



2018/33

DP

Olivier Finance, Arnaud Adam,
Jones Jonathan and Isabelle Thomas

Révéler la polarisation
économique d'une ville
à partir de traces GPS de
camions. Le cas de Liège



CORE

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: immaq-library@uclouvain.be

[https://uclouvain.be/en/research-institutes/
lidam/core/discussion-papers.html](https://uclouvain.be/en/research-institutes/lidam/core/discussion-papers.html)

Révéler la polarisation économique d'une ville à partir de traces GPS de camions. Le cas de Liège

FINANCE Olivier ^{1 2}, ADAM Arnaud ¹, JONES Jonathan ³,
THOMAS Isabelle ¹

¹ CORE, Centre of Operations Research and Econometrics, UCLouvain, Belgique

² UMR 8504 Géographie-cités, France

³ Viapass EIR / IRE

Résumé :

Avec la multiplication des capteurs, nous disposons désormais de quantités massives de données à l'échelon individuel. Dans cette contribution, nous illustrons comment une nouvelle source de données issue du système de prélèvement kilométrique des camions en Belgique peut présenter une plus-value pour le géographe. Dans ce cas précis, le suivi spatio-temporel quasi-exhaustif des camions est utilisé afin de dessiner une géographie précise des circulations de camions en Belgique et d'explorer la place spécifique de Liège dans ce réseau. La quasi-exhaustivité de ces données nous permet-elle réellement de dépasser les biais classiques telles l'agrégation ou la représentativité des données ? En quoi révèlent-elles l'ADN des territoires ?

Tout en discutant l'apport des "big-data", en particulier en géographie des transports, nous présentons la méthodologie mise en œuvre pour passer de données GPS brutes à une matrice origines-destinations plus classique. Plusieurs méthodologies sont appliquées à cette matrice pour révéler comment les big-data peuvent nous aider à révéler une facette de la polarisation économique d'une ville.

Mots-clés

Big-data ; Interactions ; Transport ; Traces spatio-temporelles ; Hinterland ; Polarisation.

Abstract

With the multiplication of sensors, we now have access to a large volume of individual data. In this contribution, we illustrate how a new source of data – the kilometre charge system of trucks in Belgium – can be useful for geographers. More deeply, the spatio-temporal tracking of nearly every truck is used to elaborate an accurate geography of the trucks movements within Belgium, as well as exploring the specific position of the city of Liège within this

network. Can this massive dataset help us to overcome classical biases in geography as aggregation or representativeness? Do these data reveal the “DNA” of territories?

By discussing the contribution of “big-data”, and more particularly in geography, we present the methodology applied to transform the original raw GPS data into a more classical origin-destination matrix. Diverse methodologies are applied on this matrix to reveal how big-data are useful to explore an aspect of the economic polarization of a city.

Key-words

Big-data; Interactions; Transport; Spatio-temporal traces; Hinterland; Polarization.

"Big data" et géomatique au service de la géographie des transports

La géomatique permet – depuis des décennies – le croisement entre des questionnements géographiques et les possibilités techniques offertes par l’informatique. Elle semble prendre aujourd’hui un nouvel élan tout particulier, lié à la masse de données spatialisées inédites nouvellement collectées et aux avancées technologiques permettant leur acquisition, leur stockage, leur traitement et leur diffusion. Les objets connectés, le suivi en temps réel de personnes ou de véhicules, l’enregistrement de leurs traces spatio-temporelles (qu’il soit ou non volontaire et conscient du point de vue de l’utilisateur), sont autant de nouvelles sources d’informations spatialisées utiles au géographe et au géomaticien. Ces données semblent enfin permettre des analyses à l’échelle individuelle (personne, véhicule, objet) et sont vivement attendues notamment en géographie des transports et mobilités où les données individuelles n’étaient, jusqu’il y a peu, fournies que pour de petits échantillons.

La constitution de la géomatique comme discipline à part entière est indissociable de l’histoire des sciences et des technologies et en particulier de l’informatique géographique (Joliveau, 2004). Après la construction de premiers outils spécifiques dès les années 1960 par certains pionniers de la révolution quantitative aux États-Unis ou en Suède, la diffusion des ordinateurs et des programmes géographiques s’effectue bien au-delà de ces foyers initiaux. La géomatique prend réellement son envol autour de 1985, en lien étroit avec les systèmes d’information géographique. Désormais, nous sommes entrés dans l’« *âge d’une quatrième informatique* » (Joliveau, 2004), autorisant la collecte et l’enregistrement d’informations qui étaient jusque-là "inaccessibles" du fait d’incapacités techniques ou financières, ouvrant la porte à des analyses bien plus exhaustives et désagrégées.

À l’heure de l’essor des "data sciences", de la "fouille des données", voire des "connected geomatics" intégrant des senseurs terrestres et aériens au sein de la géomatique (Kitchin, 2013 ; Miller, 2017a, 2017b ; Li *et al.*, 2018), de nouvelles opportunités s’ouvrent tout en impliquant

parfois de coûteuses étapes de nettoyage et de traitement des données (Goodchild, 2013). Nous passons d'une période où les données étaient rares à une nouvelle ère où l'information irait jusqu'à écraser nos techniques et compétences (voir le concept d'infobésité ; Pornion, 2017). Dans le cas de la géographie des transports et des mobilités, très peu de domaines offraient la possibilité de travaux empiriques sur base d'observations individuelles, exhaustives (ou du moins représentatives), et spatialisées. Les déplacements domicile-travail en sont probablement le principal contre-exemple, même si les données ne sont dans ce cas pas systématiquement désagrégées et souffrent de nombreux écueils (Commenges, 2013). Concernant le transport de marchandises, là où « *de plus en plus de données concernant les infrastructures, l'offre de transports réguliers et les territoires deviennent disponibles* » (Dobruszkes, 2012), il est en revanche toujours assez difficile de s'intéresser de manière exhaustive et désagrégée aux services de transport non-réguliers, menant à une frustration largement rappelée par Dobruszkes (2012).

Nous abordons ici une nouvelle possibilité d'analyse spatiale des échanges réalisés dans le cadre de transports de marchandises par la route en Belgique, via le suivi spatio-temporel continu des camions y circulant. L'exploitation de données nouvelles, proches de l'exhaustivité, lèvera ici les difficultés de mise en œuvre d'enquêtes et le doute possible sur la représentativité des données qui en sont obtenues. Mais elle n'est pas sans poser des questions relativement classiques quant à la qualité des données et les techniques de visualisation à mobiliser. Toute la frustration de l'absence de données est-elle alors levée par cette manne que semble constituer le suivi spatio-temporel exhaustif des camions circulant à travers le pays ?

La connaissance des circulations de camions entre les différentes villes au niveau national est primordiale dans l'optique de contrôler les diverses externalités liées au transport de marchandises (Macharis & Melo, 2011 ; Rodrigue *et al.*, 2013), dans une démarche de durabilité en termes de planification urbaine et d'aménagement du territoire. En complément ici à ces travaux, nous utilisons cette connaissance des circulations de camions afin de révéler la polarisation économique d'une ville. Pour cela, nous dessinons tout d'abord la géographie du transport routier de marchandises en Belgique avant de nous focaliser sur la zone de Liège, dont le tissu économique est connu pour être tourné de manière conséquente vers les activités logistiques, industrielles et commerciales génératrices de mouvements de camions importants (Strale, 2009 ; Merenne-Schoumaker *et al.*, 2015). L'objectif sous-jacent est de tenter de circonscrire l'hinterland de Liège (Donnay, 1995) vu ici à travers les circulations de camions, qui révèlent une des dimensions de la polarisation économique de la ville. Quels sont les territoires environnants ou à distance, qui se trouvent polarisés par la région liégeoise ?

Nous présenterons dans un premier temps les données desquelles découlent nos analyses, et les différentes étapes de nettoyage et préparation de ces données, de données brutes massives à une information spatio-temporelle construite et fiable (Section I). Nous élaborerons ensuite une géographie des transports routiers en Belgique puis autour de Liège afin d'identifier quelle position elle occupe dans le système de transport routier belge, ainsi qu'en lien avec les pays limitrophes (Section II). Enfin, nous discutons la pertinence de l'extension de ces travaux à l'ensemble de la Belgique et l'apport de ces données dans la compréhension du système de fret routier belge. Tout au long de cette contribution, nous veillerons également à identifier à la fois

les biais nouveaux et les problèmes classiques qui peuvent émerger de ces analyses basées sur des données spatio-temporelles individuelles – certes riches – mais non conçues dans une optique de recherche.

I. Des données GPS brutes à une géo-information

Le point clé de cette section est d'illustrer la difficulté de passer de données massives brutes à des informations spatio-temporelles pertinentes et exploitables (Thakur *et al.*, 2015 ; Gingerich *et al.*, 2016). Après une description des données brutes (Section I.A.), la Section I.B. explique la démarche suivie pour passer des points GPS à des données géo-intelligentes (de la trace GPS à une succession de trajets ou segments). La Section I.C. explique le choix du milieu d'étude et particulièrement la délimitation de la "région liégeoise" qui permettra ainsi de révéler le poids qu'occupe Liège dans l'ensemble des circulations de camions effectuées en Belgique.

A. Suivre les camions dans l'espace et le temps

Depuis le 1^{er} avril 2016, la Belgique a instauré une redevance kilométrique pour l'utilisation de 6 500 kilomètres de son réseau routier ; elle concerne les véhicules d'une masse maximale autorisée supérieure à 3,5 tonnes destinés au transport de biens¹ que nous désignerons ici par le terme *camions*. Contrairement à des systèmes de péages "physiques" en pleine voie, ce prélèvement (système dit Viapass) s'opère de manière "embarquée", grâce à des capteurs GPS (appelés OBU pour "On Board Units") qui doivent obligatoirement être installés et activés lors de tout déplacement sur le territoire belge. Ces OBU individuels, destinés à détecter à chaque instant l'utilisation ou non du réseau routier taxé, pistent les véhicules, calculent et transmettent la taxation éventuelle associée et de fait enregistrent à des pas de temps courts (toutes les 30 secondes) les positions GPS précises des camions en mouvement. Ce sont ces positions successives de la quasi-totalité des camions circulant durant une semaine de novembre 2016 que nous explorons dans cette contribution², semaine complète que nous ramenons aux seuls jours ouvrables. La base de données brutes mise à disposition est constituée des enregistrements précis de chaque point GPS (latitude, longitude, temps) et des caractéristiques du camion concerné (identifiant, masse, pays d'immatriculation, classe Euro d'émission de polluants, etc.).

