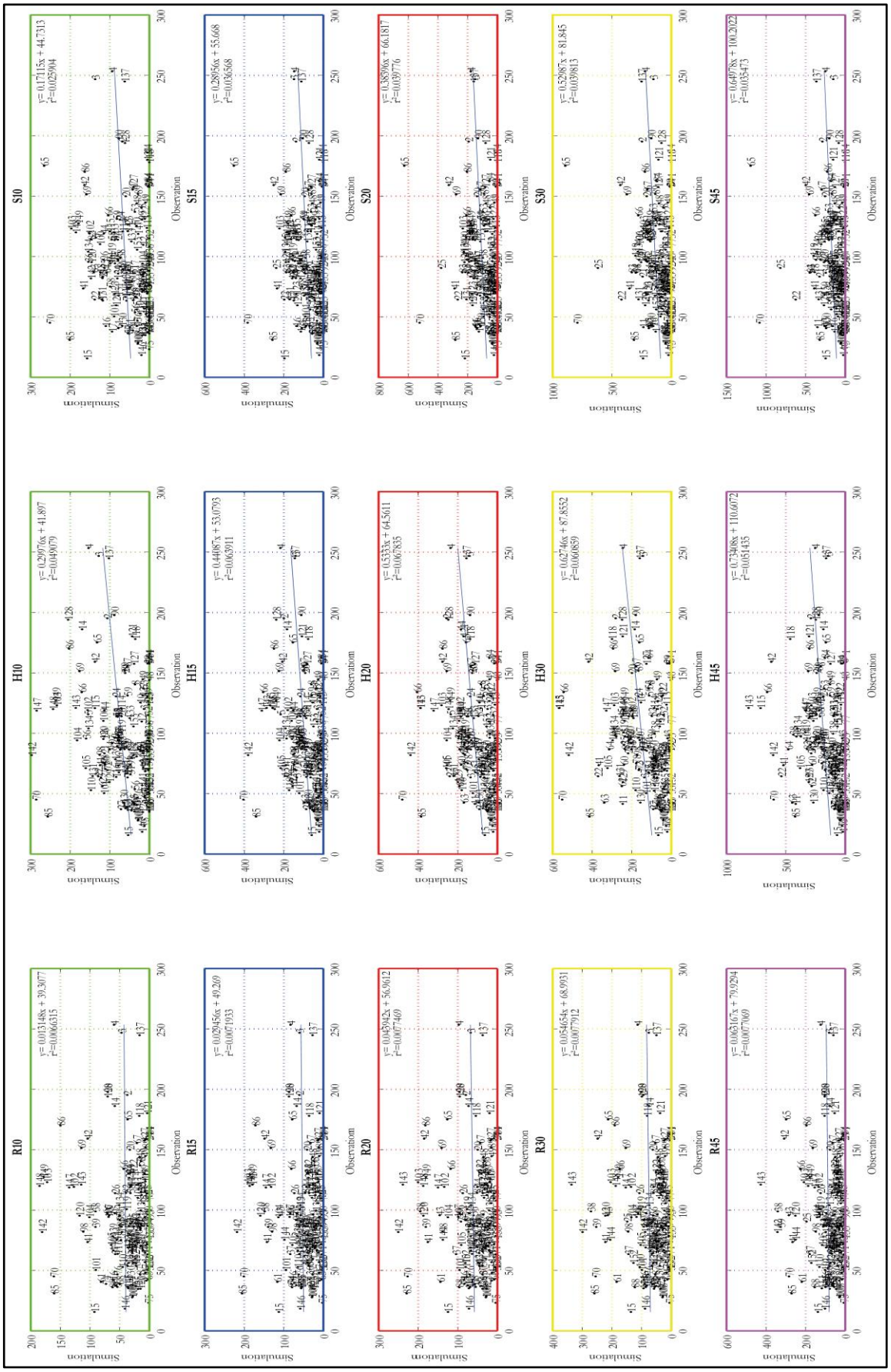



**Figure I : Relation entre le nombre de véhicules simulés et observés par station de comptage pour les 156 stations d'observations (les points correspondent aux stations)**

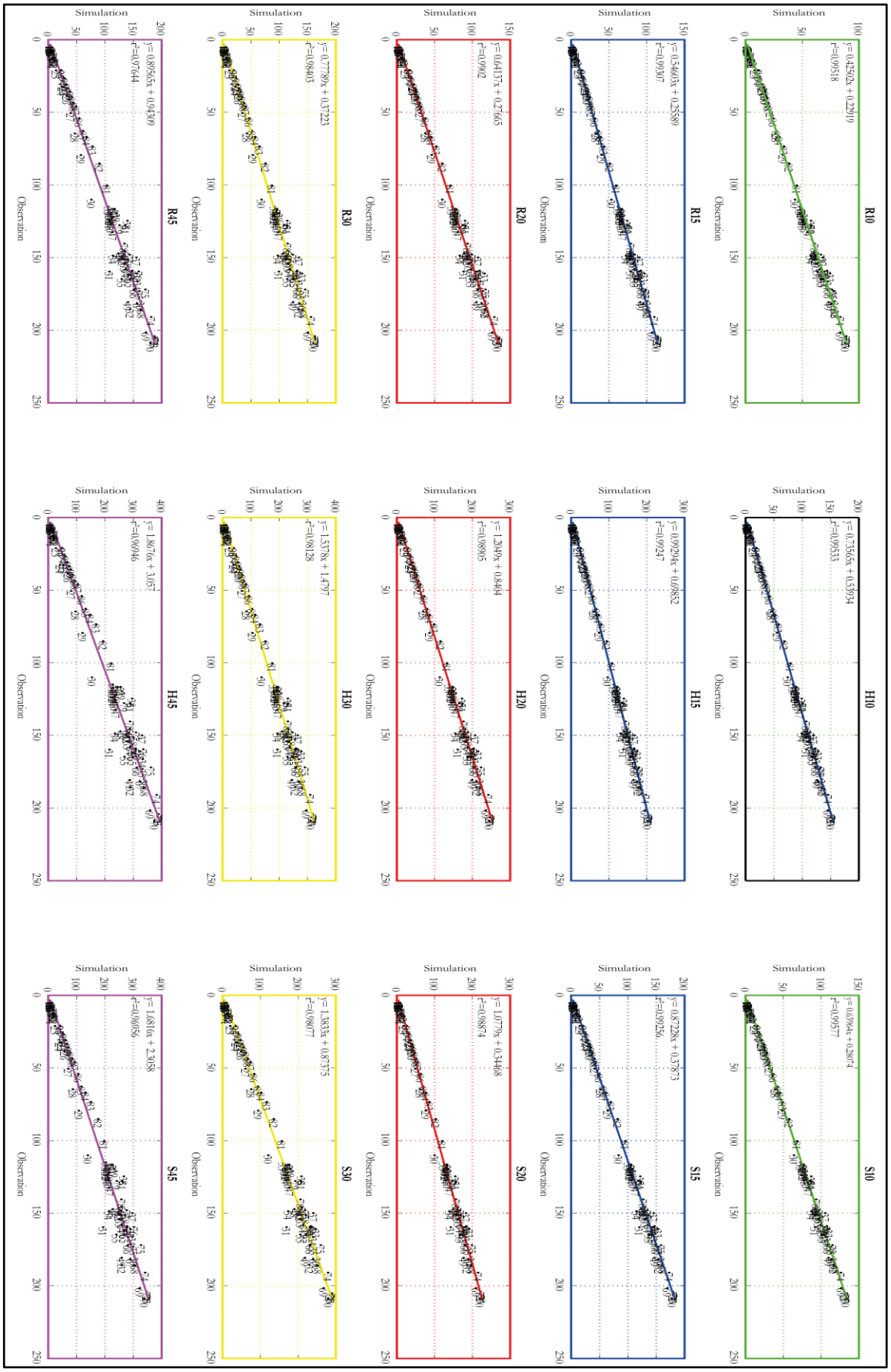



**Figure J : Régressions linéaires en moyenne temporel des 156 stations d'observations (Nb de véhicules moyen/Quart d'heure) (les points correspondent au quart d'heure)**

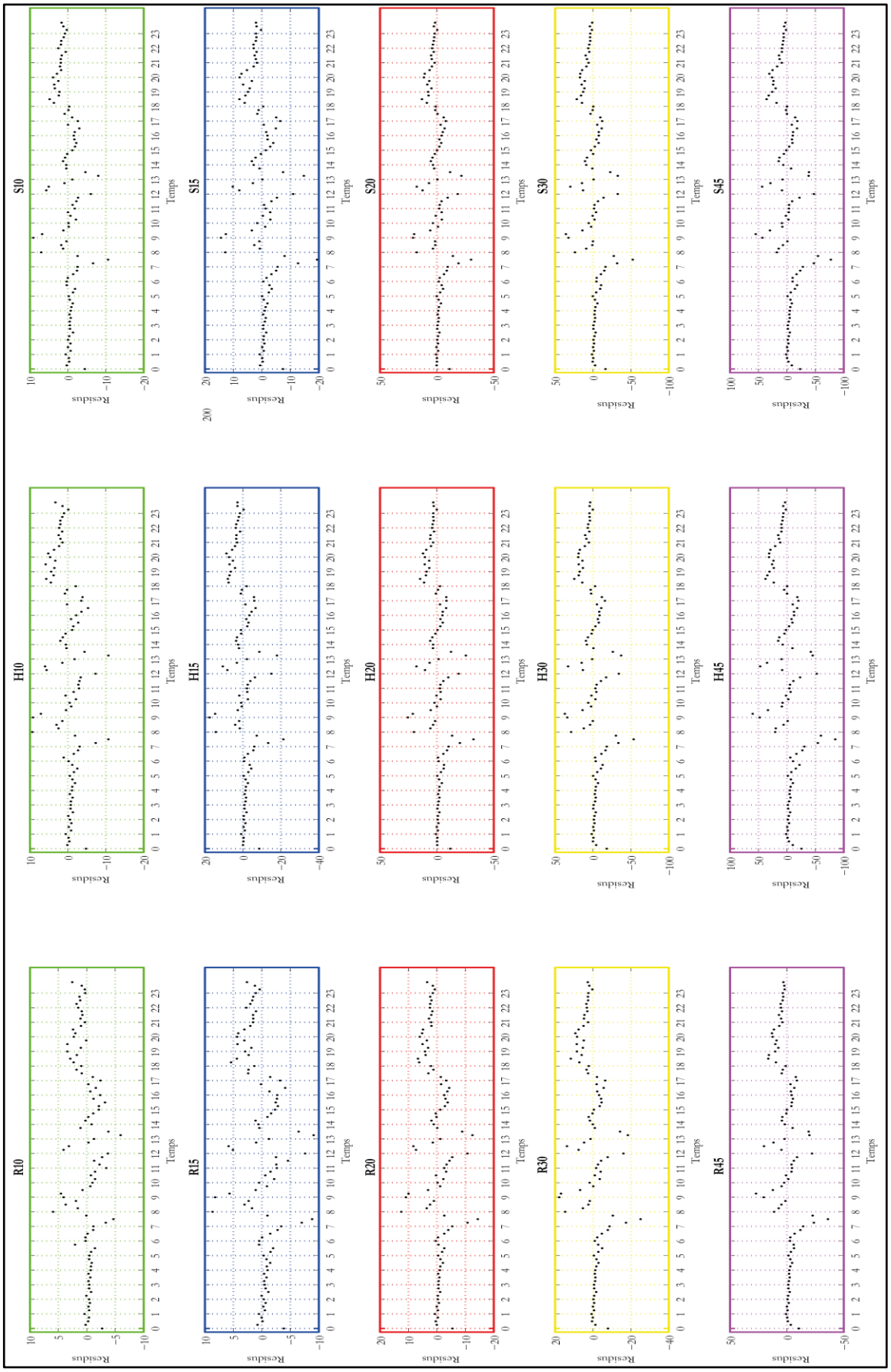



**Figure K : Analyse des résidus en moyenne temporel des 156 stations d'observations (Nb de véhicules moyen/Quart d'heure) (les points correspondent au quart d'heure)**

