



HAL
open science

Modélisation de structures habitées à partir d'une approche de graphe

Mathieu Domingo

► **To cite this version:**

Mathieu Domingo. Modélisation de structures habitées à partir d'une approche de graphe. Géographie. Université de Bretagne occidentale - Brest, 2016. Français. NNT : 2016BRES0121 . tel-01492531v2

HAL Id: tel-01492531

<https://hal.science/tel-01492531v2>

Submitted on 30 May 2017

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.



université de bretagne
occidentale

UNIVERSITE
BRETAGNE
LOIRE

THÈSE / UNIVERSITÉ DE BRETAGNE OCCIDENTALE

sous le sceau de l'Université Bretagne Loire

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE

Mention : Géomatique

École Doctorale des Sciences de la Mer

présentée par

Mathieu Domingo

Préparée à l'Institut de Recherche de l'École
Navale (EA3634)

Modélisation de structures habitées à partir d'une approche de graphe

Thèse soutenue le 14 décembre 2016

devant le jury composé de :

Jérôme GENSEL

Professeur des universités, Université Grenoble Alpes / *rapporteur*

Marie-Hélène DE SEDE MARCEAU

Professeur des universités, Université de Franche-Comté / *rapporteur*

Françoise GOURMELON

Directeur de recherche CNRS, IUEM - BREST / *Examinatrice*

Géraldine DEL MONDO

Maitre de conférences, INSA Rouen / *Examinatrice*

Christophe CLARAMUNT

Professeur des universités, École Navale / *Directeur de thèse*

Rémy THIBAUD

Enseignant chercheur, École Navale / *Co-Directeur de thèse*

UNIVERSITÉ DE BRETAGNE OCCIDENTALE
ECOLE DOCTORALE DES SCIENCES DE LA MER

THESE

pour obtenir le titre de

Docteur en sciences

De l'Université de Bretagne Occidentale

Mention : Géomatique

Présentée et soutenue par

Mathieu Domingo

Modélisation de structures habitées à partir d'une approche de graphe

Thèse dirigée par Christophe Claramunt

Préparée à l'institut de Recherche de l'École Navale (EA3634)

Soutenue le : **14 décembre 2016**

Jury :

Rapporteurs :	Jérôme Gensel	Université Grenoble Alpes
	Marie-Hélène De Sede Marceau	Université de Franche-Comté
Examineurs :	Françoise Gourmelon	CNRS UMR 6554 - IUEM Brest
	Géraldine Del Mondo	INSA de Rouen
Directeur :	Christophe Claramunt	École Navale
Co-Directeur :	Remy Thibaud	École Navale

Remerciements

Table des matières

Remerciements	i
Table des matières	iii
Liste des figures.....	vii
Liste des tableaux	xi
Liste des requêtes	xiii
Introduction	1
Contexte.....	1
Verrous liés à la structuration des territoires	2
Les objectifs et l'organisation du manuscrit.....	4
Chapitre I État de l'art	9
I.1 Éléments de modélisation et de représentation	10
I.1.1 Les objets géographiques	10
I.1.2 Structuration des espaces habités	16
I.2 Éléments de caractérisation de la structure spatiale.....	21
I.2.1 Mesure de Densité.....	21
I.2.2 Mesure de distance	22
I.2.3 Mesure de centralité	23
I.2.4 Mesure d'interconnectivité.....	25
I.2.5 Extraction de sous-ensembles	26
I.3 Modèles d'analyse d'espaces habités	30
I.3.1 Approche primaire ou duale.....	30
I.3.2 Syntaxe Spatiale	31

