

# UNE EXPÉRIENCE DE GÉNÉRALISATION DE RÉSEAU ROUTIER DANS UN DISPOSITIF DE PRODUCTION

par *Jérémy Renard*,

Laboratoire COGIT - Institut national de l'information géographique et forestière  
73 avenue de Paris 94165 Saint Mandé cedex  
jeremy.renard@ign.fr

et *Silvio Rousic*

Centre d'études techniques de l'équipement Méditerranée  
Pôle d'activités - 30, avenue Albert Einstein  
CS 70499 - 13593 Aix-en-Provence cedex 3  
silvio.rousic@developpement-durable.gouv.fr

---

*Cet article repose sur un travail collaboratif entre deux instituts publics (l'IGN et le CETE Méditerranée) dans le but de tester des processus de recherche sur la généralisation de réseau routier dans un contexte réel de production de données. Les premiers algorithmes testés sont exclusivement basés sur la géométrie et la topologie des données, comme généralement décrits dans la littérature, et ils donnent des résultats jugés insuffisants. Des améliorations majeures ont été, par la suite, apportées à ces algorithmes en prenant en compte l'information sémantique contenue dans les données, avec des résultats finaux bien meilleurs. Cette conclusion pose la question de la dualité entre généralité et dépendance au schéma dans l'approche de l'automatisation de la généralisation.*

## 1 Introduction

La généralisation de données géographiques est un processus complexe qui a pour but d'adapter le contenu d'une base de données géographiques en fonction de spécifications prédéfinies ou de règles cartographiques, pour s'assurer que la base de données généralisée réponde au besoin formulé (et que la carte produite en bout de chaîne soit lisible) tout en préservant ses caractéristiques principales. Le but principal de ce processus cartographique est de réduire le nombre d'objets géographiques dans la base pour éviter les surdensités et de schématiser ces objets pour simplifier leur structure et améliorer leur lisibilité. L'automatisation de la généralisation est un large sujet de recherche, étudié depuis une vingtaine d'années, dont les travaux passés ont permis la mise en place d'un grand nombre de méthodes implémentées de généralisation, allant des algorithmes simples aux processus les plus complexes.

Dans le cas d'un réseau routier complet, la réduction de l'échelle de sortie d'une carte implique de mettre en œuvre plusieurs opérations de généralisation pour rendre le réseau cohérent et lisible une fois mis à l'échelle finale. Parmi ces opérations, on peut évoquer la détection de structures routières complexes comme les ronds-points ou les échangeurs, la

simplification de ces structures, la sélection de routes pour ne conserver que les axes routiers les plus importants tout en supprimant ceux qui ne sont pas significatifs, la schématisation ou le déplacement de routes pour éviter les superpositions de symboles particulièrement dans le cas de routes proches ou de virages de montagnes, etc... La plupart de ces opérations sur le réseau routier ont déjà été étudiées et décrites par la littérature puis implémentées dans des plates-formes et des logiciels dédiés à la généralisation, et ils semblent donner de bons résultats dans un contexte de recherche. Néanmoins, il pourrait être intéressant de tester ces opérations dans le contexte plus strict d'un dispositif de production, avec des contraintes différentes et souvent plus fortes, comme l'adaptation à des schémas de données particuliers ou le traitement intensif sur de gros volumes de données, des problématiques qui sont plutôt marginales en recherche.

Le but de cet article est d'illustrer le travail collaboratif entre deux instituts français, le Centre d'études techniques de l'équipement [CETE] Méditerranée d'une part et l'Institut national pour l'information géographique et forestière [IGN] par l'intermédiaire de son laboratoire de recherche en cartographie et géomatique [COGIT] d'autre part. Cette collaboration vise à adapter les travaux de recherche en générali-

sation automatique de réseau routier du COGIT à un contexte de production réelle de données au CETE. La partie 2 détaille le contexte et les objectifs de ce travail collaboratif. La partie 3 se concentre sur la méthodologie algorithmique utilisée pour généraliser le réseau routier, puis la partie 4 dresse un aperçu des premiers résultats obtenus et des premières conclusions qui en découlent. La partie 5 propose d'aller plus loin en démontrant comment les résultats de généralisation peuvent être améliorés en prenant en considération le schéma de données. Enfin, la partie 6 souligne les conclusions qui peuvent être tirées de l'ensemble de ce travail, en posant des questions plus générales sur la façon optimale de décrire et de présenter les algorithmes de généralisation pour parvenir à les optimiser sans perdre en généralité.

## 2 Contexte et objectifs

### 2.1 Un contexte de collaboration entre deux instituts

Le CETE Méditerranée est un des centres d'études techniques du Ministère pour l'écologie, le développement durable et l'énergie [MEDDE]. Il propose des services d'ingénierie sur différents sujets dans le champ d'action du ministère. Parmi ces sujets, le CETE Méditerranée est en charge de questions relatives à la planification et à la gestion des transports, et dans cette optique il maintient – entre autres – une base de données routières appelée Routes à grande circulation [RGC] qui répertorie toutes les routes les plus empruntées par les utilisateurs du réseau.

La base de données RGC a longtemps été maintenue à un faible niveau de détail, correspondant à peu près à des produits cartographiques à l'échelle nationale ou régionale (entre 1 : 250 000 et 1 : 500 000). En application des nouveaux standards français en termes d'information géographique, la base de données RGC doit désormais être maintenue à un niveau de détail bien plus élevé en utilisant le Référentiel à grande échelle [RGE] pour correspondre à des produits cartographiques à l'échelle locale (autour de 1 : 50 000). Cette mise à jour nécessaire de la base de données implique deux problématiques :

- Appariement de données pour rapatrier les attributs de la base RGC originale sur le RGE.
- Généralisation cartographique pour continuer à proposer des produits nationaux ou régionaux à petite échelle à partir du RGE enrichi par la base RGC.