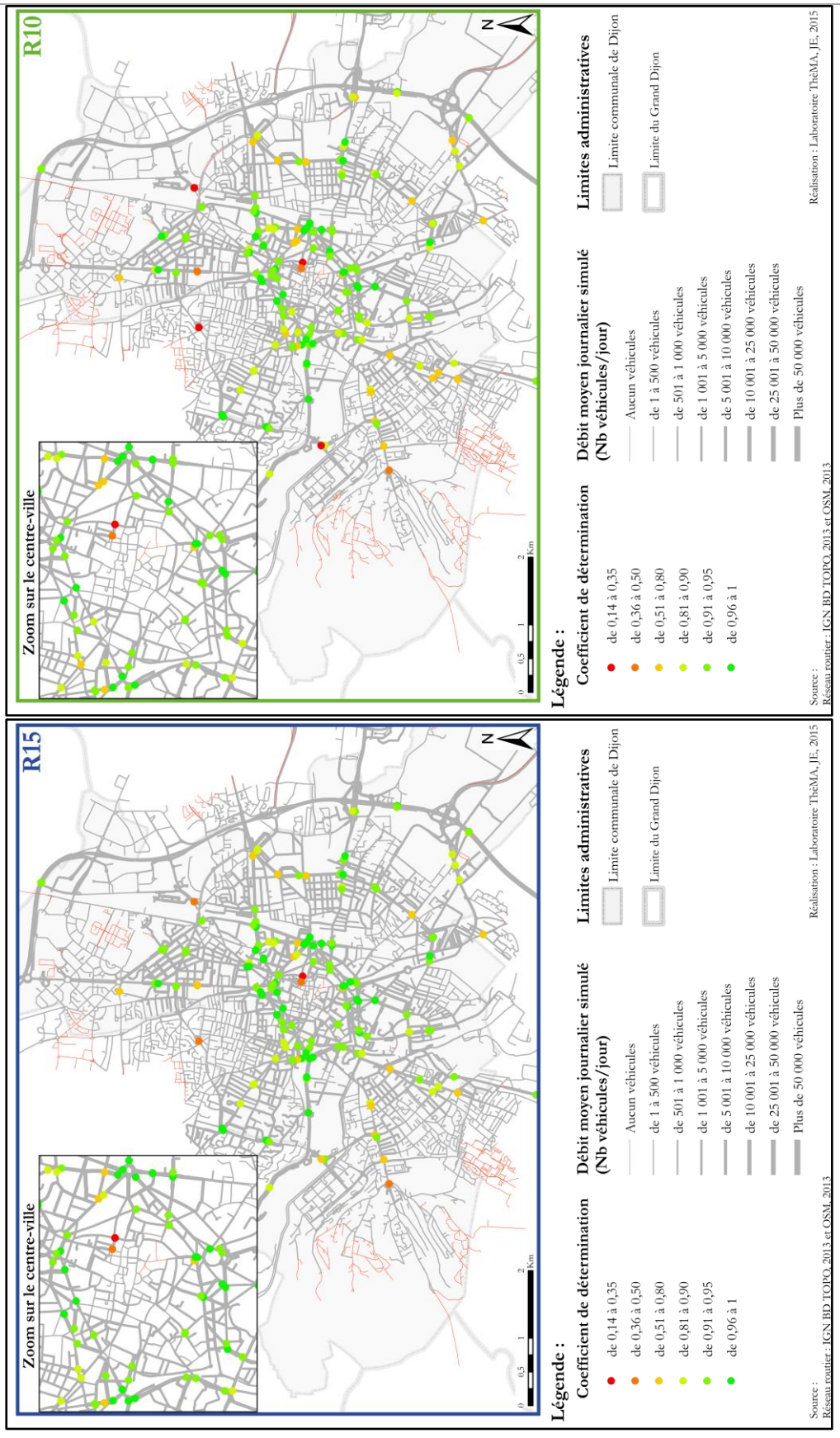
# ANNEXES CARTOGRAPHIQUES

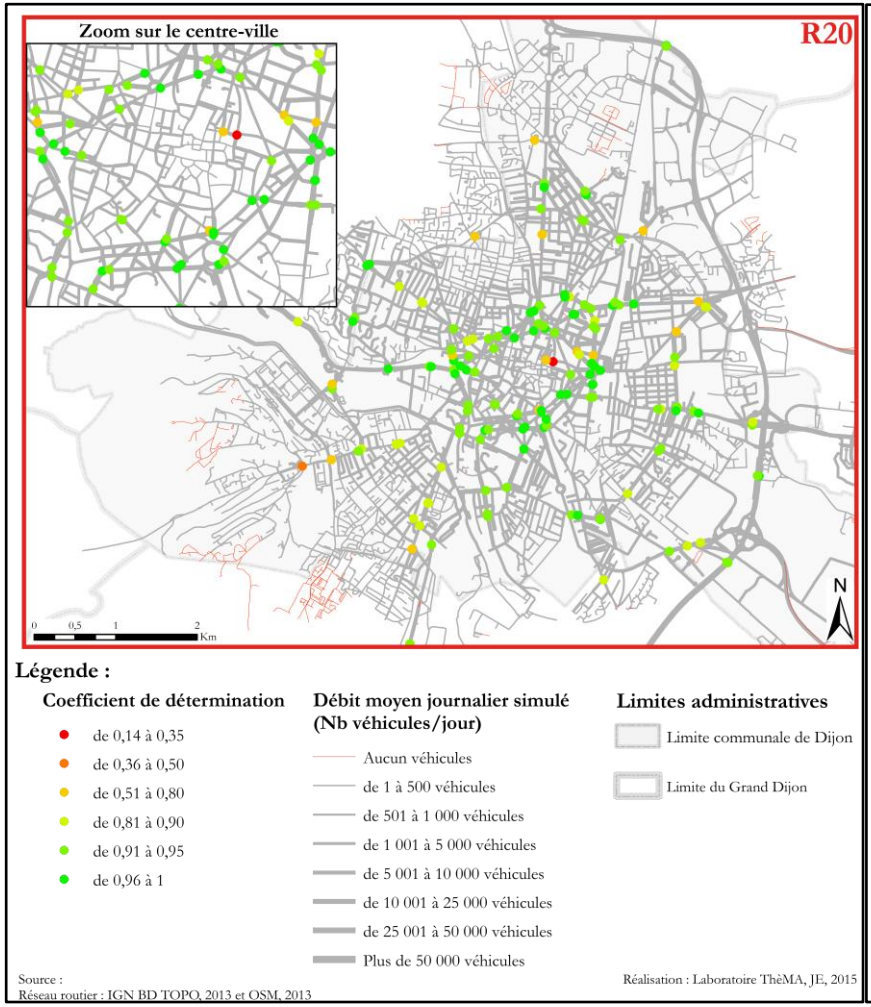

ANNEXE CARTOGRAPHIQUE 1. SIMULATION DU TRAFIC MOYEN JOURNALIER À PARTIR DE SCAUP ..... VII
ANNEXE CARTOGRAPHIQUE 2. INDICE DE REPRODUCTIBILITÉ DE SCAUP ..... XV
ANNEXE CARTOGRAPHIQUE 3. CARTOGRAPHIE DU GEH À PARTIR DE SCAUP ..... XXIII

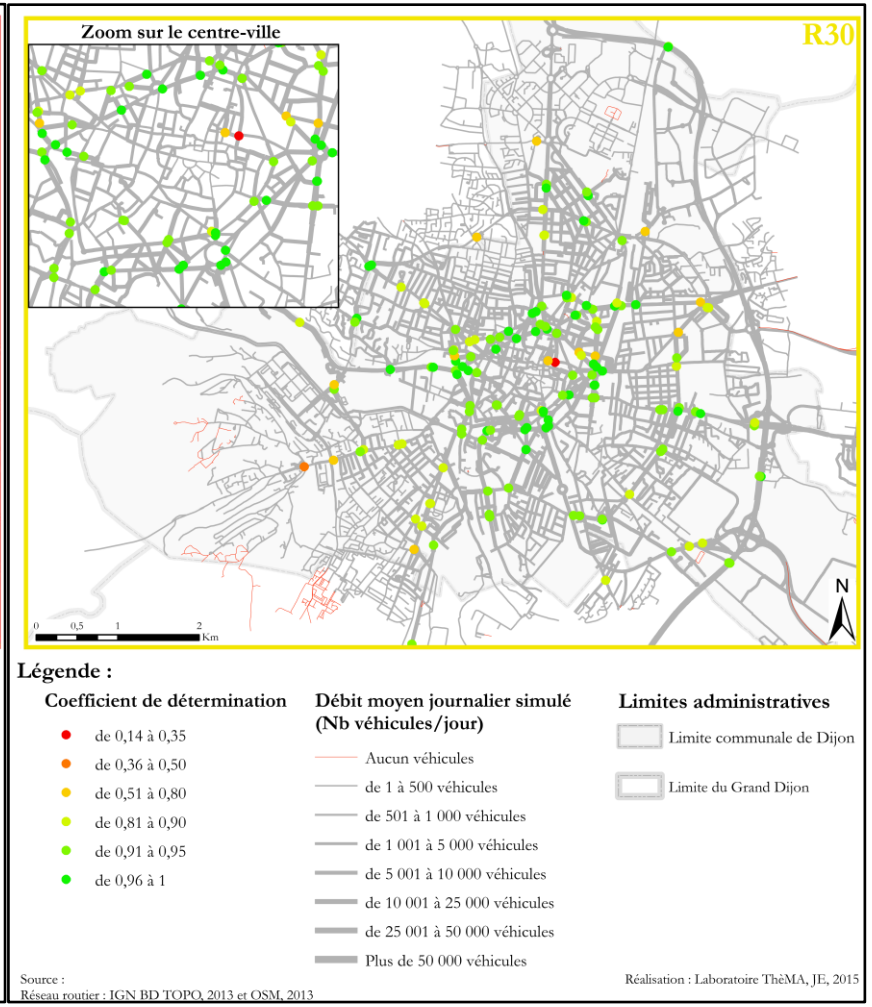

| | Temps de stationnement | 10 min | 15 min | 20 min | 30 min | 45 min |
|---|---|---|---|---|---|---|
| Choix routiers | Aléatoire (R) | R10 | R15 | R20 | R30 | R45 |
| | Hierarchie (H)* | H10 | H15 | H20 | H30 | H45 |
| | Vitesse (S) | S10 | S15 | S20 | S30 | S45 |

* Nomendature de la hiérarchie des routes issue de OSM

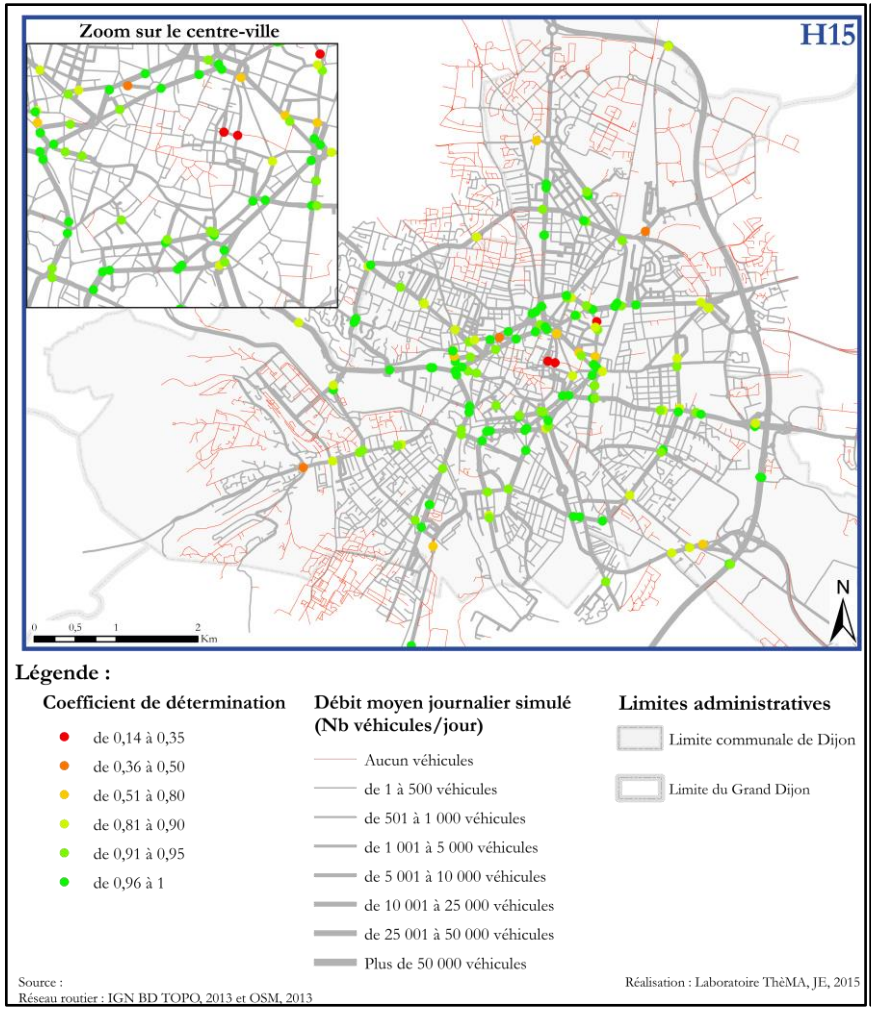
## Annexe cartographique 1. Simulation du trafic moyen journalier à partir de SCAUP