Ainsi les données exploitées ici ne sont en rien destinées à de la recherche scientifique. Conçues en vue de la taxation d'une partie du réseau routier avec un objectif d'internalisation des coûts externes du transport routier de marchandises, nous détournons ici leur usage pour répondre à nos questionnements géographiques. Elles nous permettent, après adaptation, nettoyage et filtrage, de disposer de manière inédite dans le cas de la Belgique, d'informations spatiales et temporelles individuelles et quasi-exhaustives sur les déplacements des camions. Que ce soit en quantité ou en qualité, les données disponibles préalablement (collectées via stations de comptage ou enquêtes) étaient bien plus limitées. Plus précisément, les stations de comptage sur certaines portions routières n'autorisaient qu'une mesure ponctuelle du trafic – soit le flux de camions transitant par ces portions routières – sans aucune généralisation à l'échelle d'un

territoire. Aucune information supplémentaire n'était disponible sur les véhicules, ni leur provenance, ni leur destination, ni à propos de leur éventuel chargement. De la même façon, raisonner par enquête auprès de transporteurs routiers et d'entreprises générant des flux de marchandises par la route nécessite d'abandonner l'exhaustivité voire toute représentativité par contraintes temporelles et financières. Il est en effet nécessaire de travailler sur une sélection préalable d'entreprises, dont les taux variables de réponse engendrent des échantillons relativement faibles et à partir desquels il peut être difficile de monter en généralité (voir par exemple Lombard, 1999 dans le cas du transport routier dans le nord de la France). Le questionnement autour de la qualité et la quantité des données utilisées ici n'est pour autant pas totalement levé et sera discuté ici.

Ce n'est pas la première fois que des traces de véhicules sont analysées : piétons, cyclistes et taxis ont déjà fait l'objet d'analyse spatio-temporelles dans d'autres contextes géographiques, avant tout à l'échelle intra-urbaine (voir par exemple Laurila *et al.*, 2012 ; Thomopoulos *et al.*, 2015). D'autres travaux ont permis d'explorer le système de fret routier par l'intermédiaire de différentes sources de données, les capteurs étant placés tantôt sur la route (stations de comptage, caméras de surveillance...), tantôt embarqués (transpondeurs, chronotachygraphes, boîtiers GPS, smartphones...). Une synthèse en est proposée par Antoniou *et al.* (2011). Les travaux basés sur un suivi continu des camions s'appuient souvent sur les données provenant d'un unique fabricant de systèmes de GPS ou d'un échantillon d'entreprises de transport. Ils répondent bien souvent à des questionnements particuliers liés à la performance du réseau routier (Flaskou *et al.*, 2015), dans des contextes géographiques très spécifiques – des états américains ou des régions de pays émergents, Chine ou Afrique du Sud par exemple (Kuppam *et al.*, 2013 ; Joubert & Meintjes, 2015 ; Ma *et al.*, 2016).

L'élément principal qui se dégage de cette revue de la littérature existante est l'absence de convergence dans les critères, seuils et opérations de filtrage opérés pour passer des données GPS brutes à une information construite, fiable et exploitable géographiquement. L'étape commune à la plupart des travaux est l'identification de trajets et de segments origine-destination se succédant dans l'espace et dans le temps. Shen et Stopher (2014) désignent cette étape comme un "challenge" tant les difficultés liées aux pertes de signal GPS et au bruit dans les données enregistrées peuvent être importantes. Afin de déterminer comment s'enchaînent les différents trajets effectués par un camion au cours d'une journée, les auteurs des travaux cités précédemment déterminent des seuils soit temporels, soit spatio-temporels afin de découper la succession de points GPS d'un camion en trajets distincts. Des seuils temporels variant de quelques dizaines de secondes à plusieurs dizaines de minutes sont évoqués sans qu'aucun ne soit a priori meilleur qu'un autre (Shen & Stopher, 2014 ; Zanjani *et al.*, 2015 ; Thakur *et al.*, 2015), d'autant que leur signification diffère à la fois selon le contexte géographique, les objectifs de recherche, et la fréquence d'enregistrement de la donnée (régulière ou non, intervalle variable d'un cas à un autre, enregistrement uniquement en circulation ou en continu, etc.). Parmi les critiques majeures que l'on peut formuler face à cette sélection de travaux figurent la faible prise en compte des questions d'exhaustivité et représentativité et l'absence fréquente de détails à propos des méthodes de validation des données.

B. De la trace GPS à une succession de trajets et segments

La question du nettoyage des données revient régulièrement en creux dans les contributions où des données massives sont mobilisées ; il est parfois considéré que si des données aberrantes ou peu pertinentes sont présentes dans le jeu de données, elles se retrouveraient noyées dans la masse et ne seraient ainsi pas gênantes. Ici nous prenons le parti-pris de nettoyer les données mais aussi de transformer leur structure afin de répondre à nos objectifs. Ces différentes étapes sont présentées ci-après.

1. Définition des lieux d'arrêt des camions

L'objectif est de transformer la multitude de coordonnées GPS enregistrées toutes les 30 secondes par les OBU embarqués dans les camions, en passant de traces GPS complètes à des lieux d'origine et de destination successifs (Figure 1). Chaque point GPS est tout d'abord remplacé dans la succession spatio-temporelle des points émis par un camion (sa *trace*) puis chaque trace est transformée pour aboutir à des *trajets* (succession de points GPS entre deux arrêts) et ensuite à des *segments O-D* (couples origine-destination résumant les trajets). La détection de ces trajets et segments O-D repose sur l'identification des arrêts et donc de points d'origines et de destinations supplémentaires par rapport aux seuls premier et dernier points GPS d'une journée d'un camion.

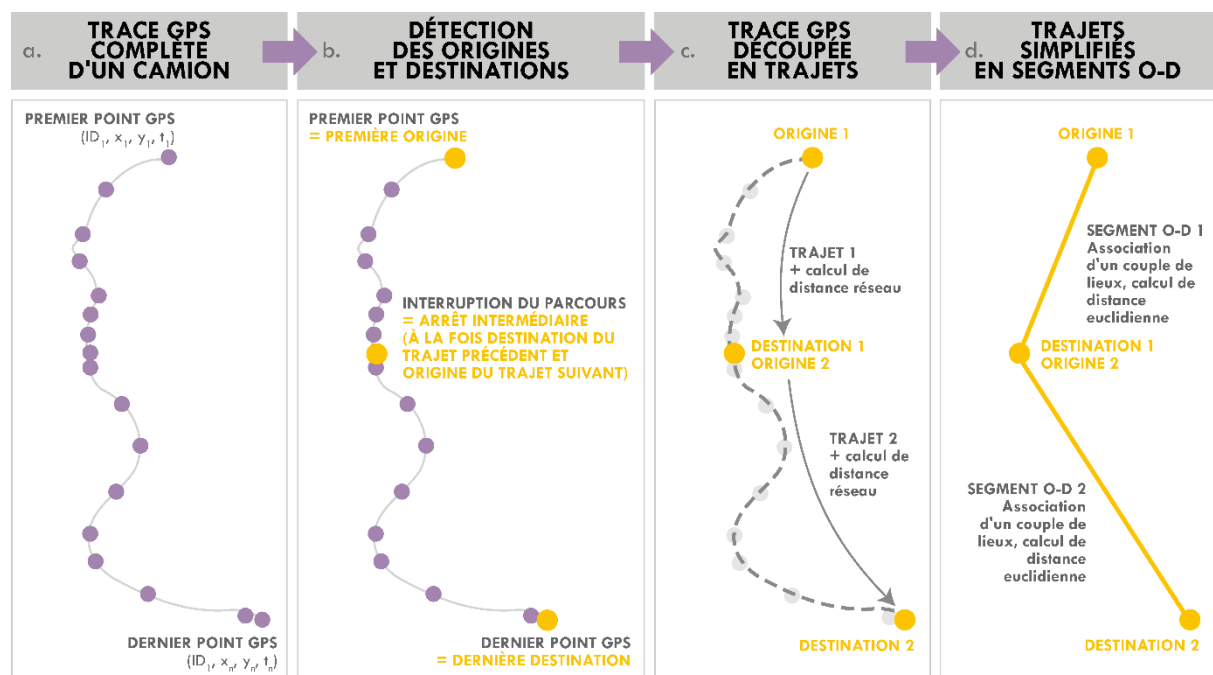


Figure 1 | Transformation des traces GPS brutes en trajets et segments origine-destination

Un arrêt est défini ici selon un critère temporel comme une interruption de parcours d'au moins 10 minutes (intervalle de temps entre deux points GPS successifs). Ce seuil, certes discutable, a été choisi a priori pour être suffisamment bas pour capter différentes natures d'arrêts (chargement, déchargement, pause...) sans pour autant intégrer des arrêts parasites (congestion, feux de signalisation, redémarrages intempestifs de l'OBU...). Une analyse systématique de la variation du nombre d'arrêts détectés selon divers seuils temporels n'a pas

révélé de rupture majeure permettant de justifier ou rejeter le choix des 10 minutes. Ainsi, chaque trace GPS complète est découpée en trajets pour lesquels des distances réseau sont calculées (distances euclidiennes entre couples de points successifs cumulées sur l'ensemble du trajet), eux-mêmes simplifiés en segments O-D reliant directement chaque origine et destination successives. Une distance euclidienne calculée est associée à chaque segment O-D.