Afin de résoudre ces deux problématiques, le CETE Méditerranée s'est adressé à l'IGN pour tirer profit de

son expertise dans le domaine. En effet, l'appariement de deux bases de données similaires a fait l'objet de plusieurs travaux de recherche (Mustière et Devogele, 2008 ; Olteanu-Raimond, 2008) qui ont été industrialisés pour donner des processus de production efficaces. Cette première problématique a donc fait l'objet d'une collaboration avec le Département vecteur du service des développements [SD] de l'IGN. Parallèlement, la question de la généralisation de données géographiques est un des axes de recherche historiques du laboratoire COGIT de l'IGN. Cette seconde problématique a donc été le centre d'un travail collaboratif entre le CETE Méditerranée et le COGIT, travail qui constitue le sujet de cet article.

### 2.2 Objectifs pour les deux parties

La généralisation d'un réseau routier complet implique différentes méthodes et algorithmes qui ont déjà fait l'objet de publications – la partie 3 dresse un panorama des opérations concernées. La plupart d'entre elles ont été implémentées dans CartAGen, la plate-forme de généralisation du COGIT (Renard et al., 2010 ; Renard et al., 2011), et testées avec succès sur des échantillons-tests de données. Le but de la collaboration entre le CETE Méditerranée et le COGIT est donc d'utiliser les algorithmes existants de la plate-forme du COGIT dans le but de généraliser les données RGC du CETE. À travers cette collaboration, des objectifs différents sont envisagés pour chaque intervenant.

Le but recherché par le CETE Méditerranée est de tester et d'acquérir la maîtrise des outils de généralisation de réseau routier pour être ensuite capable de lancer une généralisation complète et adaptée du réseau routier RGC. Pour atteindre ce but, la partie open-source de CartAGen est utilisée comme une librairie de traitements sur laquelle un utilisateur peut brancher ses propres données avant de lancer des processus de généralisation. L'objectif final est d'arriver à adapter les algorithmes existants pour proposer au CETE une solution clef en main de généralisation automatique de la base de données RGC.

Du côté du COGIT, la principale contribution de ce travail repose sur l'opportunité de tester les algorithmes implémentés dans un contexte de production réelle de données. Ainsi, les principes de base des algorithmes peuvent être discutés, adaptés en fonction des données, et améliorés pour s'intégrer dans un dispositif de production. Une certaine validation des travaux de recherche peut ainsi être offerte aux méthodes et aux algorithmes de généralisation.

### 3 Méthodologie pour la généralisation du réseau routier RGC

#### 3.1 Description des données

La base de données routières RGC est partitionnée sur l'ensemble du territoire français en se basant sur les contours des départements : une base par département. En considérant la base RGC sur un département comme celui des Bouches-du-Rhône (fig. 1), on remarque que le réseau n'est pas très dense, ce qui est logique puisqu'il ne contient que des routes à grande circulation. Certaines zones du territoire peuvent toutefois conserver une densité assez élevée de routes (les abords d'agglomérations notamment). Néanmoins, puisque la base ne concerne que des axes routiers massivement utilisés pour les transports, elle ne contient pas de petites routes sinueuses (typiquement les routes de montagne), ce qui rend inutile l'utilisation de méthodes de généralisation adaptées à ce type de routes pour éviter les empâtements de symbole comme GALBE ou AGENT (Mustière et Duchêne, 2001). En dehors de cette particularité, certaines caractéristiques des routes du réseau RGC sont très significatives : par exemple il n'y existe aucune petite impasse qui aurait besoin d'être éliminée, ou aucune différenciation contextuelle entre les zones urbaines et les zones rurales. Ces caractéristiques permettent d'office de mettre de côté des méthodes de généralisation habituellement utilisées sur les réseaux routiers mais qui deviennent caduques dans le cas de la base RGC.

Les routes du réseau RGC sont riches en informations sémantiques. Plus de quarante attributs sont attachés à ces routes : nature, numéro, classification, etc... Dans un premier temps, tous ces attributs ne seront pas utilisés par la généralisation qui ne tiendra exclusivement compte que des aspects géométriques et topologiques du réseau routier, comme c'est l'usage dans la littérature sur le sujet. En s'affranchissant ainsi de l'aspect sémantique des données, les algorithmes de généralisation peuvent être en théorie appliqués sur tout type de données quels que soient leurs schémas. Cela assure généricité et portabilité aux traitements concernés.

L'objectif avancé est de mettre en place un processus complet de généralisation de l'ensemble du réseau routier. Un tel processus a déjà été proposé au COGIT (Touya, 2010), mais dans le cas présent il a besoin d'être simplifié et adapté pour tirer profit des caractéristiques du réseau RGC. Le processus utilisé dans ce cas d'étude précis est basé sur plusieurs étapes : d'abord détection et simplification de structures complexes comme les ronds-points ou les voies à chaussées séparées, ensuite sélection de

tronçons routiers, enfin quelques questions plus complexes comme le déplacement de routes ou la simplification des échangeurs.

#### 3.2 Détection et simplification des ronds-points

Les ronds-points sont des structures particulières du réseau qui doivent être simplifiées en carrefour simple lors de la généralisation vers des échelles inférieures. Avant la simplification, une première étape nécessaire consiste à détecter ces ronds-points dans le jeu de données, en utilisant des outils de mesure et d'analyse spatiale. La méthode consiste à analyser les faces d'un graphe topologique construit sur la base des tronçons routiers du réseau (Sheeren et al., 2004). En effet, un rond-point peut être défini à partir d'une face topologique dont la forme est quasi-parfaitement circulaire, cette mesure de forme étant traduite par l'indice de Miller sur la compacité d'une surface ( $C = 4\pi \cdot \text{area} / \text{perimeter}^2$ ). Selon les expérimentations menées par Sheeren, un seuil minimal de 0.97 sur l'indice de Miller permet de discriminer si une face est constitutive d'un rond-point ou non. Une fois une face de rond-point détectée, les routes adjacentes ou connectées à cette face sont extraites : ce sont les routes internes et externes du rond-point. Une étape supplémentaire de détection est appliquée pour déterminer si les faces adjacentes au rond-point sont des pattes d'oie : une patte d'oie est une face triangulaire de petite surface adjacente à un rond-point.

