HAL open science

Caractérisation de la porosité frontalière par le calcul de densité de points de passage

Matthieu Viry, Timothée Giraud, Marianne Guérois, Ronan Ysebaert, Nicolas Lambert, Amel Feredj

To cite this version:

Matthieu Viry, Timothée Giraud, Marianne Guérois, Ronan Ysebaert, Nicolas Lambert, et al.. Caractérisation de la porosité frontalière par le calcul de densité de points de passage. [Rapport de recherche] RIATE - Réseau interdisciplinaire pour l'Aménagement et la Cohésion des Territoires de l'Europe et de ses voisinages CNRS - CGET - Université Paris Diderot. 2016, 39 p. hal-03589034

HAL Id: hal-03589034

https://hal.science/hal-03589034v1

Submitted on 25 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

RAPPORT TECHNIQUE

6. Caractérisation de la porosité frontalière par le calcul de densité de points de passage

RÉSUMÉ

Ce rapport technique présente le jeu de données OpenStreetMap et les méthodes d'extraction existantes en se concentrant sur les données relatives aux réseaux de transport et aux limites administratives.

Ces méthodes d'extractions seront utilisées pour caractériser la porosité des frontières étudiées via la densité de points de passage (route, rail et ferry) qu'elles présentent.

La densité de points de passage est restituée, pour chaque tronçon de frontière et pour chaque mode de transport, sous deux formes : en fonction de la longueur de frontière (nombre de point de passage pour 100km de frontière) et en fonction du volume de population contenu dans un rayon de 60km autour du tronçon frontalier (nombre de point de passage pour 1 million d'habitants).

Une dernière partie présente différents éléments de la base de données OSM pouvant être utilisés en analyse spatiale et en particulier sur les questions d'accessibilité.

AUTEURS

Matthieu Viry, Timothée Giraud, Marianne Guérois, Ronan Ysebaert, Nicolas Lambert, Amel Feredj (UMS RIATE)

EXPERTS

Sophie de Ruffray (UMR IDEES), Grégory Hamez (LOTERR)

TYPOLOGIE SOCIO-ÉCONOMIQUE DES RÉGIONS FRONTALIÈRES DE L'UNION EUROPÉENNE (2000-2012)

Étude commanditée par le Commissariat Général à l'Égalité des Territoires

TABLE DES MATIÉRES

1 Éléments préalables 3
1.1 Données nécessaires 3
1.2 Outils utilisés 3
1.3 Informations sur les outils utilisés et sur les données OSM 3
2 Préparation et extraction des données 5
2.1 Extraction des réseaux de transport (routier, ferroviaire et ferry) 5
2.2 Le choix d'utiliser la géométrie des frontières issue d'OpenStreetMap plutôt que l'une des deux déjà à disposition 6
2.3 Extraction du découpage OpenStreetMap des entités NUTS2/3 de référence *à partir des géométries RIATE* 9
2.4 Extraction de la frontière internationale terrestre 9
2.5 Extraction de la frontière internationale *maritime* 10
*2.6 (*Optionnel) - Redécoupage des réseaux (conservation des parties situées dans les NUTS2/3 transfrontaliers uniquement) 10
2.7 Extraction des points de passage frontaliers pour les différents réseaux 11
2.8 Nettoyage des géométries créées et pondération des points de passage routiers 12
2.9 Vérification de l'importance des points de passage retenus en utilisant les points de passage créés en utilisant les relations OSM décrivant des itinéraires 16
3 Calcul des indicateurs et mise en forme des données de sortie 18
3.1 Calcul de la densité de points de passage 18
3.2 Vérification de la répartition des points de passage par classe au sein de chaque frontière 22
3.3 Les autres informations disponibles sur ces points de passage 22
3.4 Choix effectués pour la représentation cartographique 24
3.5 Caractériser les frontières ne disposant d'aucun point de passage ? 30
4 Aller plus loin dans l'utilisation d'informations issues d'OpenStreetMap 32
4.1 Avec les données déjà évoquées : comparaison de différents cas d'accessibilité frontalière à d'autres échelles 32
4.2 Avec d'autres informations présentes dans OSM pouvant permettre d'appréhender les questions d'accessibilité des régions frontalières 32
4.3 L'exploitation des métadonnées pour en savoir plus (sur l'historique et, potentiellement, sur la complétude) des principales données utilisées pour ce rapport 34
Bibliographie indicative 35
Annexes 36
Annexe 1 : Arbre de décision correspondant à l'étape de pondération des voies routières 36
Annexe 2 : Les différents types d'actions effectués lors du nettoyage de la table des points de passage routier 37

TABLE DES ILLUSTRATIONS

Figure 1 Exemple de commande ogr2ogr sur laquelle se base l'extraction des réseaux (ici le réseau routier principal) 5
Figure 2 Problème de précision posé par l'utilisation de géométries généralisées / externes à OpenStreetMap 7
Figure 3 Exemple de points de passage créés à tort lors de l'utilisation des géométries EuroGeographics ..... **Erreur ! Signet non défini.**
Figure 4 Exemple de points de passage créés à tort lors de l'utilisation des géométries EuroGeographics 8
Figure 5 Exemple de commande (en « *overpass QL* ») permettant de récupérer la géométrie correspondant aux 4 frontières maritimes évoquées 10
Figure 6 *Exemple de commande SQL sur laquelle repose l'extraction des points de passage (ici pour le réseau routier)* 11
Figure 7 Point de passage à la frontière Franco-Belge. Les différents points (ici en jaune) ne sont comptés que comme un seul point de passage (une entité "multipoint") grâce au regroupement effectué par couple « identifiant de frontière/identifiant d'objet OSM ». 12

Figure 8 Représentation de la distribution suivant le paramètre utilisé et modèle de régression linéaire visant à valider l'opération de nettoyage et pondération des points de passage 15
Figure 9 Représentation de la distribution suivant le niveau hiérarchique du réseau routier utilisé et modèle de régression linéaire 16
Figure 10 Exemple de commande SQL permettant de connaitre les différents types de parcours ferroviaires traversant les frontières. 17
Figure 11 Représentation de la distribution des tronçons de frontière lors de l'utilisation du nombre de voie par million d'habitant dans un rayon de 60km 21
Figure 12 Représentation de la distribution des tronçons de frontière lors de l'utilisation du nombre de voie par 100km de frontière..... 21
Figure 13 Les différents types de voies routières permettant de franchir la frontière 22
Figure 14 Les points de passage ferroviaire disposant d'informations sur le type de service proposé..... 23
Figure 15 Densité de voies routières transfrontalières à l'échelle de l'Europe 24
Figure 16 Densité de voies de passage routières par million d'habitant dans un rayon de 60km (discrétisation en classes d'effectifs égaux). 25
Figure 17 Densité de voies de passage routières par million d'habitant dans un rayon de 60km (discrétisation selon l'algorithme de Jenks). 26
Figure 18 Représentation de la densité de points de passage transfrontaliers par voie ferroviaire à l'échelle de l'Europe 27
Figure 19 Densité de voies routières transfrontalières à l'échelle nationale (France métropolitaine) 28
Figure 20 Densité de voies routières transfrontalières à l'échelle régionale 29
Figure 21 Localisation des points de passage frontalier (automobile) à l'échelle régionale 29
Figure 22 Comparaison de la longueur de frontière entre les tronçons offrant des points de passage et les autres..... 30
Figure 23 Les tronçons de la *polyligne* frontalière n'offrant pas de point de passage 31
Figure 24 *Méthode suivit pour obtenir le nombre de voie correspondant à chaque point de passage* 36
Figure 25 A*utre méthode proposée pour obtenir le nombre de voies correspondant à chaque point de passage* 36
Figure 26 Cas le plus courant : lecture du nombre de voies et regroupement de points proches possédant les mêmes attributs..... 37
Figure 27 Nettoyage nécessaire lorsqu'une route coupe plusieurs fois la frontière ou la longe 38

1 Éléments préalables

1.1 Données nécessaires

Cadre général

- Une extraction des données OSM au format *.osm.pbf* ou *.osm* d'une emprise supérieure ou égale à l'espace d'étude.
- Un fichier contenant les entités géographiques de référence de l'espace d'étude (fichier shapefile de polygones).

Données utilisées dans l'exploration présentée

- Une extraction des données OSM de l'emprise de l'Europe (téléchargeable sur http://download.geofabrik.de/europe.html, le fichier utilisé pour les manipulations à suivre a été récupéré au format *osm.pbf* et porte le *timestamp* du 05-07-2015).
- Le découpage des entités NUTS2/3 de référence produit par le RIATE.
- Le tracé des frontières terrestres et maritimes étudiées, produit par le RIATE (en guise de référence et pour vérification, toutefois le tracé OSM est celui utilisé pour les calculs).

1.2 Outils utilisés

- osmconvert et osmfilter *(→ préparation des données OSM)*
- ogr2ogr *(→ extraction des données OSM)*
- SpatiaLite et sa bibliothèque partagée *(→ stockage des données, traitements SIG)*
- Un environnement Python 3 *(→ interface avec la BD, traitements SIG et statistiques)*

Ces outils fonctionnent indifféremment sur Windows et GNU/Linux et l'ensemble de la chaîne de traitements a été testée avec les deux systèmes.

L'utilisation de SpatiaLite n'est pas une obligation en soit. Son utilisation permet ici de centraliser les données utilisées tout en offrant l'accès à ses fonctions spatiales, notamment un puissant index spatial. Les données qui y sont insérées (sous forme de tables) auraient pu être stockées sous forme de Shapefiles (avec toutefois une première limitation si les fichiers *.dbf* ou *.shp* étaient amenés à dépasser 2 Go, et une seconde concernant le nombre de fichiers produit par les extractions, notamment celles supplémentaires proposées en partie 3) et les fonctions spatiales mises en œuvres ont les mêmes effets avec les différents outils respectant les standards de l'OGC (PostgreSQL, Qgis, OpenJump, OGR ses bindings dans les principaux langages, etc.).

1.3 Informations sur les outils utilisés et sur les données OSM

Les informations nécessaires à la compréhension des spécificités des données OSM ainsi que celles relatives à leur manipulation ont été consignées dans deux documents :

→ Fichier ***TR_Etat_de_l_art_completude_OSM.doc (***état de l'art et informations générales sur OSM et les réseaux de transport qui y sont présents),
Fiches_outils_partie1_outils_generaux.odt et ***Fiches_outils_partie2_outils_python.html***

Le premier vise tout d'abord à décrire le projet OSM, la structure des données et les applications qui en sont faites. Ce document dresse ensuite un aperçu des recherches académiques ayant porté sur OpenStreetMap et plus particulièrement sur la thématique de la « qualité des données », cet aspect ayant souvent été évalué sur le réseau routier.
Le second document, structuré sous forme de fiches, présente les principaux outils ayant été utilisé pour manipuler les données issues d'OpenStreetMap (dont ceux cités dans la partie précédente), en mettant notamment l'accent sur l'utilisation du langage Python pour servir d'interface aux différents outils.

Les données OpenStreetMap sont sous licence OdbL, afin se conformer aux conditions de la licence les données qui en sont issues doivent mentionner « OpenStreetMap et les contributeurs » (licence OdbL). La mention « licence OdbL » peut être remplacée par un lien vers www.openstreetmap.org/copyrigth et on note que cette licence n'empêche pas l'utilisation des données (transformées ou non) dans le cadre d'une autre licence.

2 Préparation et extraction des données

Les étapes ici décrites correspondent aux étapes réalisées dans le cadre de l'exploration de différents aspects des données OpenStreetMap et toutes ne nécessitent pas d'être effectuées :

- les étapes **2.2 à 2.5** qui ne s'appliquent que dans le cas où le tracé des entités de références n'est pas fourni ou ne correspond pas au besoin (en termes de précision par exemple),
- l'étape **2.6** n'est quant à elle utile que pour gagner de l'espace de stockage si le réseau situé dans les entités de référence à vocation à être conservé (pour représentation cartographique ultérieure par exemple) sans volonté de conserver le reste du réseau extrait,
- l'étape **2.9** a seulement pour but de vérifier les résultats obtenus lors du calcul des points de passage par voies ferrées, les résultats de cette étape n'étant pas intégrés par la suite aux résultats fournis.

Les différentes étapes ici détaillées sont chaînées dans un script python permettant de répéter facilement le schéma de traitement en modifiant un ou plusieurs paramètres (le tracé des entités d'étude ou les niveaux hiérarchiques du réseau routier à prendre en considération par exemple).

Note sur la terminologie utilisée :

- Lorsqu'il est fait mention du *réseau routier principal* (ou du *niveau supérieur du réseau routier*) il est fait référence aux voie taguées *highway = motorway | trunk | primary | secondary | tertiary* (appelée *catégorie 1* dans la base de données).
- Lorsqu'il est fait mention du *niveau intermédiaire* du réseau routier il est fait référence aux voies tagués *motorway = unclassified | road* (appelée *catégorie 2* dans la base de données), la mention *réseau routier étendu* signifiant alors la prise en compte du total de ces deux catégories.

2.1 Extraction des réseaux de transport (routier, ferroviaire et ferry)

L'extraction des réseaux est réalisée avec l'outil *ogr2ogr* qui repose sur l'utilisation d'un fichier de configuration (*osmconf.ini*) qui permet de sélectionner les tags à lire. Ces fichiers de configurations sont fournis dans le dossier des scripts d'extraction/préparation des données.

Les données sont filtrées au préalable avec *osmfilter/osmconvert*, cette opération elle-même chronophage va permettre de réduire le temps nécessaire à ogr2ogr pour extraire les informations. Cette étape de filtrage n'est pas indispensable et ne s'avère utile que dans le cas où des nombreux objets issues de *relations* OSM sont à extraire d'un fichier volumineux (alors qu'il garde de bonnes performances lors de l'extraction d'objets simples (*node/way*) depuis un fichier volumineux, ogr2ogr semble s'avérer dans ce cas particulièrement lent à extraire certains types de relations ; le temps utilisé pour un filtrage préalable, rendant ainsi le fichier OSM d'une taille plus raisonnable, sera alors récupéré lors de l'extraction des relations par ogr2ogr).

```
$ ogr2ogr -F SQLite bd.sqlite -dsco SPATIALITE -nln lines_highway
-sql "SELECT * FROM lines WHERE highway LIKE 'motorway%' or highwa
y LIKE 'trunk%' OR highway LIKE 'primary%' OR highway LIKE 'second
ary%' OR highway LIKE 'tertiary%'" -lco ENCODING=UTF-8 --config OG
R_INTERLEAVED_READING YES --config OSM_CONFIG_FILE "fileconf/osmco
nf.ini-HIGHWAY" -gt 65536 --config OSM_MAX_TMPFILE 2048 --config O
GR_SQLITE_CACHE 2048 --config OGR_SQLITE_SYNCHRONOUS OFF europe-la
test.osm.pbf
```

Figure 1 - Exemple de commande ogr2ogr sur laquelle se base l'extraction des réseaux (ici le réseau routier principal)

Les différentes étapes (avec ou sans filtrage) de l'extraction des données se réalisent dans une console (Figure 1). Ils sont regroupés ici dans un script qui prend en argument le nom du fichier

OSM de référence ainsi qu'éventuellement le nom des entités polygones de référence, qu'il se charge d'importer dans la base de données (dont le nom peut également être choisi).

→ Script **extraction.py**

2.2 Le choix d'utiliser la géométrie des frontières issue d'OpenStreetMap plutôt que l'une des deux déjà à disposition

Plusieurs options se présentent lorsqu'il s'agit de confronter les réseaux de transport avec le tracé des frontières. Cette *polyligne* transfrontalière, qui se doit de comporter les identifiants des entités qu'elle sépare, peut en effet être obtenue à partir des entités NUTS2/3 utilisées au RIATE pour les représentations cartographiques (fichiers *Shapefile* nommés *GEOM_EU32_NUTS23* pour la version 'RIATE' généralisée et *GEOM_EU32_NUTS23_ng* pour la version non-généralisée fournie par EuroGeographics, versions retenues pour les calculs SIG dans les TR1 et TR2). Ces géométries peuvent également être obtenues à partir des limites administratives présentes dans OpenStreetMap, présentant alors la difficulté éventuelle de devoir faire correspondre des entités administratives OSM (ou éventuellement un regroupement d'entre-elles) avec le découpage des entités disponibles.

Un extrait de l'espace d'étude a été défini pour réaliser un test comparatif d'extraction des points de passage basé sur la *polyligne* trans-frontalière issue des 3 différentes géométries évoquées.
Cet essai montre, de manière prévisible, que l'utilisation de frontières généralisées (la frontière issue d'OpenStreetMap étant ici celle qui présente le niveau de détails le plus important) engendre des différences dans le nombre de point de passage obtenus (Tableau 1). Afin de comparer l'intérêt des géométries généralisées ou OpenStreetMap (Figure 2, Figure 3 et Figure 4), l'opération a été répétée à l'ensemble de la zone d'étude pour comparer les temps de traitements. Exprimées pour l'ensemble de l'Europe ces différences sont élevées (en ne prenant en compte que les routes principales[1] ce sont 46% de point de passage supplémentaires qui sont créés lors de l'utilisation des géométries *EuroGeographics* et près du double avec les géométries généralisées RIATE). Ces différences peuvent être moins perceptibles pour certaines frontières une fois que ce nombre de point de passage est rapporté au linéaire de frontière, mais elles masquent parfois de fortes différences locales. De plus ces différences vont s'accroître lors de la prise en compte des niveaux hiérarchiques moins important du réseau routier ou lors de la prise en compte du réseau ferroviaire.

En raison de ces différents résultats, et en espérant ainsi restituer au mieux la réalité du terrain, le choix s'est porté sur l'utilisation des géométries issues d'OpenStreetMap. Ce choix n'est toutefois pas anodin puisqu'il va nécessiter plusieurs étapes supplémentaires pour la préparation des données (les 3 étapes suivantes) et que la durée d'exécution du calcul des intersections entre la frontière et le réseau peut être fortement allongée.

On note que, s'il est conseillé de conserver ces géométries très détaillées pour l'étape du calcul des points de passage, il est néanmoins possible de les généraliser (tout en préservant la topologie si nécessaire), permettant ainsi de les utiliser pour des représentations cartographiques à petite échelle sans avoir à manipuler de fichiers volumineux[2].