2. Changements journaliers des identifiants

Différentes opérations de filtrage simples ont été appliquées, comme par exemple la non-sélection, des camions émettant moins de 10 points GPS. Ensuite, une autre opération a été rendue nécessaire du fait du changement systématique d'identifiant des camions (ID) programmé chaque jour pour des raisons de confidentialité. Ceci rend impossible le suivi d'un même camion durant l'ensemble de la semaine. Ce changement d'ID ayant lieu autour de 2 heures du matin heure locale (période de plus faible intensité des circulations), nous optons pour un découpage de la semaine en journées indépendantes du jour n à 02h00 au jour $n+1$ à 1h59. Si les journées étaient définies classiquement (0h00-23h59), un camion en circulation à cet horaire serait comptabilisé deux fois (une fois pour chaque ID)

Ceci explique l'enregistrement d'un grand nombre de premières origines et dernières destinations entre 1h55 et 2h, bien au-delà du nombre habituellement enregistré dans les minutes précédentes et suivantes (Figure 2). La très grande majorité de ces origines et destinations est à attribuer à ce changement systématique d'ID. Une fois les jours redéfinis, les ID apparaissant uniquement après 1h55 sont exclus de l'analyse (probable poursuite du parcours d'un camion ayant changé d'ID), et les destinations enregistrées après 1h55 tout comme les origines enregistrées avant 2h00'30 sont notées comme floues (probable dernier/premier point avant/après changement d'ID).

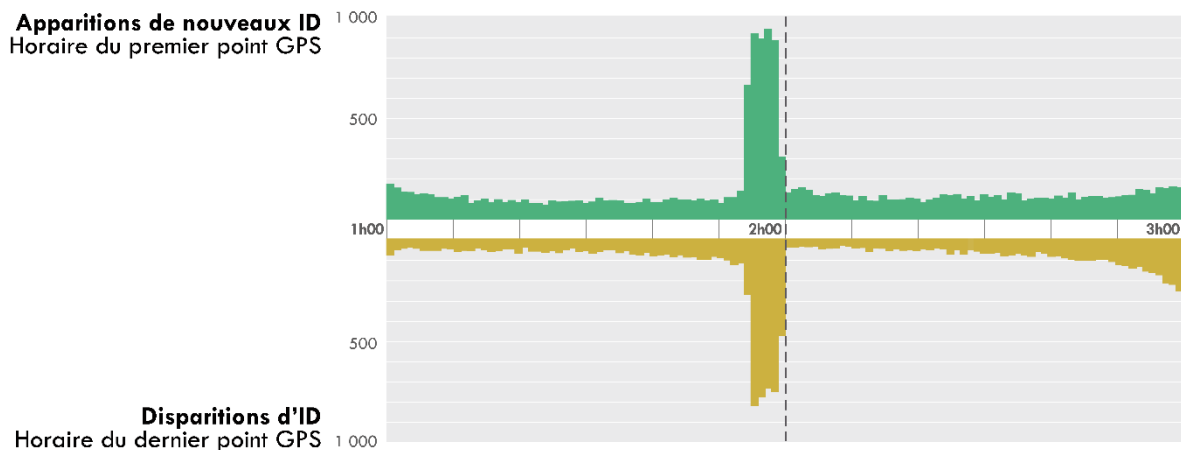


Figure 2 | Apparitions et disparitions d'ID de camions autour de 2 heures du matin

3. Validation et correction des traces, trajets et segments

La base de données restructurée est ensuite corrigée et validée, notamment par comparaison des distances réseau et euclidienne, ainsi que des temps et vitesses entre couples de points GPS

successifs et entre couples origine-destination. Trois points majeurs résument les corrections et suppressions effectuées lorsque des incohérences ont été identifiées (Figure 3) :

- Lorsqu'entre 3 points GPS successifs, les temps sont courts (pas d'arrêt détecté) mais les distances conséquentes (plus de 3 km), le point intermédiaire aberrant est supprimé et les distances sont recalculées (Figure 3a). La mauvaise réception du signal GPS par l'OBU ou une difficulté de transmission de l'information sont des causes probables à ces points aberrants exceptionnels. À l'inverse, si l'erreur est observée à plusieurs reprises pour un même ID, la totalité de la trace est supprimée de notre base de données (l'attribution d'un ID étant gérée directement par les OBUs, il n'est pas impossible qu'un même ID soit attribué à deux camions différents).
- Lorsqu'il y a une faible distance entre une destination détectée et l'origine suivante (seuil fixé arbitrairement à 3 kilomètres), mais qu'elles se trouvent dans deux unités spatiales d'analyse différentes (mailles d'1 km², maillage présenté en Section 2.3), elles sont attribuées à la même maille (Figure 3b). La sélection de la maille retenue s'appuie sur les vitesses instantanées à destination et origine, en considérant qu'il y a un signe plus évident de la localisation précise de l'arrêt au point parmi les deux présentant une vitesse instantanée nulle. Cette erreur est vraisemblablement largement causée par la réactivation tardive de l'OBU après le démarrage du camion. La correction est ici appliquée autant de fois que l'erreur est détectée.
- Enfin, lorsque la distance entre la destination et l'origine suivante est conséquente (plus de 3 kilomètres) alors qu'un arrêt est détecté, la fiabilité des origines et destinations est remise en cause (Figure 3c) : il peut en réalité ne pas s'agir d'un arrêt mais d'une perte de signal GPS pendant plus de 10 minutes. Il n'est alors pas certain que le camion ait effectué un arrêt. La totalité de la trace du camion est alors supprimée, sauf lorsque cette erreur n'est détectée qu'à une seule reprise pour un camion, pour ne pas supprimer une proportion trop importante de camions. Dans ce cas précis, les trajets et segments ceinturant l'arrêt incertain sont fusionnés. Les causes potentielles de cette erreur sont nombreuses : la perte du signal GPS par l'OBU, une panne ou un arrêt volontaire de l'OBU, une erreur de transmission, de collecte ou d'enregistrement des données par les serveurs.

À eux seuls, ces trois points sont à l'origine de la suppression de près de 15 millions de points GPS (Tableau 1).

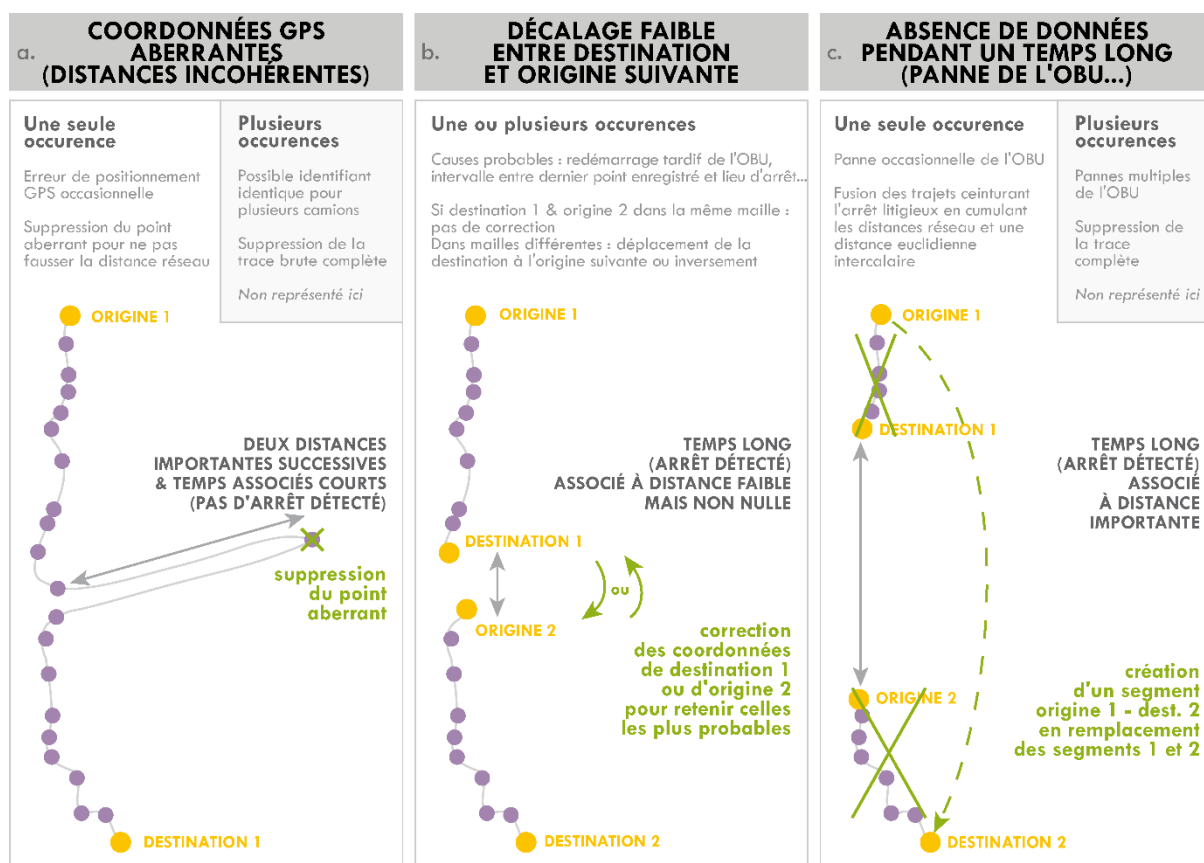


Figure 3 | Étapes principales de détection et correction des erreurs, selon le nombre d'occurrences des erreurs détectées pour chaque ID

	ID de camions		Points GPS		Segments O-D	
	N	%	N	%	N	%
Base de données initiale	798 453	100,0	269 194 440	100,0		
Après suppression des IDs avec moins de 10 points	719 425	90,1	265 564 849	98,6		
Après redéfinition des jours et détection des arrêts	655 701	82,1	253 684 428	94,2	3 746 263	100,0
Base de données finale après les étapes de la Figure 3	628 769	78,7	238 927 089	88,8	2 758 940	73,6

Tableau 1 | Part d'information conservée après les principales opérations de filtrage, validation et correction

Au final, environ 11% des points GPS et 21% des identifiants de camions sont supprimés de la base de données initiale, soit une part bien trop importante pour que ces informations considérées ici comme erronées soient simplement "noyées dans la masse" en cas d'absence de nettoyage des données. Il convient donc de rester critique face à la qualité des données fournies par les capteurs et générateurs de données de masse.