En fin de processus de détection, un rond-point détecté est décrit par une face centrale (en rouge sur la fig. 2), des pattes d'oie éventuelles (en vert sur la fig. 2), des routes internes qui délimitent toutes ces faces constitutives du rond-point, et des routes externes connectées à ces faces.

La simplification d'un rond-point est ensuite assez simple (fig. 2) et peut-être décrite en trois étapes :

- Calcul du centre du rond-point.
- Suppression des routes internes (et par là même des faces du rond-point).
- Reconnexion des routes externes au centre du rond-point.

#### 3.3 Détection et simplification des voies à chaussées séparées

Les voies à chaussées séparées sont représentées dans les données avec chacune de leurs chaus-

sées de façon indépendante, ce qui donne en pratique deux voies parallèles. Cela pose problème car lorsque ces voies seront symbolisées, chacune aura son propre symbole et donc la route sera représentée par deux symboles au lieu d'un seul, symboles qui risquent de plus d'être en superposition l'un avec l'autre. Les voies à chaussées séparées doivent donc être simplifiées pour ne garder qu'une seule voie qui doit idéalement se situer juste entre les deux voies originales. Des propositions ont déjà été faites pour détecter et simplifier les voies à chaussées séparées (Thom, 2005), mais elles sont relativement dépendantes d'un certain type de données, c'est pourquoi nous utilisons ici une autre approche.

Comme pour les ronds-points, la détection des voies à chaussées séparées est basée sur l'analyse des propriétés géométriques des faces d'un graphe topologique construit sur la base des tronçons routiers du réseau. L'idée est de détecter les faces fines et allongées en utilisant des mesures d'élongation, de convexité et de compacité, car ces faces topologiques particulières représentent des séparateurs de voies à chaussées séparées (i.e. l'espace qui se trouve entre les deux voies). Cette méthode de détection est la même que celle utilisée pour détecter des haies dans une couche de végétation (Touya et al., 2010). Dans un premier temps, la convexité des faces est calculée en effectuant le ratio entre l'aire de la face et l'aire de son enveloppe convexe. L'expérience montre qu'un seuil de 0.8 permet de déterminer si une face réseau est plutôt convexe ou non. À partir de là, suivant la valeur de convexité, deux possibilités existent :

- Soit la face est considérée comme convexe (convexité > 0.8), alors une mesure d'élongation peut être considérée comme significative car la forme de la face est relativement régulière. L'élongation est calculée comme étant le rapport entre la longueur et la largeur du plus petit rectangle englobant [PPRE] de la face. Si l'élongation est supérieure à 5.0, la face est considérée comme un séparateur de voies à chaussées séparées.
- Soit la face est considérée comme concave (convexité < 0.8), alors la mesure d'élongation n'a plus de sens car la forme irrégulière de la face ne s'y prête pas. À la place, on utilise une mesure de compacité calculée par l'indice de Miller. Si la compacité est inférieure à 0.2, la face est considérée comme un séparateur de voies à chaussées séparées.

Les voies à chaussées séparées sont ensuite extraites comme étant les routes délimitant les sépa-

rateurs ainsi détectés. Pour la simplification, l'idée est de squelettiser les séparateurs de voies, de la même manière que l'on peut squelettiser en fleuve linéaire des fleuves surfaciques représentés avec leurs deux rives (Regnauld et Mackaness, 2006). Une triangulation de Delaunay est construite sur la face, et les triangles de cette triangulation sont utilisés pour reconstruire de proche en proche le squelette de la surface allongée. Ce squelette est par la suite filtré (suppression des petits décrochages) et lissé (amélioration de la forme) pour donner une ligne centrale matérialisant la nouvelle route qui doit prendre place entre les deux voies à chaussées séparées initiales. Pour conclure, les voies initiales sont supprimées de manière à ne garder que le tronçon central fraîchement construit, et la reconnexion du réseau à ce nouveau tronçon central est assurée, particulièrement pour le cas des bretelles dans les échangeurs (fig. 3).

### 3.4 Sélection de routes

Une fois la détection et la simplification des structures routières complexes effectuées, une nouvelle étape du processus de généralisation consiste à éliminer un certain nombre de routes du réseau de manière à éviter les surdensités et à ne garder que les routes les plus importantes. Deux algorithmes distincts sont utilisés pour effectuer la sélection des routes.

Le premier se base sur la construction de strokes routiers sur des principes de bonne continuité des axes majeurs (Thomson et Richardson, 1999). Les strokes les plus longs sont considérés comme les routes les plus importantes et sont toujours conservés sans modification lors de l'étape de sélection.

Le second algorithme utilisé pour la sélection de routes s'appuie sur des calculs de plus courts chemins (Dijkstra, 1959) entre une sélection de points d'attraction définis sur le réseau (Richardson et Thomson, 1996). Dans notre cas d'étude, étant entendu que nous ne disposons pas d'information contextuelle attachée aux données (exemple : où se trouvent les centres-villes, les aires d'activité, etc...), les points d'attraction sont définis aléatoirement de manière à créer une grille régulière maillant le réseau. L'idée de l'algorithme est ensuite de calculer tous les plus courts chemins entre tous les points d'attraction, d'attribuer un point à chaque fois qu'un tronçon routier est parcouru par un plus court chemin, d'en déduire ainsi une valeur sur chaque tronçon (aussi élevée que le nombre de fois où il est parcouru), et d'éliminer les routes les moins importantes en fonction des valeurs ainsi calculées.