[1] *highway = motorway|trunk|primary|secondary|tertiary*
[2] Le poids du fichier shapefile contenant les 290 entités transfrontalières pèse environ 50Mo avec les géométries OSM contre environ 1,6Mo avec les géométries EuroGeographics.

	Frontières issues **géométries généralisées RIATE**	Frontière issue des **géométries Eurographics**	Frontières issues d'**OpenStreetMap**
Nombre de point de passage (niveau hiérarchique supérieur[3])	**3134*** ***(3145)***	**2459** ***(2461)***	**1681** ***(1692)***
Nombre de point de passage (niveau hiérarchique intermédiaire[4])	**2411** ***(2417)***	**2180** ***(2182)***	**1030** ***(1030)***
Total	**5545** ***(5562)***	**4639** ***(4613)***	**2711** ***(2722)***
Nombre de couple de NUTS2/3 concernés	**225**	**220**	**213**

Tableau 1 - Comparaison[5] du nombre de points de passage obtenus selon le niveau de généralisation choisi pour les frontières (les 4 frontières maritimes ne sont pas prises en compte)

* Le premier chiffre désigne le nombre de point de passage (sans tenir compte des NUTS2/3 qu'ils séparent, le second chiffre correspond aux même points, affectés des identifiants des NUTS2/3 qu'ils séparent, la différence entre ces deux chiffres correspondant ainsi à des points situés au croisement d'au moins 3 de ces entités).

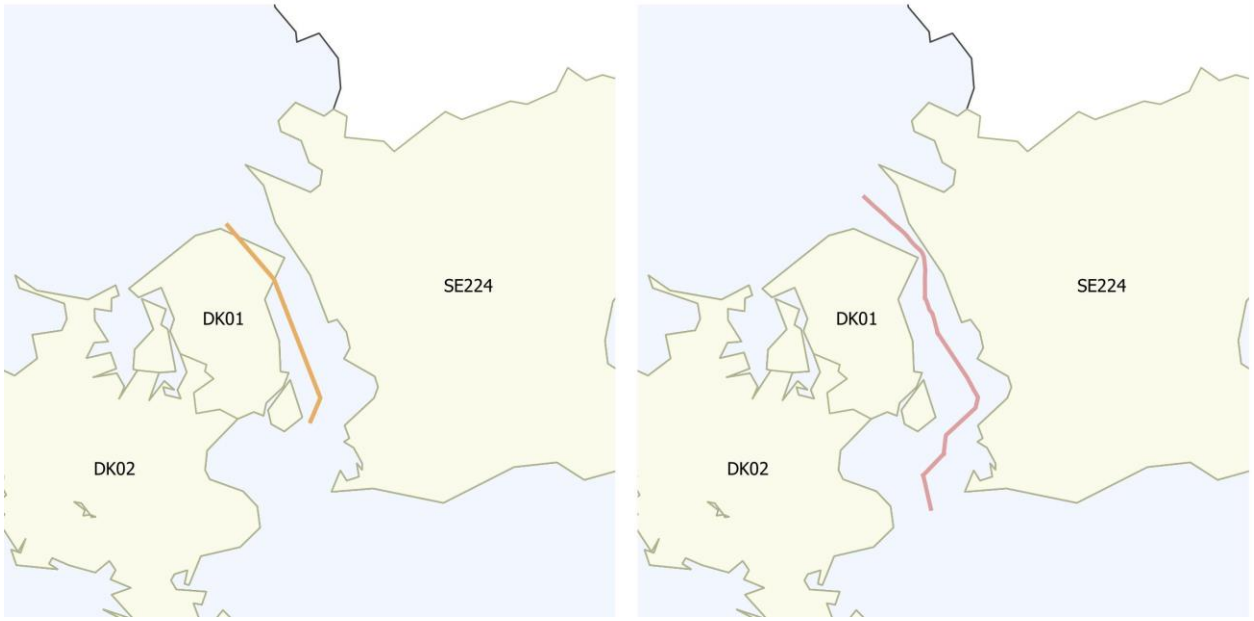


Figure 2 - Problème de précision posé par l'utilisation de géométries généralisées / externes à OpenStreetMap

À gauche : tracé obtenu lors du calcul des discontinuités maritimes (géométries généralisées RIATE)
À droite : tracé retenu pour le calcul des points de passage (géométrie OpenStreetMap)

[3] *highway = motorway|trunk|primary|secondary|tertiary*
[4] *highway = unclassified|road*
[5] Les requêtes SQL correspondantes sont fournies dans le fichier « *comparaison_frontiere_extraction.sql* ».

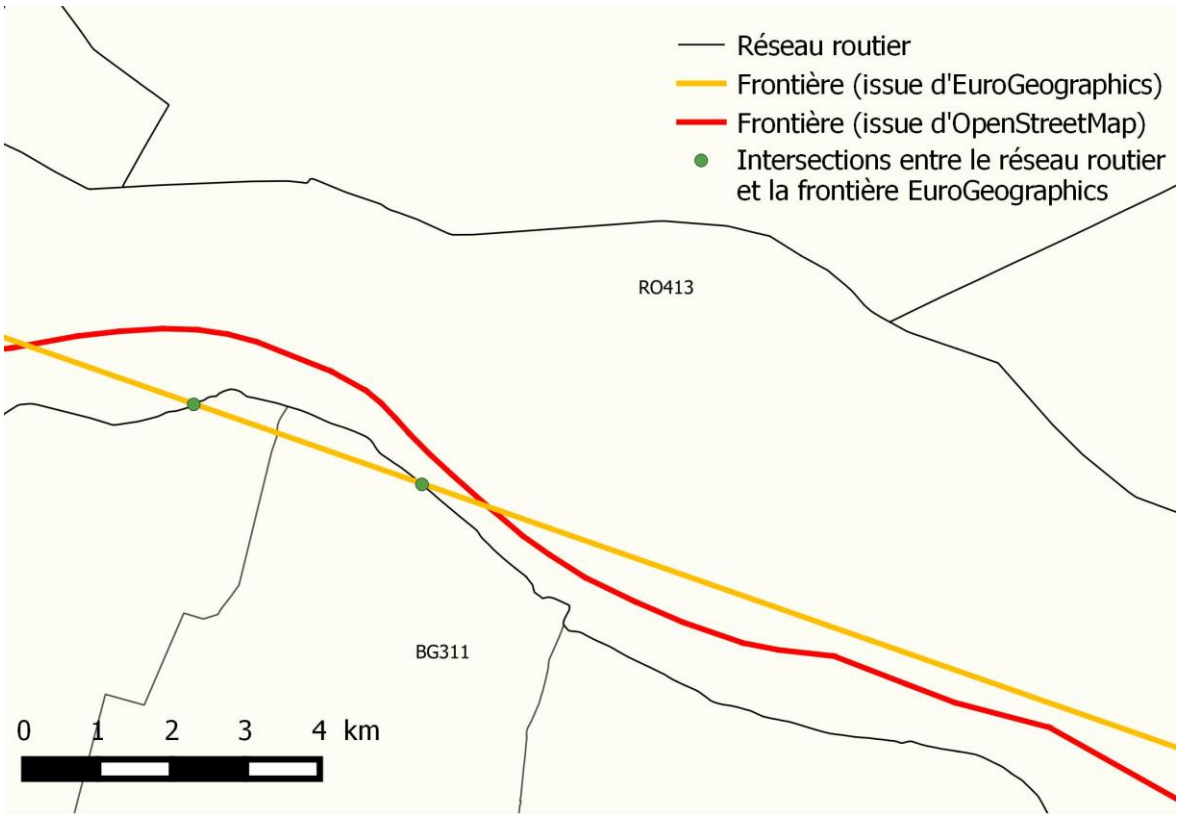


Figure 3 - Exemple de points de passage créés à tort lors de l'utilisation des géométries EuroGeographics

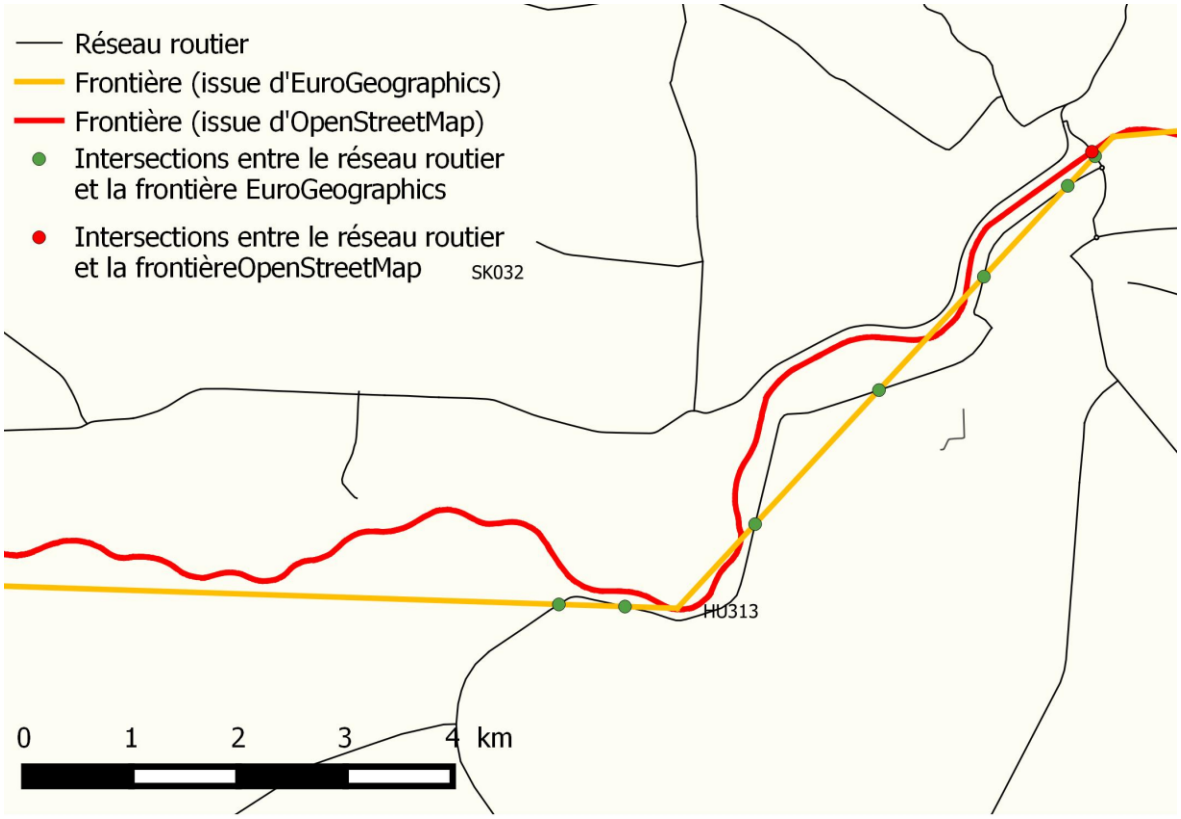


Figure 4 - Exemple de points de passage créés à tort lors de l'utilisation des géométries EuroGeographics

2.3 Extraction du découpage OpenStreetMap des entités NUTS2/3 de référence *à partir des géométries RIATE*

Cette étape vise à obtenir la géométrie OSM correspondant à chaque entités NUTS2/3 définie comme étant transfrontalière (290 entités NUTS2/3, cf. *TR_bande_transfrontaliere.doc*).
Comme l'extraction des réseaux, un fichier de configuration *osmconf.ini* personnalisé est utilisé et les différentes étapes sont regroupées dans un script**.**

→ *Script* ***extraction_limites_admin.py***

Suite à cette extraction un script python a été écrit et vise, dans un premier temps, à chercher parmi ces géométries OSM les entités pouvant correspondre aux entités transfrontalières NUTS2/3 RIATE (cette étape a ainsi été valable pour les pays dans lesquels un niveau hiérarchique OSM correspond au niveau NUTS2/3 retenu) puis dans un second temps, pour les pays dans lesquels il n'existe pas de niveau hiérarchique correspondant, à proposer un regroupement d'entités OSM de niveau inférieur afin de reconstituer l'entité NUTS2/3 en question (cette étape a concerné les entités NUTS2/3 situées en Pologne et en Allemagne).
Cette étape se base uniquement sur la ressemblance géométrique entre les entités de référence et les entités à faire correspondre et n'a pas pu répondre entièrement à sa tâche, en effet quelques entités restantes (moins d'une dizaine) n'ayant pas trouvé de correspondance via ces méthodes ont été sélectionnées à la main.

→ *Script* ***matching_polygon.py***

Plusieurs options se présentent lorsqu'il s'agit de confronter les réseaux de transport avec le tracé des frontières. Cette *polyligne* transfrontalière, qui se doit de comporter les identifiants des entités qu'elle sépare, peut en effet être obtenue à partir des entités NUTS2/3 utilisées au RIATE pour les représentations cartographiques (fichiers ShapeFile nommés *GEOM_EU32_NUTS23* pour la version 'RIATE' généralisée et *GEOM_EU32_NUTS23_ng* pour la version non-généralisée fournie par EuroGeographics, versions retenues pour les calculs SIG dans les TR1 et TR2). Ces géométries peuvent également être obtenues à partir des limites administratives présentes dans OpenStreetMap, présentant alors la difficulté éventuelle de devoir faire correspondre des entités administratives OSM (ou éventuellement un regroupement d'entre-elles) avec le découpage des entités disponibles.

2.4 Extraction de la frontière internationale terrestre

C'est à partir de la géométrie de ces 290 entités transfrontalières que va pouvoir être extrait la *polyligne* (affectée des identifiants des deux entités qu'elle sépare) représentant la frontière internationale entre les différents NUTS2/3.

Le calcul peut prendre un temps important en raison du fort niveau de détails des frontières OpenStreetMap. Le bon déroulement de l'opération peut être testé au préalable sur les géométries moins détaillées disponibles (ici les fichiers *GEOM_EU32_NUTS23_ng.shp* et *GEOM_EU32_NUTS23.shp*).

Cette étape est réalisée en python et utilise les capacités spatiales des bases de données précédemment créées pour transformer le fichier des polygones des entités étudiées en *polylignes* et en extraire la *polyligne* frontalière (cette méthode est basée à la fois sur la topologie du fichier d'origine qui se doit d'être correcte ainsi que sur les valeurs attributaires des polygones de référence qui sont censés contenir un identifiant unique (par exemple un code NUTS) et un identifiant correspondant au pays (ou à un autre échelon supérieur) et elle est donc reproductible selon ces conditions).

Cette opération peut être réalisée dans le cadre de la chaîne de traitement présentée ici, sur la couche insérée dans la base de données (étape 2.1), ou être réalisée de manière indépendante, sur un fichier Shapefile fournit en entrée.

→ Script ***extract_borders.py***

2.5 Extraction de la frontière internationale *maritime*

Le calcul des discontinuités maritime a été réalisé au préalable avec des géométries présentant une généralisation plus appropriée à la taille du territoire d'étude (cf. fichiers *TR_discontinuite.doc* et la sortie *contig_n23_maritime.shp*).
Afin de ne pas générer d'erreurs en utilisant des géométries pouvant utiliser un niveau de généralisation différent, des géométries OSM correspondants à ces discontinuités maritimes ont été récupérées via l'API overpass (voir la commande utilisée en Figure 5). Ces géométries correspondent ici à limite des eaux territoriales et/ou le cas échéant à une ligne équidistante de la côte des deux pays.

```
[out:json];
(
way["boundary"~"administrative"]
  ["maritime"~"yes"]["left:country"="Sweden"];
way["boundary"~"administrative"]
  ["name"="France - United Kingdom"]
["border_type"="territorial"];
way(65710125);
way(49732934);
way(32041859);
);
out body;>;out geom;
```

Figure 5 - Exemple de commande (en « overpass QL ») permettant de récupérer la géométrie correspondant aux 4 frontières maritimes évoquées

*2.6 (*Optionnel) - Redécoupage des réseaux (conservation des parties situées dans les NUTS2/3 transfrontaliers uniquement)

Cette étape a pour but de diminuer le volume des données stockées dans le cas où la base de données contenant les réseaux aurait vocation à être conservée (pour des représentations cartographiques ultérieures ou pour le calcul des points de passage). Ce découpage s'effectue dans *SpatiaLite* et peut nécessiter quelques dizaines de minutes en raison de la taille des réseaux et en fonction du niveau de détails choisi pour le contour des NUTS2/3 (d'autant que l'index spatial est alors à recréer). On note toutefois que ce découpage (entraînant ainsi la suppression des données hors de l'espace d'étude) va nécessiter une opération de nettoyage des pages inutilisées (commande SQL *vacuum*), pouvant elle aussi s'avérer chronophage. Bien que ce découpage permette d'utiliser moins d'espace de stockage, la durée du calcul des points de passage ne s'en trouve que peu améliorée (car elle repose en grande partie sur l'index spatial de la polyligne frontalière). De plus les données résultant de cette étape ne sont pas adaptées au futur calcul des points de passage frontaliers à l'emplacement des frontières maritimes.