C. Délimitation de la région d'étude

Après des premières analyses au niveau national (Section 3.1), nous nous focaliserons dans cette contribution sur la "région liégeoise" et sur sa connexion au reste du territoire belge et étranger par l'intermédiaire des circulations de camions. La délimitation de cette zone – que nous désignerons parfois directement par "Liège" – a été effectuée à partir d'une classification de l'occupation du sol réalisée sur un maillage complet du territoire belge à une résolution d'1 km² (approximativement 32 000 mailles), résolution conservée pour l'ensemble des analyses de cette contribution (source : Eurostat, maillage "Geostat"). Cette classification mixte prend appui à la fois sur les données d'occupation du sol CORINE Land Cover (dans la lignée de Donnay, 1995), la digitalisation des points d'entrée/sortie majeurs du territoire à partir du réseau routier issu d'OpenStreetMap, celle des aires d'autoroute, et enfin sur la densité de population résidant dans chaque maille (Eurostat).

Le cœur de l'aire urbaine liégeoise (320 mailles) est extrait en considérant l'ensemble des mailles contiguës classifiées dans les classes les plus "urbaines" autant en termes d'usages résidentiel, économique ou par la présence de grands équipements (Figure 4). Ainsi 31% des mailles retenues correspondent à l'urbain dense, 30% à une présence industrielle remarquable ou encore 4% à des zones portuaires ou aéroportuaires (incluant l'aéroport de Bierset et le port fluvial). Cette zone inclut l'intégralité des communes de Liège, Herstal, Saint-Nicolas et Beyne-Heusay, et couvre partiellement 21 autres communes pour un total de 489 000 habitants. Cette définition de Liège dépasse à la fois la simple prise en compte de la commune de Liège, tout en s'affranchissant de limites administratives. Il en résulte une zone urbaine proche des limites de l'agglomération liégeoise telle que définie par Van Hecke *et al.* (2009), qui ne s'affranchit cependant pas des limites communales et qui malheureusement n'a pas été mise à jour.

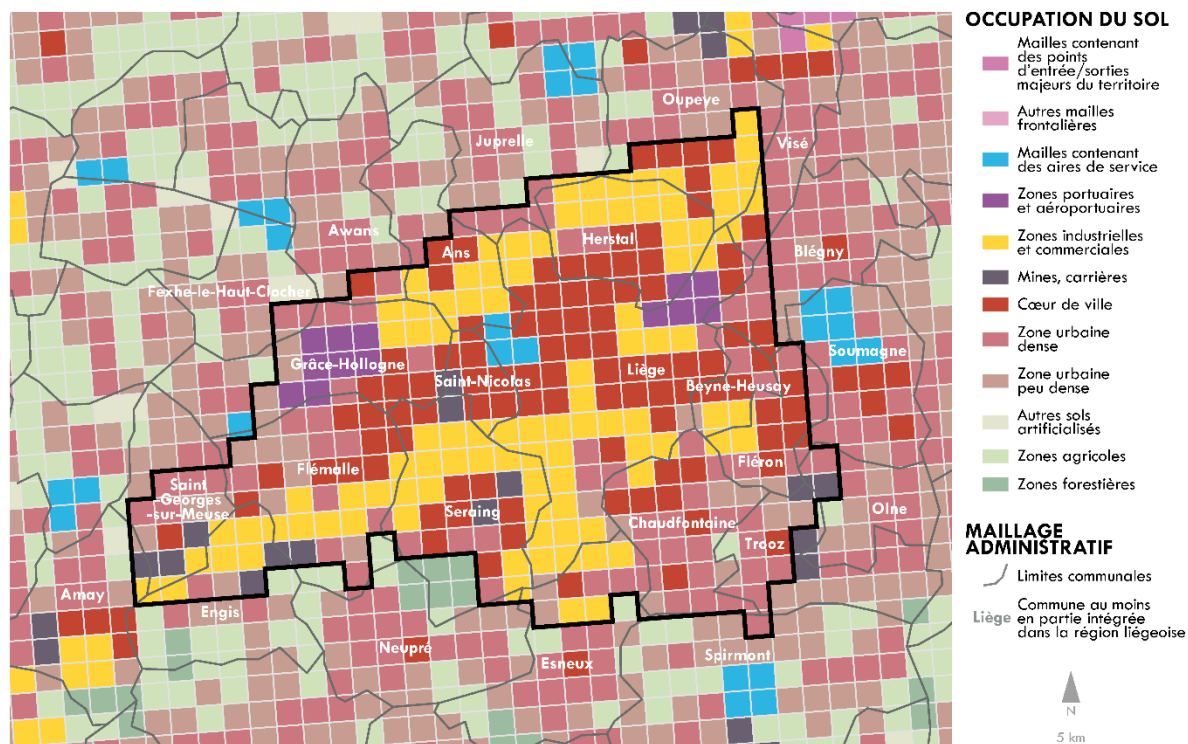


Figure 4 | Délimitation de la zone d'étude (la "région liégeoise") à partir de l'analyse de l'occupation du sol

Il est alors possible de construire un sous-ensemble de la base de données en ne considérant que les liens entrants ou sortants de Liège, et dont les caractéristiques sont comparées au jeu de données complet dans le Tableau 2. Ces liens entrants et sortants de Liège (plus de 12 000 par jour) apparaissent plus longs que la moyenne belge, ce qui s'explique par une proportion par définition plus conséquente de trajets inter-régionaux et inter-urbains car les trajets intra-urbains sont ici exclus (Liège-Liège en particulier).

		Belgique (jeu de données complet)	Sous-ensemble Liège (liens entrants et sortants de la zone Liège)
Camions par jour ouvrable	Nombre	125 753	5 581
	%	100,00	4,44
Trajets par jour ouvrable	Nombre	551 796	12 204
	%	100,00	2,21
Distance moyenne des trajets ou segments	Réseau (m)	40 892	65 757
	Euclidienne (m)	27 845	46 647
	Temps (minutes)	51	70

Tableau 2 | Comparaison des circulations de camions : base complète Belgique et extraction Liège

II. Liège dans le réseau des échanges de marchandises par camion

Nous explorons dans cette section les circulations des camions en Belgique afin de tenter de révéler et circonscrire l'hinterland de Liège. En observant ce réseau de circulations dans son ensemble (Section II.A.), avant de le circonscrire aux trajets de camions effectués en provenance de Liège et à destination de Liège (Section II.B.), nous approchons ici la délimitation de la zone que Liège polarise économiquement.

A. Circulations des camions à l'échelle de la Belgique et à Liège

Observer la place de Liège dans le réseau des échanges de marchandises par camions passe dans un premier temps par une analyse à l'échelle nationale. Nous décomposons ces circulations en deux composantes complémentaires que nous développons ci-après : le "trafic" et les "connexions".

1. Le trafic

Premièrement, nous définissons *le trafic*, afin de visualiser les disparités d'*intensité de passages de camions* sur l'ensemble du territoire. Plutôt que de cartographier le nombre de points GPS enregistrés par maille d'1 km² (qui aurait été très dépendant des vitesses différentes entre types de routes, d'événements entraînant une congestion forte, ou de déplacements internes à des parkings ou emprises d'entreprises), nous dénombrons le nombre de camions différents (ID) en mouvement dans chaque maille d'1 km², heure par heure. Le fait de dénombrer le nombre de camions uniques par maille chaque heure permet d'éviter les effets d'accumulation de points GPS émis par des camions se déplaçant à très faible vitesse (zones de congestion, chargement, déchargement, etc.). Les informations horaires obtenues à partir des traces complètes des camions (Figure 1a) sont sommées sur 24 heures pour obtenir un trafic journalier ; elles sont ensuite moyennées dans la Figure 5 sur l'ensemble de la semaine (jours ouvrables) pour obtenir un trafic journalier moyen³.

Certes, les principes de base de cartographie tendent à prôner un trafic exprimé en nombre de véhicules par tronçon de route, l'épaisseur du tronçon rendant alors compte de l'importance du trafic (Donnay, 2013). Ici, nous optons pour une solution indépendante d'un tracé de route préalablement défini : une densité de camions par maille. Ceci permet par ailleurs de rendre les différentes analyses comparables en optant pour le même type d'unités spatiales.

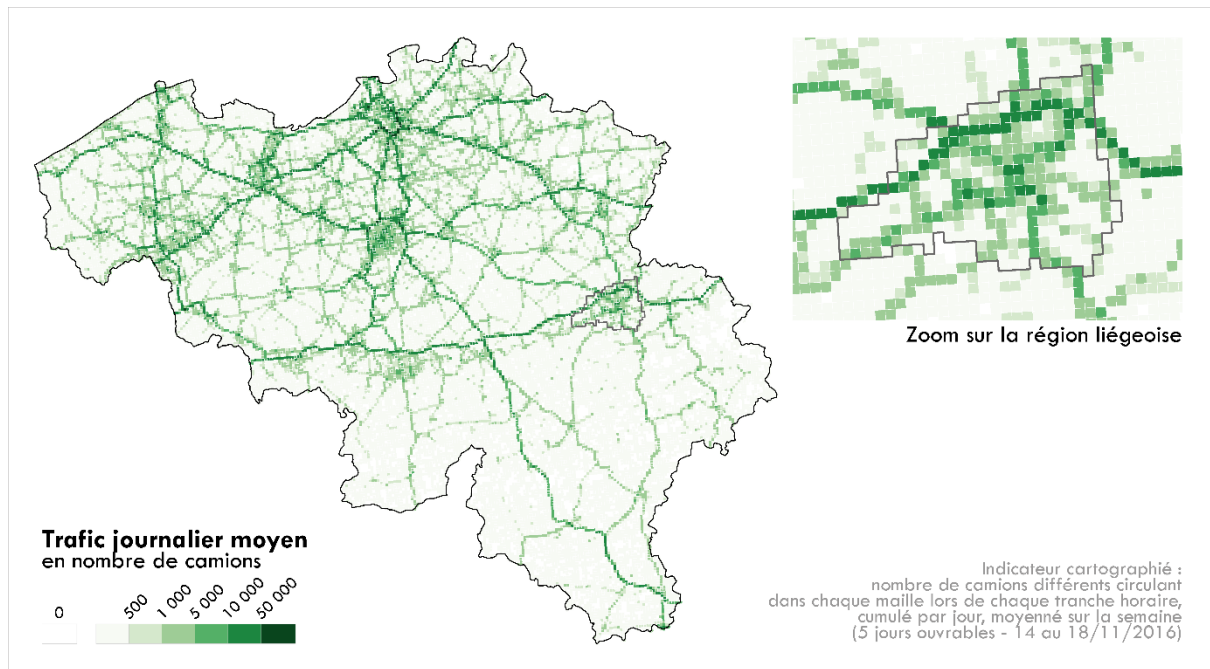


Figure 5 | Circulations de camions en Belgique : trafic journalier moyen

La Figure 5 révèle clairement un réseau routier très hiérarchisé, présentant des intensités de circulation remarquables, notamment aux abords des grandes villes ; les rings de Bruxelles et d'Antwerpen se distinguent tout particulièrement, tout comme dans les environs de Liège l'autoroute de contournement nord. Sans surprise, après les rings de contournement, les autoroutes reliant les pôles économiques majeurs entre eux et aux pays limitrophes sont les lieux où le trafic est le plus important.