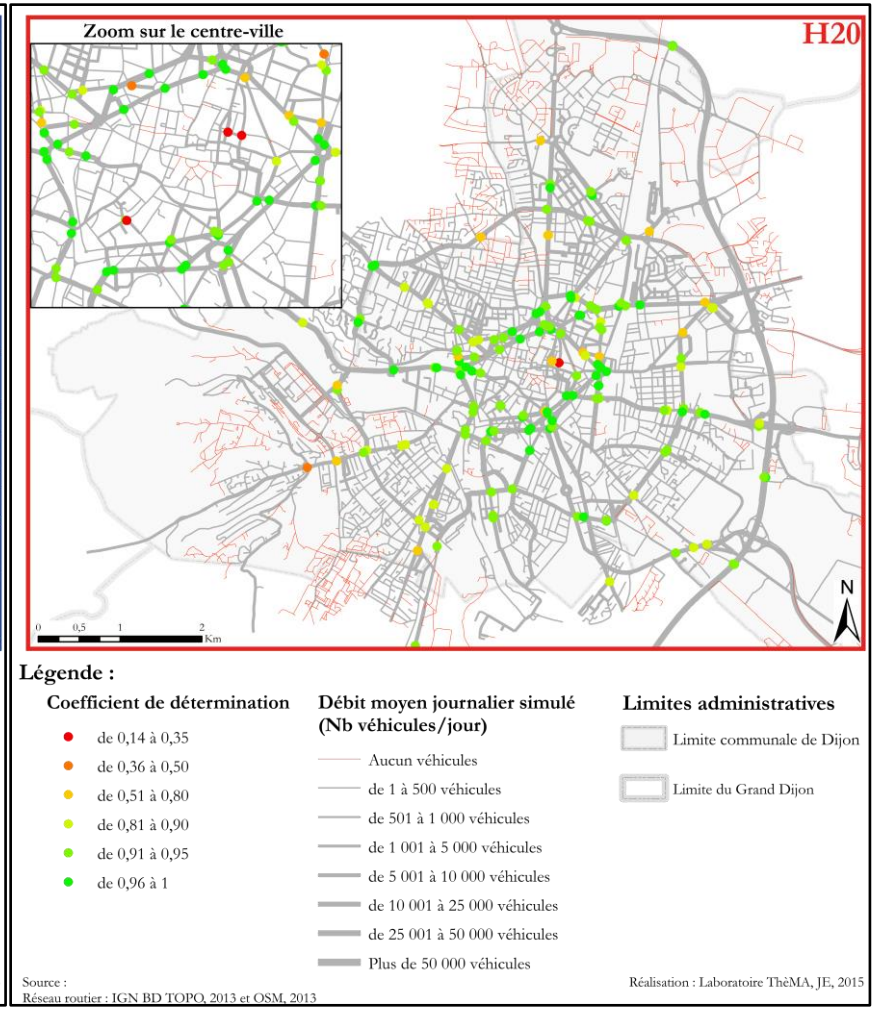



Zoom sur le centre-ville
R20
N
0 0,5 1 2 Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015

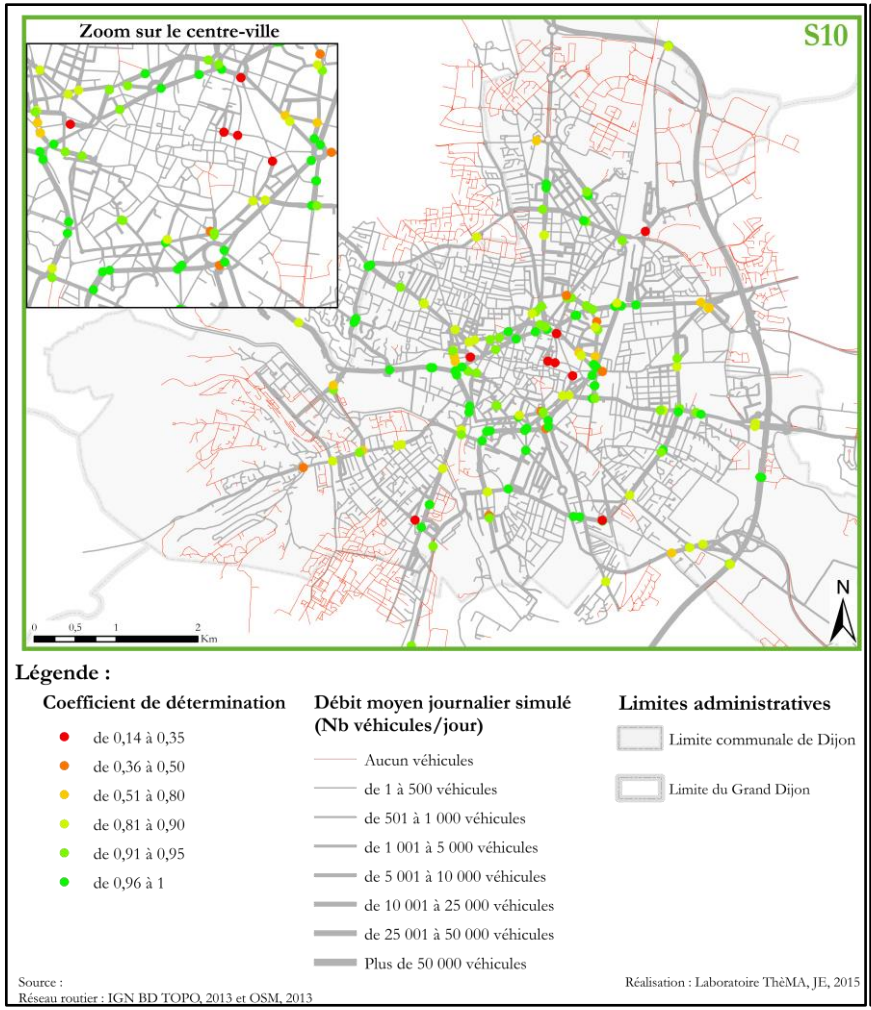

Zoom sur le centre-ville
R30
N
0 0,5 1 2 Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015

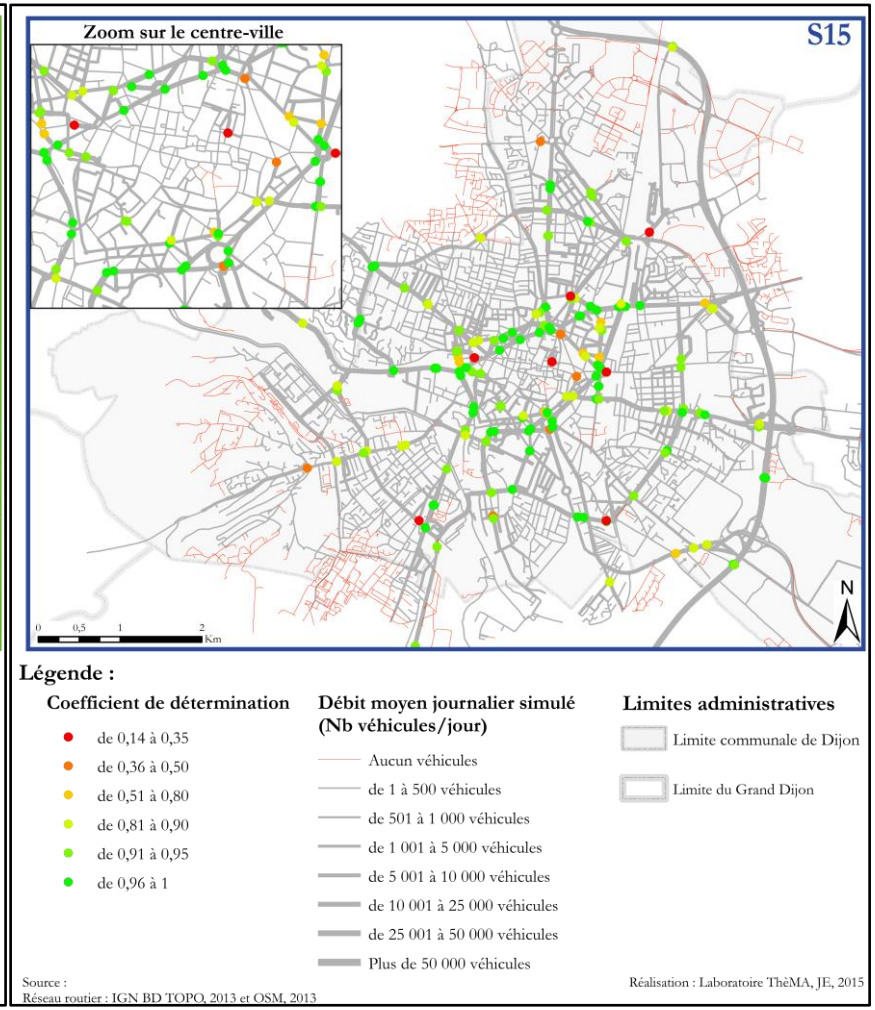

H10
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015
R45
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015

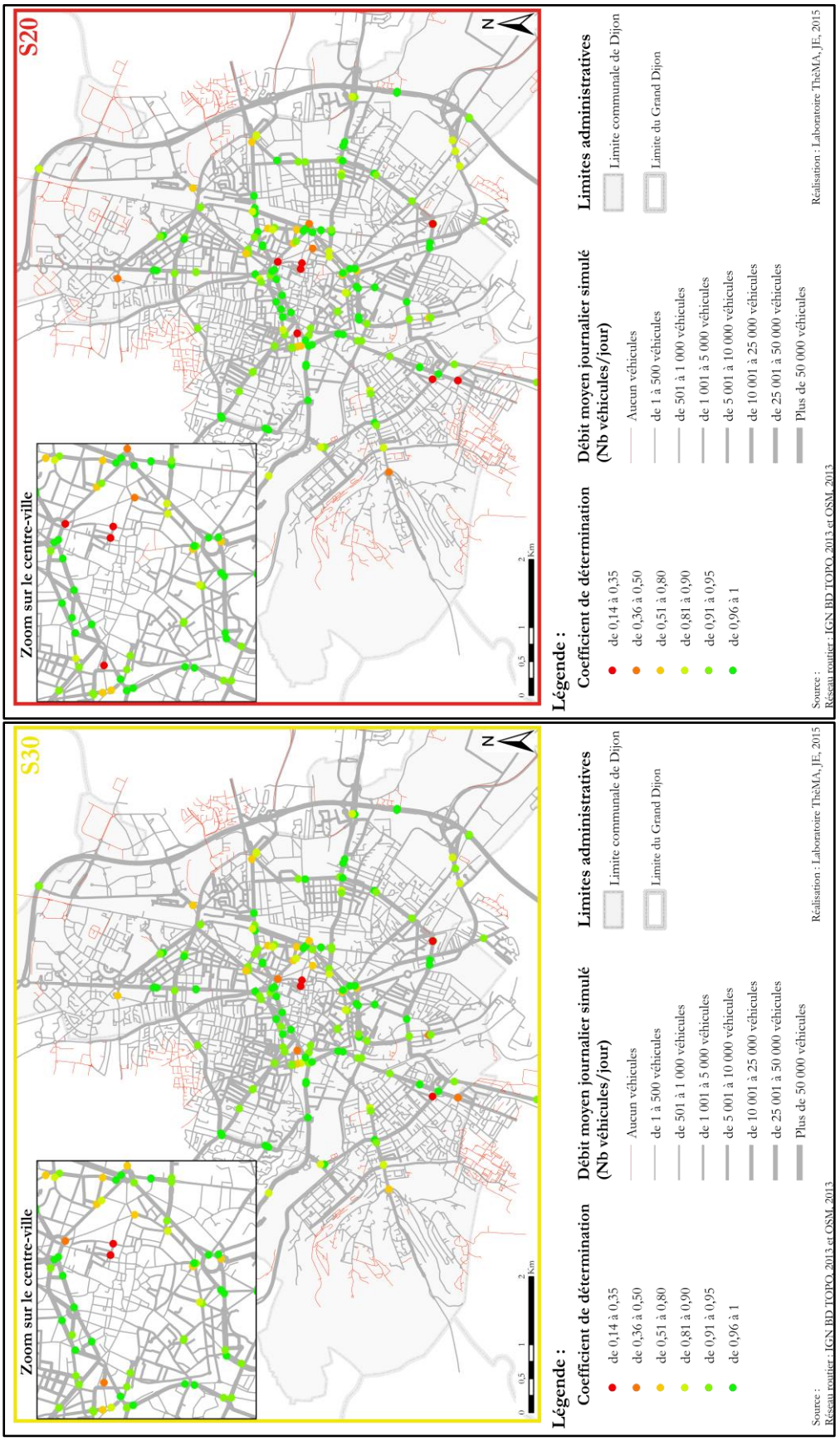

Zoom sur le centre-ville
H15
N
0 0,5 1 2 Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015

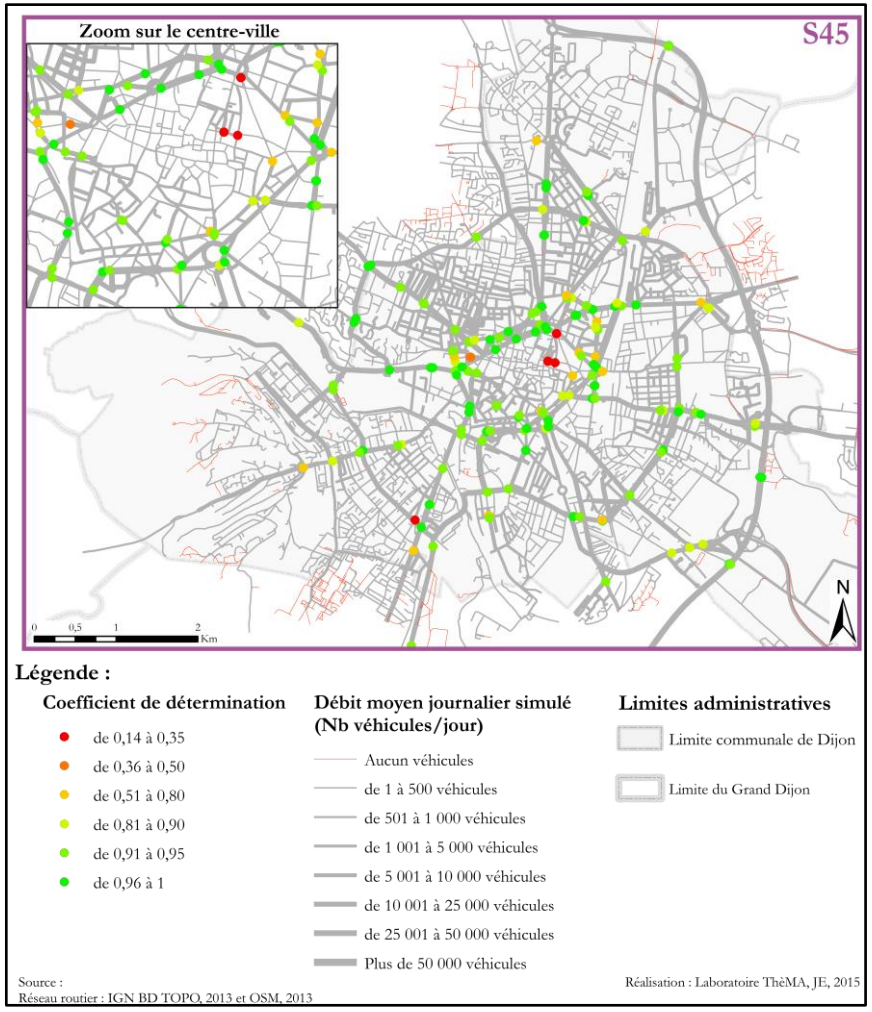

Zoom sur le centre-ville
H20
N
0 0,5 1 2 Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015