Toutefois, il est essentiel de noter que le réseau RGC est déjà assez largement épuré puisqu'il ne contient uniquement que les routes à grande circulation. De fait, il n'est pas réellement nécessaire d'éliminer encore plus de routes, sauf éventuellement en zones de forte densité. L'étape de sélection des routes n'est donc pas primordiale dans l'application du processus complet sur la base de données RGC.

### 3.5 Aller plus loin : déplacements de routes, gestion des échangeurs...

En plus des traitements détaillés précédemment, quelques algorithmes additionnels pourraient être appliqués en fin de processus sur le réseau RGC pour le généraliser de façon complète. Par exemple, dans le cas de superpositions généralisées de symboles à cause d'une trop grande densité de tronçons routiers dans une zone réduite, il peut être nécessaire de pratiquer des déplacements sur les routes incriminées pour éviter les superpositions. La meilleure solution théorique pour orchestrer ces déplacements est d'utiliser le principe des beams élastiques (Bader, 2001), chaque route étant modélisée comme une "poutre" possédant un pouvoir de répulsion des objets environnants, l'ensemble de ces répulsions conduisant toutes les routes à s'éloigner progressivement les unes des autres. Malheureusement, cette méthode relativement complexe n'est pas implémentée dans la plate-forme CartAGen.

Pour aller encore plus loin, une des questions encore mal traitées dans la généralisation de réseau routier concerne la détection et la simplification de carrefours complexes comme les échangeurs. Quelques travaux intéressants existent sur le sujet (Mackaness et Mackechnie, 1999 ; Dogru et al., 2009), mais il ne reste que des travaux de recherche exploratoires qui n'ont jamais été réellement industrialisés et qui ne sont pas considérés à l'heure actuelle comme assez matures pour être intégrés dans CartAGen en tant que solution pleinement automatique.

## 4 Résultats et premières conclusions

Les premières expérimentations ont été menées en utilisant les algorithmes existants sans les modifier, tels que décrits dans la partie 3, en ne prenant en compte que les aspects géométriques et topologiques du réseau routier. Les résultats n'ont pas été considérés suffisamment bons pour être intégrés dans un dispositif de production. Le défaut de qualité dans les données généralisées provient principale-

ment de la détection des structures routières qui n'est pas exhaustive, que ce soit pour les ronds-points (fig. 4) ou pour les voies à chaussées séparées (fig. 5). Plusieurs raisons à cela. D'abord, en ce qui concerne les ronds-points, le critère de compacité exclut les ronds-points qui n'ont pas une forme régulière parfaite, comme les ronds-points ovales. Ensuite, en ce qui concerne les voies à chaussées séparées, la multiplicité des interconnexions au niveau des échangeurs casse la continuité des séparateurs de voies. Surtout, pour les ronds-points comme pour les voies à chaussées séparées, la détection est basée sur les faces d'un graphe topologique construit à partir de l'ensemble des routes du réseau, et toutes les routes sont traitées à niveau égal durant cette étape de détection. La conséquence immédiate tient en l'existence de routes "parasites" qui traversent une structure sans en faire réellement partie, et qui perturbent le processus de détection.

En termes d'évaluation des résultats, on observe la bonne détection d'environ 80% des ronds-points. Les voies à chaussées séparées ne sont elles approximativement que 60% à être correctement détectées, avec de nombreux problèmes de continuité et de connexions.

En dehors des problèmes rencontrés lors de la détection, la simplification des structures donne en revanche des résultats satisfaisants à partir du moment où ces structures ont été correctement détectées. Seulement quelques optimisations pourraient être apportées lors de cette étape, par exemple dans la bonne reconnexion des extrémités des routes généralisées au reste du réseau, et surtout dans la généralisation des échangeurs en tant que structures complexes (problème non traité dans le processus). Mais évidemment, si les structures sont mal détectées en prétraitement, elles ne peuvent pas être correctement généralisées par la suite. Concernant la sélection de routes par calcul des strokes et de plus courts chemins, elle semble donner de très bons résultats, mais les tests n'ont pas été poussés aussi loin que possible.

Ces résultats nous amènent à une première conclusion : les algorithmes de détection tels qu'ils sont décrits dans la littérature ne sont pas assez robustes pour supporter un contexte de production avec des données réelles et des contraintes fortes. Le problème vient de la philosophie des algorithmes qui consiste à ne tirer profit que des informations géométriques et topologiques des objets géographiques. Bien sûr, cette idée est extrêmement sensée pour assurer la généricité et la portabilité des algorithmes puisqu'ils peuvent ainsi être appliqués

sur tous types de données quel que soit leur schéma, mais dans un dispositif de production réelle, avec toutes les contraintes que cela implique par rapport à un contexte de pure recherche, les premiers résultats obtenus mettent en évidence la nécessité de tirer profit du schéma des données et des informations sémantiques comme leviers pour améliorer les algorithmes et les processus de généralisation.