→ Requêtes correspondantes à l'opération dans le script **clip_network.sql**

2.7 Extraction des points de passage frontaliers pour les différents réseaux

Ces points de passage frontaliers correspondent aux intersections entre la frontière internationale terrestre et les réseaux précédemment définis. La même opération a été réalisée dans l'étape suivante pour les frontières maritimes. Le cœur de cette étape est réalisé en SQL (

```
- Réseau routier: Niveau hiérarchique supérieur (cf. notice)
CREATE TABLE intersection_highway_geom_osm_t1 AS
      SELECT AsText(CastToMulti(ST_DissolvePoints(ST_Intersection(a.GEOMETRY, b.GEOMETRY)))) as GEOM_TXT,
            CastToMulti(ST_DissolvePoints(ST_Intersection(a.GEOMETRY , b.GEOMETRY))) as GEOM,
            a.FRONT as frontiere_id, b.osm_id as highway_osm_id, b.highway as highway_type,
            b.maxspeed as highway_maxspeed, b.lanes as highway_lanes,
            b.bridge as highway_bridge, b.tunnel as highway_tunnel,
            b.name as highway_name, b.toll as highway_toll
      FROM front_internat_osm as a, lines_highway as b
      WHERE b.highway NOT LIKE '%link' AND ST_Intersects(a.GEOMETRY , b.GEOMETRY) AND
            b.ROWID IN (
                  SELECT ROWID
                  FROM SpatialIndex
                  WHERE f_table_name = 'lines_highway' AND search_frame = a.GEOMETRY )
      GROUP BY highway_osm_id, frontiere_id;
```

Figure 6).


```
- Réseau routier: Niveau hiérarchique supérieur (cf. notice)
CREATE TABLE intersection_highway_geom_osm_t1 AS
      SELECT AsText(CastToMulti(ST_DissolvePoints(ST_Intersection(a.GEOMETRY, b.GEOMETRY)))) as GEOM_TXT,
            CastToMulti(ST_DissolvePoints(ST_Intersection(a.GEOMETRY , b.GEOMETRY))) as GEOM,
            a.FRONT as frontiere_id, b.osm_id as highway_osm_id, b.highway as highway_type,
            b.maxspeed as highway_maxspeed, b.lanes as highway_lanes,
            b.bridge as highway_bridge, b.tunnel as highway_tunnel,
            b.name as highway_name, b.toll as highway_toll
      FROM front_internat_osm as a, lines_highway as b
      WHERE b.highway NOT LIKE '%link' AND ST_Intersects(a.GEOMETRY , b.GEOMETRY) AND
            b.ROWID IN (
                  SELECT ROWID
                  FROM SpatialIndex
                  WHERE f_table_name = 'lines_highway' AND search_frame = a.GEOMETRY )
      GROUP BY highway_osm_id, frontiere_id;
```


Figure 6 - Exemple de commande SQL sur laquelle repose l'extraction des points de passage (ici pour le réseau routier)

Lors de ces requêtes SQL une seule entité "multipoint" est créée (**GROUP BY osm_id, frontiere_id**) lorsqu'un même segment du réseau routier intersecte plusieurs fois une même ligne de frontière (mais les différents point créés restent visualisables au sein de cette entité). Ceci a pour but d'éviter le comptage de nombreux points de passage lorsqu'une route et la frontière partagent leurs parcours ou se croisent à de nombreuses reprises en raison de leur proximité (Figure 7).

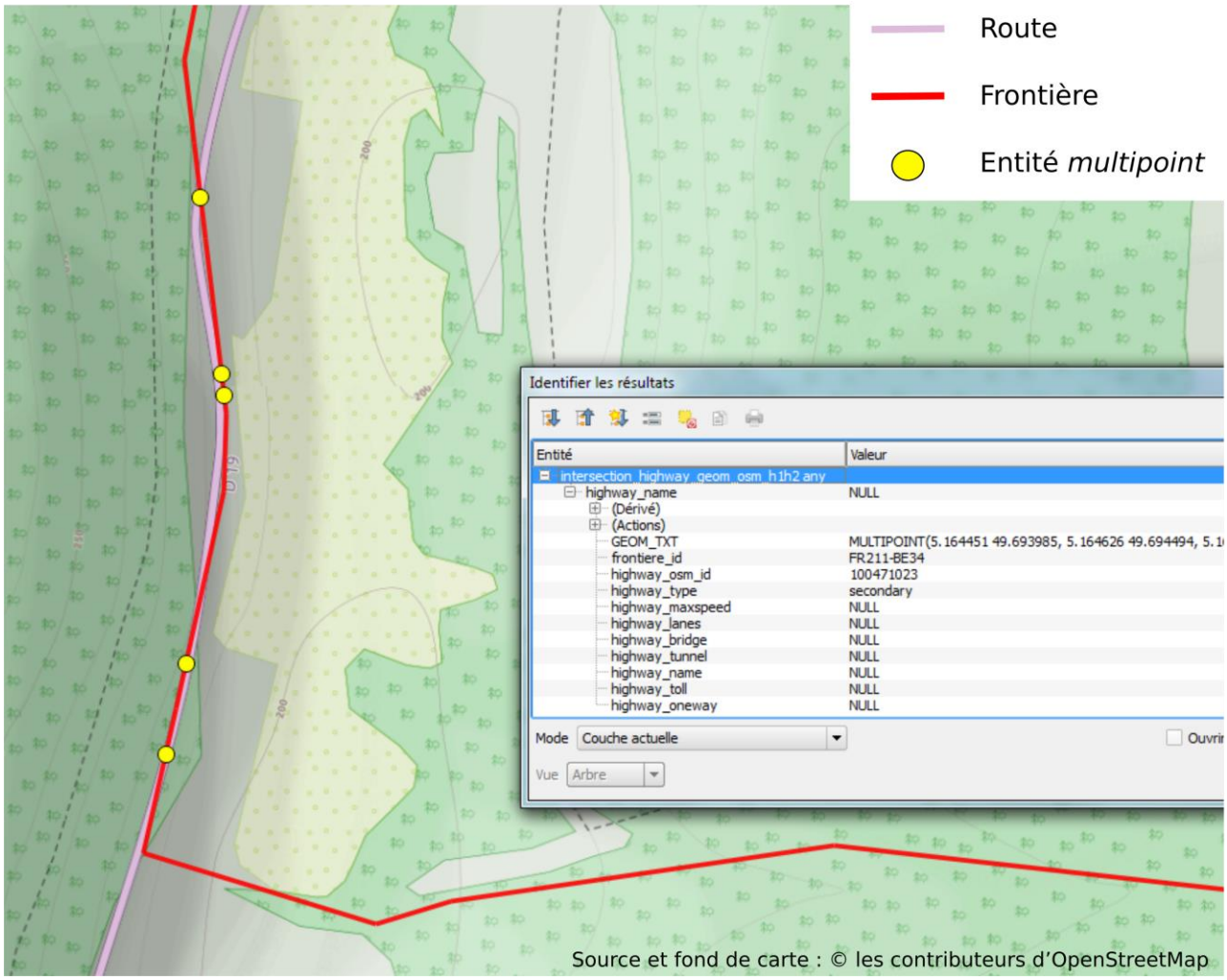


Figure 7 - Point de passage à la frontière Franco-Belge. Les différents points (ici en jaune) ne sont comptés que comme un seul point de passage (une entité "multipoint") grâce au regroupement effectué par couple « identifiant de frontière/identifiant d'objet OSM ».

Ces requêtes SQL permettant l'extraction des différents types de points de passage sont rédigées dans le fichier *extraction_intersection.sql* et exécutées via le script principal en python, qui effectue également par la suite les actions de nettoyage/pondération décrites dans la partie suivante. La relation entre, d'une part, la répartition des points de passage avant et après nettoyage (illustration 8) et, d'autre part, selon l'utilisation de l'un ou des deux niveaux hiérarchiques extraits (illustration 9) est représentée par via leurs droites de régression linéaire. Les principaux résidus correspondent alors respectivement aux frontières comportant de nombreux doublons et aux frontières ne comportant pas de point de passage sur des voies « intermédiaires » (*unclassified* | *road*).

→Script ***extraction_nettoyage_pt_passage.py***

2.8 Nettoyage des géométries créées et pondération des points de passage routiers

Différentes spécificités liées au jeu de données pourraient fausser le calcul de la densité de points de passage et les couches précédemment obtenues nécessitent donc un nettoyage. En effet, que ce soit pour les **voies ferroviaires ou routières**, OSM autorise aussi bien la création d'une géométrie pour chacun des sens de circulation que la création d'une seule géométrie, taguée de manière à indiquer qu'elle est à double sens.

Les **voies routières** présentent également la spécificité de différencier les bretelles des tronçons principaux auxquelles elles se rattachent. Ces bretelles ne seront pas prises en compte dans le calcul des points de passage routier. L'utilisation des informations attributaires de chacun de ces points de passage (type de la voie, sens unique ou non, nombre de voies spécifié ou non, etc.) va permettre de transformer le nombre de géométries obtenues en nombre de voies (directement à partir du nombre de voie renseigné en attributs ou à partir de la catégorie de la voie le cas échéant, voir Annexe 1 pour le détail de l'opération de pondération/comptage des voies) et éventuellement de supprimer des doublons (voir le Tableau 2 pour la répartition selon la catégorie de la voie ainsi que l'Annexe 2 pour le détail des opérations de nettoyage réalisées).

	Type de voie	Nombre de points de passage (avant nettoyage / pondération)	Nombre de voies (après nettoyage / pondération)
Niveau hiérarchique supérieur	*motorway*	126	271
	trunk	78	172
	primary	433	783
	secondary	481	920
	tertiary	574	1093
Niveau hiérarchique intermédiaire	*unclassified*	1017	1007
	road	13	13

Tableau 2 - Répartition des points de passage routier selon le type de voie

Le nombre de voies obtenu lors de cette étape est toutefois sujet à discussion : le choix a été fait d'attribuer une valeur simple (un nombre entier) lorsque le nombre de voie n'était pas présent. Il est également possible d'utiliser l'ensemble des voies présentes dans le jeu de données afin de prédire ce nombre de voies. Cette méthode n'a toutefois pas été retenue dans un premier temps, notamment car il reste possible que les attributs renseignant sur le nombre de voies, tels que '*lanes,*' ne soient présent que lorsqu'ils indiquent un nombre de voie exceptionnellement élevé ou faible par rapport aux voies alentours pour lesquelles cette information n'était pas précisée (de manière analogue à l'arbre de décision suivi, cette option est expliquée en Annexe 1). À l'issue de cette opération, chaque point de passage représente un nombre déterminé de voies.

Les **voies ferroviaires** présentent pour leur part la spécificité de se multiplier à l'approche des gares et notamment à l'approche des plus importantes d'entre elles. Bien qu'il s'agisse théoriquement de voies de service, disposant donc d'un tag approprié, ce n'est pas le cas de toutes. Afin de contourner ce problème ainsi que celui du dédoublement des géométries pour les sens allers/retours, le choix a alors été fait de regrouper en un seul point de passage les différents points et possédant des attributs similaires et situés dans un rayon de recherche défini selon les type de voies (500m pour le réseau ferroviaire, entre 25m et 100m, selon le type de voie, pour le réseau routier). Les points de passage ainsi obtenues sont classés selon qu'il s'agisse d'une ligne de tramway, d'une ligne de train métrique ou d'une ligne de train traditionnelle (voir Tableau 3).
À l'exception des frontières maritimes, on note qu'aucune frontière n'est rendue accessible seulement grâce à l'existence de point(s) de passage ferroviaire.
De plus il est possible de constater qu'il n'existe pas de relation forte entre la densité de points de passage ferroviaire et la densité de points de passages routiers. Cette relation dépend toutefois des catégories du réseau routier qui sont considérées. En effet, la présence de voies ferroviaires semblerait plus en lien avec des catégories intermédiaires du réseau routier (*highway* = *primary* | *secondary*) qu'avec le réseau d'importance locale (*highway* = *tertiary* | *unclassified* | *road*).

Type de la voie (clé *railway*)	Catégorie de voie (clé *usage*)	Nombre de points de passage (après nettoyage)
rail	*main*	122
	branch	37
	NULL	33
	freight	2
	industrial	1
	mainline	1
narrow_gauge	*branch*	4
	NULL	2
	tourism	1
tram	*NULL*	3
light_rail	*NULL*	1

Tableau 3 - Répartition des points de passage ferroviaire selon le type de voie

Les intersections entre **lignes de ferry/bacs de traversée** et la frontière internationale ont fait l'objet du même type de nettoyage visant à enlever les géométries en double lorsque l'aller et le retour sont symbolisés. Une attention particulière a également été apportée afin de ne sélectionner que les points de passage correspondant effectivement à un parcours ayant un embarquement dans les deux pays séparés par la frontière en question (et ne correspondant ainsi pas à un parcours, effectué sur le cours d'eau et traversant virtuellement la frontière, mais dont l'embarquement et le débarquement se situent dans le même pays).

A l'exception des frontières maritimes, on note l'existence de 2 frontières (deux tronçons non-continus de la frontière Roumanie-Bulgarie, d'une longueur totale d'environ 100km) pour lesquelles les liaisons de ce type constituent les seuls points de passage obtenus lors de cette étude.

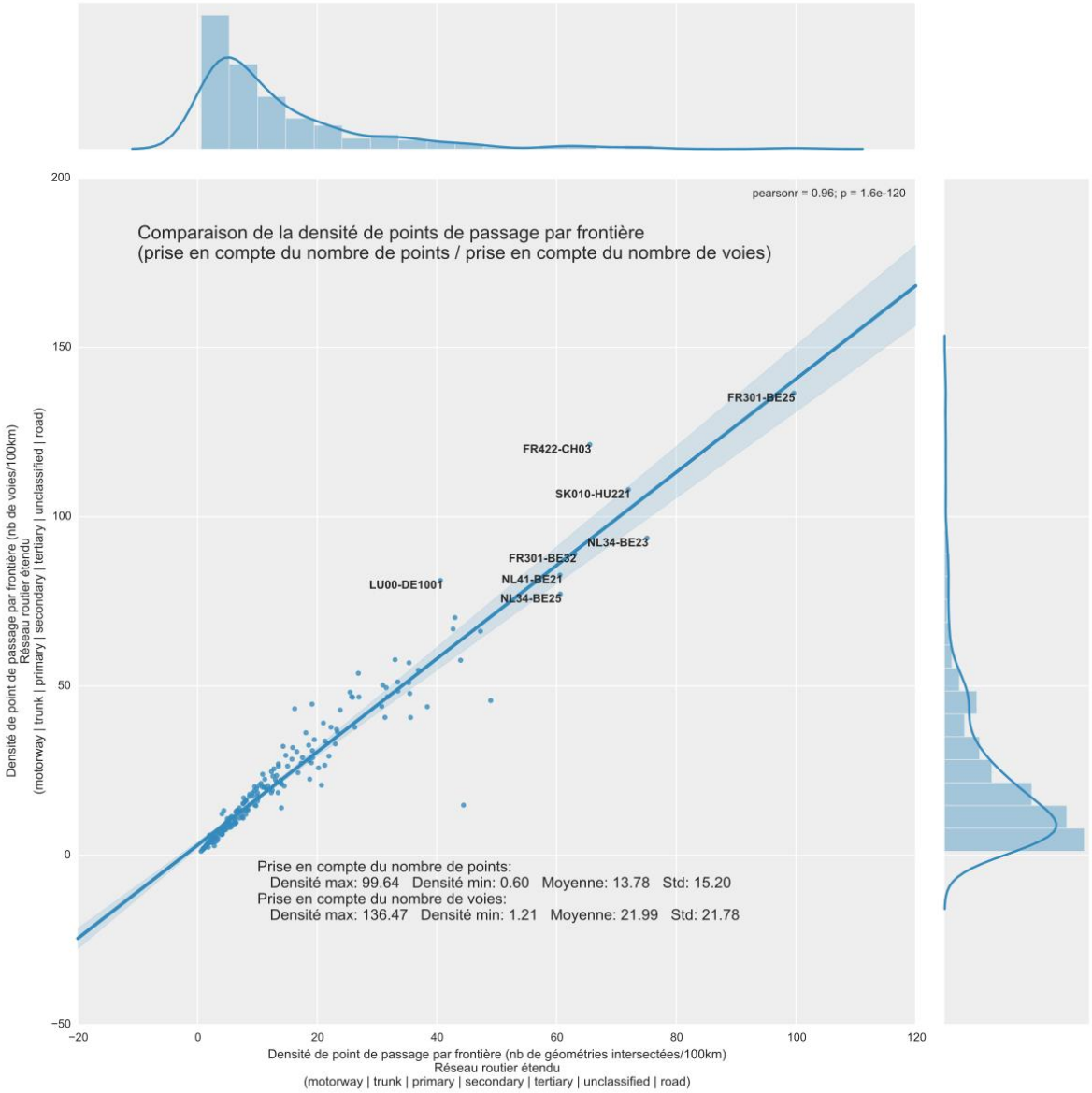


Figure 8 - Représentation de la distribution suivant le paramètre utilisé et modèle de régression linéaire visant à valider l'opération de nettoyage et pondération des points de passage

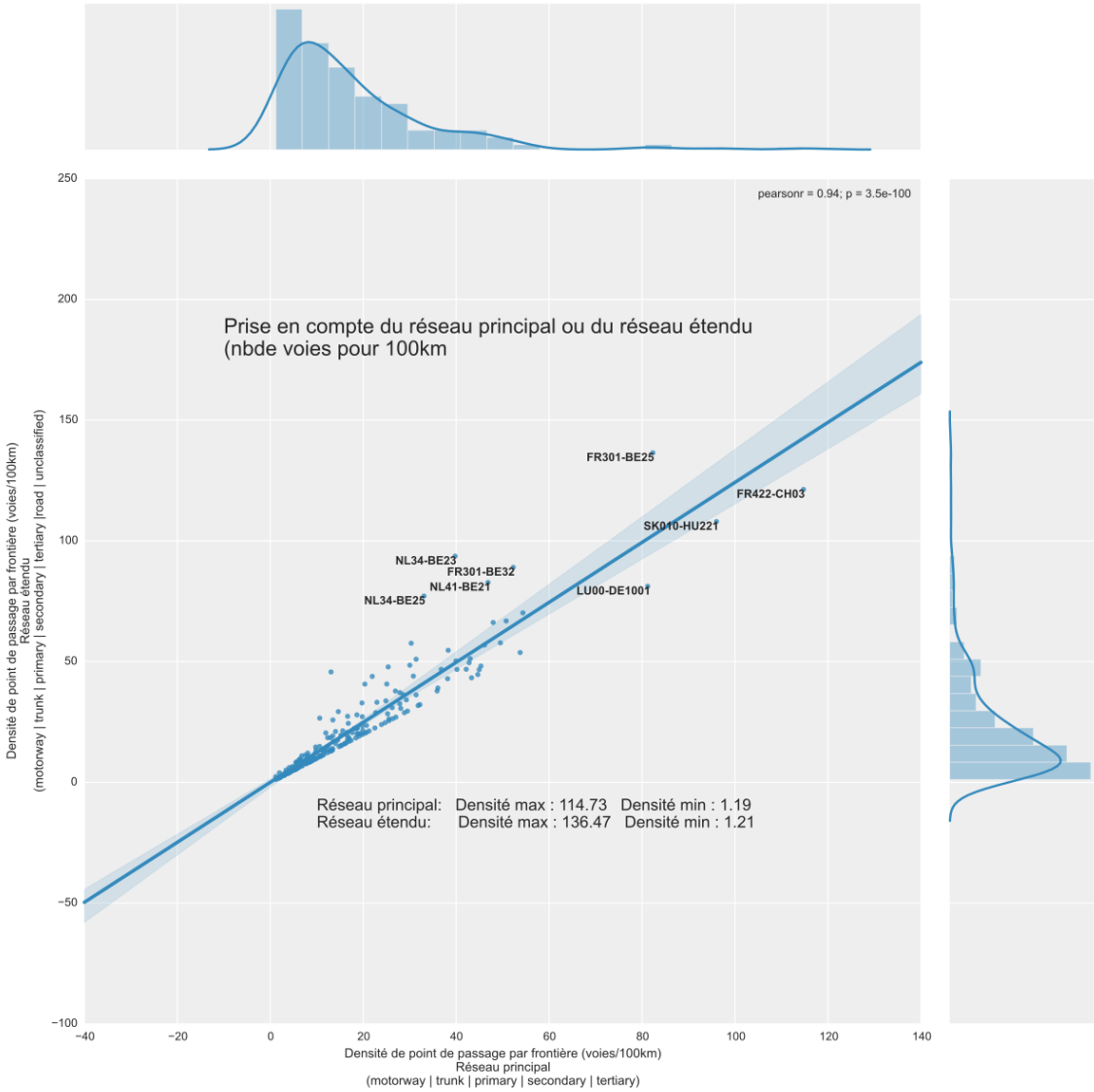


Figure 9 - Représentation de la distribution suivant le niveau hiérarchique du réseau routier utilisé et modèle de régression linéaire

2.9 Vérification de l'importance des points de passage retenus en utilisant les points de passage créés en utilisant les relations OSM décrivant des itinéraires

- Les **relations décrivant des routes destinées aux automobiles** permettent notamment de notifier l'appartenance de certaines routes (telles que l'autoroute française A5) à un réseau particulier ou d'échelle plus importance (tel que le réseau des routes européennes *International E-road network*, dont les routes E54, E511 et E17 utilisent des portions de l'A5).
Ces relations n'apportent que peu d'informations supplémentaires et ne sont pas nécessaire pour calculer les points de passage routiers.