H45
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015
H30
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015

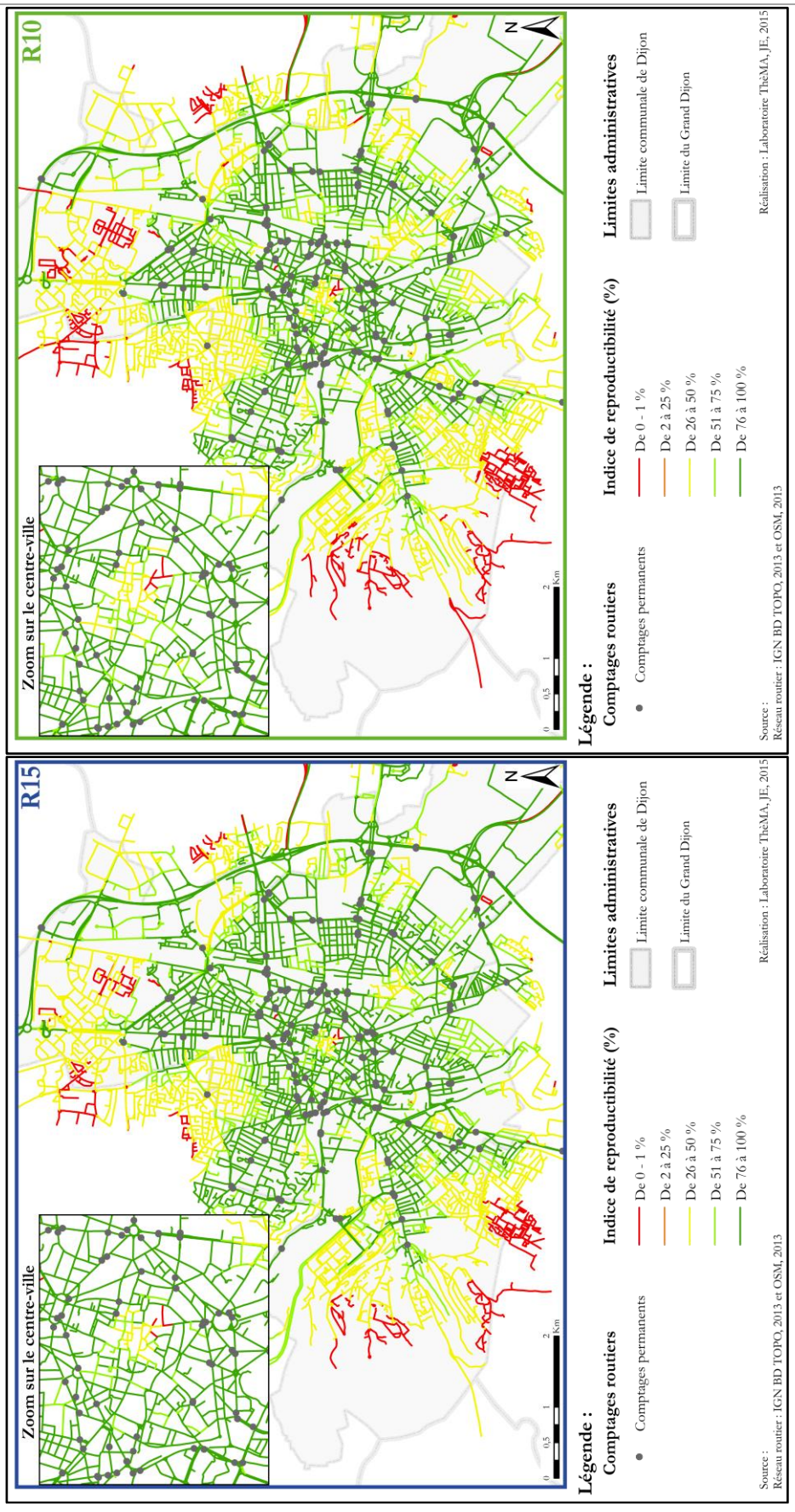

Zoom sur le centre-ville
S10
N
0 0,5 1 2 Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015

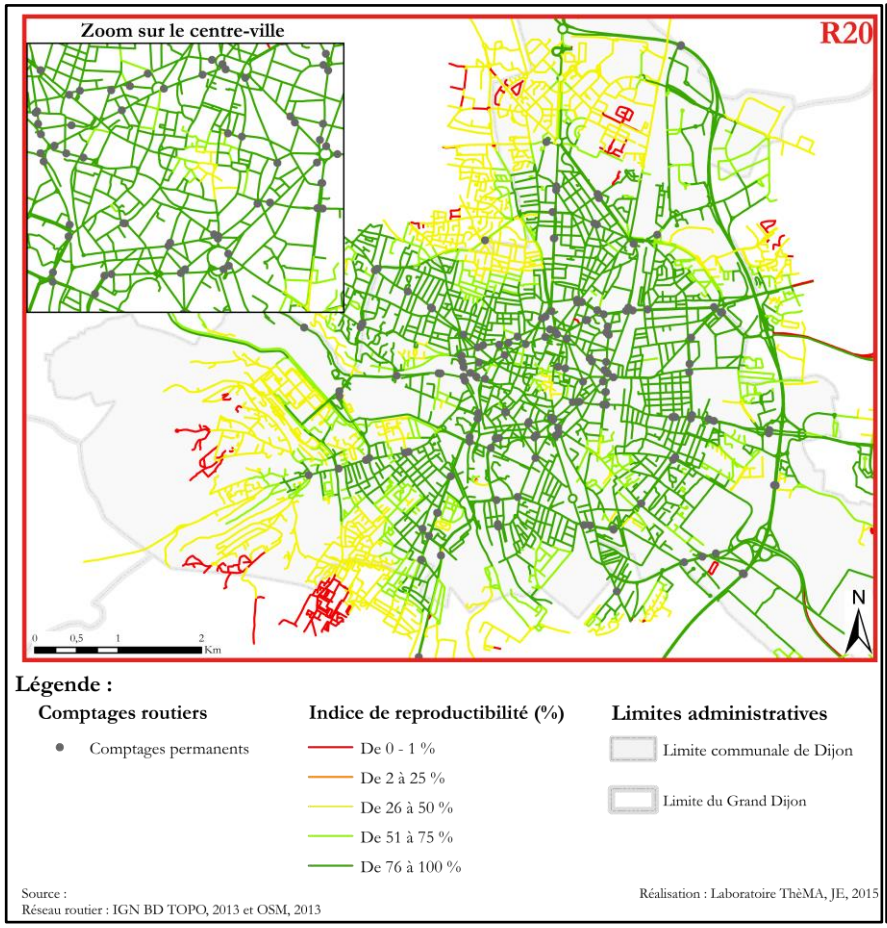

Zoom sur le centre-ville
S15
N
0 0,5 1 2 Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015

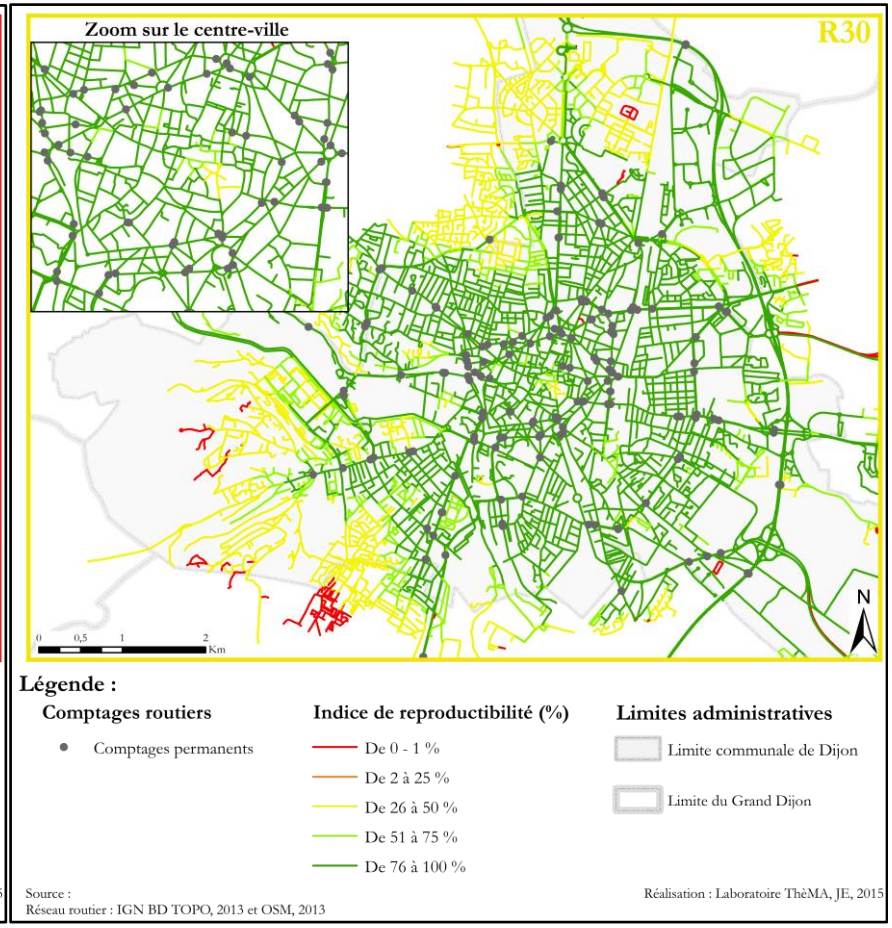

S20
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015
S30
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015

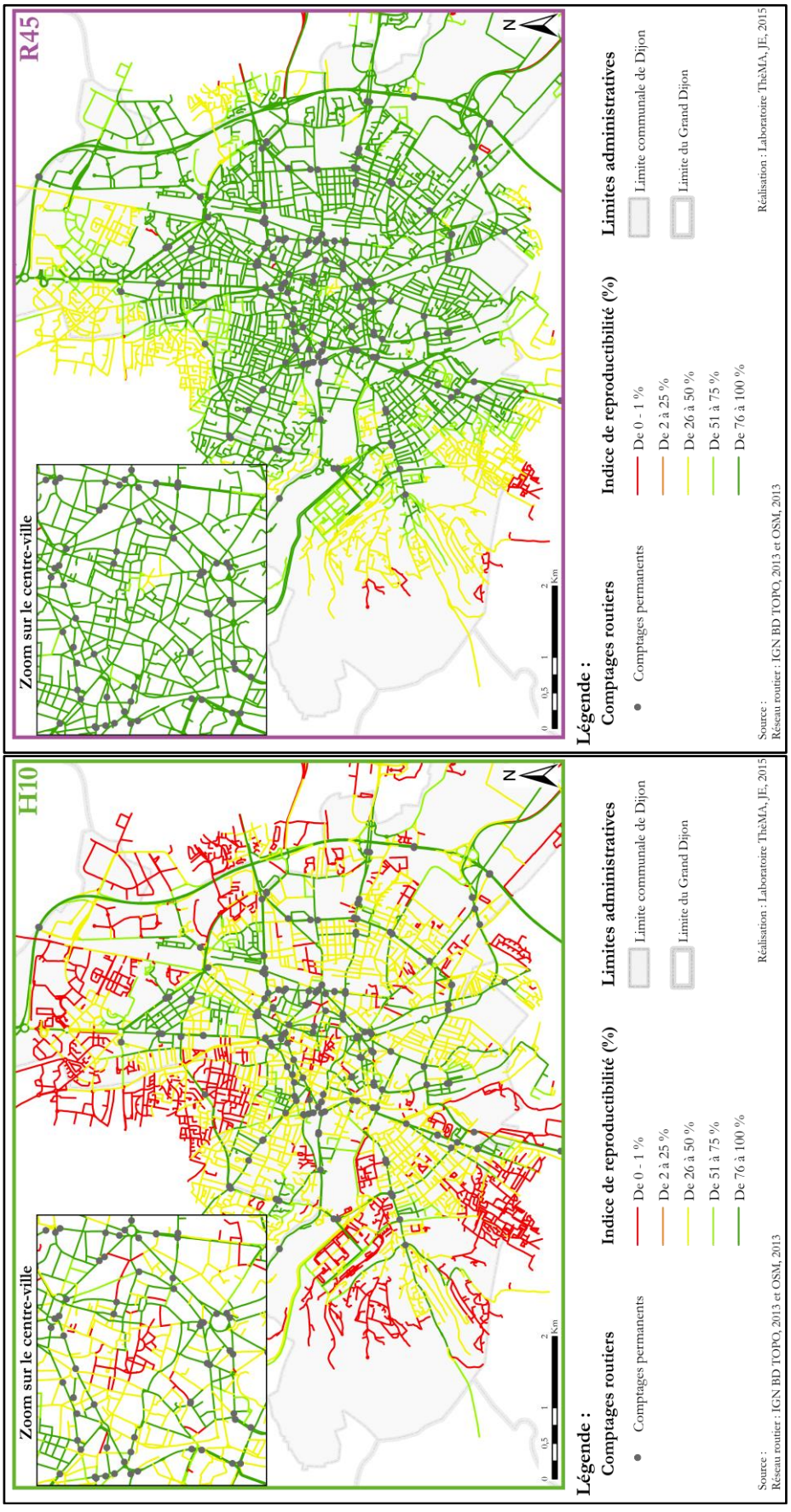

Zoom sur le centre-ville
S45
N
0 0,5 1 2 Km
Légende :
Coefficient de détermination
de 0,14 à 0,35
de 0,36 à 0,50
de 0,51 à 0,80
de 0,81 à 0,90
de 0,91 à 0,95
de 0,96 à 1
Débit moyen journalier simulé (Nb véhicules/jour)
Aucun véhicules
de 1 à 500 véhicules
de 501 à 1 000 véhicules
de 1 001 à 5 000 véhicules
de 5 001 à 10 000 véhicules
de 10 001 à 25 000 véhicules
de 25 001 à 50 000 véhicules
Plus de 50 000 véhicules
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015