## 5 Adaptation des traitements au schéma de données

### 5.1 Comment tirer profit du schéma de données ?

La plupart des problèmes dans la détection des structures routières vient de la création du graphe topologique à partir de l'ensemble des routes du réseau, avec pour conséquence l'apparition localement de routes "parasites" traversant les structures. Il existe une solution pour résoudre ces problèmes : construire à chaque étape de détection un graphe topologique basé uniquement sur les routes potentiellement concernées par la structure que l'on souhaite détecter, plutôt que sur l'ensemble des routes. Par exemple, si l'on regarde le cas 2 de rond-point non détecté de la fig. 4, la route parasite passant au-dessus du rond-point ne devrait pas être prise en compte dans la construction du graphe topologique (elle ne peut de toute façon pas faire partie du rond-point). Pour atteindre ce but, deux attributs des routes sont examinés : la valeur de l'attribut NATURE doit être différente des valeurs "Autoroute" ou "Route à 2 chaussées", et la valeur de l'attribut DIRECTION doit être différente de la valeur "Double". En effet, un rond-point est composé de routes secondaires simples (pas d'autoroutes ou de voies à chaussées séparées) que l'on ne peut emprunter que dans une direction et pas les deux. En fait les seules routes dont l'attribut DIRECTION a une valeur différente de "Double" (valeurs possibles "Direct" et "Inverse") sont de façon exclusive soit des autoroutes, soit des voies à chaussées séparées, soit des ronds-points (toutes les autres routes du réseau peuvent être empruntées dans les deux sens). Avec une telle requête attributive, toutes les routes potentiellement parasites ne sont pas prises en compte lors de la construction du graphe topologique, et la détection du cas 2 de la fig. 4 se passe bien (fig. 6). Pour le premier cas de figure, la solution consiste à abaisser le seuil de compacité de l'indice de Miller à 0.80 (au lieu de 0.97), de cette manière les ronds-points n'ayant pas une forme parfaitement circulaire sont tout de même détectés sans introduire de détection abusive (fig. 6).

Concernant la détection des voies à chaussées séparées, les problèmes principaux sont dus aux échangeurs et aux interconnexions entre plusieurs axes routiers majeurs. La solution idéale serait de traiter chaque axe principal un par un pour s'affranchir des connexions potentielles avec les autres routes. C'est rendu possible par l'utilisation de l'attribut NUMERO : tous les tronçons routiers possédant le même numéro font partie du même axe routier et sont pris en compte ensemble pour calculer un graphe topologique sans s'occuper des autres routes. Ensuite ce graphe topologique restreint est utilisé pour détecter les voies à chaussées séparées de l'axe en cours de traitement, et l'opération est répétée pour chaque numéro de route rencontré dans les données. Au final, les axes routiers sont traités un par un pour la détection puis la simplification. Cette méthode résout quasiment tous les problèmes que l'on avait initialement soulevés sur la détection et la simplification des voies à chaussées séparées (fig. 7).

La façon dont on a utilisé les informations sémantiques pour améliorer les algorithmes originaux n'est pas la plus importante en soi, ce n'est qu'une illustration sur un cas d'étude. En revanche, l'idée essentielle sous-jacente à cela, c'est d'admettre que considérer les attributs des données comme un moyen d'optimiser et d'adapter les processus de généralisation est une idée très significative pour améliorer les résultats finaux.

### 5.2 Résultats améliorés

Les améliorations obtenues par les algorithmes prenant en compte le schéma de données sont très significatives en termes de résultats. 95% des ronds-points sont désormais correctement détectés et simplifiés. Quasiment 100% des voies à chaussées séparées sont correctement détectées, et la plupart d'entre elles sont simplifiées de façon satisfaisante. Si nous avons perdu en généricité en introduisant une dépendance au schéma de données, nous avons gagné de bien meilleurs résultats dans le processus de généralisation. En fait, nous avons réussi à adapter un processus très générique et théorique (Touya, 2010) au cas particulier d'un dispositif réel de production sur des données possédant des caractéristiques très particulières, avec en sortie un processus global adapté et optimisé (fig. 8).

Les problèmes restants sont principalement dus à un certain manque de précision et de qualité dans les données : attributs incorrects (un tronçon autoroutier que l'attribut NATURE désigne comme une bretelle)

ou manquants (pas de numéro sur le tronçon routier), voire des tronçons absents qui créent des trous dans le réseau. Sans ces erreurs qui incombent au producteur de données en amont, les algorithmes améliorés de détection de structures donneraient quasiment 100% de bons résultats quels que soient les cas de figure rencontrés. Le principal effort à mener par la suite concerne la simplification des voies à chaussées séparées qui produit encore un certain nombre d'imperfections notoires dans les cas les plus complexes (typiquement au croisement d'un échangeur).

Parallèlement aux résultats bruts, on remarque également que les améliorations apportées par la prise en compte du schéma de données diminuent considérablement le temps de calcul du processus complet. Cette optimisation temporelle vient du fait que les graphes topologiques créés au fur et à mesure du processus – graphes dont le calcul est le facteur le plus important en termes de temps de traitement – sont construits en se basant à chaque fois sur un nombre restreint de tronçons routiers plutôt que sur l'ensemble du réseau.

## 6 Conclusions

Dans ce travail, nous avons essayé d'appliquer des algorithmes existant en recherche dans un dispositif réel de production par l'intermédiaire d'un processus complet de généralisation automatique de réseau routier. Les premiers résultats ont prouvé que les algorithmes tels qu'ils sont décrits dans la littérature ne sont pas assez robustes pour être considérés comme pleinement opérationnels en production. Des adaptations simples ont donc été apportées à ces algorithmes en prenant en compte la sémantique des données généralisées. Le processus ainsi optimisé a été testé et semble assez solide pour supporter un contexte de production de données, mais il mérite encore d'être amélioré pour pouvoir proposer une généralisation 100% automatique du réseau routier, particulièrement sur la simplification de voies à

chaussées séparées qui n'est pas totalement satisfaisante.