- **L'importance des points de passage** routier pourra toutefois être testée en utilisant l'engin de *routing* OSRM et en générant des parcours transfrontaliers aléatoires. Un script a été écrit permettant de récupérer la géométrie de chaque parcours, permettant de calculer par la suite le nombre de parcours ayant utilisé chacun des points de passage.

→ Plugin QGIS **OSRM**

- Les **relations OSM décrivant des lignes de train particulières** (ligne étant ici entendu comme une succession connectée de rails, pouvant généralement être identifiée par un nom, telle que la ligne *EuroTunnel* et étant tagué avec *route=railway*) ou des **parcours particuliers**[6] (parcours étant ici entendu comme une déserte de train, incluant notamment ses arrêts éventuels, dont les lignes sont ici taguées avec *route=train*) vont pouvoir être utilisés pour mieux connaître l'importance des points de passage précédemment identifiés. En effet, grâce aux informations apportées par les parcours décrits dans ces relations OSM, il va ainsi être possible de connaître approximativement quelles sont les zones et les localités desservies par un point de passage et ainsi d'en estimer le potentiel en terme de voyageurs. De plus la superposition de nombreux parcours pour un point de passage (éventuellement mis en relation avec la faible/forte perméabilité de la frontière sur laquelle il se situe) permet de mettre en évidence son importance dans les éventuels échanges transfrontaliers.

Le comptage du nombre de parcours de train passant par un point peut être effectué avec une requête du type :

```
CREATE TABLE intersection_routes_rail_superposition as SELECT GEOM,
type, route, frontiere_id, count() as nb_ligne

FROM intersection_routes_rail_geom_osm GROUP BY GEOM_TXT, route ORDER
BY nb_ligne DESC, route DESC, GEOM_TXT DESC;
```

Figure 10 - Exemple de commande SQL permettant de connaitre les différents types de parcours ferroviaires traversant les frontières.

Les points pour lesquels le nombre de ces parcours est élevé correspondent ainsi à des lignes importantes et l'on peut dire, d'autant plus si ces points sont situés sur des frontières disposant d'une faible densité de points de passage, que c'est de ces points que dépend une partie des trajets ferroviaires transfrontaliers (il est en effet possible de savoir que c'est grâce à ce point de passage, situé sur la frontière Germano-Suisse, que circulent les trains *'Hamburg – Zurich'*, *'Offenburg – Basel'*, *'Berlin Ostbahnhof - Interlaken Ost'*, *'Berlin – Zurich'* et *'Koln – Basel'* par exemple).
Toutefois, la solidité des informations disponibles par ce biais ne permet pas réellement d'utiliser cet élément comme base pour un éventuel indicateur d'accessibilité, notamment en raison du risque élevé que l'absence d'information ne relève pas d'une absence d'un parcours en train mais d'une omission de saisie des données (en effet les points pour lesquels peuvent être comptés le plus de parcours de ce type sont tous situés dans des zones pour lesquelles de nombreuses contributions ont été effectuées et il est possible que ces points nous informent en réalité plus sur le niveau de complétude des données que sur le potentiel réel en terme de parcours par voie ferrée).

[6] La distinction entre ces deux termes est explicitée plus en détails sur le wiki OSM : http://wiki.openstreetmap.org/wiki/Train_routes et http://wiki.openstreetmap.org/wiki/Tag:route%3Drailway.

3 Calcul des indicateurs et mise en forme des données de sortie

3.1 Calcul de la densité de points de passage

Les étapes finales permettent d'obtenir les densités de points de passage attendues. Lors de cette étape les frontière maritimes (ici au nombre de 4) présentant des points de passage (routiers, ferrés, nautiques) sont également considérées. Les valeurs obtenues pour ces frontières en termes de densité de point de passage restent toutefois peu concrètes car elles dépendent de la longueur de la *polyligne* frontalière choisie (la précision de cette dernière pouvant être variable en fonction des méthodes utilisées pour obtenir son tracé et/ou des géométries duquel elle est issue).

Les densités de points de passage sont calculées de manière à prendre en compte les spécificités des différents réseaux :

- Le nombre de voies routières traversant chaque tronçon de frontière internationale est ramené à la longueur du tronçon en question, afin d'obtenir une donnée comparable dans l'ensemble de l'aire d'étude : un **nombre de voies routières transfrontalières pour 100km de frontière**.
  Cette densité a ainsi été calculée selon que soit pris en compte le réseau routier principal ou le réseau routier étendu).

- Le nombre de voies ferroviaires traversant chaque tronçon de frontière internationale étant plus difficile à appréhender (et moins représentatif du flux pouvant s'y exercer que pour les voitures) ce sont ici les points de passages qui sont comptés et ramenés à la longueur du tronçon de frontière en question, afin d'obtenir une données comparable dans l'ensemble de l'aire d'étude : un **nombre de point de passage transfrontaliers par le réseau ferroviaire pour 100km de frontière**. (indépendant du nombre de voies qu'ils représentent et de l'importance des trains qui y circulent)

- Le nombre de **lignes de ferry**/bac de traversée franchissant chaque frontière ne donne pas lieu à la création d'un indicateur à part entière puisqu'il ne concerne que les frontières (partiellement ou entièrement) situées sur un cours d'eau ou un plan d'eau ainsi que les frontières maritimes. Cet élément vient toutefois compléter les informations apportées par OpenStreetMap en terme de calcul des potentialités d'accès et peut être utile afin de mieux comprendre certaines des densités de points de passages routiers/ferroviaires précédemment calculées. De plus cette information est utilisée par différents engins de routage, prenant ainsi en compte la possibilité (lorsque renseignée dans OpenStreetMap) d'y charger des véhicule ou d'utiliser une ligne régulière lors d'un calcul d'itinéraire piéton (le calcul d'un itinéraire piéton de Lausanne à Évian-les-Bains, effectué avec *GraphHopper*, conseille par exemple de traverser le lac Léman avec la ligne n°1 du CGN, dont la durée renseignée dans OSM est de 35min, alors que le calcul d'un itinéraire routier de Béthunes à Canterbury avec *MapQuest* conseillera de prendre le ferry Calais/Douvres).
  La présence de points de passage de ce type, combinée à la présence de points de passage situés sur des ponts (ou combinée à l'absence de point de passage routier par exemple) pourra laisser présumer une frontière essentiellement située le long d'un cours d'eau.

La table obtenue en sortie de cette étape contient une ligne pour chaque frontière considérée, les différentes cellules représentant les variables précédemment calculées (cf.Tableau 4). Une autre table, moins synthétique et permettant de refaire varier les densités de point de passage obtenues

(par exemple en restreignant le calcul aux autoroute/voie express), est également insérée dans la base de données (elle contient notamment le nombre de points de passage pour chaque catégorie et sous-catégorie (*motorway*, *trunk*, etc.), avant et après l'opération de nettoyage/pondération) et est accompagnée, en sortie, de graphiques simples représentant la forme de la distributions des variables considérées ainsi que la droite de régression linéaire entre les densités de points de passage selon que soit pris en compte le nombre de points créés ou le nombre de voies calculées.

frontiere_id	Densite_route_cat_1	Densite_route_cat_1_2	Densite_rail	Densite_ferry	Q_densite_route_cat_1	Q_densite_route_cat_1_2
FR301-BE25	82.315407	136.470280	NULL	NULL	Groupe5	Groupe5
FR422-CH03	114.732705	121.288859	3.278077	NULL	Groupe5	Groupe5

Tableau 4 - Extrait de la table synthétique obtenue en sortie

Cette table donne les informations au niveau de chaque frontière (217 frontières concernées). Des tables agrégeant ces informations au niveau des entités NUTS2/3 de référence (222 entités NUTS2/3 concernées) et au niveau des États (formant 46 couples d'Etats, cf. Tableau 5, pour 20 pays concernés, cf. Tableau 6) sont également insérées.

Pays 1	Pays 2	Densité de voies routières transfrontalières – Réseau étendu (Nombre de voies / 100km de frontière)	Densité de voies routières transfrontalières – Réseau principal (Nombre de voies / 100km de frontière)	Densité de points de passage ferroviaire (pour 100km de frontière)	Longueur de frontière concernée (km)
CH	AT	13,07	13,07	0,50	198,93
CZ	AT	15,05	13,79	0,84	478,52
DE	AT	14,66	13,93	0,82	1 098,13
DE	BE	46,36	37,98	0,99	202,75
DE	CH	43,08	37,26	2,91	343,57
DE	CZ	9,90	9,18	1,43	1 121,44
DK	DE	43,99	30,79	2,93	68,20
EL	BG	2,75	2,75	0,39	508,82
FR	BE	71,11	44,77	0,97	721,46
FR	CH	38,44	31,05	1,68	595,80
FR	DE	29,47	26,10	1,68	475,05
FR	ES	17,21	12,53	0,84	598,42
FR	UK	0,00	0,00	1,59	62,73
HU	AT	18,92	14,74	1,67	359,47
HU	HR	5,79	5,79	1,24	241,61
IT	AT	5,01	5,01	0,60	498,57
IT	CH	11,28	9,84	0,66	762,42
IT	FR	6,40	6,40	0,71	562,72
LI	AT	27,03	27,03	4,50	22,20
LI	CH	28,40	25,25	3,16	31,69
LU	BE	62,54	46,52	1,55	193,48
LU	DE	30,07	29,37	0,70	143,00
LU	FR	35,91	28,72	7,18	97,48
LV	EE	7,97	6,83	0,23	439,14
LV	LT	8,39	6,10	0,61	655,38
NL	BE	72,75	37,67	1,03	581,41
NL	DE	43,80	32,34	1,03	680,32
NO	FI	1,64	1,64	0,00	733,32
PL	CZ	18,36	15,83	1,39	789,88

PL	DE	13,74	13,51	2,75	436,64
PL	LT	15,96	13,44	0,84	119,01
PT	ES	12,63	10,51	0,23	1 322,05
RO	BG	3,46	3,20	0,80	375,21
RO	HU	8,82	8,60	1,13	441,93
SE	DK	3,44	0,00	0,86	116,22
SE	FI	2,96	2,79	0,17	574,13
SE	NO	7,16	5,78	0,28	1 453,50
SI	AT	24,15	17,97	0,84	356,12
SI	HR	23,75	15,79	1,24	728,45
SI	HU	25,24	20,56	0,93	106,98
SI	IT	30,57	26,75	1,15	261,71
SK	AT	9,26	8,33	1,85	108,04
SK	CZ	17,15	15,29	2,61	268,18
SK	HU	10,16	9,26	1,49	669,41
SK	PL	6,60	6,26	0,52	575,48
UK	IE	47,80	25,37	0,18	543,90

Tableau 5 - Densités de voies/points de passage frontalier par couple de pays

Pays	Réseau routier étendu [7] *(nb de voies / 100km)*	Réseau routier principal [8] *(nb de voies / 100km)*	Réseau ferroviaire *(nb de pt de passage / 100km)*	Ferry et bac de traversée *(nb de pt de passage / 100km)*
CZ	56,16	35,59	1,01	3,12
UK	42,86	22,75	0,33	0,00
NL	40,68	25,45	1,12	4,62
FR	36,55	27,56	1,43	5,40
DK	27,86	23,15	1,55	16,54
PL	25,62	17,89	1,32	2,70
SI	16,17	12,21	0,69	3,47
PT	15,60	13,20	0,38	6,19
DE	14,82	13,34	1,15	8,38
CH	14,46	12,90	0,90	5,36
LI	14,44	11,55	0,51	7,70
LU	13,09	10,80	0,69	2,30
HU	11,99	9,34	1,39	2,79
EL	11,75	10,53	1,34	0,82
LV	9,83	8,18	0,98	0,00
SK	9,10	8,38	1,14	3,64
IT	8,48	7,73	0,61	0,98
NO	7,48	6,87	0,82	1,90
SE	6,82	5,35	0,38	2,47
RO	6,36	6,12	0,98	7,47

Tableau 6 - Densités de voies et de points de passage à la frontière internationale des États européens

→ Script **output_densite_pt_passage.py**

[7] *highway = motorway | trunk | primary | secondary | tertiary*
[8] *highway = motorway | trunk | primary | secondary | tertiary | unclassified | road*

D'autre part il est proposé d'évaluer **le nombre de voies traversant les différents tronçons de la polyligne frontalière en le rapportant au stock de population vivant de part et d'autre du tronçon, dans un rayon de 60km**. La donnée obtenue tient ainsi mieux compte de la répartition de la population dans l'aire d'étude et cette donnée est donc comparable au sein de l'espace étudié. Il est ainsi possible de comparer (notamment via leurs représentations graphiques) les résultats obtenus lors du calcul de la densité de points de passage par rapport à la longueur de la frontière ou par rapport à la population environnante. L'illustration de ce résultat est également présentée en **partie 3.4** (Figure 16 et Figure 17). Toutefois la forme de la distribution obtenue rend plus difficile la création de classe en vue de représenter l'information (voir la comparaison Figure 11 et Figure 12).

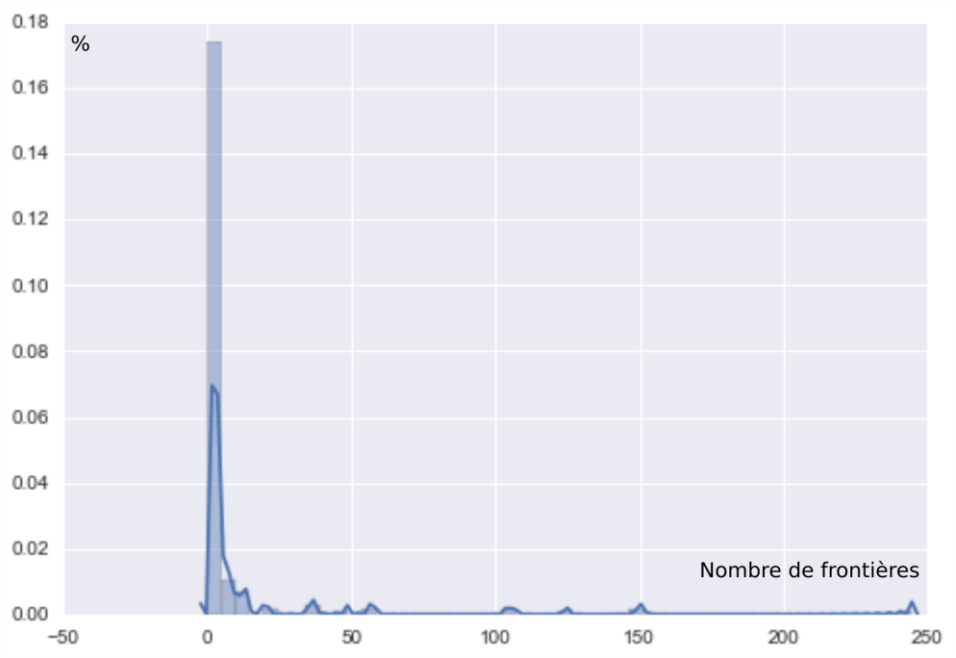


Figure 11 - Représentation de la distribution des tronçons de frontière lors de l'utilisation du nombre de voie par million d'habitant dans un rayon de 60km

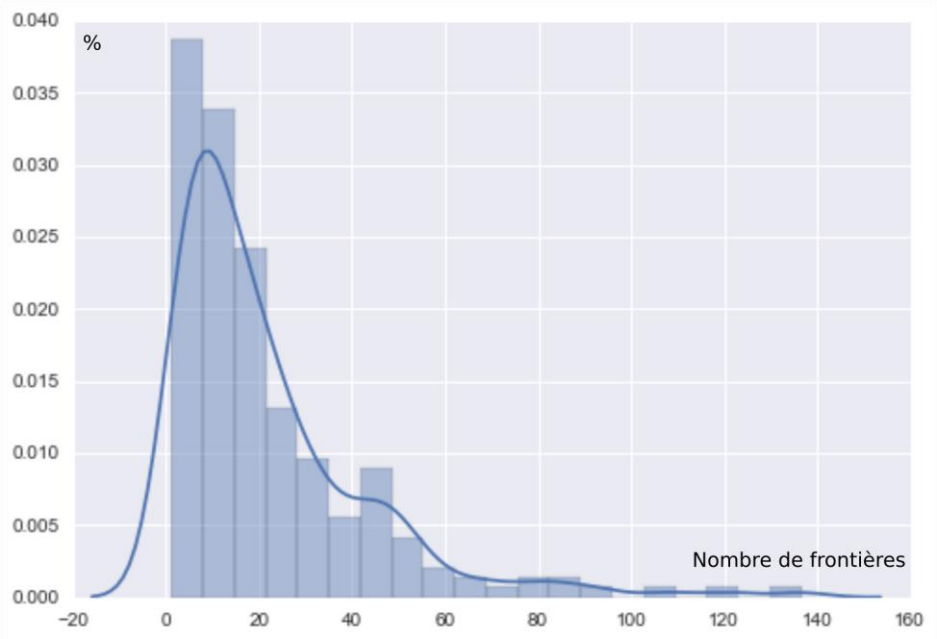


Figure 12 - Représentation de la distribution des tronçons de frontière lors de l'utilisation du nombre de voie par 100km de frontière.