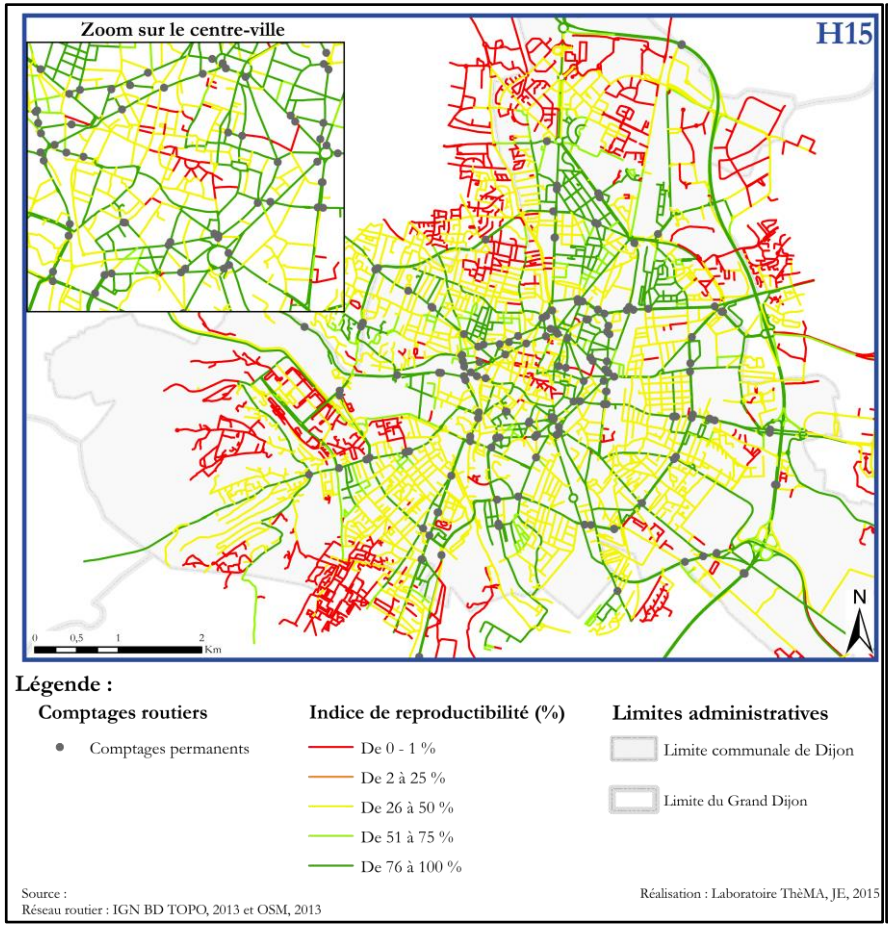
## Annexe cartographique 2. Indice de reproductibilité de SCAUP

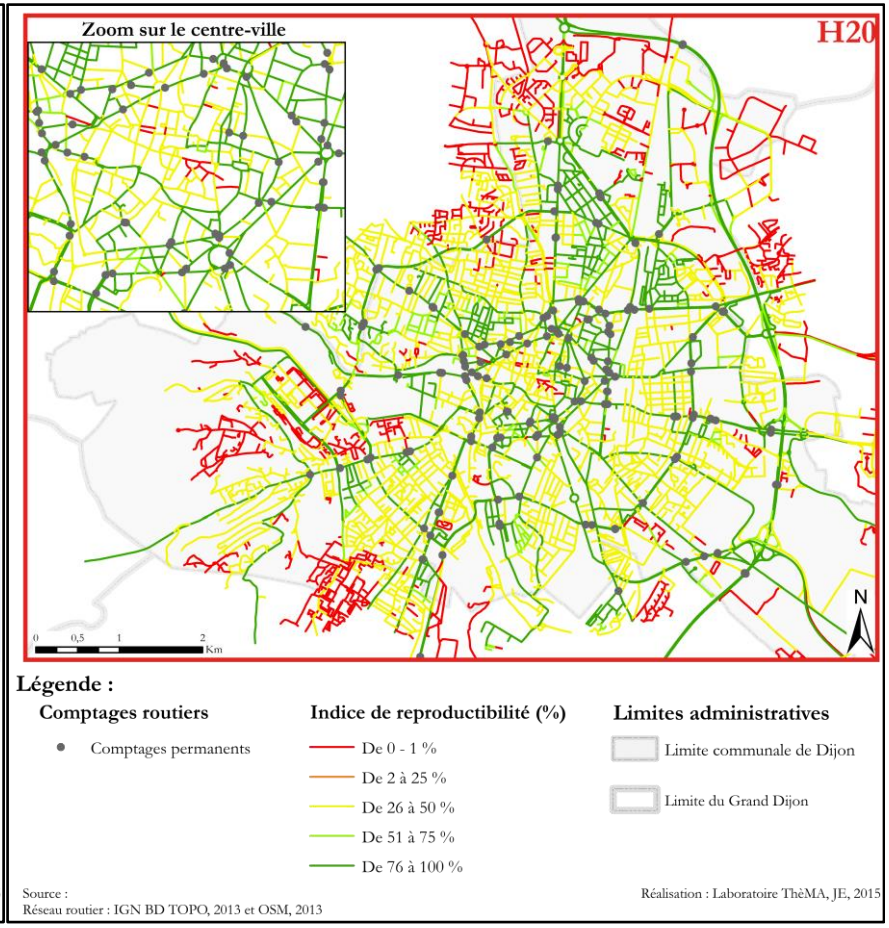


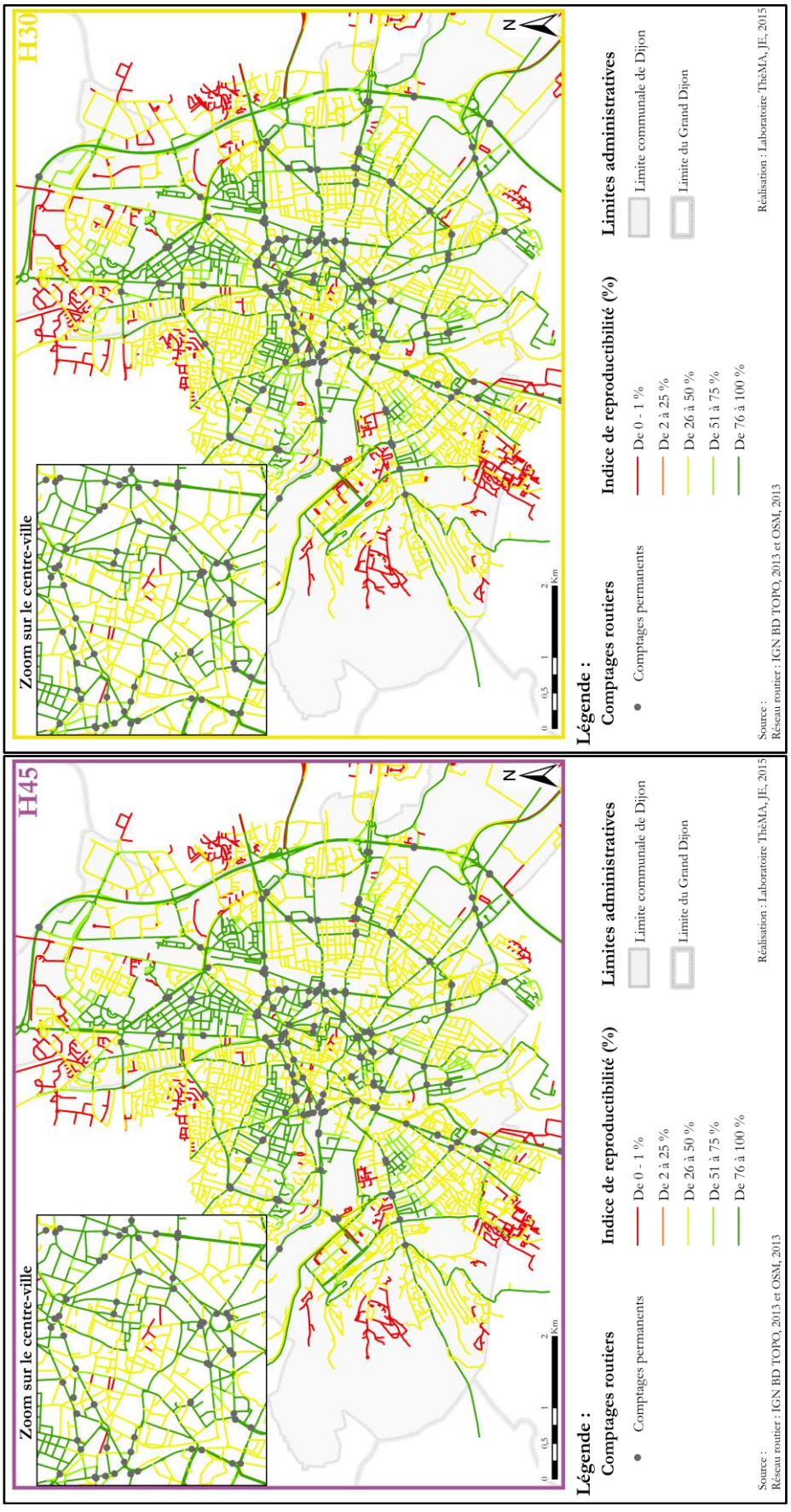

Zoom sur le centre-ville
R20
N
0 0,5 1 2 Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015


Zoom sur le centre-ville
R30
N
0 0,5 1 2 Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015

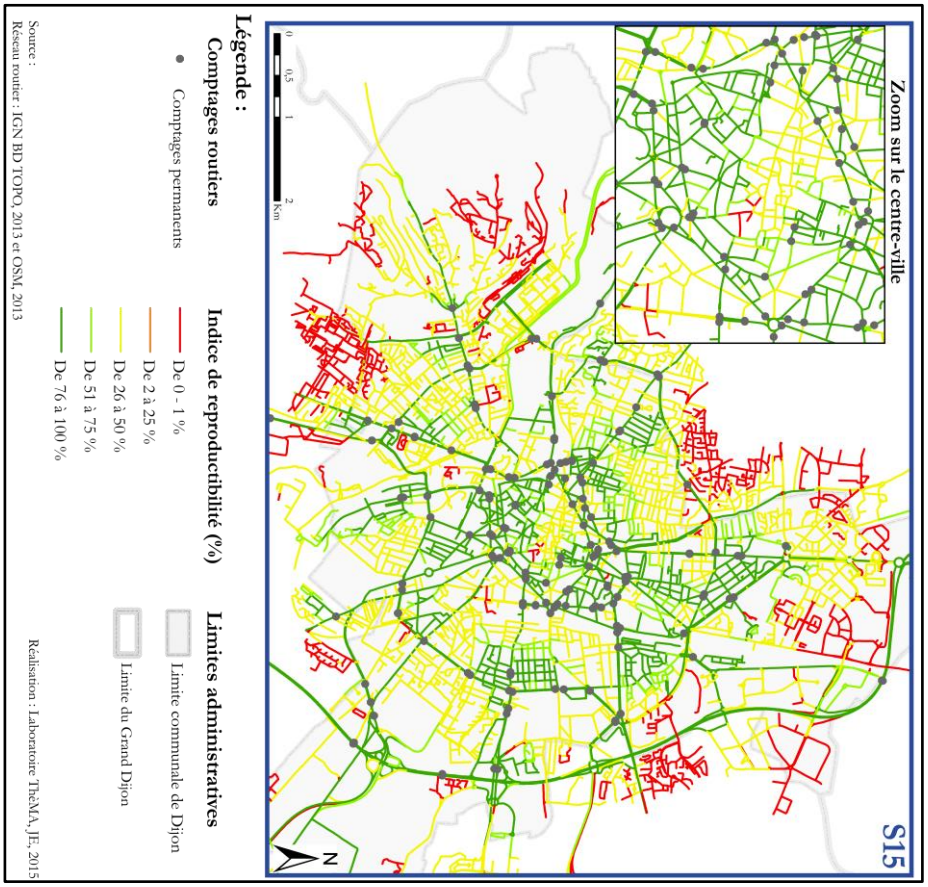

R45
Zoom sur le centre-ville
N
0
0,5
1
2 Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015
H10
Zoom sur le centre-ville
N
0
0,5
1
2 Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015