Au-delà des aspects purement pragmatiques qui découlent du cas d'étude, une conclusion essentielle plus générale peut être tirée de ce travail : les algorithmes de généralisation automatique de réseau routier basés exclusivement sur des considérations géométriques et topologiques ont besoin d'être adaptés à la structure sémantique des données pour donner toute la mesure de leurs possibilités. Des conclusions similaires pourraient probablement être tirées en testant d'autres algorithmes sur d'autres thèmes cartographiques, certains travaux de recherche ayant déjà proposé quelques pistes de réflexion assez concordants (Revell et al., 2006 ; West-Nielsen et Meyer, 2007). Cette observation générale doit logiquement nous amener à nous poser quelques questions sur la façon dont doivent être menées les recherches en automatisation de la généralisation.

Les algorithmes et processus de généralisation sont la plupart du temps mis au point pour être le plus génériques possible en ne prenant en compte que la géométrie des objets généralisés, ce qui autorise la portabilité des traitements sur tous types de données. Pourtant, l'expérience pratique décrite dans cet article souligne la nécessaire dépendance au schéma de données pour proposer des solutions de généralisation adaptées et optimisées dans le contexte d'une chaîne de production. Cette conclusion est finalement relativement paradoxale, et elle pose un problème sous-jacent : quel est l'objectif final recherché dans la création de processus de généralisation automatique ? Servir de base aux unités de recherche pour tenter de développer de nouvelles solutions génériques de plus en plus novatrices et efficaces, ou bien proposer des solutions 100% automatisées sur des cas particuliers dans des contextes réels de production ? De la réponse à cette question dépendra notre façon de conduire les orientations de recherche dans le domaine de l'automatisation de la généralisation.

## Bibliographie

- Bader M.**, 2001, « Cartographic Displacement in Generalization : Introducing Elastic Beams », *4<sup>th</sup> ICA Workshop on Generalisation and Multiple Representation*, Beijing (Chine), 2-4 août 2001.
- Dogru O., Duchêne C., Van de Weghe N., Mustière S., Ulugtekin N.**, 2009, « Generalization Approaches for Car Navigation Systems », *24<sup>th</sup> International Cartographic Conference (ICC'09)*, Santiago (Chili), 15-21 novembre 2009.
- Mackness W. A., Mackechnie G.**, 1999, « Detection and Simplification of Road Junctions in Automated Map Generalization », *GeoInformatica*, vol. 3, p. 185-200.
- Mustière S., Duchêne C.**, 2001, « Comparison of Different Approaches to Combine Road Generalization Algorithms: GALBE, AGENT and CartoLearn », *4<sup>th</sup> ICA Workshop on Generalisation and Multiple Representation*, Beijing (Chine), 2-4 août 2001).
- Mustière S., Devogele T.**, 2008, « Matching Networks with different Levels of Detail », *GeoInformatica*, vol. 12, n°4, p. 435-453.
- Olteanu-Raimond A.-M.**, 2008, « A Multi Criteria Fusion Approach for Geographical Data », dans *Quality Aspects in Spatial Data Mining*, Taylor and Francis, p. 45-56.
- Regnaud N., Mackness W.A.**, 2006, « Creating a Hydrographic Network from its Cartographic Representation: A Case Study Using Ordnance Survey MasterMap Data », *International Journal of Geographical Information Science*, vol. 20, n°6, p. 611-631.
- Renard J., Gaffuri J., Duchêne C.**, 2010, « Capitalisation Problem in Research - Example of a New Platform for Generalisation: CartAGen », *12<sup>th</sup> ICA Workshop on Generalisation and Multiple Representation*, Zürich (Suisse), 12-13 septembre 2010.
- Renard J., Gaffuri J., Duchêne C., Touya G.**, 2011, « Automated Generalisation Results using the Agent-Based Platform CartAGen », *25<sup>th</sup> International Cartographic Conference (ICC'11)*, Paris (France), 3-8 juillet 2011.
- Revell P., Regnaud N., Thom S.**, 2007, « Generalising and Symbolising Ordnance Survey Base Scale Data to Create a Prototype 1:50 000 Scale Vector Map », *The Cartographic Journal*, vol. 44, n°3, p. 251-257.
- Richardson D.E., Thomson R.C.**, 1996, « Integrating Thematic, Geometric and Topologic Information in the Generalization of Road Networks », *Cartographica*, vol. 33, p. 75-83.
- Sheeren D., Mustière S., Zucker J.-D.**, 2004, « How to Integrate Heterogeneous Spatial Databases in a Consistent Way ? », dans *Proceedings of Conference on Advanced Databases and Information Systems (ADBIS'04)*, Budapest (Hongrie), Springer, p. 364-78.
- Thom S.**, 2005, « A Strategy for Collapsing OS Integrated Transport Network Dual Carriageways », *8<sup>th</sup> ICA Workshop on Generalisation and Multiple Representation*, La Corogne (Espagne), 7-8 juillet 2005.
- Thomson R.C., Richardson D.E.**, 1999, « The "Good Continuation" Principle of Perceptual Organization applied to the Generalization of Road Networks », dans *Proceedings of 19<sup>th</sup> International Cartographic Conference (ICC'99)*, Ottawa (Canada), p. 1215-1223.
- Touya G.**, 2010, « A Road Network Selection Process Based on Data Enrichment and Structure Detection », *Transactions in GIS*, vol. 14, n°5, p. 595-614.
- Touya G., Duchêne C., Mustière S.**, 2010, « Généralisation et intégration pour un fond vert commun entre l'IGN et l'IFN », *Revue internationale de géomatique*, vol. 20, n°1, p. 65-86.
- West-Nielsen P., Meyer M.**, 2007, « Automated Generalisation in a Map Production Environment – the KMS Experience », dans *Generalisation of Geographic Information: Cartographic Modelling and Applications*, Elsevier, p. 301-313.