3.2 Vérification de la répartition des points de passage par classe au sein de chaque frontière

La répartition des points de passage par catégories (route, rail, ferry) et par sous-catégories (valeurs prises par les clés *highway* et *railway*) est observée pour chaque tronçon frontière ainsi qu'au niveau de l'ensemble de l'espace d'étude, notamment dans le but de faire émerger des groupes de frontières ayant un comportement similaire.

3.3 Les autres informations disponibles sur ces points de passage

Différents attributs sont présents sur les voies routières et ferrées en plus de ceux permettant de les distinguer en tant que telles et de les hiérarchiser. Les autres attributs ayant été observées (voir notamment Figure 13) pour le réseau routier concernent le fait que la voie soit un pont (*bridge = yes*) ou un tunnel (*tunnel = yes*) et son caractère payant (*toll = yes | hgv | ...*).

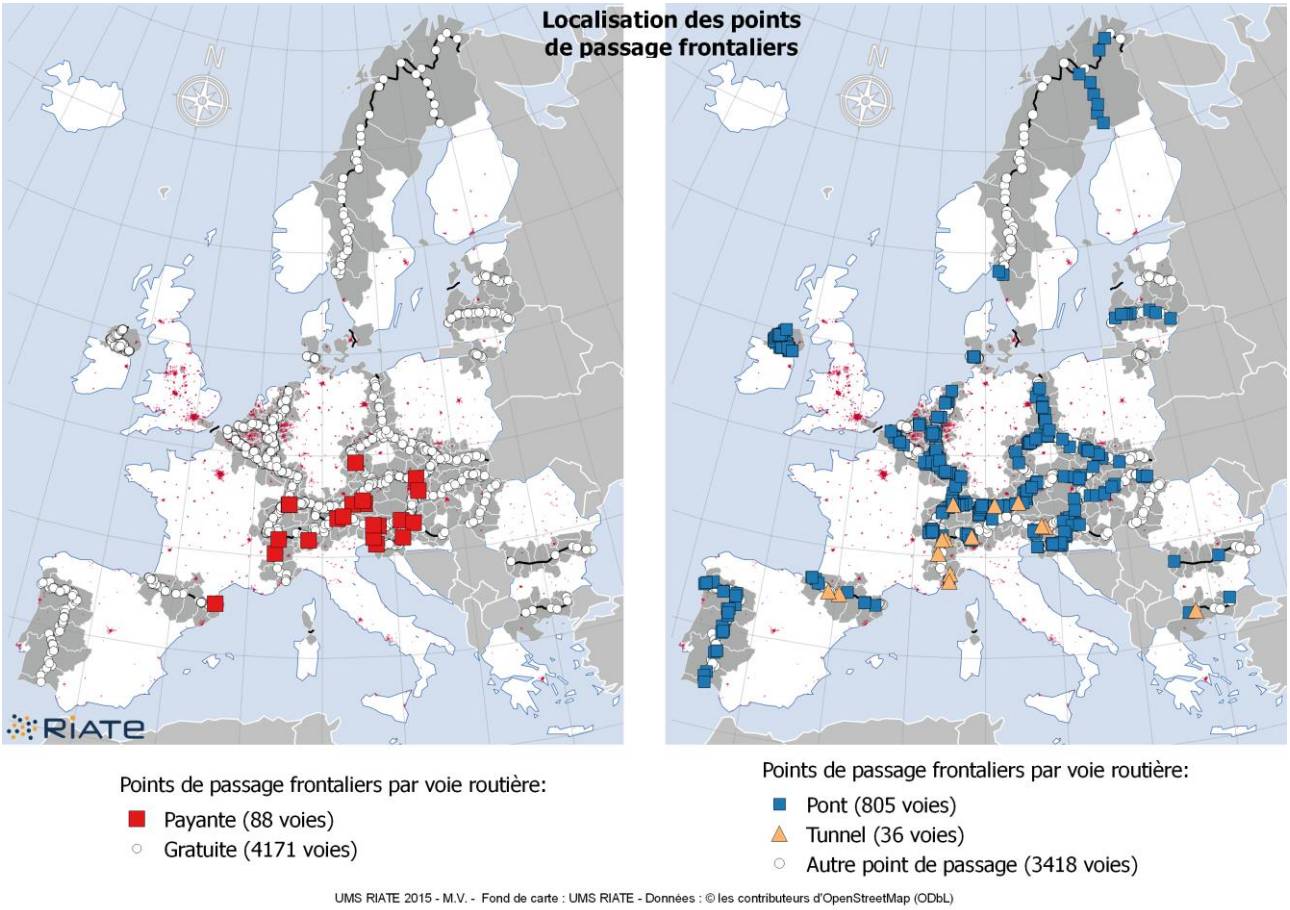


Figure 13 - Les différents types de voies routières permettant de franchir la frontière

Il s'agit d'attributs pour lesquels l'absence de valeur signifie par défaut l'absence du caractère en question (ces informations étant ajoutées par les contributeurs, il n'est toutefois par possible d'avoir la certitude qu'une voie n'ayant pas de clé *toll* n'est pas payante ou s'il s'agit d'une omission, contrairement à celles portant explicitement la valeur *'no'*), de plus certaines valeurs sortent du schéma binaire *yes/no* et décrivent directement l'élément en question (c'est par exemple le cas du tag *tunnel=culvert* qui désigne le fait que la route passe au-dessus d'un cours d'eau busé ou *bridge=movable* qui désigne un pont mobile). Deux autres attributs ont été observés, l'un définissant le nom de la voie (*name=...*) n'ayant été utilisé que lors de l'étape de nettoyage et l'autre définissant la vitesse maximum autorisée sur le tronçon (cette clé peut paraître peu utilisée,

toutefois il est d'usage de se reporter à la table de correspondance[9] lorsqu'aucune valeur n'est spécifiée ; à l'inverse il est recommandé de remplir explicitement ce champ lorsque la vitesse du tronçon est différente de la valeur attendue).

Les attributs ayant été observés pour le **réseau ferré** (Figure 14) concernent l'usage qui peut être fait de la ligne et nous renseignent ici sur la présence de ligne à grande vitesse, de parcours « longue distance »[10] ou au contraire de parcours « régionaux »[11].

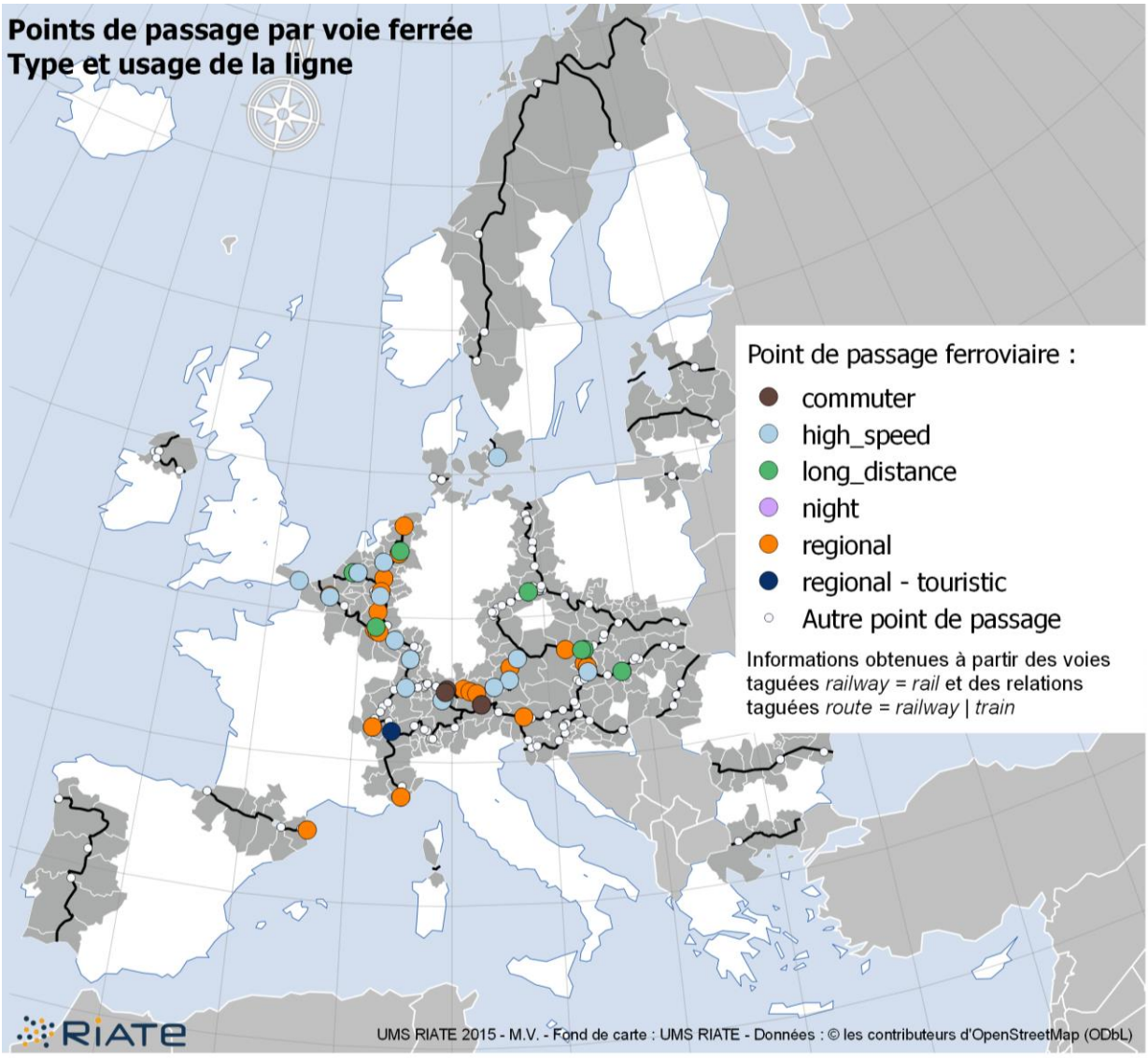


Figure 14 - Les points de passage ferroviaire disposant d'informations sur le type de service proposé

Les points de passage issus des lignes de ferry/bac de traversée disposent également d'attributs permettant d'en connaître les spécificités. En effet il est couramment renseigné l'éventuelle possibilité d'emprunter l'engin avec une automobile (clés *motorcar* et *motor_vehicle*), un vélo (clé *bicycle*), à pied (clé *foot*) ainsi que la durée du parcours de la ligne décrite (clé *duration*). Ces informations n'ont pas été traitées/représentées.

[9] http://wiki.openstreetmap.org/wiki/OSM_tags_for_routing/Maxspeed (consulté le 27/08/2015)
[10] L'un des points de la carte tagué *service='long_distance'* correspond par exemple au parcours Berlin-Budapest
[11] Les TER (et notamment ceux qui franchissent la frontière française en direction de l'Espagne ou de la Suisse) correspondent à ce tag *service = 'regional'.*

3.4 Choix effectués pour la représentation cartographique

Les informations précédemment calculées sont représentées pour 3 espaces d'étude, du global au local :

- L'Union Européenne (Figure 15 et Figure 16),
- La France métropolitaine (Figure 19),
- L'échelon infra-national/régional (croisement des frontières belge, française et luxembourgeoise, voir Figure 20 et Figure 21).

Cette zone frontalière située au croisement de la Belgique de la France et du Luxembourg présente la particularité d'être la zone de cette dimension qui concentre les plus fortes densités (voir Tableau 5 et Tableau 6).

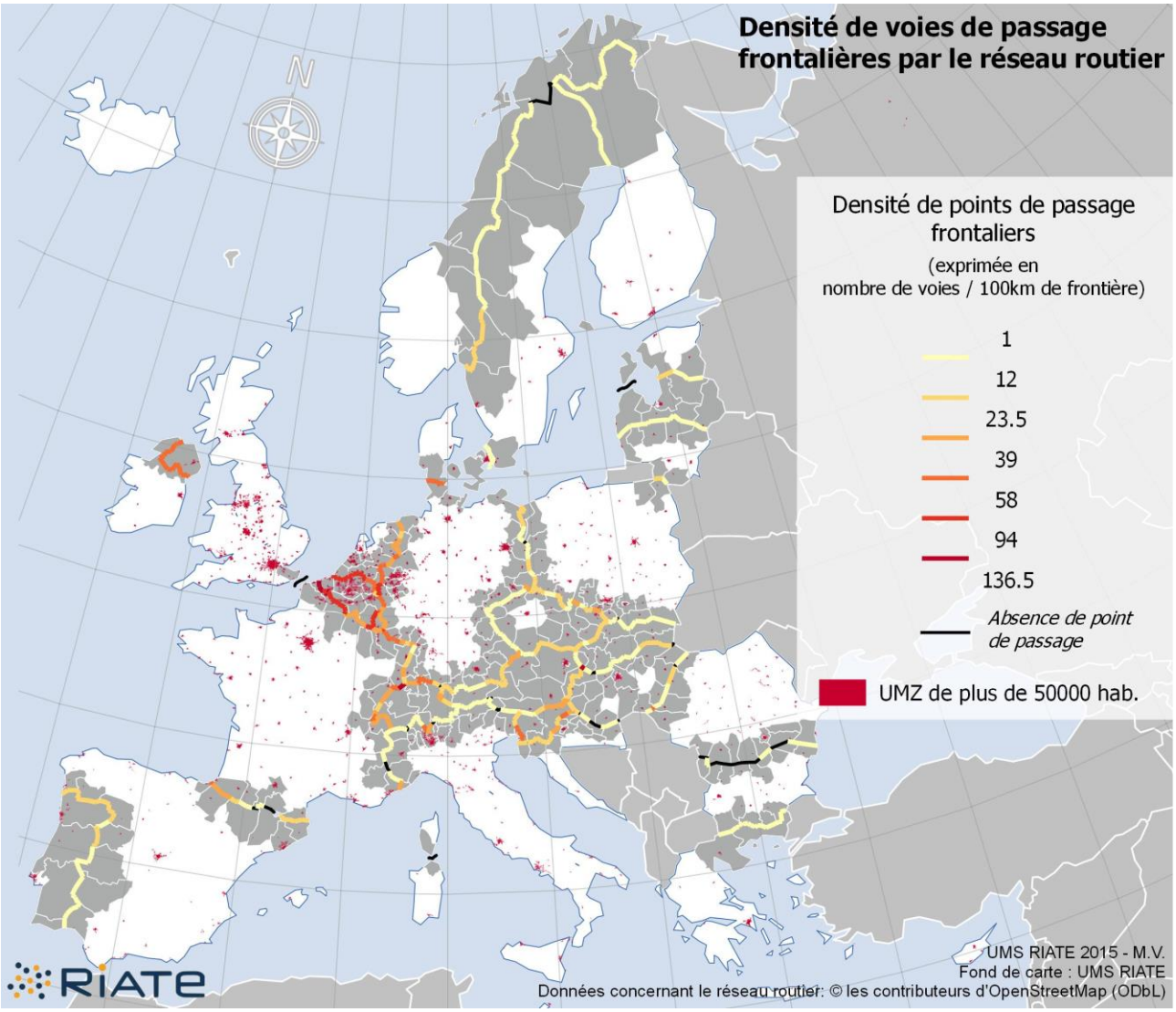


Figure 15 - Densité de points de passages routiers transfrontalières à l'échelle de l'Europe

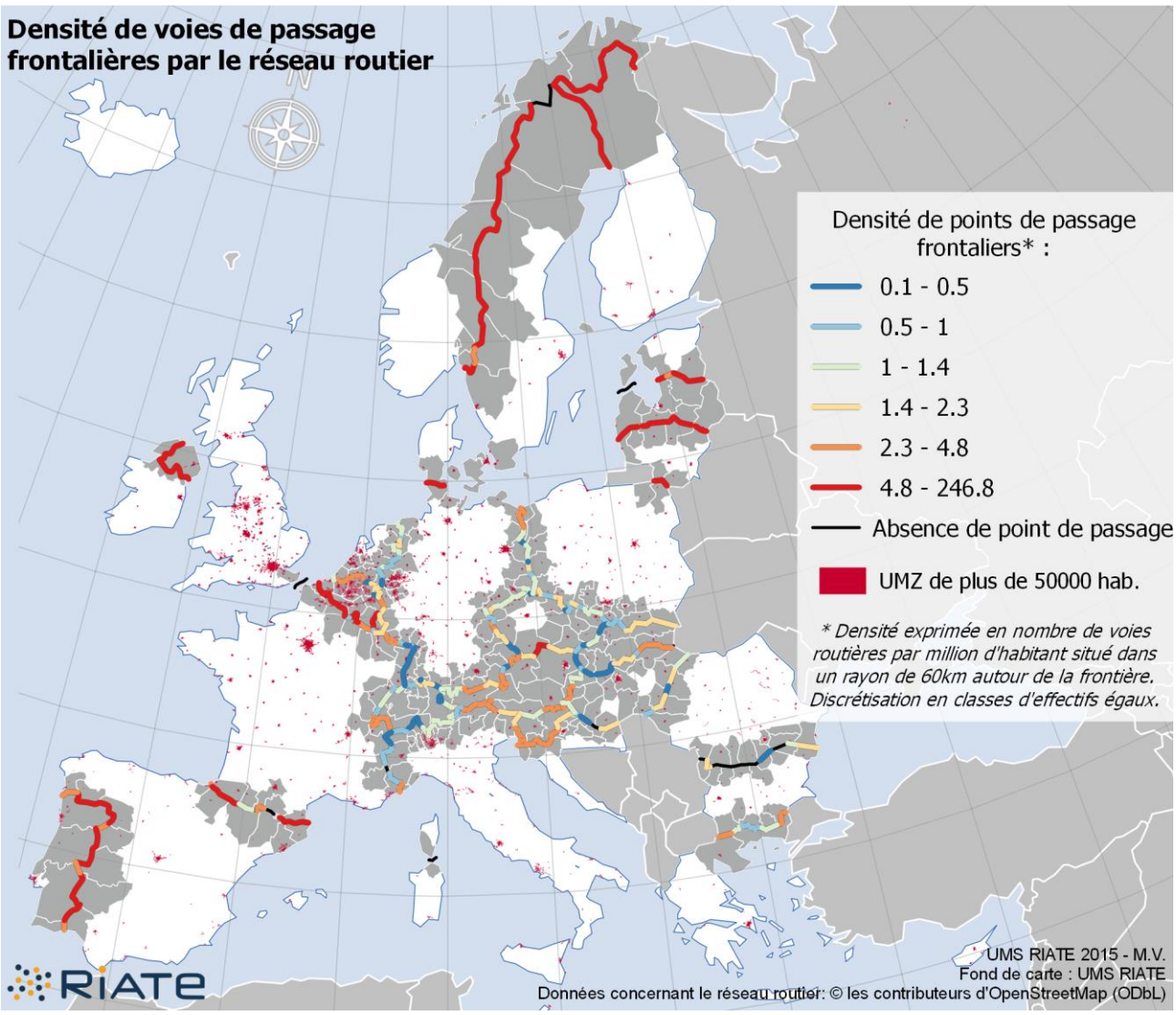


Figure 16 - Densité de voies de passage routières par million d'habitant dans un rayon de 60km (discrétisation en classes d'effectifs égaux).