Zoom sur le centre-ville
H15
N
0 0,5 1 2 Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015


Zoom sur le centre-ville
H20
N
0 0,5 1 2 Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015


H45
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015
H30
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015

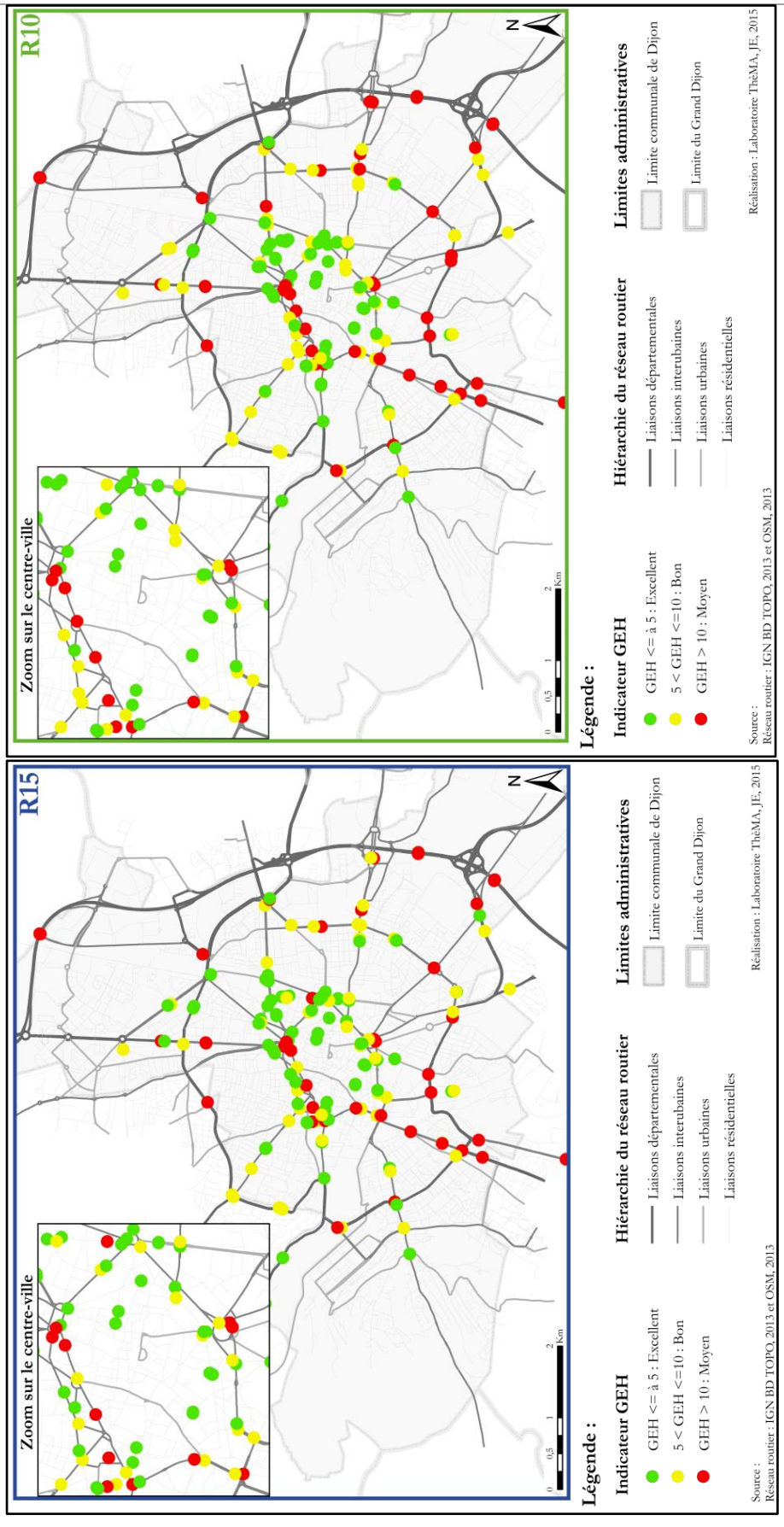

S10
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015

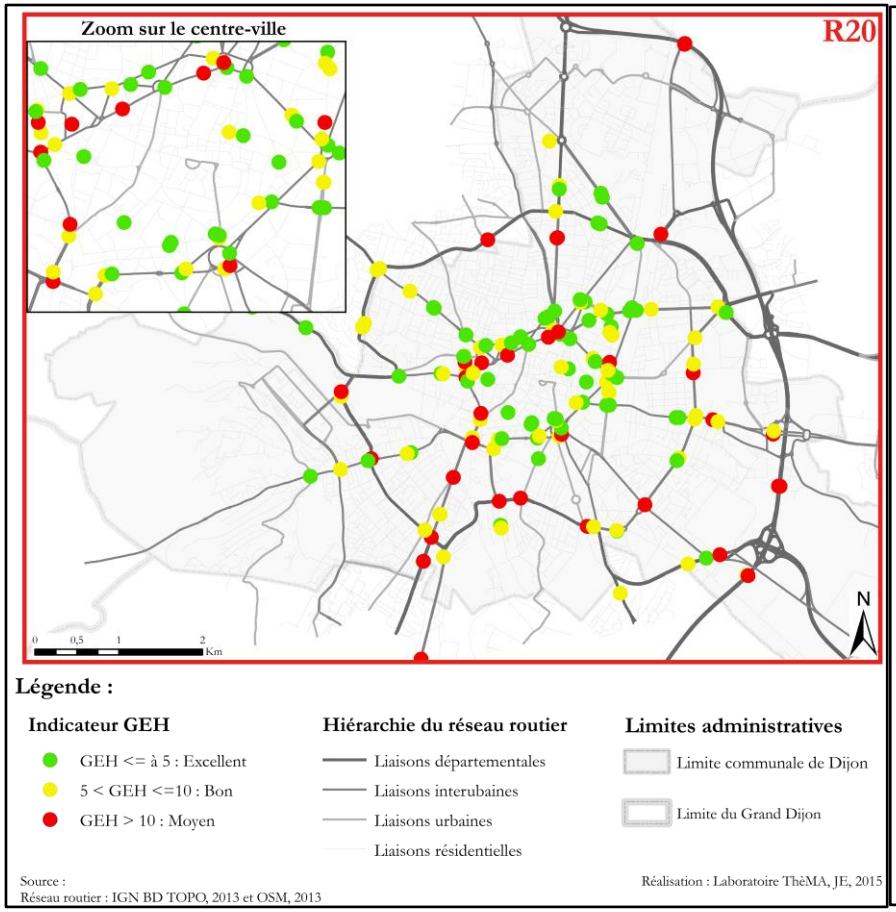

S15
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015

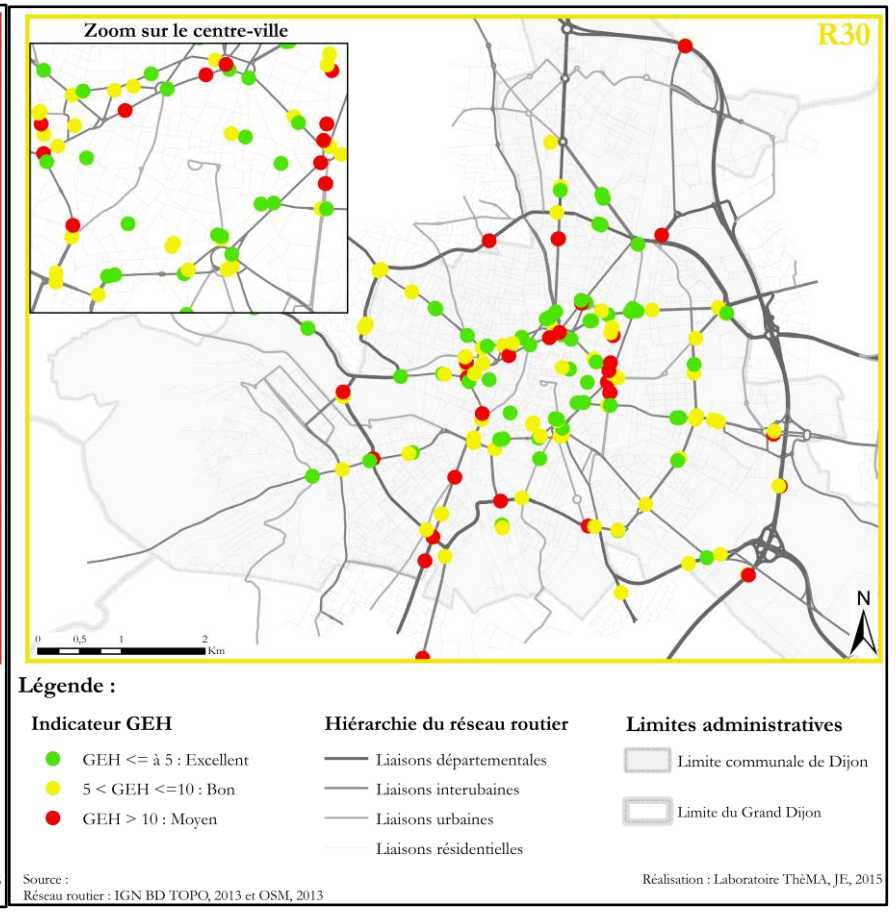

S20
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015
S30
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015

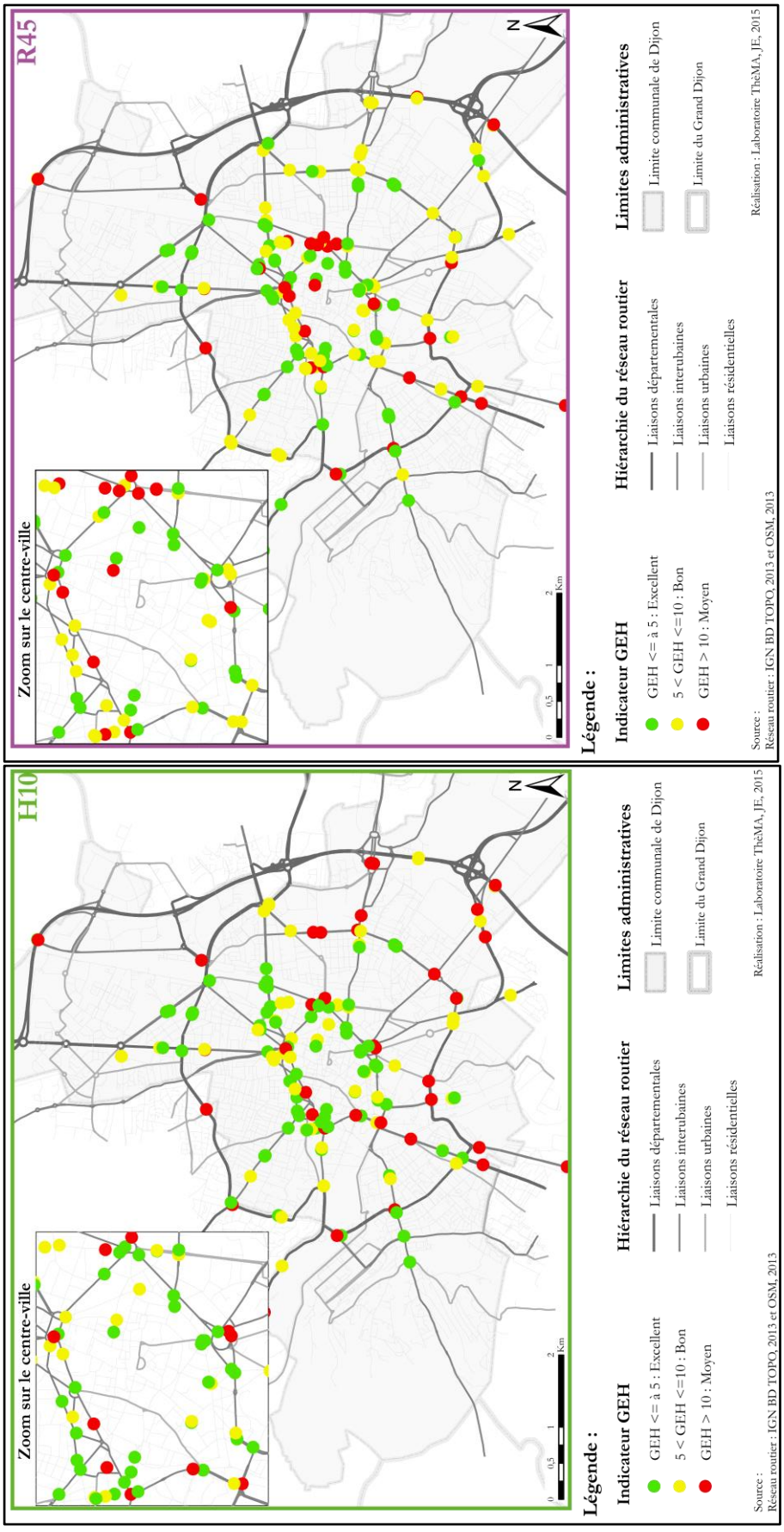

Zoom sur le centre-ville
S45
N
0
0,5
1
2
Km
Légende :
Comptages routiers
Comptages permanents
Indice de reproductibilité (%)
De 0 - 1 %
De 2 à 25 %
De 26 à 50 %
De 51 à 75 %
De 76 à 100 %
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015