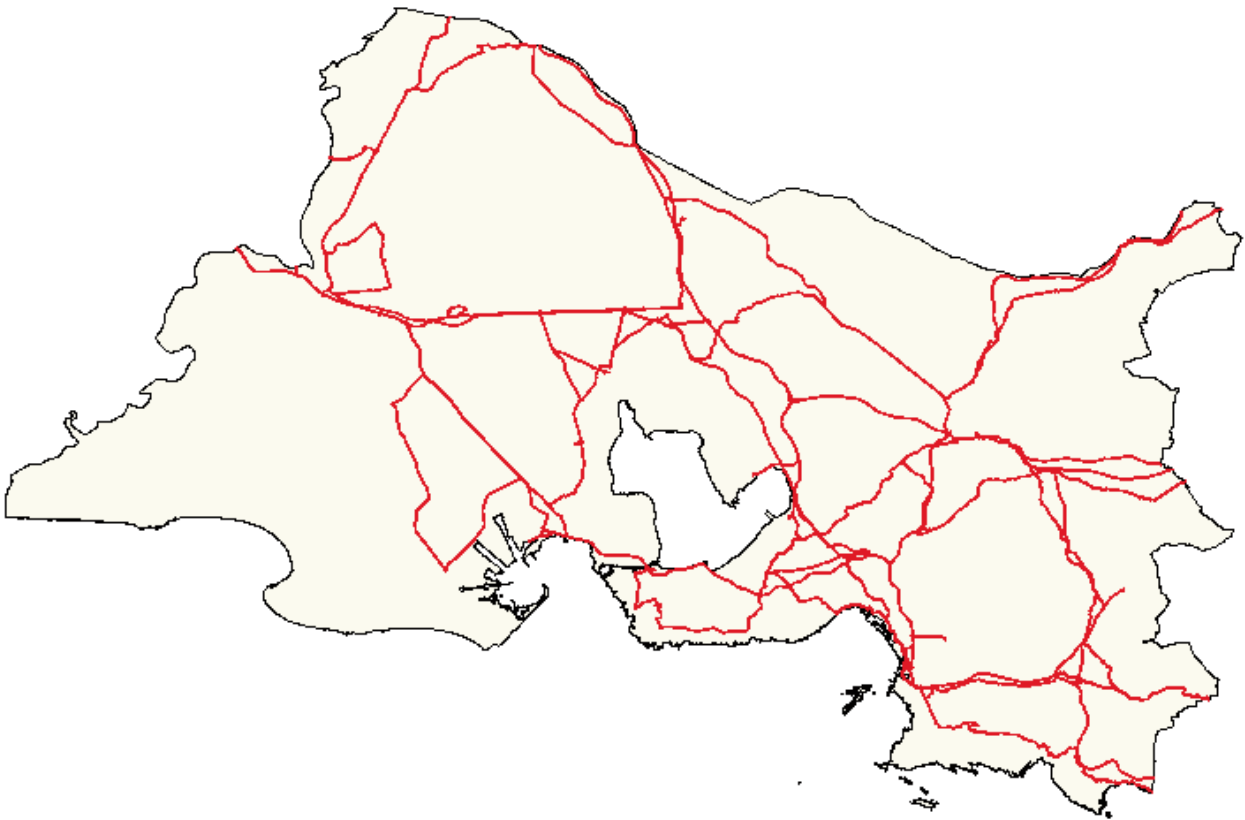


Figure 1 : Aperçu dans QGIS de l'ensemble des routes du réseau RGC sur un département entier  
 (Source : CETE Méditerranée, ©RGC)

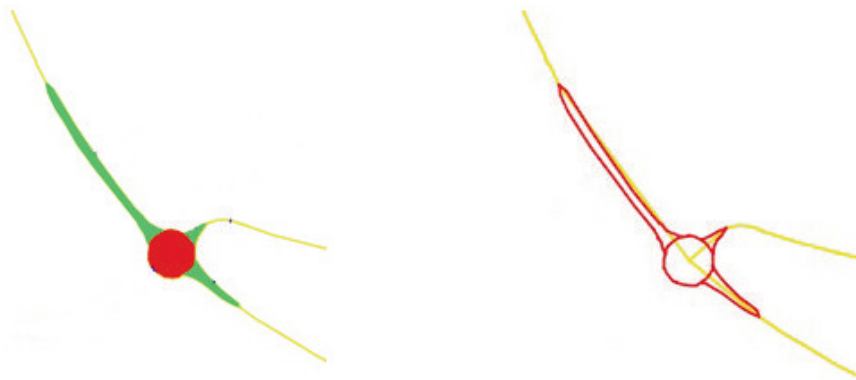


Figure 2 : À gauche, détection d'un rond-point (face centrale en rouge, pattes d'oie en vert)  
 À droite : simplification du rond-point (original en rouge, généralisé en orange)

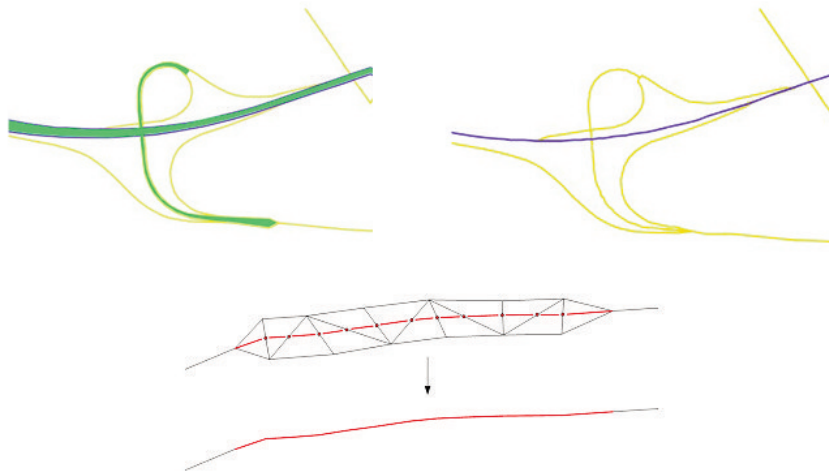


Figure 3 : Détection (en haut à gauche) et simplification (en haut à droite) de voies à chaussées séparées, par création d'un squelette central sur le séparateur grâce à une triangulation (en bas)