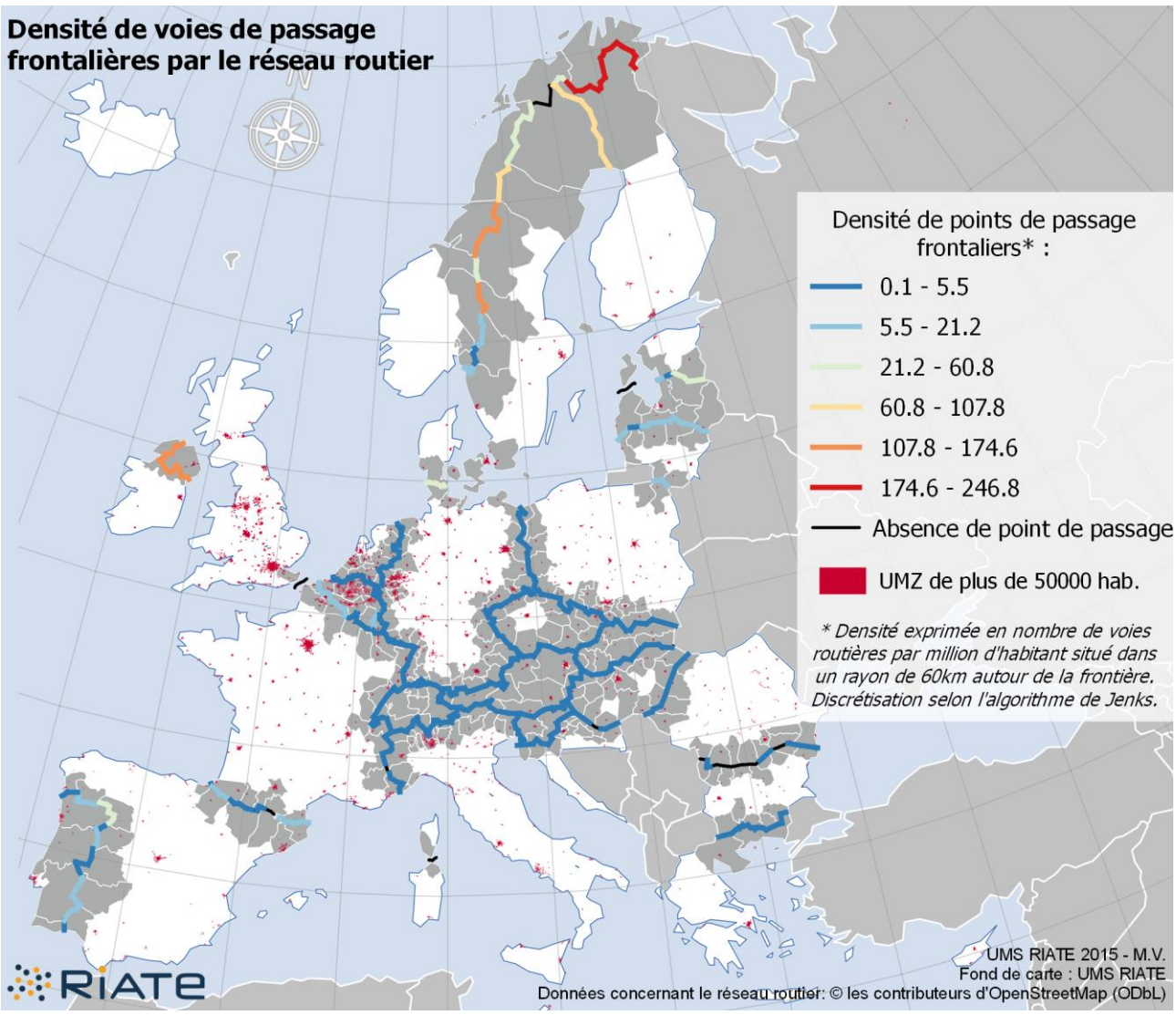


Figure 17 - Densité de voies de passage routières par million d'habitant dans un rayon de 60km (discrétisation selon l'algorithme de Jenks).

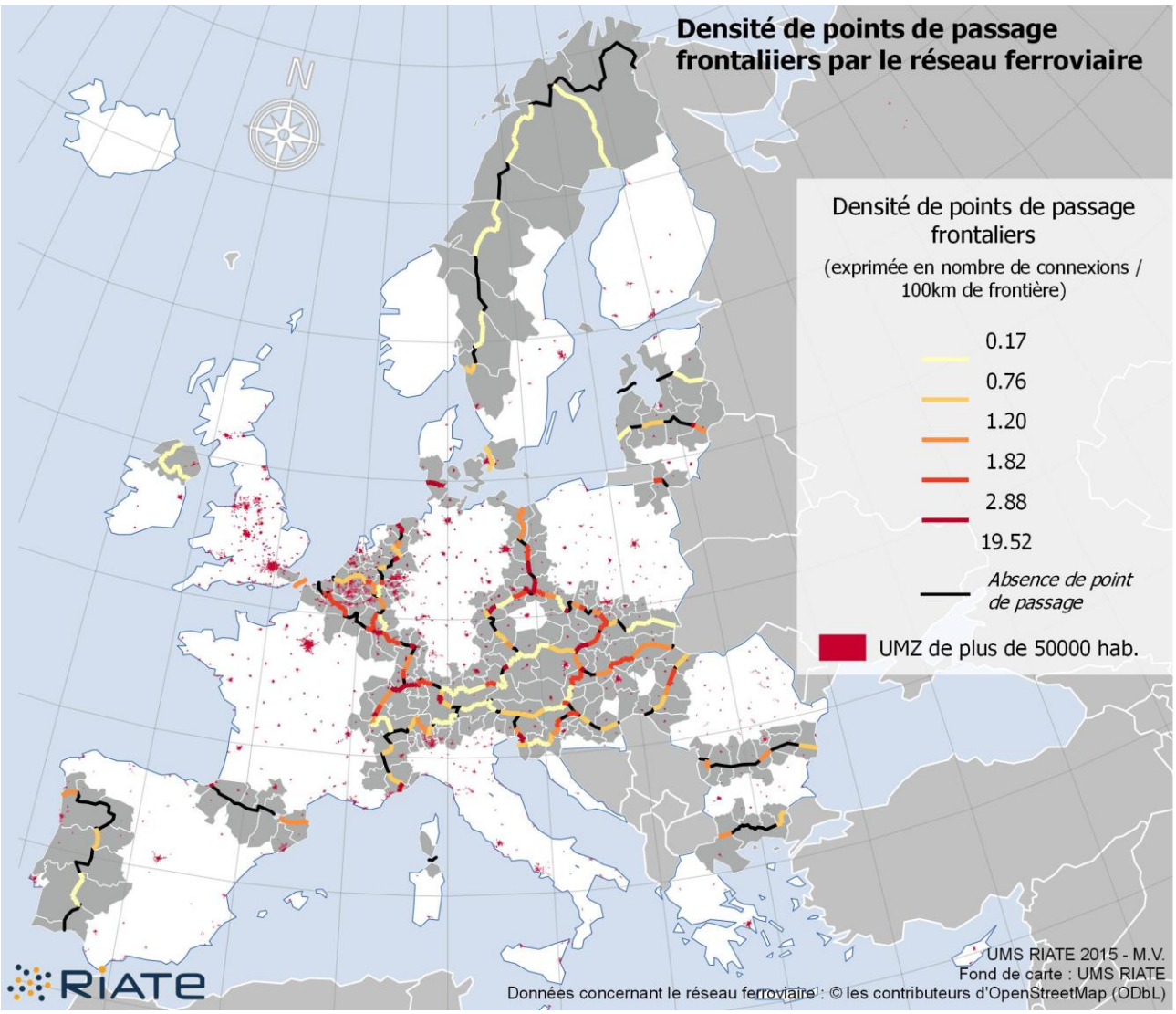


Figure 18 - Représentation de la densité de points de passage transfrontaliers par voie ferroviaire à l'échelle de l'Europe

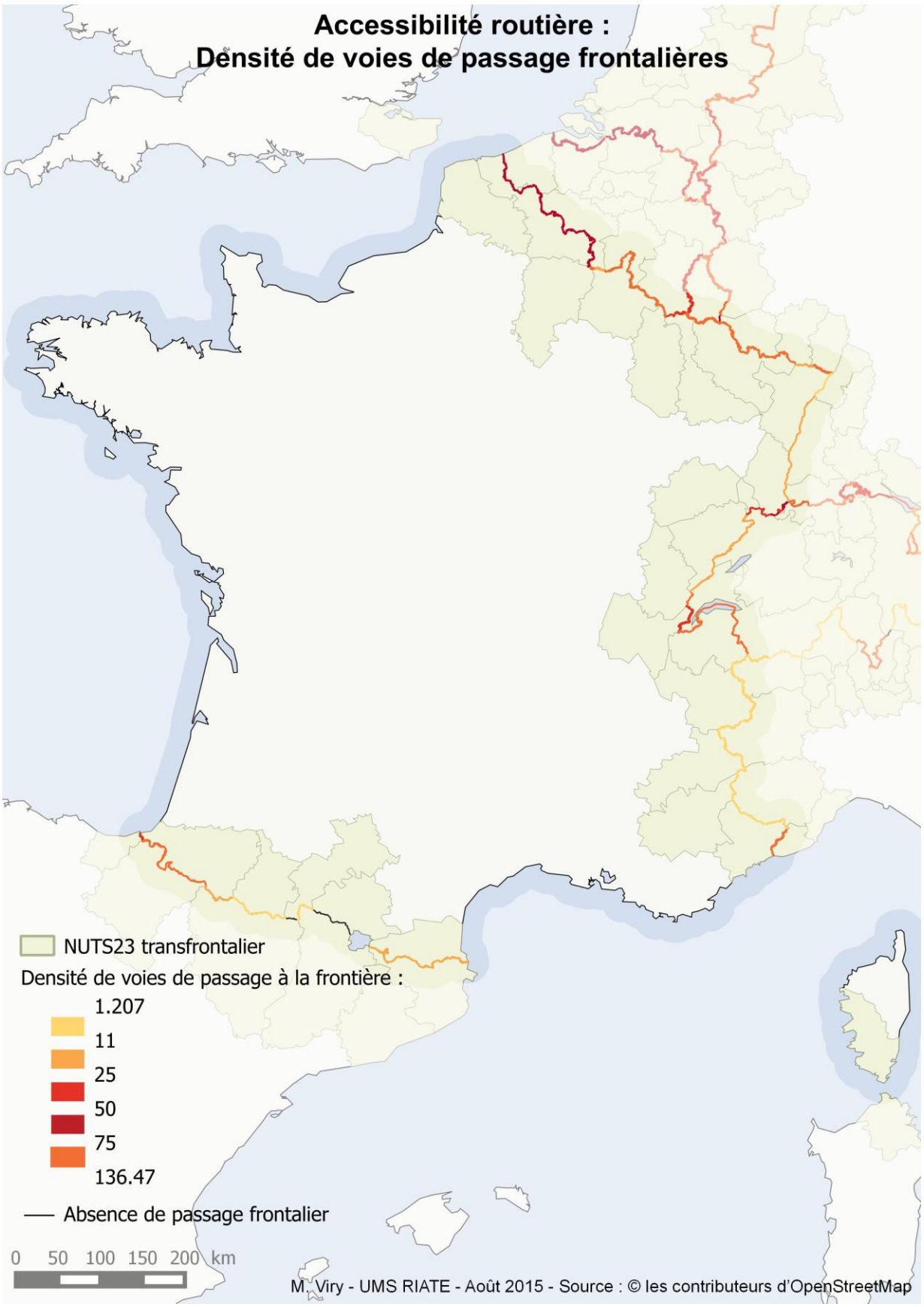


Figure 19 - Densité de voies routières transfrontalières à l'échelle nationale (France métropolitaine)