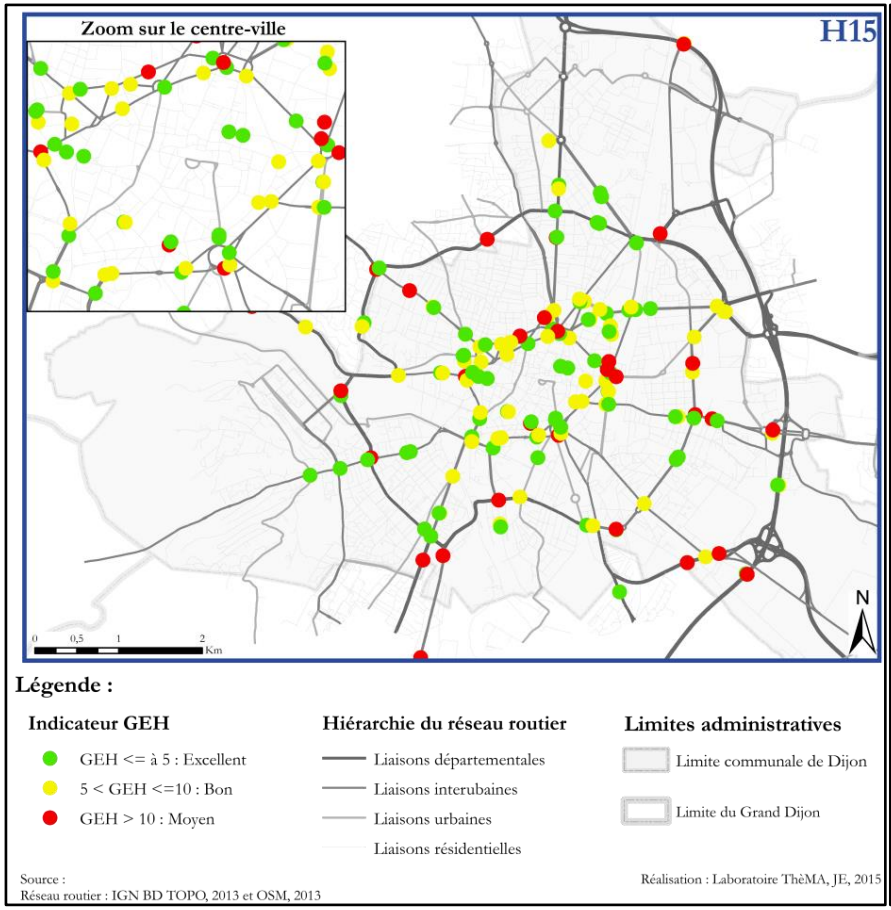
## Annexe cartographique 3. Cartographie du GEH à partir de SCAUP

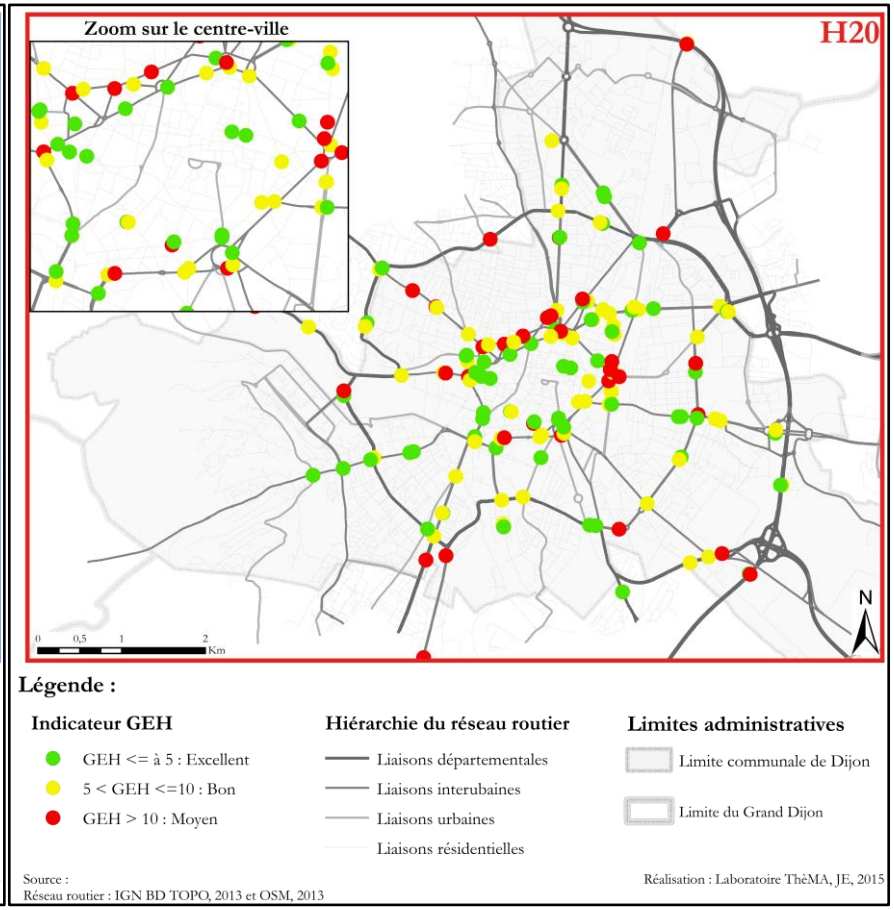


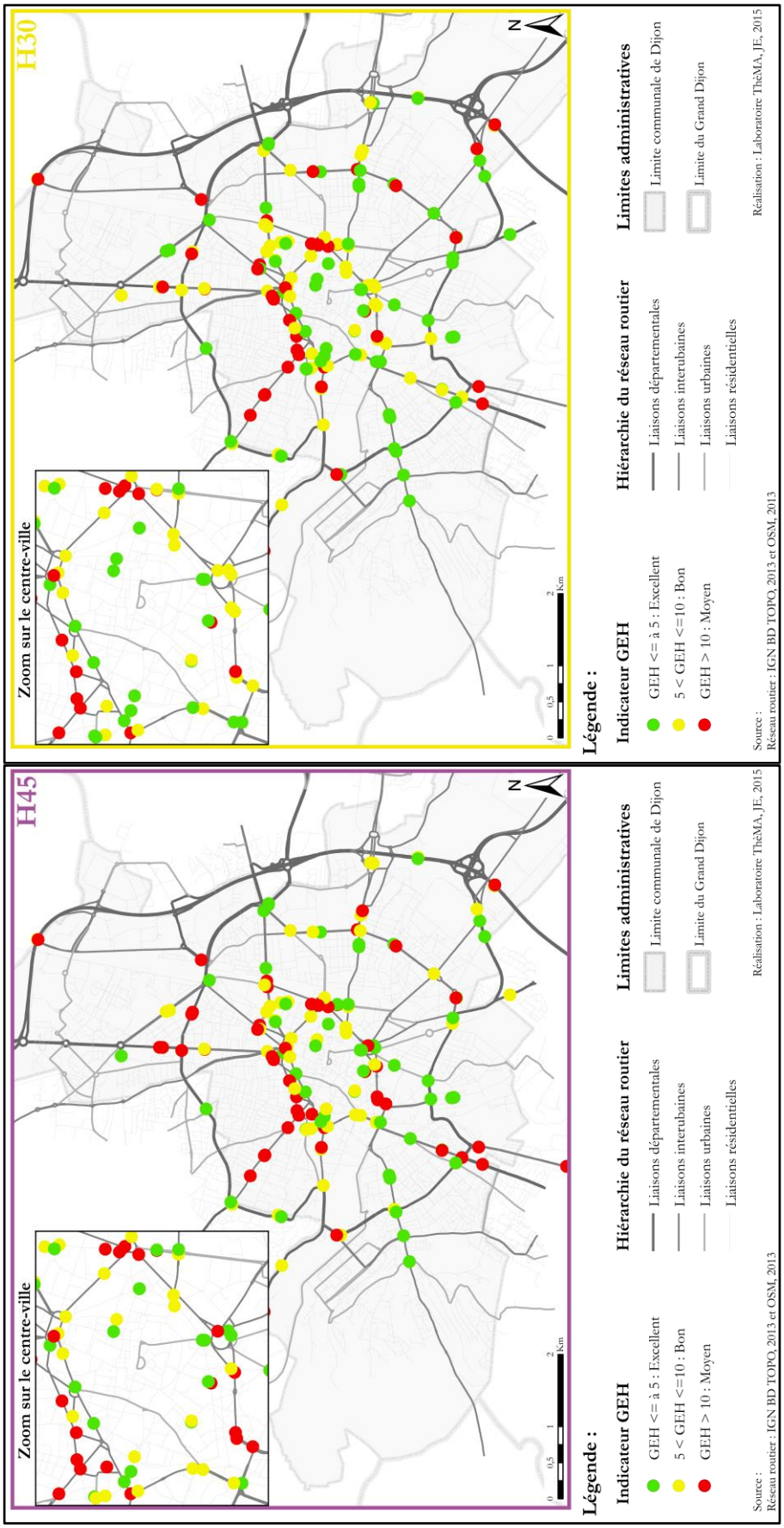

Zoom sur le centre-ville
R20
N
0 0,5 1 2 Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interubaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015

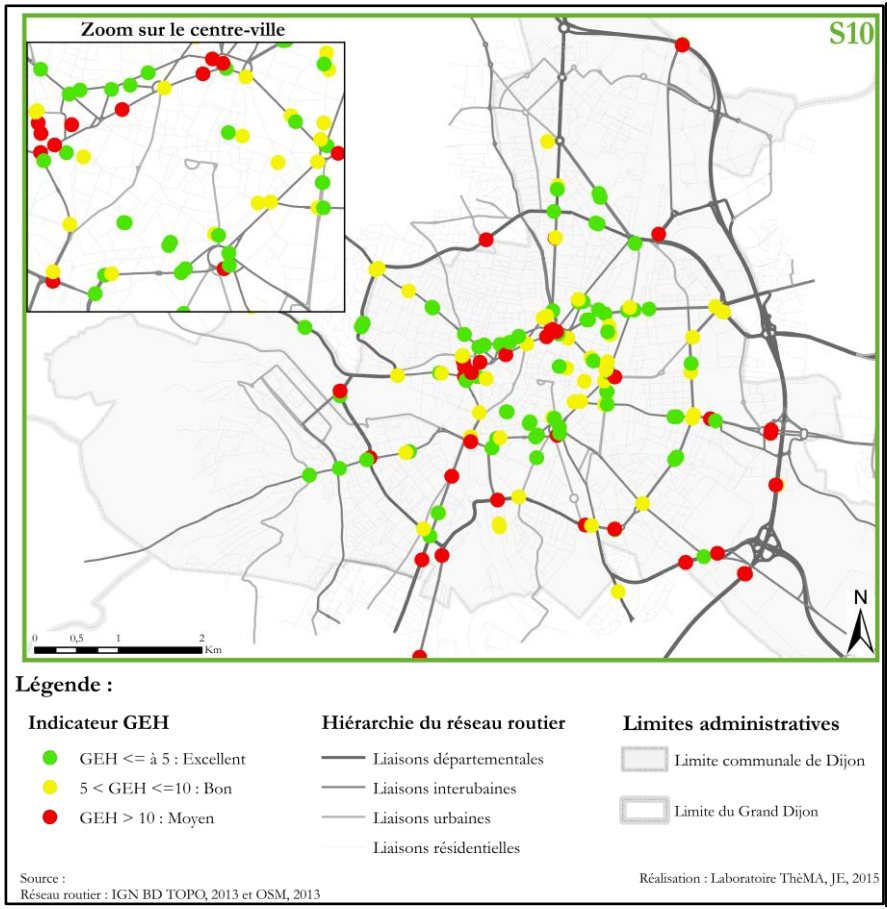

Zoom sur le centre-ville
R30
N
0 0,5 1 2 Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interubaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015