Il apparaît ici un des avantages indirects de la mise en place du système Viapass : la possibilité de mesurer avec précision le trafic routier de camions en chaque point du territoire. Le système Viapass permet d'éviter de conserver une infrastructure dédiée (points de comptages en bordure de route), et d'extrapoler sur les tronçons routiers non couverts l'information recueillie de manière ponctuelle. À relatif court-terme, l'exploitation détournée du système Viapass autoriserait les gouvernements à proposer des pistes d'amélioration du réseau routier et, à plus long terme, à détecter en temps réel des zones de congestion inhabituelles afin de prendre des mesures de détournement de trafic.

Cartographier les lieux où *passent les camions* ne révèle cependant pas les lieux visités par ces camions, que nous cernons par les connexions.

2. Les connexions

Au-delà de la visualisation des lieux où *passent les camions*, nous cartographions aussi les *connexions*, soit le cumul du nombre d'origines et de destinations par maille à partir des segments O-D détectés précédemment (Figure 1d). Ce sont alors les *lieux d'où partent et où se rendent les camions* qui sont révélés, indépendamment des routes suivies pour rallier un point d'arrêt (ici, un camion émettant des points dans une maille sans jamais y effectuer au moins une origine ou une destination n'y sera pas comptabilisé). De la même façon que pour le trafic, ces connexions sont initialement construites heure par heure, cumulées sur la journée et

moyennées sur la semaine (jours ouvrables) pour obtenir une intensité de connexion journalière moyenne (Figure 6).

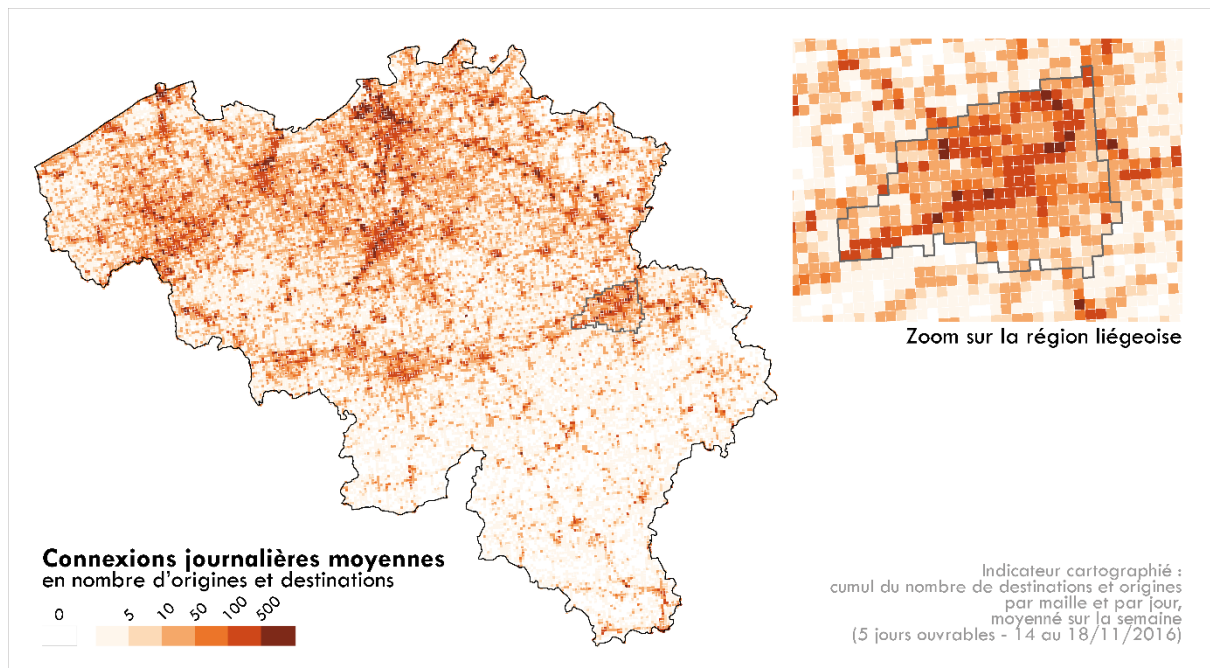


Figure 6 | Circulations de camions en Belgique: connexions journalières moyennes

À travers cette seconde approche, ce sont les lieux d'activité économique reposant sur des échanges de marchandises par la route qui sont révélés. Des pôles majeurs (aéro)portuaires, industriels et urbains se distinguent nettement. On notera premièrement la forte intensité de connexions dans la région de Bruxelles et notamment dans un axe Nord-Est – Sud-Ouest le long du canal Bruxelles-Charleroi, ou autour de l'aéroport de Zaventem. Se détachent également les ports d'Antwerpen, Gent ou Zeebrugge, ou les axes industriels Antwerpen-Hasselt (le long du Canal Albert et de l'autoroute E313) ou Liège-Namur-Charleroi-Mons (le long du sillon Sambre-et-Meuse et de l'autoroute E42). Là aussi les niveaux de connexion sont très hiérarchisés et les deux approches révèlent des réalités bien distinctes mais complémentaires, observables facilement en zoomant sur Liège.

L'apport de cette seconde approche est particulièrement net lorsque l'on compare les connexions au trafic dans la zone de Liège. Si le trafic le plus intense s'observe en bordure nord de l'agglomération (contournement autoroutier) complété par un réseau hiérarchisé d'axes organisés de manière radiale autour de la ville, les connexions (arrêts) concernent avant tout les bordures de Meuse qui correspondent aux localisations économiques industrielles, logistiques et commerciales. Découper les traces GPS complètes des camions en segments origine-destination aura donc permis de révéler ces lieux *d'où viennent* et ces lieux *où se rendent les camions*, au-delà des zones *où ils passent*, que ce soit ici en situation moyenne mais aussi heure par heure (non illustré ici). Au prix d'une lourde préparation et correction des données, nous disposons d'un outil de mesure des circulations en Belgique. Notons toutefois que la base de données ne nous permet pas de connaître l'état de chargement des camions, ni en quantité ni en nature.

B. Révéler la polarisation économique de Liège par les circulations de camions

Plusieurs méthodes complémentaires sont proposées ici afin d'identifier l'aire d'influence de Liège en termes d'échanges de marchandises, approximée par l'analyse des circulations de camions. Dans un premier temps, il est proposé d'évaluer l'emprise spatiale de l'arrière-pensée de Liège en ne sélectionnant que les trajets de camions ayant la zone Liège pour origine ou destinations, en suivant à nouveau les approches trafic (Section II.3.2.a) et connexions (Section II.3.2.b). Ensuite, ce sont l'ensemble des segments O-D réalisés par les camions circulant en Belgique qui sont considérés simultanément dans une matrice partitionnée en communautés de lieux qui interagissent de manière privilégiée (Section II.3.2.c). Les relations internationales de Liège seront enfin révélées dans la section II.3.2.d.

1. Trafic entre la région liégeoise et le reste de la Belgique

Le trafic journalier brut vers ou depuis Liège est tout d'abord comparé au trafic journalier relatif vers ou depuis la même zone (Figure 7). La première grandeur est similaire au trafic national mais n'est mesurée que sur les trajets ayant Liège pour origine ou destination. La seconde ramène cette grandeur au trafic national observé dans chaque maille pour construire une proportion de trafic en lien avec Liège (autrement dit la probabilité qu'un camion circulant dans une maille se dirige vers Liège ou en reparte).