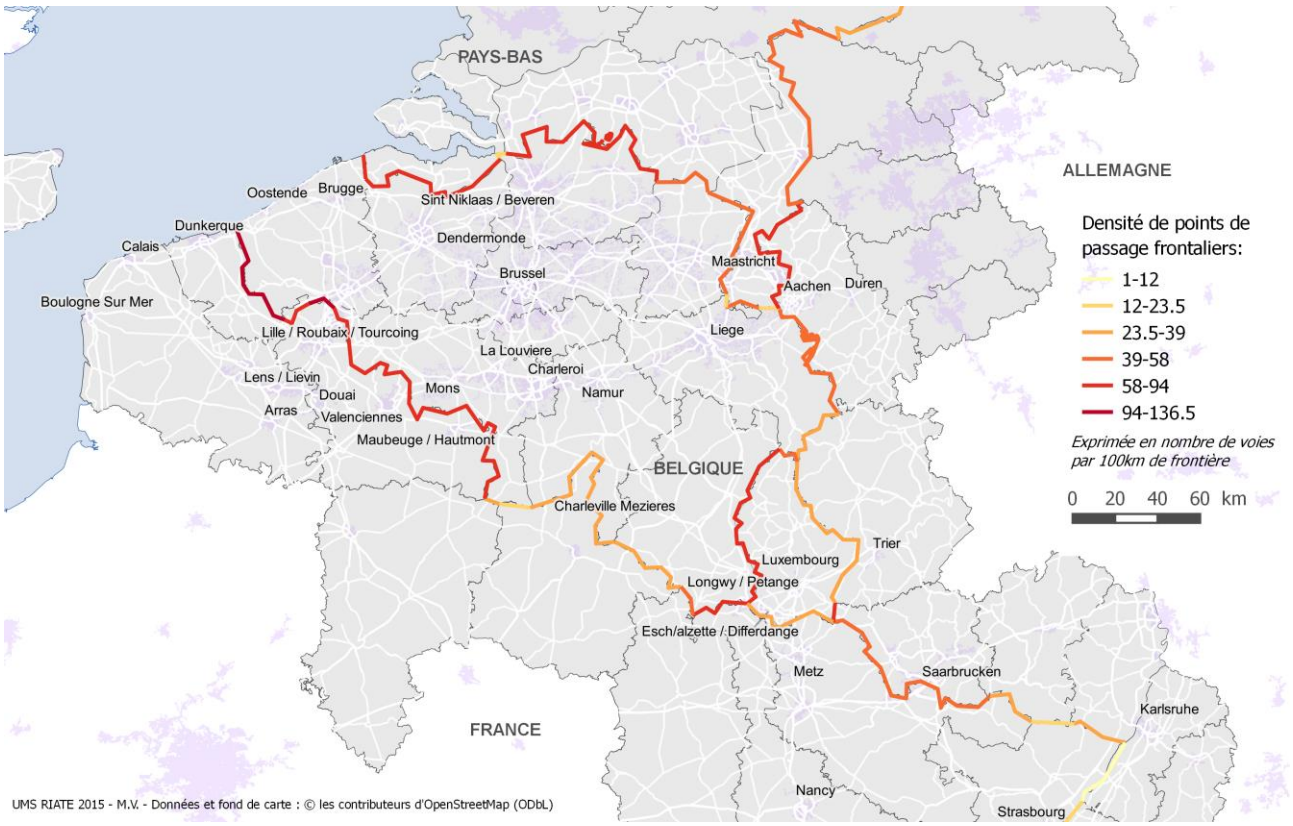


Figure 20 - Densité de voies routières transfrontalières à l'échelle régionale

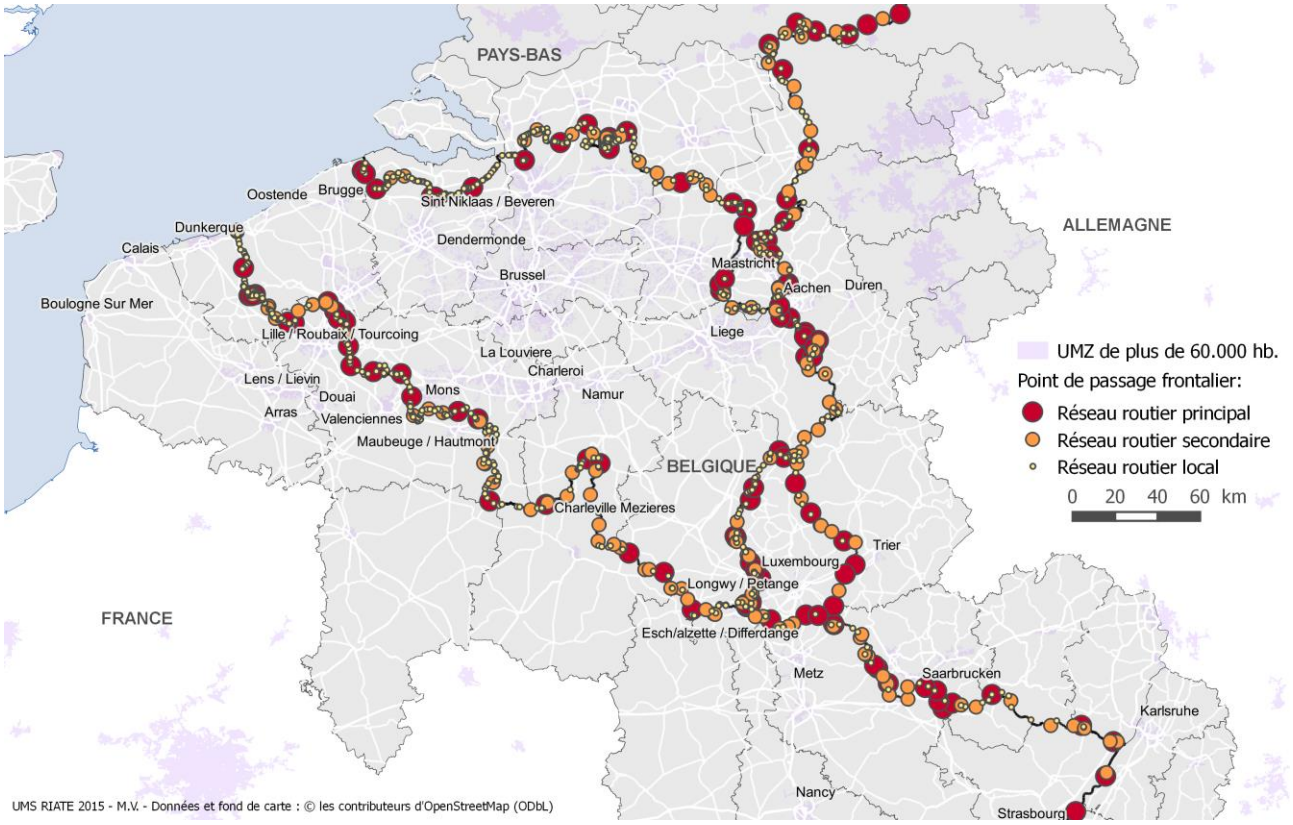


Figure 21 - Localisation des points de passage frontalier (automobile) à l'échelle régionale

3.5 Caractériser les frontières ne disposant d'aucun point de passage ?

Certains tronçons de la *polyligne* frontalière n'offrent pas de points de passage (Figure 23). Il s'agit d'une minorité (28 tronçons sur les 247 tronçons considérés, soit 3.48 % de sa longueur) qui sont répartis dans un nombre limité de pays (15) de l'espace d'étude. Ces frontières entre deux États ne sont pas affectées dans la même mesure par le phénomène (ce sont les frontières Bulgarie/Roumanie et Hongrie/Croatie qui sont les plus concernées, puisque respectivement 38,3 % et 30,2 % de leurs frontières étudiées sont constituées de tronçons sans point de passage). Il est possible de constater des similarités entre les différents tronçons, en effet il s'agit de tronçons en moyenne plus courts que les tronçons possédant des points de passage (27 km contre 99 km, voir Figure 22), le plus long mesurant tout de même près de 200km, et on constate que cette absence d'accès transfrontalier peut dans plusieurs cas être mise en relation avec les contraintes du milieu : montagne (par exemple à la frontière Italie/Suisse) ou cours d'eau (par exemple à la frontière Bulgarie/Roumanie).

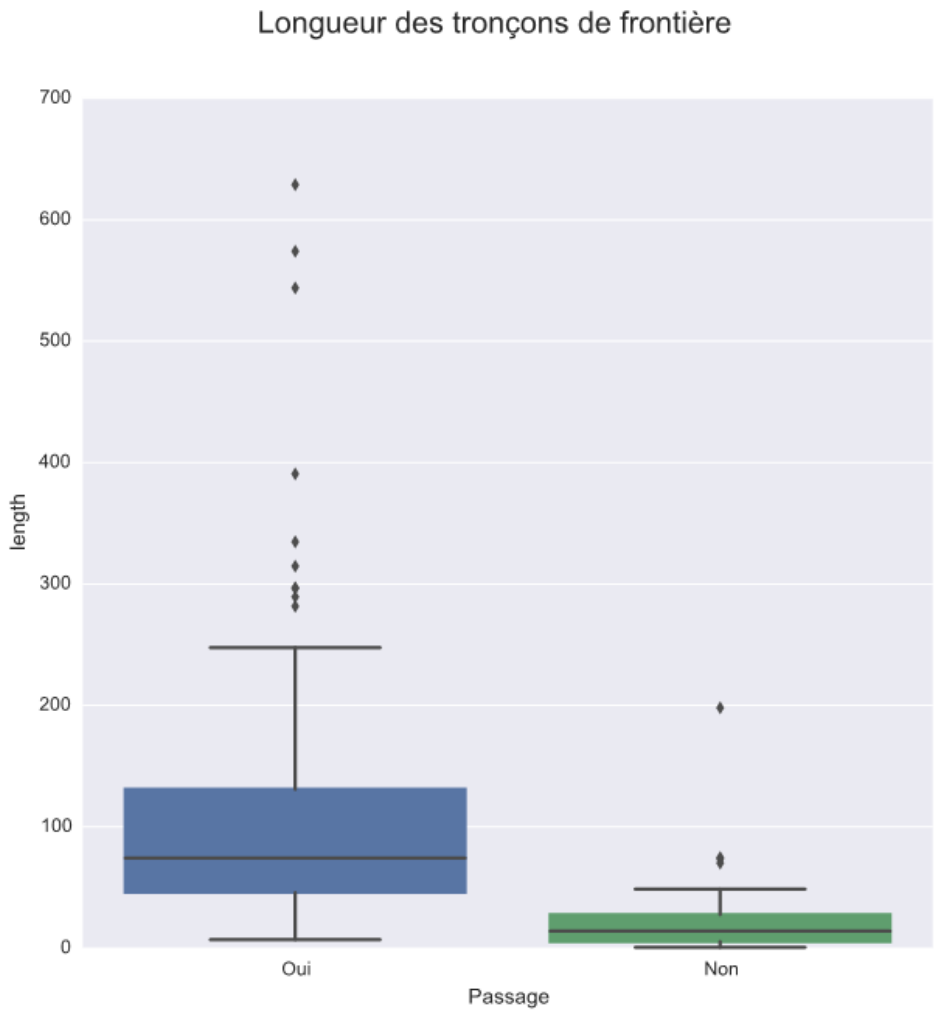
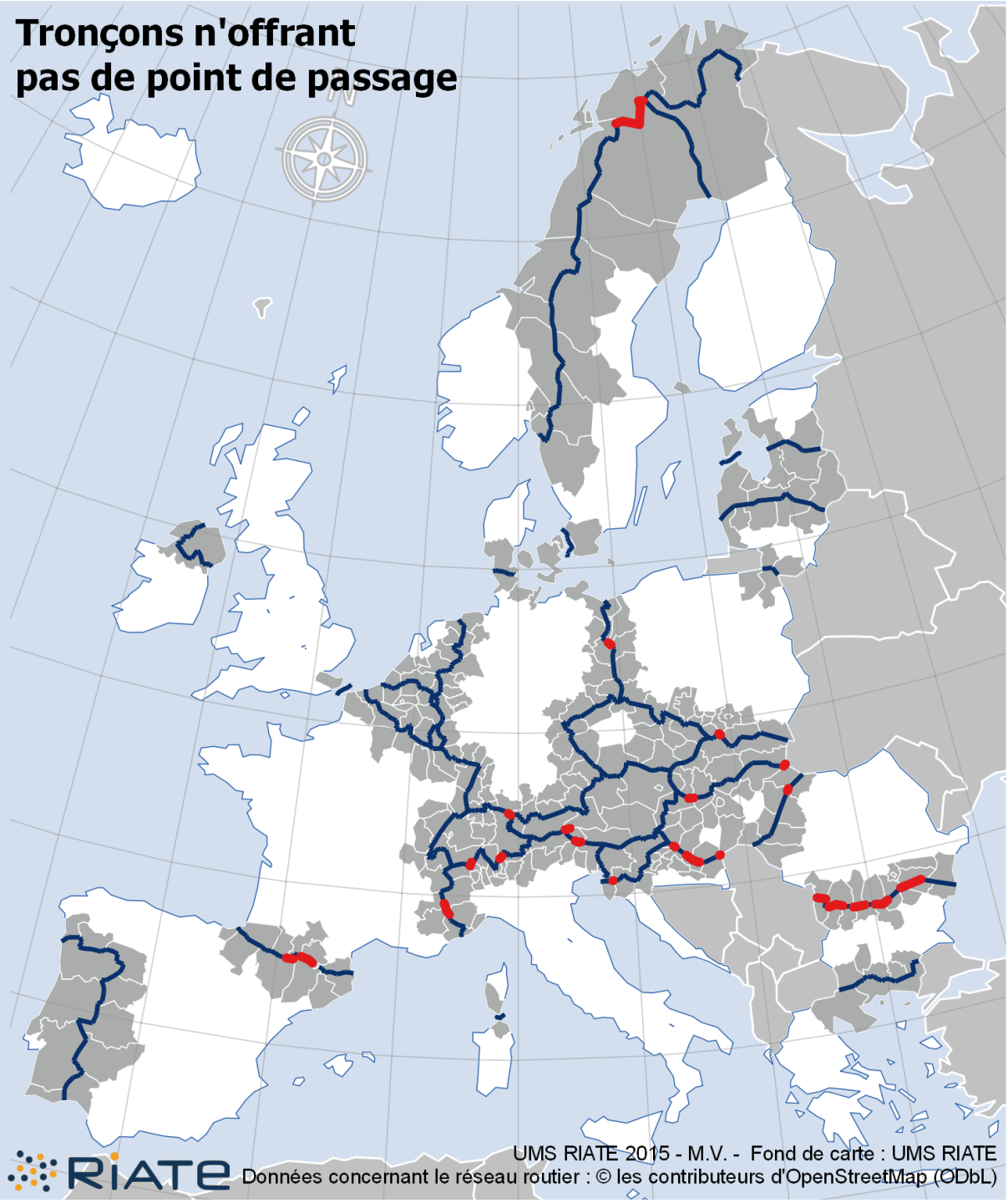


Figure 22 - Comparaison de la longueur de frontière
entre les tronçons offrant des points de passage et les autres



Les moyens de déplacement pris en compte sont:
- l'automobile
  (voies OSM *highway* = *motorway* | *trunk* | *primary* | *secondary* | *tertiary* | *unclassified* | *road*)
- le train et le tramway
  (voies OSM *railway* = *rail* | *narrow_gauge* | *tram* | *light_rail* hors voies de service)
- le ferry / bac de traversé
  (voies OSM *route* = *ferry*)

Figure 23 - Les tronçons de la polyligne frontalière
n'offrant pas de point de passage

4 Aller plus loin dans l'utilisation d'informations issues d'OpenStreetMap

4.1 Avec les données déjà évoquées : comparaison de différents cas d'accessibilité frontalière à d'autres échelles

D'autres données OSM répondant aux mêmes clés que celles précédemment utilisées (*highway* et *railway*) peuvent également être mise en place à une échelle plus fine, par exemple afin d'observer les potentialités d'échanges dans une zone particulière ou afin de caractériser des frontières particulièrement perméables (tourisme, ville installée à proximité immédiate du tracé de la frontière, etc.) ou à l'inverse particulièrement peu perméables (obstacles naturels, héritage historique, choix politiques, etc.) via l'extraction des voies empruntables à pied et/ou en vélo ainsi qu'en regardant les réseaux de transports en commun de dimension urbaine (notamment avec le tag *railway=subway* ou *railway=tram* ainsi que les différents itinéraires recensés avec des tags comme *route=bus*, *route=subway*, etc.)

En effet on peut faire l'hypothèse que des frontières très perméables le seraient aussi pour les piétons (chemins piétons/touristiques, villes à proximité de la frontière et présence de nombreuses rues situées à cheval sur les deux pays, etc.). D'autre part, des frontières définies comme peu perméables en raison d'obstacles naturels pourraient dans ce cas ne pas voir leur nombre de point de passage évolué, celui-ci étant limité par les infrastructures (pont, tunnel, etc.) en place.

4.2 Avec d'autres informations présentes dans OSM pouvant permettre d'appréhender les questions d'accessibilité des régions frontalières

Différents autres éléments présents dans la base de données OpenStreetMap permettent d'appréhender les questions d'accessibilité. C'est par exemple le cas de différentes informations supplémentaires apportées sur les réseaux via les différentes paires *key:value* utilisées (exemple vu en partie 3 avec le nombre de parcours empruntant un point de passage ferroviaire ou avec la prise en considération des attributs *'tunnel'*, *'bridge'* et *'toll'*). Plusieurs possibilités existent pour vérifier les combinaisons de *key:value* possible pour un type d'élément et visualiser leurs localisations avant de se lancer dans le processus d'extraction (notamment via l'API *TagInfo*, décrite dans les fiches jointes).
Au-delà des réseaux de transports routiers et ferroviaires en eux-mêmes, de nombreux autres éléments, ponctuels ou linéaires, et relatifs à l'accessibilité et aux déplacements sont présents dans la base de données. Ces différents éléments sont énumérés ici et ont également fait l'objet de la rédaction d'un script permettant de les extraire/préparer simplement.

Emplacement des aéroports :
La localisation des aéroports est contenue dans OpenStreetMap et peut donc en être extrait. Différentes informations sont présentes en attributs et permettent de hiérarchiser ces aéroports.
(→ Le script d'extraction sélectionne les aéroports ouverts au public et d'une importance au moins nationale et/ou au aéroport disposant d'un code IATA ; ces différentes entités (qu'elles soient à l'origine surfacique ou ponctuelles) sont enregistrées sous forme de points)

Emplacement des gares et des arrêts de transport en communs :
Il va être possible d'extraire les gares (pour l'ensemble des voies ferrées) mais également l'ensemble des arrêts de transport en commun, par exemple pour une approche plus fine sur l'étude des potentialités d'accès aux infrastructures de transport en communs.
*(→ L'extraction proposée correspond aux gares de train (*railway=station *et* railway=halt*) ainsi qu'aux arrêts de tramway ; comme pour les aéroports, ces entités sont sauvegardées sous forme de points)*

Extraction des itinéraires empruntés par les bus/car (et notamment ceux parcourant de longues distances) :
D'autres informations de parcours sont disponibles dans OpenStreetMap et c'est exemple le cas des parcours empruntés par les lignes régulières de bus/car. Comme pour les informations précédemment énoncées, il est également possible de douter la validité totale de ces informations. En effet cette validité dépendra d'une part du « taux de contribution » de la zone où elle se situe (de nombreuses omissions sont donc à craindre sur ce type d'élément) mais également des choix effectués par l'opérateur (ajout/suppression de lignes, changement d'itinéraire, etc.) qui peuvent varier dans le temps. Comme d'autres informations d'OpenStreetMap, ces itinéraires tendent à s'enrichir, notamment grâce à la mise à disposition de tables GTFS[12] par certains opérateurs, complétée par le développement d'engins de routage basés sur OSM et intégrant les données de ces tables horaires[13].
(→ Les itinéraires correspondant à route=bus *ou* route=coach *sont ici candidats à l'extraction)*

Extraction de POI relatifs aux déplacements automobiles (stations essences ou de rechargement électrique, parkings dédiés au covoiturage, aires de repos, péages autoroutiers, radars de contrôle de vitesse) :
Différents points d'intérêts spécifiques au transport automobile sont présents dans OSM et peuvent être identifiés et extraits. Bien qu'ils ne renseignent pas directement sur les potentialités en termes d'échanges transfrontaliers, la densité de ces équipements pourra être utilisée pour décrire des itinéraires/frontières/lieux particulièrement propices aux échanges. La présence de ces différents points d'intérêt peut également être confrontée à leur répartition sur le reste du territoire (en regardant par exemple l'impact de la frontière sur leur répartition par rapport au reste du territoire). On note que selon les applications, une grande variété d'informations peut être extrait d'OpenStreetMap sur des thématiques spécifiques et/ou à fine échelle, notamment en vue de la préparation d'itinéraires spécifiques (localisation des éléments de signalétique tels que les feux de signalisation, les passages piétons ; localisation des trottoirs, des pistes cyclables ; localisation des bancs et tables de pique-nique ; localisation des aménagements/lieux ouverts au public ; description des lieux accessibles en fauteuil roulant ; etc.)
(→ Les points d'intérêts retenus ici sont les stations essences et de rechargement électrique ainsi que les aires de repos et les parkings proposant un espace covoiturage)

Extraction d'informations relatives à l'occupation et à l'usage des sols :
La différence, pourtant expliquée dans le wiki OSM, entre l'occupation des sols (au sens des éléments biophysiques qui recouvrent le globe, observés à petite échelle) et l'usage des sols (au sens de l'usage premier qui est fait du territoire, notamment par les humains) n'est pas complètement intégrée dans le schéma des tags utilisés. En effet des éléments, pour la plupart comparables à ceux référencés dans le schéma NLCD92[14], sont présents dans OSM (qui a notamment profité d'import de *Corine Land Cover*). Ils sont toutefois répartis entre quartes clés : *natural*, *landuse*, *surface* et *landcover*. Les 3 premières, selon les valeurs de la clé, décrivent aussi des éléments relevant plutôt de l'usage du sol que de sa couverture et sont les plus utilisées. Le dernier, *landcover,* tend à être utilisé afin de lever l'ambiguïté entre certains éléments présents dans *natural* ou *landuse* et relevant de l'occupation du sol (selon un schéma clairement défini). Ces informations concernant l'utilisation du sol, et notamment à fine échelle peuvent être très détaillées, en effet la qualité du rendu des tuiles générées pour la visualisation en dépend, et il n'est pas rare que des arbres soit localisés à l'unité par un point pour une visualisation à fine échelle. Ainsi la présence des cours d'eau y est renseignée (sous forme de polygone pour ceux dont l'emprise au sol le nécessite) et il est possible de l'utiliser pour vérifier par exemple une hypothèse selon laquelle les frontières administratives correspondraient souvent à des éléments physiques du territoire.
(→ En raison du poids et du niveau de détails de ces informations, seules celles relatives aux cours d'eau / plan d'eau sont proposées à l'extraction, permettant ainsi de calculer la part de la frontière située sur un cours d'eau)

---

[12] GTFS : *General Transit Feed Specification.*
[13] Par exemple les projets OpenTripPlanner et TripGo.
[14] http://landcover.usgs.gov/classes.php

→ Script **extractions_supplementaires.py**

4.3 L'exploitation des métadonnées pour en savoir plus (sur l'historique et, potentiellement, sur la complétude) des principales données utilisées pour ce rapport

Au-delà de leurs attributs, différentes données ont été récupérées pour chacune des voies qui intersecte la frontière internationale. Ces données n'ont pas vocation à être utilisées pour caractériser les potentialités d'accès mais peuvent être utile pour tester certaines hypothèses formulées par différents chercheurs sur la complétude des données et les schémas de modifications des objets OSM.