I.3.3 Modélisation basée sur l'approche directe.....	35
I.3.4 Modèle basé sur le bâti	36
I.4 Synthèse.....	37
Chapitre II Modèle de graphe.....	41
II.1 Espace habité.....	42
II.1.1 Définition	42
II.1.2 Hypothèses et propriétés générales	43
II.1.3 Propriétés topologiques	43
II.2 Modèle de graphe	45
II.2.1 Associations spatiales.....	46
II.2.2 Graphes élémentaires	46
II.2.3 Graphes intermédiaires.....	48
II.2.4 Graphe intégral	48
II.2.5 Traitement d'un exemple à travers le modèle de graphe proposé.....	49
II.3 Opérateurs	50
II.3.1 Mesure de centralité du bâti	50
II.3.2 Mesure de connectivité locale du bâti	54
II.3.3 Mesure de densité de bâti	55
II.3.4 Mesure de compacité du réseau de circulation.....	56
II.3.5 Impact de la mesure de compacité sur l'analyse de structuration de l'espace.....	58
II.4 Synthèse	61
Chapitre III Implémentation et Résultats.....	63
III.1 Schéma fonctionnel.....	64
III.2 Le modèle de données.....	66
III.2.1 Modélisation des objets et des relations dans l'espace	66
III.2.2 Modèle conceptuel et modèle relationnel	66
III.3 Construction du graphe	68
III.3.1 Prétraitement	68

III.3.2 Construction des tables relationnelles de relations.....	72
III.4 Implémentation des opérateurs.....	74
III.4.1 Centralité	74
III.4.2 Mesure de connectivité locale du bâti	78
III.4.3 Mesure de densité de bâti relative à une route.....	79
III.4.4 Metric Compactness	80
III.5 Application et résultats.....	82
III.5.1 Présentation des données géographiques.....	82
III.5.2 Mesure de centralité du bâti.....	83
III.5.3 Mesure de connectivité locale du bâti	98
III.5.4 Mesure de densité de bâti	100
III.5.5 Mesure de compacité.....	101
III.5.6 Corrélation entre la mesure de densité de bâti et la mesure de compacité	109
III.6 Synthèse.....	113
Conclusion.....	115
Contribution.....	115
Perspectives	117
Bibliographie	121

Liste des figures

Figure 1 : Contextualisation de la thèse suivant les différents domaines scientifiques abordés	6
Figure 2 : Les huit relations du RCC8 avec les transitions topologiques possibles	14
Figure 3 : Exemple de typologie du tissu urbain (Puissant, et al., 2010)	17
Figure 4 : (a) Ville dispersée, (b) Ville couloir et (c) Ville compacte. (Borrego, et al., 2006)	18
Figure 5 : Taxonomie des formes de réseaux de circulation (Marshall, 2005)	19
Figure 6 : (a) Réseaux de Manhattan, (b) Réseaux en treillis et (c) Toile orbitèle. (Chekir, et al., 2014).....	19
Figure 7 : Exemple de deux structures (<i>tree</i> , <i>grid</i>) à différents niveaux de granularité (micro, macro). (Marshall, 2005).....	20
Figure 8 : Empreinte du bâti et Carte axiale résultante	32
Figure 9 : Mesure d'accessibilité obtenue à partir de la carte axiale de la Figure 8	33
Figure 10 : Mesure de centralité de la ville d'Ahmedabad sur un graphe primaire; a) centralité globale mesurée sur tout le graphe b) centralité locale avec une distance entre les nœuds inférieure à 400 mètres c) centralité locale avec une distance entre les nœuds inférieure à 200 mètres (Porta, et al., 2006).....	35
Figure 11 : Schéma de structure du graphe urbain du modèle REMUS. Modifié de (Badariotti, et al., 2007).....	37
Figure 12 : Représentation schématique d'une partie d'espace habité. Il est formé par la combinaison de bâtiments (b), de parcelles (p) et de routes (r).....	42
Figure 13 : (a) Représentation schématique d'un espace habité, (b) Graphe de bâtiments adjacents, (c) Graphe de parcelles adjacentes et (d) Graphe de routes adjacentes.....	47
Figure 14 : (a) Représentation schématique d'un espace habité, (b) Graphe de bâtiments et de parcelles, (c) Graphe de parcelles et de routes et (d) Graphe intégral combinant bâtiments, parcelles, et routes.	49
Figure 15 : Influence de l'homogénéité du réseau sur l'évolution de la <i>Minimum Ratio Minimum Distance</i> (D_{\min}^X).....	59
Figure 16 : Impact de la densité sur la <i>Minimum Ratio Minimum Distance</i>	60
Figure 17 : Schéma fonctionnel.....	65