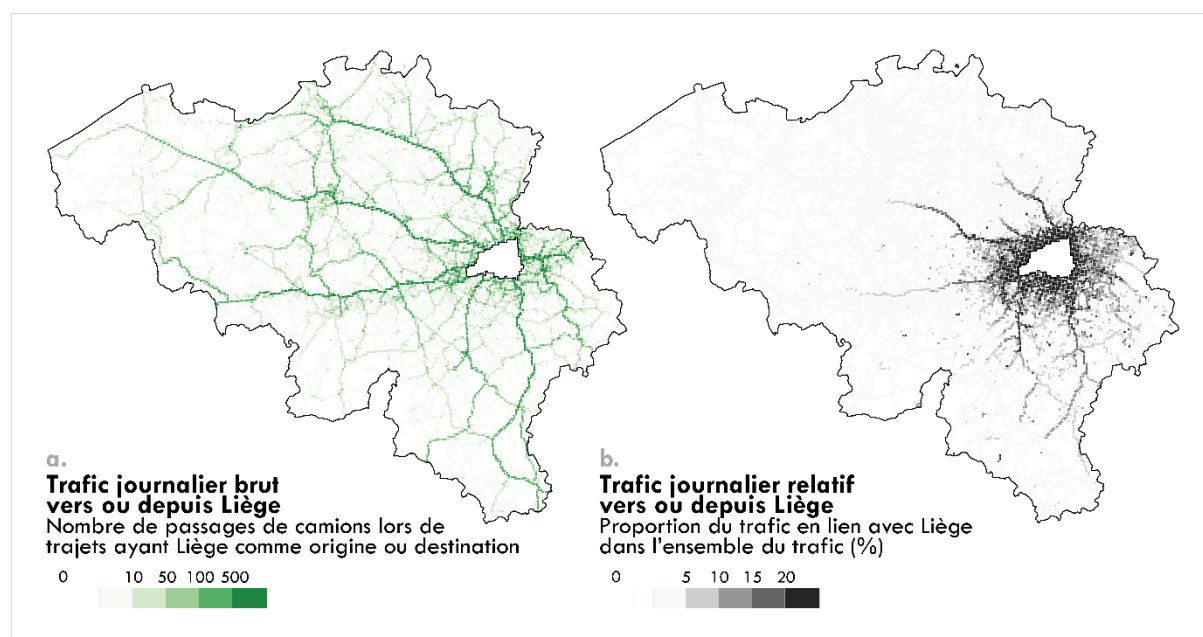


Figure 7 | Trafic journalier brut (7a) et relatif (7b) des camions ayant la zone Liège comme origine ou destination

La Figure 7 montre que le trafic des camions ayant Liège comme origine ou destination s'inscrit dans la cohérence des structures observées dans le trafic national (Figure 5). Notons néanmoins que les camions en provenance ou en route vers Liège ont une plus forte tendance à emprunter le réseau autoroutier plutôt que des routes secondaires localisées à distance de Liège

(Figure 7a). Cet effet est clairement illustré par le trafic relatif vers ou depuis Liège (Figure 7b). Il est évident que les notions de proximité et de distance sont des variables importantes pour les transporteurs routiers. Les camions en lien avec la zone de Liège ont une plus forte propension à rester à proximité de la ville, et ce, quel que soit le type de route, ou à emprunter les axes autoroutiers en direction de certaines villes belges majeures (Charleroi, Bruxelles, Antwerpen) et les pays voisins (Pays-Bas et le Grand-Duché de Luxembourg).

2. Connexions entre la région liégeoise et le reste de la Belgique

De la même façon, nous comparons les connexions journalières brutes vers et depuis Liège aux connexions journalières relatives vers et depuis la même zone (Figure 8). Les premières correspondent au nombre d'origines et destinations de segments ayant Liège comme origine ou destination, alors que les connexions relatives les ramènent aux connexions totales au niveau national. Ici, les connexions journalières relatives correspondraient à une probabilité qu'un camion réalisant une origine ou une destination dans une maille donnée arrive de Liège ou s'y dirige.

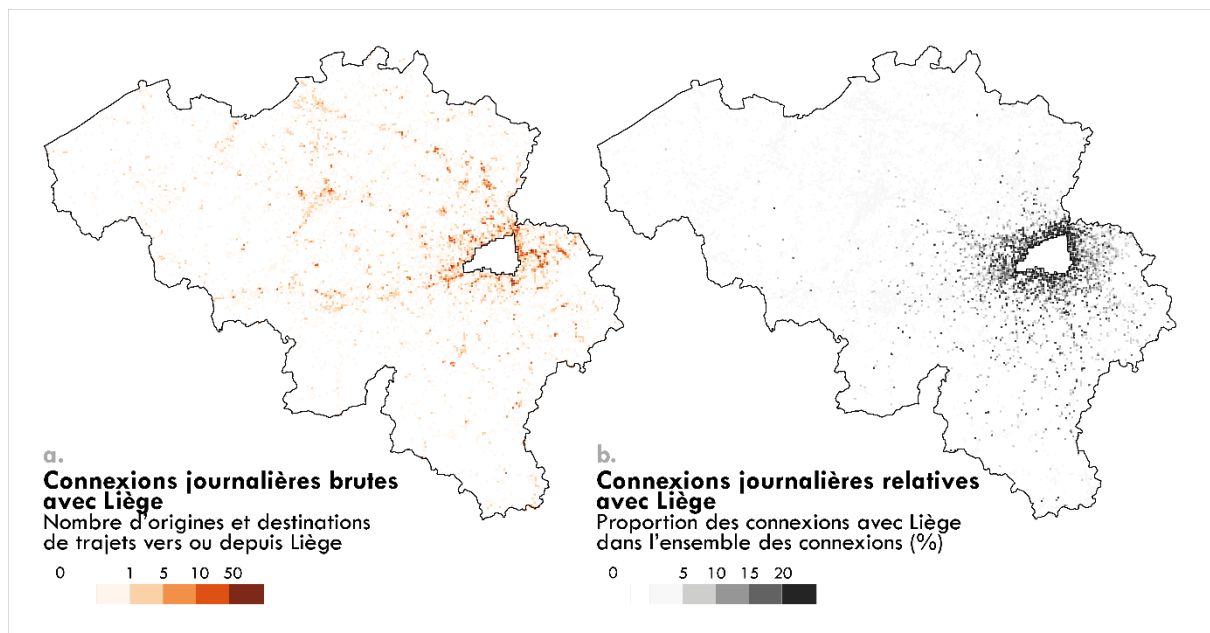


Figure 8 | Origines et destinations brutes (8a) et relatives (8b) des camions ayant la zone Liège comme origine ou destination

Au contraire des deux cartes de la Figure 7 qui mettaient en avant le réseau routier, les cartes de la Figure 8 montrent les lieux fréquentés par les camions provenant de Liège ou s'y dirigeant. Les arrêts de ces camions ont tendance à se localiser à la fois dans la proximité immédiate à Liège, et à distance plus conséquente aux alentours de zones urbaines ainsi que dans des zones industrielles, logistiques et (aéro)portuaires (Figure 8a). L'axe Bruxellois, le sillon Sambre-et-Meuse ainsi que la zone du port d'Antwerpen sont des lieux fortement connectés à Liège, même si les intensités de connexion sont plus faibles que dans sa proche périphérie. À noter qu'une partie conséquente des échanges avec le nord du pays s'effectue par canaux et voies ferrées, diminuant ainsi les échanges effectués par l'infrastructure routière. En

complément aux connexions brutes, les relatives illustrent que les connexions très intenses se réalisent avant tout dans une proximité immédiate à Liège (Figure 8b). Des lieux qui apparaissent très connectés à Liège à travers les connexions brutes (Bruxelles, Antwerpen) ne sont plus présents sur la carte des connexions relatives : au contraire des mailles à proximité immédiate de la ville, ils ne réalisent pas la majorité de leurs échanges avec Liège.

Les Figures 7 et 8 révèlent non seulement le rôle de la distance, frein dans les échanges routiers, mais probablement aussi dans les relations entre établissements d'une même entreprise ou entre clients et fournisseurs, à l'origine de ces circulations et qui pour tout un ensemble de raisons sont plus intenses à proximité directe.

3. Communautés de circulations centrées sur Liège

Afin de s'affranchir de l'étape de sélection préalable des segments commençant ou aboutissant dans la zone Liège (particulièrement sensible à la délimitation même de la zone d'étude), une méthode complémentaire est mobilisée ici afin de partitionner en communautés l'ensemble des segments O-D du réseau belge. Ces communautés sont des sous-ensembles de lieux fortement interconnectés, définis par des algorithmes qui veillent à maximiser les liens internes et minimiser les liens externes à ces sous-ensembles (voir par exemple Fortunato, 2009). Ces sous-ensembles sont ici détectés dans un réseau où les nœuds représentent les mailles d'1 km² et où les liens entre les différents nœuds sont pondérés par le nombre de trajets effectués. Dans notre cas, l'algorithme appliqué sur le réseau O-D est une variante de la "méthode de Louvain" permettant d'extraire un nombre variable de communautés tout en maîtrisant la valeur du paramètre dit de résolution (Reichardt & Bornholdt, 2004; voir Adam *et al.*, 2018 pour des analyses de la sensibilité).

La détection de communautés est appliquée à plusieurs reprises selon diverses valeurs de ce paramètre de résolution. L'arrière-pensée de Liège est alors mise en exergue par cartographie de la fréquence absolue de présence des unités spatiales dans la communauté spatialement centrée sur Liège pour les différents partitionnements (Figure 9). Ces communautés sont détectées en deux temps : d'abord sur l'ensemble du réseau et ensuite en excluant les "liens faibles", dont l'occurrence est très faible au cours de la semaine (moins de 5 segments entre un couple de lieux).

En considérant tout d'abord le réseau complet (Figure 9a), il est mis en exergue une zone liégeoise fortement connectée avec la quasi-totalité de la région wallonne : hormis la zone proche de Tournai, la grande majorité des cellules de la Région sont classées au moins une fois dans la même communauté que Liège. De manière très surprenante, les frontières régionales et provinciales semblent jouer un rôle important dans la structuration du réseau de circulation des camions (la frontière linguistique tout comme les limites de la province de Liège se devinent nettement). Ainsi les camions circulant dans et autour de Liège semblent avoir une forte tendance à se déplacer au sein de la province de Liège, dans une moindre mesure en Wallonie, et de manière très marginale en Flandre. Ces structures "régionales" peuvent notamment s'expliquer par l'organisation par Provinces ou par Régions des activités logistiques, mais aussi de services publics effectuant des tournées régulières.