La récupération de l'historique complet de ces différents objets permet de recréer différentes variables pour chaque voie : différence entre le nombre d'attributs actuels et celui à sa création, différence entre le nombre de nœuds actuels et celui à sa création, nombre de contributions, nombre de contributeurs uniques, distance moyenne entre les nœuds qui composent la voie, distance de *Damerau-Levensthein* moyenne entre les contenus successifs des attributs.

Bibliographie indicative

- Documentation de l'outil OGR2OGR : http://www.gdal.org/ogr2ogr.html

- Guide de référence SpatiaLite : http://www.gaia-gis.it/gaia-sins/spatialite-sql-4.3.0.html

- Outil en ligne Taginfo permettant de connaitre la fréquence de combinaison des tags : https://taginfo.openstreetmap.org/

- Wiki OpenStreetMap, la base de référence décrivant les tags utilisés par le projet : http://wiki.openstreetmap.org/

Annexes

Annexe 1 : Arbre de décision correspondant à l'étape de pondération des voies routières

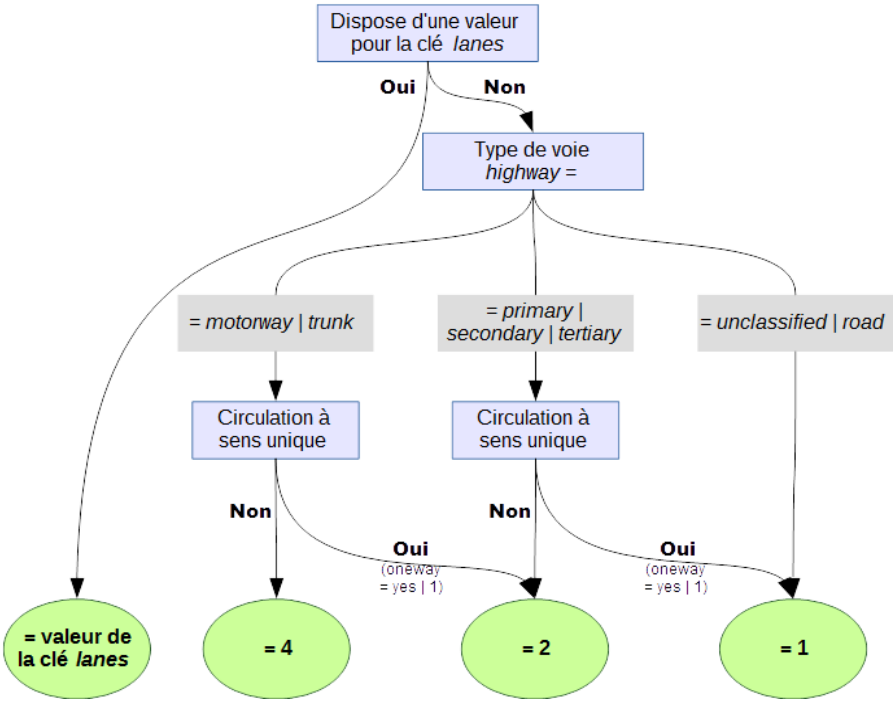


Figure 24 - Méthode suivit pour obtenir le nombre de voie correspondant à chaque point de passage

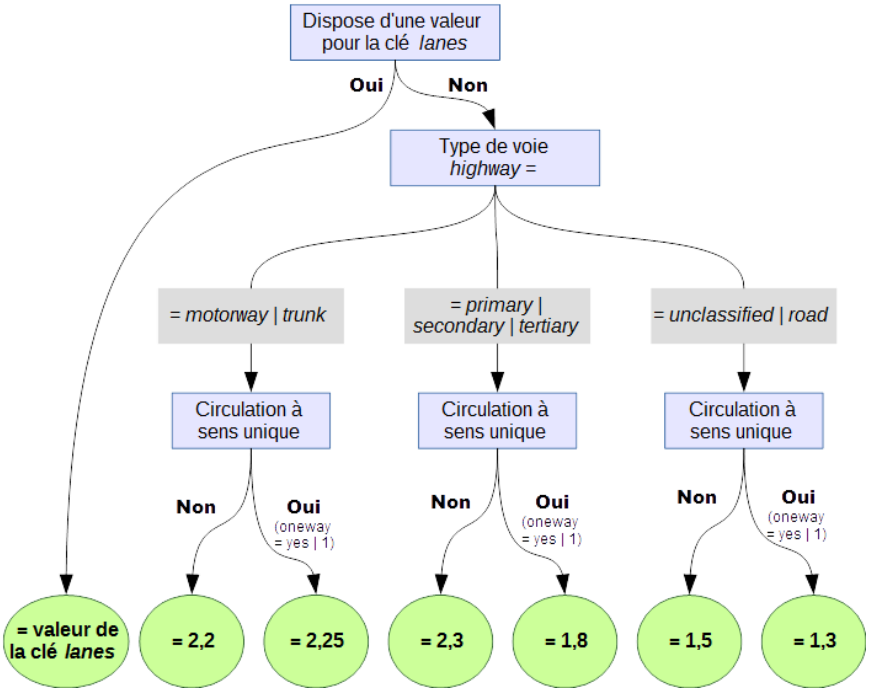


Figure 25 - Autre méthode proposée pour obtenir le nombre de voies correspondant à chaque point de passage

Annexe 2 : Les différents types d'actions effectués lors du nettoyage de la table des points de passage routier

Exemple 1

Deux points sont créés lorsque que la frontière intersecte deux voies *highway=primary* correspondant au sens aller et au sens retour (chacun tagué *oneway=yes*) d'une même route. Ces deux points vont être regroupé en un seul qui additionne (dans le champ '*poids*') le nombre de voies lu dans les attributs (ou extrapolé depuis sa catégorie d'appartenance le cas échéant) et qui conserve (dans le champ *'join_id'*) l'identifiant du point fusionné (Figure 26).

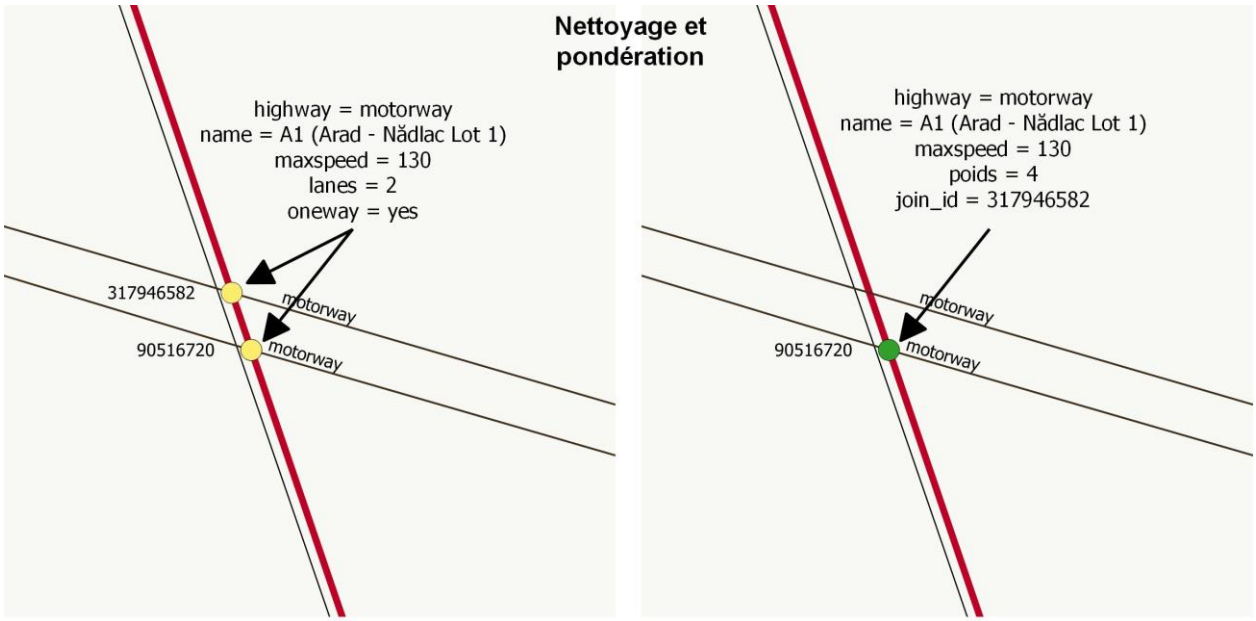


Figure 26 - Cas le plus courant : lecture du nombre de voies et regroupement de points proches possédant les mêmes attributs

Exemple 2

Une ou plusieurs entités multipoints ont été créées dans un secteur où la route longe la frontière (Figure 27). Seul le premier point de l'entité multipoint est conservé et une recherche est effectuée sur d'éventuels autres points de passage à proximité qui disposeraient strictement des mêmes caractéristiques (en particulier le nom). L'entité éventuellement trouvée peut alors être considérée comme un doublon et ne sera ni fusionnée ni conservée, comme c'est le cas (son identifiant apparaît toutefois, précédé du caractère '!' dans le champ *'join_id'* de l'entité qui a permis de le considérer comme un doublon).

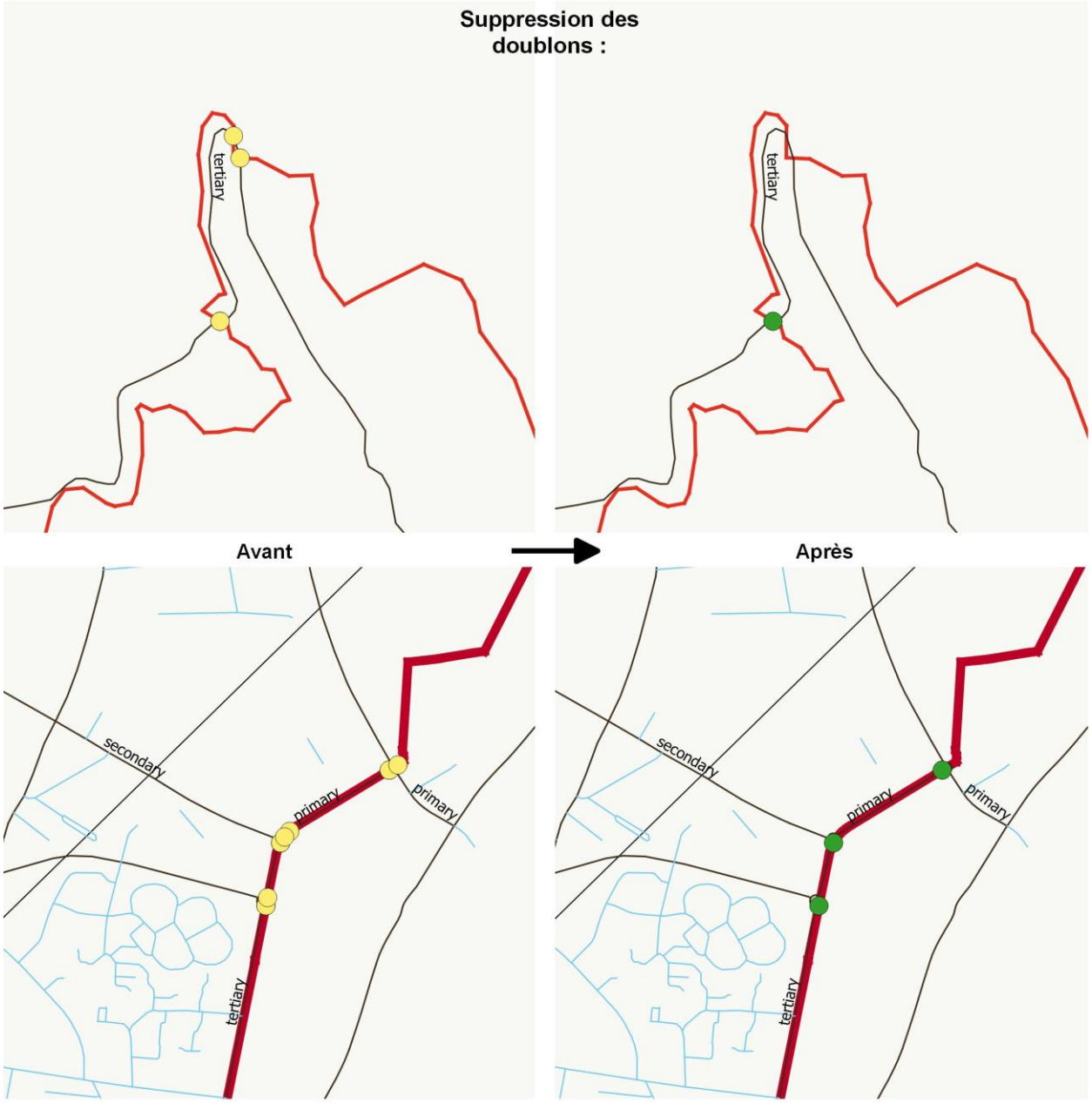


Figure 27 - Nettoyage nécessaire lorsqu'une route coupe plusieurs fois la frontière ou la longe

Exemple 3

Un point a été créé lors de l'intersection du réseau routier et de la frontière. Ce point ne dispose pas d'attributs précisant son nombre de voie. La valeur du champ *'poids'* va alors être extrapolé à partir de la valeur des clés *'highway'* et *'oneway'*.

CE QU'IL FAUT RETENIR...

… **des données OpenStreetMap :**

- Elles couvrent des thématiques très larges avec un niveau de détails attributaires pouvant être élevés en raison de la possibilité d'ajouter le nombre d'attributs voulu, de préciser la source de certaines des informations y figurant et d'y faire figurer des liens *html*.
- Ces **données sont très hétérogènes**, aussi bien dans les tags utilisés (en raison de l'absence de dictionnaire restrictif) que dans leur complétude/justesse (qui varie fortement entre certains pays et entre certaines zones urbaines et rurales, cf. document « *TR_complétude.doc* »)
- Les **données sont en évolution constante** depuis la création du projet en 2004, d'où un volume de données grandissant. Cette évolution concerne le contenu de la base, mais aussi, dans une moindre mesure, les habitudes/règles d'identification (*tag*) des objets.
- Les données sont réparties en trois types : *node*, *way* et *relation*. La structure des données qui sont issues de la base nécessite une préparation différente selon les utilisations (*routing*, géocodage, mise à disposition de tuiles, traitement SIG, etc.).

… **de l'extraction des réseaux de transports et de la création des points de passage :**

- D'une part **il n'existe pas une unique méthode** pour effectuer les opérations d'extraction / conversion / stockage des données. Les résultats attendus sont théoriquement identiques mais peuvent différer selon les outils utilisés et/ou la version des données OSM utilisées (notamment en raison de la gestion différente, selon les outils, des relations incorrectes dans OSM ainsi que des géométries incorrectes pouvant avoir été extraites), de plus la majorité des outils de filtrage/extraction des données OSM mettent en garde sur le caractère expérimental des opérations effectuées.
- D'autre part des choix ont dû être effectués lors de l'extraction des réseaux (hiérarchie des réseaux à prendre en compte), de l'extraction de la ligne transfrontalière (issue d'entités existantes ou d'OSM) ainsi que lors de la préparation des résultats (tentative d'élimination de « faux-positifs », basée sur des choix en terme de rayon de recherche et d'attributs à faire correspondre ; transformation des points créés en nombre de voies pour le réseau routier ; non-pondération des points de passage ferroviaire dans un premier temps pour faciliter la lecture des résultats).

… **des résultats obtenus :**

- Les résultats obtenus sont la conséquence des choix précédemment évoqués. Plusieurs de ces choix peuvent être modifiés facilement, selon leur nature soit en relançant la chaîne de calcul, soit en recalculant la données finale à partir d'une des tables intermédiaires conservées.
- Les 2 principaux indicateurs obtenus sont une **densité de voies routières par 100km de frontière** et **une densité de point de passage ferroviaire par 100km**. Ces informations sont données à l'échelle de chaque frontière séparant deux entités d'un État différent, permettant d'obtenir des résultats à l'échelle de chaque entité NUTS2/3 concernées puis à l'échelle de chaque pays concerné.
- Les points de passages comportent d'autres indications permettant d'aider à caractériser les frontières (regroupement ou dispersion des points le long de la frontière, passage par un pont ou un tunnel, type de voie prédominant à traverser la frontière, notamment pour le train avec la différenciation de voies à usages locaux/régionaux et des voies à grande vitesse).
- L'intégration de ces différents élément peut faire émerger une typologie des frontières permettant de qualifier de manière globale sa perméabilité / ses potentiels d'accès.
- Enfin, ces résultats étant basés sur les données OpenStreetMap il conviendra d'y faire figurer la mention adéquate lors de leur utilisation.

© UMS RIATE, Février 2016