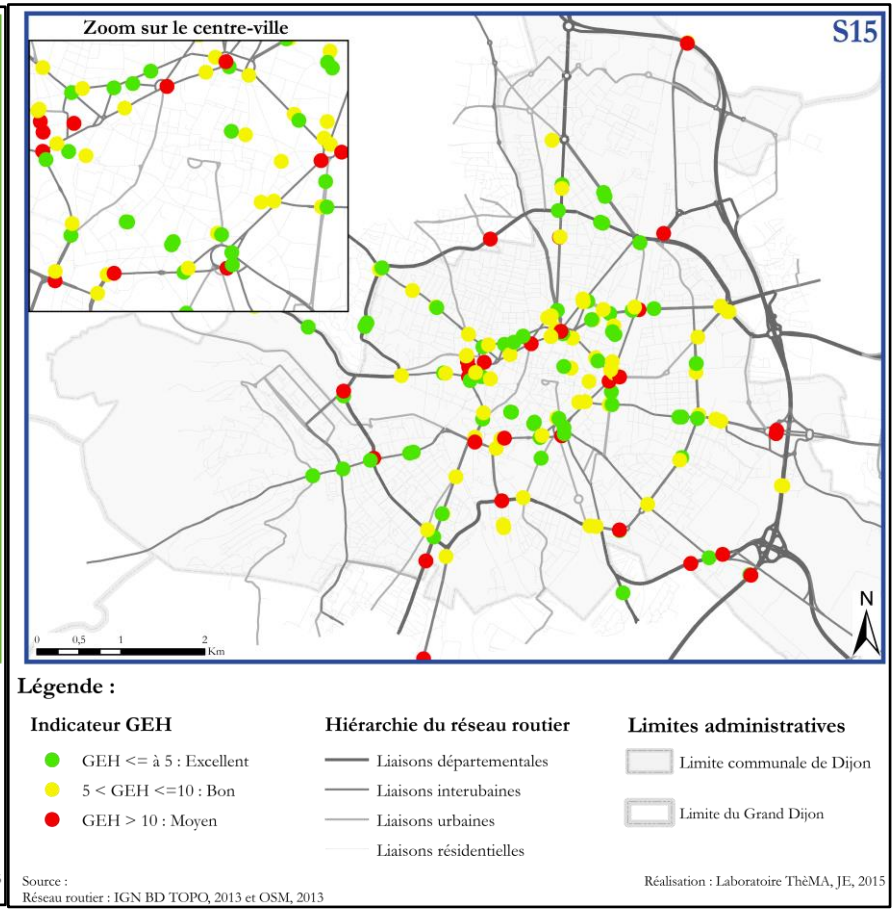

R45
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interurbaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015
H10
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interurbaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015

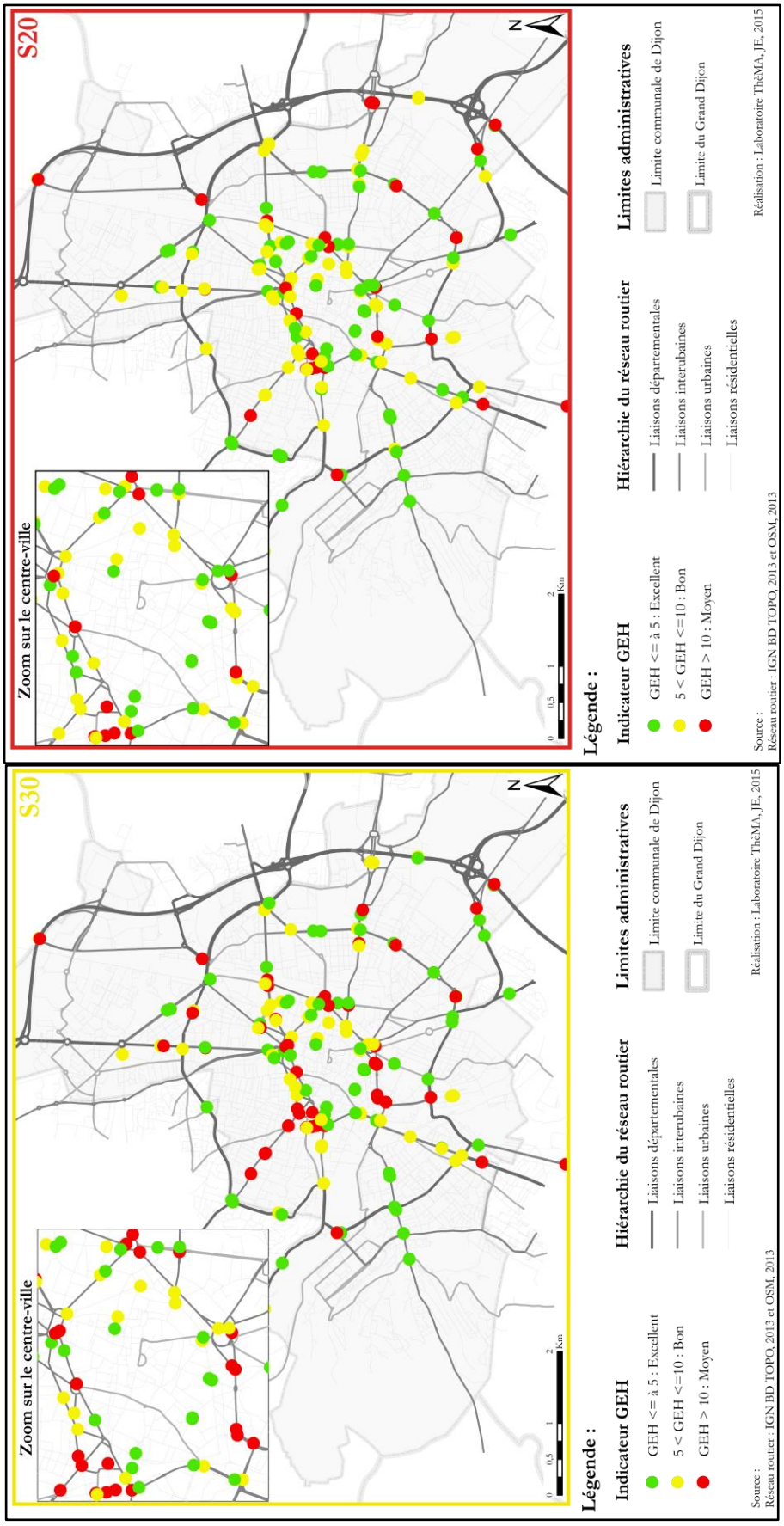

Zoom sur le centre-ville
H15
N
0 0,5 1 2 Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interubaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015

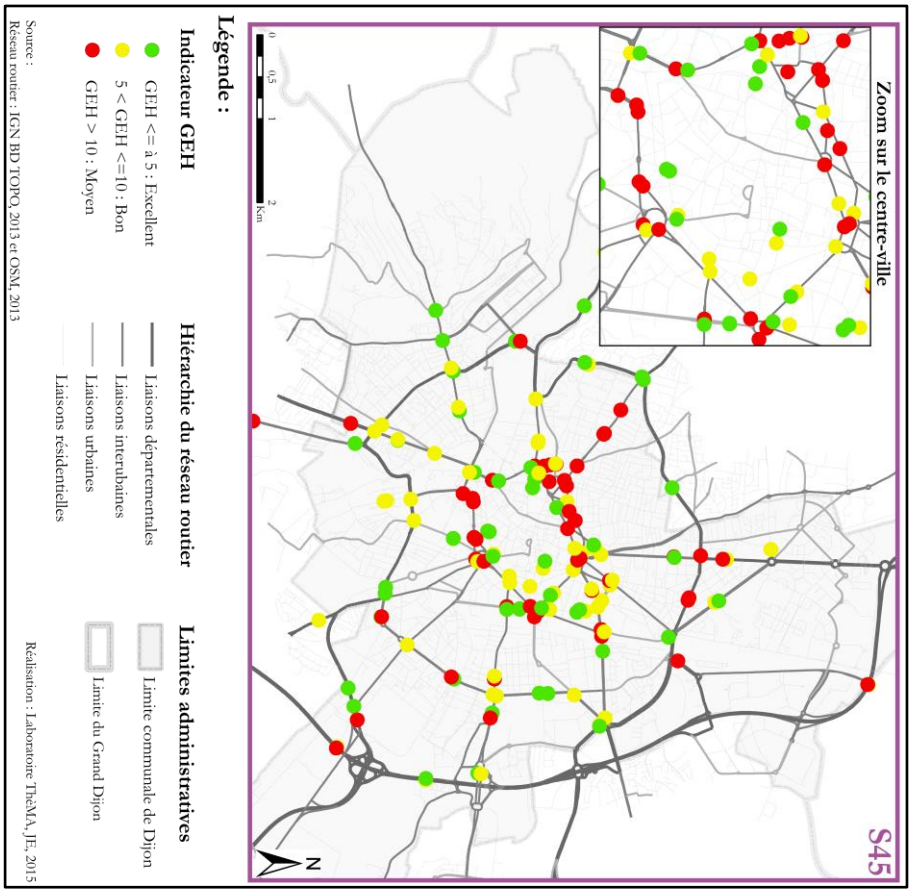

Zoom sur le centre-ville
H20
N
0 0,5 1 2 Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interubaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015


H30
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interurbaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire TheMA, JE, 2015
H45
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interurbaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire TheMA, JE, 2015


Zoom sur le centre-ville
S10
N
0 0,5 1 2 Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interubaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015


Zoom sur le centre-ville
S15
N
0 0,5 1 2 Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interubaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThèMA, JE, 2015


S20
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interurbaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire TheMA, JE, 2015
S30
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interurbaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire TheMA, JE, 2015


S45
Zoom sur le centre-ville
N
0
0,5
1
2
Km
Légende :
Indicateur GEH
GEH <= à 5 : Excellent
5 < GEH <=10 : Bon
GEH > 10 : Moyen
Hiérarchie du réseau routier
Liaisons départementales
Liaisons interurbaines
Liaisons urbaines
Liaisons résidentielles
Limites administratives
Limite communale de Dijon
Limite du Grand Dijon
Source :
Réseau routier : IGN BD TOPO, 2013 et OSM, 2013
Réalisation : Laboratoire ThéMA, JE, 2015

# RÉSUMÉ

A partir des enquêtes, le trafic n'est pas une donnée, mais il est plutôt construit à partir d'hypothèses portant sur les relations entre des origines et des destinations. En vue de reconstruire un trafic routier plus proche de la mesure, et sur un ensemble de tronçons routiers plus important, il apparait alors intéressant de partir de données de comptages issues de capteurs urbains. Notre postulat de départ part de ce constat. L'insertion de ces données d'observations du trafic routier fournit l'opportunité d'expérimenter les potentiels d'exploitations des capteurs pour estimer les niveaux de Pollution Atmosphérique Automobile (PAA) à l'échelle intraurbaine. Cependant, il est alors nécessaire de modifier la nature de la mesure en vue d'extraire une information sur la circulation routière, ce qui a été envisagé ici à travers la construction d'un modèle de simulation multi-agents. D'une manière plus générale, en partant de la donnée, c'est une démarche de construction de la connaissance sur les émissions de PAA qui est abordée tout au long de ce travail.

La mise en œuvre de la démarche de modélisation SCAUP (*Simulation multi-agents à partir de Capteurs Urbains pour la Pollution atmosphérique automobile*) a été réalisée en trois temps : 1. En se focalisant sur les dispositifs de quantification du trafic routier à travers les capteurs urbains ; 2. En proposant une démarche de modélisation et de simulation de ces données pour le trafic routier ; 3. En se rattachant aux référentiels nationaux utilisés par les AASQA pour le calcul des émissions de PAA. L'ensemble se lie et s'intègre au sein d'une *matrice technique* qui constitue la colonne vertébrale de ce manuscrit à travers trois dispositifs interdépendants : la *quantification*, la *modélisation* et l'*évaluation*. Ce travail s'inscrit dans une démarche expérimentale de simulation du trafic routier pour le calcul des émissions de PAA. Parrainé par l'AASQA locale ATMOSF'AIR BOURGOGNE, il s'inscrit aussi dans une optique de recherche appliquée en appui de ces organismes en charge de la surveillance de la qualité de l'air. À l'heure où le *big data* entre dans de nouveaux questionnements quant aux capacités des chercheurs à en extraire une connaissance, nous proposons une démarche géographique en vue de replacer la donnée au centre d'une démarche de simulation originale du trafic routier (*data-driven*).



## MOTS CLÉS

DYNAMIQUES URBAINES, TRAFIC ROUTIER, COMPTAGES, SIMULATION MULTI-AGENTS (SMA), POLLUTION ATMOSPHÉRIQUE AUTOMOBILE (PAA)



# ABSTRACT

Based on surveys, traffic is constructed from assumptions about the relationship between origins and destinations. In order to rebuild a road traffic wich would be closer to observation and on a wider set of road sections, it appears interesting to use counting data from urban sensors : this is our starting point of view. The insertion of these in-situ dataset in the road traffic measurement provides the opportunity to experience the potential of sensors to estimate Traffic Air Pollution (TAP) levels at the intraurban scale. However, this requires to change the nature of these estimation, here through the construction of a model of multi-agents simulation, in order to extract more information on the road traffic. More generally, this work can be seen as a a knowledge building approach on TAP emisssions which is discussed throughout this work.

The implementation of the SCAUP (*multi-agent simulation from Urban sensors for traffic air pollution*) approach was developped in three stages: 1. Focusing on the quantification of road traffic devices through urban sensors; 2. Proposing a modeling approach for road traffic data simulation ; 3. Using as a reference the national framework used by AASQA to calculate RTA emissions. All is integrated within a technical matrix that forms the spine of the manuscript through three interrelated systems: quantification, modeling and evaluation. This work is part of an experimental approach dedicated to the calculation of TAP emissions based on traffic simulations. Sponsored by the ATMOSF'AIR BOURGOGNE local AASQA, this work could also be used in an operational mode for these organizations in charge of the air quality monitoring. At a time when the big data enters into new questions about the ability of researchers to extract knowledge, we propose a geographical approach that enables to replace the data in the center of an original road traffic simulation approach (*data- driven*).



## KEYWORDS

URBAN DYNAMICS, ROAD TRAFFIC, COUNTS, MULTI-AGENT SIMULATION (MAS), TRAFFIC AIR POLLUTION (TAP)