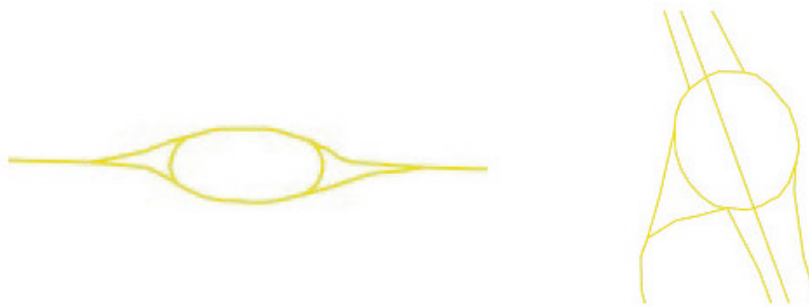


Figure 4 : ronds-points non détectés à cause d'une forme non circulaire (cas 1) ou d'une route parasite passant par-dessus le rond-point (cas 2)

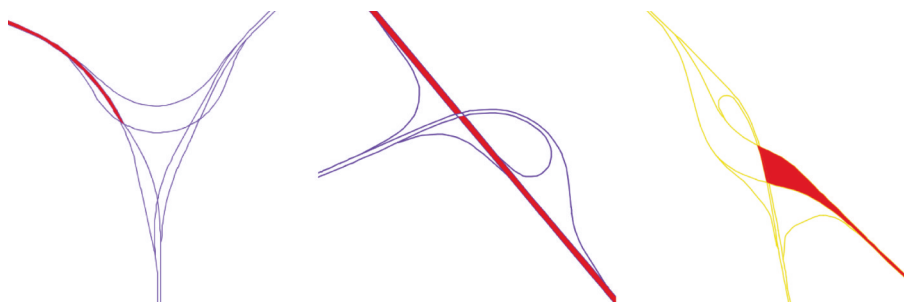


Figure 5 : Problèmes de détection de voies à chaussées séparées à cause des interconnexions

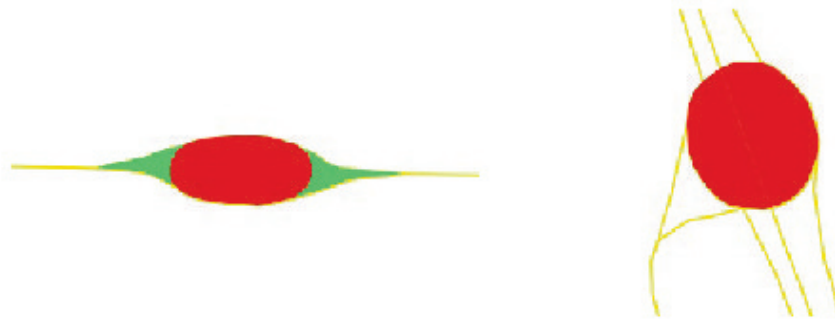


Figure 6 : Détection optimisée des ronds-points initialement non détectés de la figure 4

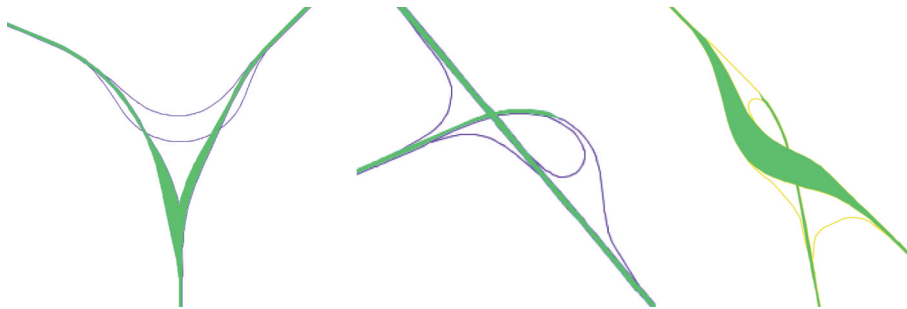


Figure 7 : Détection optimisée des voies à chaussées séparées initialement mal détectées de la figure 5

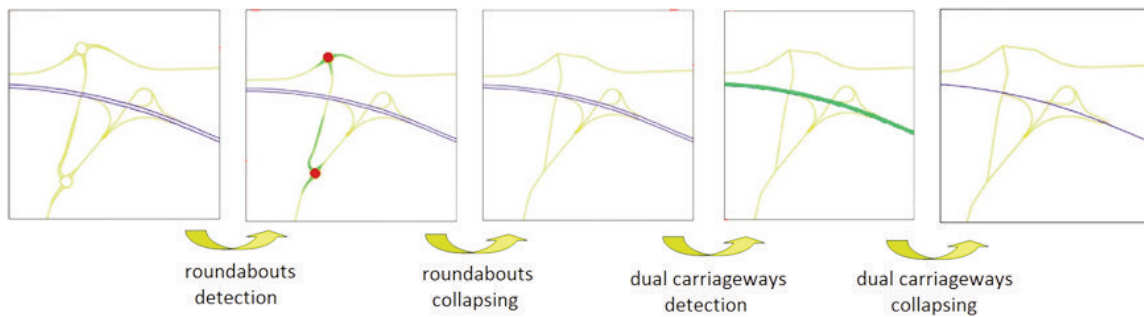


Figure 8 : Processus complet de détection et de simplification de structures routières complexes.  
La sélection et le déplacement de routes devraient intervenir dans un second temps