Lorsqu'ensuite ne sont pris en compte que les liens les plus intenses (Figure 9b), les communautés détectées n'ont plus tendance à présenter une couverture quasiment exhaustive du territoire mais uniquement des zones à la fois fortement interconnectées et fréquemment visitées (des "hotspots" principalement urbains ou industriels). La tendance générale à une plus forte connexion de Liège avec la Wallonie s'observe toujours mais sans une couverture exhaustive de la Région. Par ailleurs, des mailles situées en Flandre sont alors intégrées à la communauté Liège. Le fait ici de supprimer les déplacements sporadiques permet de mieux faire ressortir les échanges les plus robustes et évite la création de "ponts", du fait de seuls trajets exceptionnels, groupant ainsi deux communautés de lieux distincts en une seule. Malgré tout, les mailles du Nord du pays classifiées dans la même communauté que Liège (le long des axes rejoignant Bruxelles, ou au sein des ports d'Antwerpen et Zeebrugge) le sont généralement un faible nombre de fois : ces échanges, s'ils existent bien, restent peu intenses compte tenu des autres échanges. Bien que la frontière linguistique reste globalement suivie, les frontières provinciales ne semblent plus jouer un rôle aussi majeur dans la structuration du réseau ; ce sont ici des zones particulièrement restreintes dans leur emprise spatiale qui sont fortement reliées à Liège, le long des axes Est-Ouest (autoroute E42 entre la France et l'Allemagne et les Pays-Bas en passant par Charleroi et Liège) et Liège-Luxembourg.

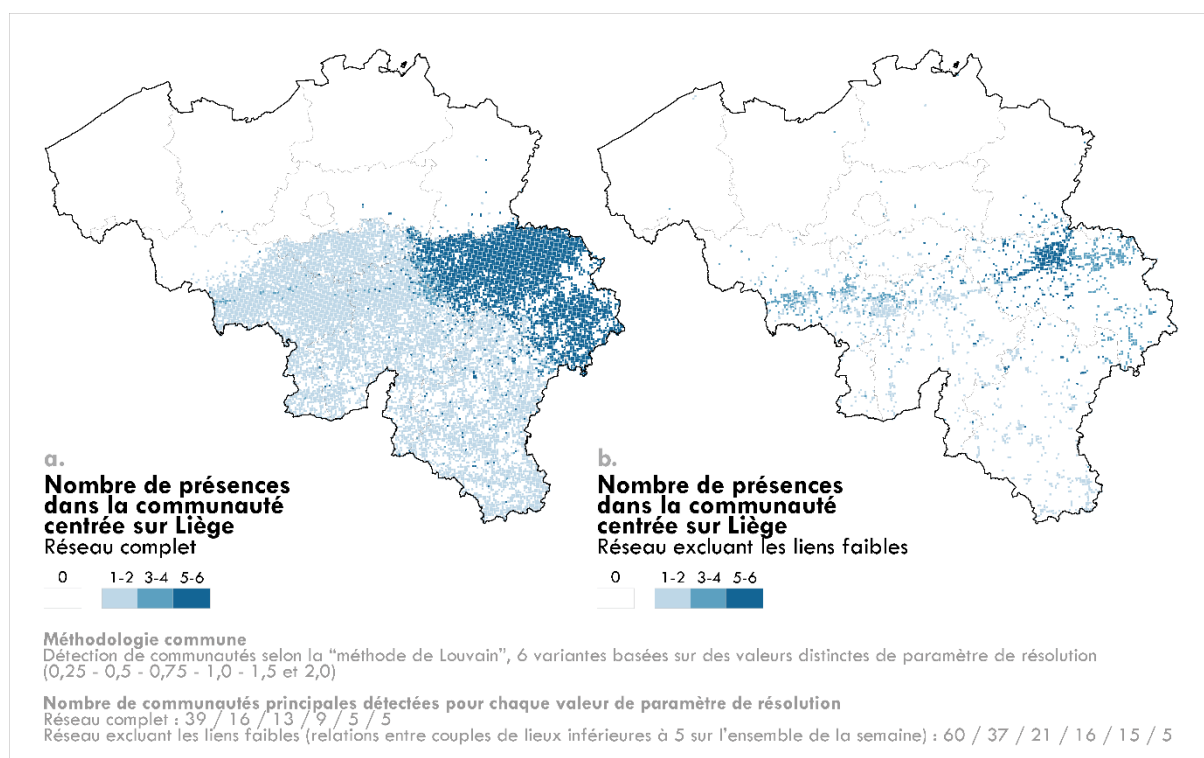


Figure 9 | Nombre de présences des unités spatiales dans la communauté centrée sur Liège, en considérant le réseau complet (9a) ou excluant les liens faibles (9b)

4. Connexions internationales avec Liège

Pour clore ces analyses de la polarisation liégeoise sur base des circulations de camions, une dernière analyse aborde succinctement sa connexion internationale, révélée par la part de

camions ayant Liège comme origine ou destination à chaque point de d'entrée ou sortie majeur du territoire (Figure 10). De même que précédemment, Liège apparaît dans ses connexions internationales tournée principalement vers le cadran Sud : une grande partie des camions du sous-ensemble Liège traversent une frontière par les zones frontalières de Wallonie (E19 vers la France, E25 et E40 vers les Pays-Bas et l'Allemagne, et les points de passage avec le Luxembourg).

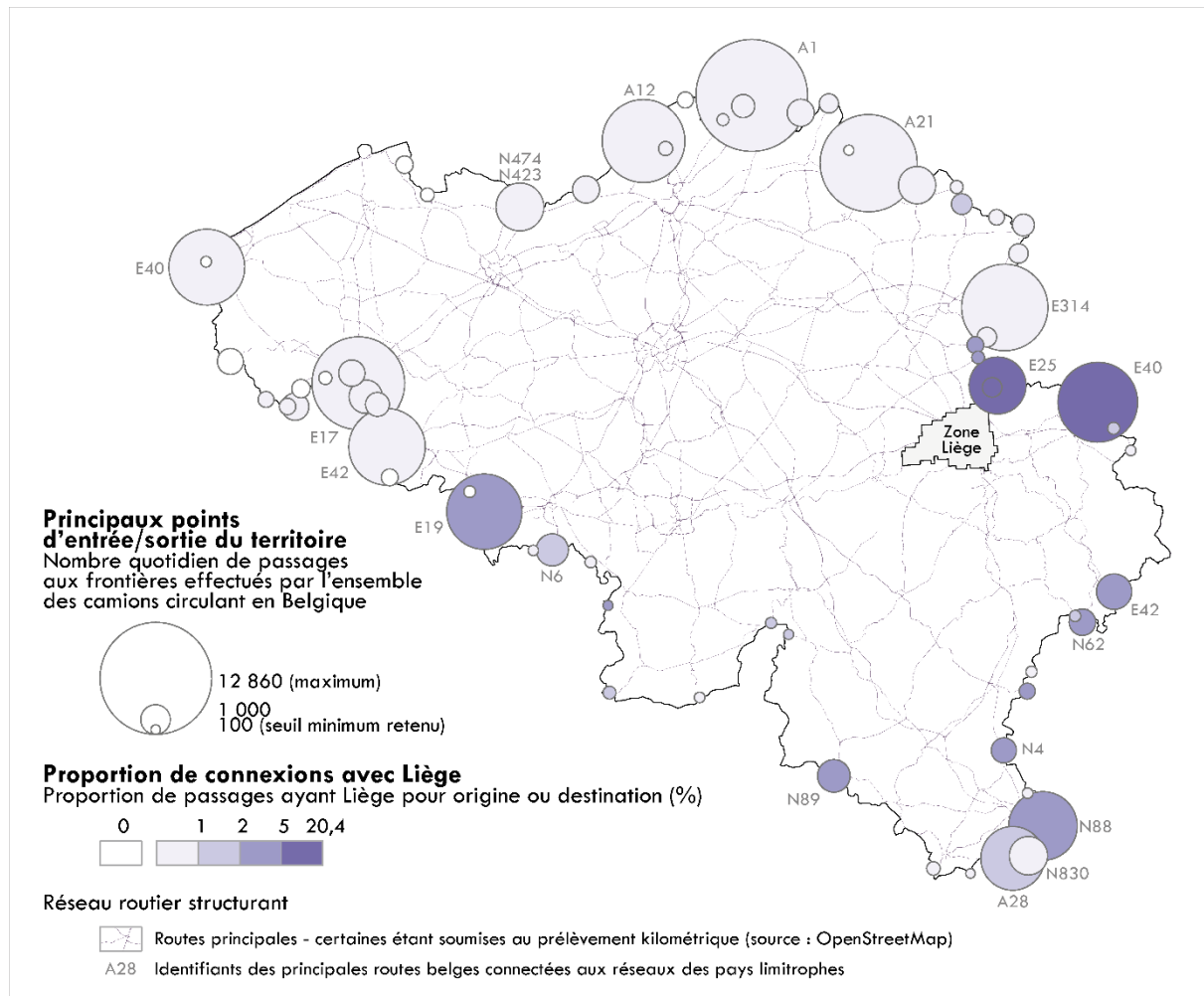


Figure 10 | Connexions internationales de Liège : points d'entrée/sortie du territoire de la totalité des camions et part des connexions avec Liège

Conclusion et discussion

La géographie des transports a toujours cruellement manqué de données spatiales à propos des déplacements des poids lourds et des marchandises. Au terme de cette contribution, les GPS embarqués dans les camions à des fins de prélèvement kilométrique se révèlent être une source de données très riche pour la recherche en géographie des transports et mobilités. Détournées de leur objectif initial et à l'aide de multiples approches (trafic, connexions, communautés), ces données révèlent ici à la fois les structures spatiales de ces circulations de camions à l'échelle de la Belgique et une facette de la polarisation économique de Liège sur son environnement.

Les échanges entre Liège et son hinterland tels que révélés dans cette contribution sont façonnés à la fois par la distance et l'emboîtement de structures spatiales diverses (limites linguistique et provinciales). Liège apparaît avant tout tournée vers le sud du pays, à la croisée des axes Liège-Luxembourg et Namur-Luxembourg et de l'axe industriel Sambre-et-Meuse. Cette tendance à polariser préférentiellement le sud s'explique partiellement aussi par le manque d'alternatives à la route dans le sud du pays alors que les relations avec les ports de mer du Nord s'effectuent aussi par canaux et voies ferrées (Strale, 2009 ; Pekin *et al.*, 2013). Une partie conséquente des importations et exportations de Wallonie s'effectuent probablement via Liège, son port fluvial et son aéroport largement tourné vers le fret. Liège jouerait ainsi le rôle de relais dans la redistribution des marchandises entre le fret aérien, fluvial et ferroviaire et le transport par camion.

Des précautions doivent malgré tout tempérer une partie de nos interprétations. Malgré l'analogie effectuée entre circulations des camions et échanges de marchandises, rien de nous permet de connaître le chargement éventuel d'un camion (ni en quantité ni en nature, seule sa masse maximale autorisée figurant dans les données originales), ni la fonction qu'occupe chaque lieu visité par ce même camion (zone de chargement, de livraison, pause légale ou de confort, etc.). Le postulat selon lequel suivre les circulations de camions permet d'analyser la géographie du transport de marchandises n'est qu'en partie approprié (Pluvinet *et al.*, 2012).

De même, les analyses par détection de communautés menées à la fois sur le réseau O-D complet et sur les seuls liens les plus forts ont montré une forte sensibilité des résultats à la structure du réseau. En effet, réduire le réseau considéré en ne conservant que les échanges fréquents a fait basculer le découpage obtenu d'une couverture spatiale quasiment exhaustive et de communautés "régionales" à une couverture plus restreinte et des communautés de "hotspots". Il s'avère que la conservation ou non des liens faibles produit des communautés aux emprises spatiales sensiblement différentes et offrant peu de recouvrements entre les multiples détections.

Par ailleurs, de lourdes étapes préalables de filtrage et nettoyage ont été rendues nécessaires pour s'assurer de la fiabilité des données à notre disposition. Ces opérations elles-mêmes posent question, les choix effectués impactant inévitablement la base de données et les résultats qui en découlent (Quesnot, 2016). La détermination des arrêts sur base d'un critère temporel arbitraire est par exemple discutable, au vu de la sensibilité de la détection de trajets et segments à ce seuil. Il aurait pu être pertinent d'adopter des critères spatio-temporels tenant compte de la fonction des lieux visités. Nous aurions pu exclure des arrêts détectés, ceux effectués sur des aires de repos, en considérant qu'ils ne constituent pas des lieux de chargement ou déchargement de marchandises. De futures recherches devraient également porter sur la détermination optimale de ces critères d'arrêt par une approche individualiste. Le potentiel du *machine learning* pourrait nous y aider en vue d'une compréhension plus complète des traces spatiales individuelles. Quoi qu'il en soit, le nettoyage s'est ici avéré crucial afin d'établir des analyses robustes tout en évitant les problèmes liés aux erreurs de GPS et aux effets de petits nombres.

Remerciements

Les auteurs remercient l'Institut Bruxellois pour la Recherche et l'Innovation de la Région de Bruxelles-Capitale (Innoviris) pour le financement du projet de recherche (Bru-Net) qui est à l'origine des traitements rapportés dans cet article.

Notes

1 | Certaines catégories (camions-grues, véhicules auto-école, etc.) sont exemptées ; voir <https://www.viapass.be/fr/informations-pratiques/>

2 | Le jeu de données ne couvre que les camions équipés des OBU de l'un des deux prestataires de services accrédités en novembre 2016 (ils sont aujourd'hui au nombre de 5). La politique de Viapass est de ne pas communiquer les parts de marchés respectives de ces prestataires de services. Nous pouvons néanmoins affirmer que l'échantillon utilisé ici ne souffre d'aucun problème de représentativité et peut être considéré comme "quasi-exhaustif".

3 | Pour illustrer ce point d'un exemple, considérons un même camion émettant des points GPS à trois reprises dans la même maille au cours d'une journée : 5 points autour de 8h25, 3 points autour de 8h40 et 5 points autour de 9h35. Dans notre mesure du trafic, ce camion sera comptabilisé une fois dans la tranche 8h-9h et une fois dans la tranche 9h-10h, soit deux fois au cours de la journée.

Bibliographie

Adam, A., Delvenne, J.-C. & Thomas, I. (2018). Detecting communities with the multi-scale Louvain method: robustness test on the metropolitan area of Brussels. *Journal of Geographical Systems*, 1-24.

Antoniou, C., Balakrishna, R. & Koutsopoulos, H. N. (2011). A synthesis of emerging data collection technologies and their impact on traffic management applications. *European Transport Research Review*, 3(3), 139-148.

Commenges, H. (2013). *L'invention de la mobilité quotidienne. Aspects performatifs des instruments de la socio-économie des transports* (Thèse de Doctorat, Université Paris-Diderot-Paris VII).

Dobruszkes, F. (2012). Stimulating or frustrating research? Transport geography and (un) available data. *Belgeo. Revue belge de géographie*, 1-2, 1-15.

Donnay, J.-P. (1995). Delineation of the hinterland of urban agglomerations from a remotely sensed image. *Revue Belge de Géographie*, 119, 3-4, 325-331.

Donnay, J.-P. (2013). Guide de rédaction des cartes thématiques. Méthodes et consignes. Liège : Université de Liège - Unité de Géomatique, 197.

- Flaskou, M., Dulebenets, M. A., Golias, M. M., Mishra, S. & Rock, R. M. (2015). Analysis of Freight Corridors Using GPS Data on Trucks. *Transportation Research Record: Journal of the Transportation Research Board*, 2478, 113-122.
- Gingerich, K., Maoh, H. & Anderson, W. (2016). Classifying the purpose of stopped truck events: An application of entropy to GPS data. *Transportation Research Part C: Emerging Technologies*, 64, 17-27.
- Goodchild, M. F. (2013). The quality of big (geo) data. *Dialogues in Human Geography*, 3(3), 280-284.
- Joliveau, T. (2004). *Géomatique et gestion environnementale du territoire: recherche sur un usage géographique des SIG* (Mémoire d'Habilitation à Diriger des Recherches, Université de Rouen).
- Joubert, J. W. & Meintjes, S. (2015). Repeatability & reproducibility: Implications of using GPS data for freight activity chains. *Transportation Research Part B: Methodological*, 76, 81-92.
- Kuppam, A., Lemp, J., Beagan, D., Livshits, V., Vallabhaneni, L. & Nippani, S. (2014). Development of a tour-based truck travel demand model using truck GPS data. In *93rd Annual Meeting of the Transportation Research Board, Washington, DC*.
- Kitchin, R. (2013). Big data and human geography: Opportunities, challenges and risks. *Dialogues in human geography*, 3(3), 262-267.
- Laurila, J. K., Gatica-Perez, D., Aad, I., Blom, J., Bornet, O., Do, T. M. T., Dousse, O., Eberle, J. & Miettinen, M. (2012). The mobile data challenge: Big data for mobile computing research. In *Pervasive Computing* (No. EPFL-CONF-192489).
- Li, D., Shen, X. & Wang, L. (2018). Connected Geomatics in the big data era. *International Journal of Digital Earth*, 11(2), 139-153.
- Lombard, J. (1999). Territoires, lieux et liens. Relations et savoir-faire de transporteurs de Saint-Omer (France). *Les Cahiers Scientifiques du Transport*, (36), 11-41.
- Ma, X., Wang, Y., McCormack, E. & Wang, Y. (2016). Understanding Freight Trip-Chaining Behavior Using a Spatial Data-Mining Approach with GPS Data. *Transportation Research Record: Journal of the Transportation Research Board*, (2596), 44-54.
- Macharis, C. & Melo, S. (Eds.). (2011). *City distribution and urban freight transport: multiple perspectives*. Edward Elgar Publishing.
- Merenne-Schoumaker, B., Vandermotten, C., Van Hecke, E., Decroly, J. M., Vanneste, D. & Verhetsel, A. (2015). *Atlas de Belgique. Tome 5: Activités économiques*. Academia Press.
- Miller, H. J. (2017a). Geographic information science I: Geographic information observatories and opportunistic GIScience. *Progress in Human Geography*, 41(4), 489-500.

- Miller, H. J. (2017b). Geographic information science II: Mesogeography: Social physics, GIScience and the quest for geographic knowledge. *Progress in Human Geography*, 42(4) 600-609.
- Pluvinet, P., Gonzalez-Feliu, J. & Ambrosini, C. (2012). GPS data analysis for understanding urban goods movement. *Procedia-Social and Behavioral Sciences*, 39, 450-462.
- Quesnot, T. (2016). L'involution géographique: des données géosociales aux algorithmes. *Netcom. Réseaux, communication et territoires*, 30(3/4), 281-304.
- Pekin, E., Macharis, C., Meers, D. & Rietveld, P. (2013). Location Analysis Model for Belgian Intermodal Terminals: Importance of the value of time in the intermodal transport chain. *Computers in Industry*, 64(2), 113-120.
- Pornion, H. (2014). La géomatique, pour le meilleur et pour le pire ? *Géomatique Expert*, 100.
- Reichardt, J. & Bornholdt, S. (2004). Detecting fuzzy community structures in complex networks with a Potts model. *Physical Review Letters*, 93(21), 218701.
- Rodrigue, J.-P., Comtois, C. & Slack, B. (2013). *The Geography of Transport Systems*. Oxon : Routledge, 441 p.
- Shen, L. & Stopher, P. R. (2014). Review of GPS travel survey and GPS data-processing methods. *Transport Reviews: A Transnational Transdisciplinary Journal*, 34(3), 316-334.
- Strale, M. (2009). La mise en place d'une politique wallonne de promotion des activités logistiques; quels enjeux pour le territoire régional?. *Territoire (s) Wallon (s)*, (3), 191-202.
- Thakur, A., Pinjari, A. R., Zanjani, A. B., Short, J., Mysore, V. & Tabatabaee, S. F. (2015). Development of algorithms to convert large streams of truck GPS data into truck trips. *Transportation Research Record: Journal of the Transportation Research Board*, (2529), 66-73.
- Thomopoulos, N. & Givoni, M. (Eds.). (2015). *ICT for transport: Opportunities and threats*. Edward Elgar Publishing.
- Zanjani, A. B., Pinjari, A. R., Kamali, M., Thakur, A., Short, J., Mysore, V. & Tabatabaee, S. F. (2015). Estimation of statewide origin–destination truck flows from large streams of GPS data: Application for Florida statewide model. *Transportation Research Record: Journal of the Transportation Research Board*, 2(2494), 87-96.