Figure 18 : Modèle conceptuel de données.....	67
Figure 19 : Création de routes en polygones.....	70
Figure 20 : Territoire de Guisseny	82
Figure 21 : Centralité de degré appliquée au graphe de bâti.....	83
Figure 22 : C_{BPR} avec $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 1$	85
Figure 23 : Rapport entre centralité et longueur de route	86
Figure 24 : C_{BPR} avec $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$	87
Figure 25 : Histogrammes des valeurs de C_{BPR} pour $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 1$ et $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$	87
Figure 26 : Histogramme des différences de valeurs de C_{BPR}	89
Figure 27 : Différence des valeurs de C_{BPR} obtenus avec $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$ et avec $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 1$	88
Figure 28 : Localisation des zones 1, 2 et 3.....	89
Figure 29 : C_{BPR} avec tous les coefficients à 1 à gauche et C_{BPR} avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$ à droite	90
Figure 30 : Différence des scores de C_{BPR}	90
Figure 31 : Valeur de centralité de degré.....	92
Figure 32 : C_{BPR} avec tous les coefficients α_i égaux à 1	93
Figure 33 : C_{BPR} avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$	93
Figure 34 : C_{BPR} avec tous les coefficients α_i égales à 1	95
Figure 35 : C_{BPR} le long d'une route avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$	95
Figure 36 : Localisation de la zone 4.....	96
Figure 37 : C_{BPR} avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$ appliqué à l'ensemble du graphe intégral.....	97
Figure 38 : C_{BPR} avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$ appliqué sur un sous-graphe	97
Figure 39 : Mesure de connectivité locale du bâti	98
Figure 40 : Mesure de connectivité locale du bâti	99
Figure 41 : Mesure de connectivité locale du bâti et longueur de route	99
Figure 42 : Building Road Density	100
Figure 43 : Score de centralité et longueur de route par score de densité (idem Figure 23).....	101
Figure 44 : Metric Compactness avec $X = 10\%$	103
Figure 45 : Metric Compactness avec $X = 25\%$	103
Figure 46 : Metric Compactness avec $X = 45\%$	104

Figure 47 : Metric Compactness avec $X = 100\%$	104
Figure 48 : Distribution des valeurs de Metric Compactness.....	105
Figure 49 : Variations de D_{\min}^X en fonction de X sur le territoire de Guisseny	106
Figure 50 : Metric Compactness avec $X = 45\%$ pour le jeu de données modifié.....	107
Figure 51 : Variations de D_{\min}^X en fonction de X pour le jeu de données d'origine et modifié.....	107
Figure 52 : Metric Compactness avec $X = 85\%$ pour le jeu de données d'origine.....	108
Figure 53 : Metric Compactness avec $X = 85\%$ pour le jeu de données modifié.....	108
Figure 54 : Corrélation entre densité de bâti et compacité de route	112

Liste des tableaux

Tableau 1 : Modèle 4IM	14
Tableau 2 : Correspondance entre RCC8 et Modèle 4IM	15
Tableau 3 : Modèle 9IM	15
Tableau 4 : Comparatif des approches de modélisation pour leur application aux espaces habités	39
Tableau 5 : Extrait des relations du RCC8 (Randell & Cohn, 1989)	43
Tableau 6 : Classification des tronçons de circulation	111

Liste des requêtes

<i>Requête 1 : Liste des paires de bâtiments adjacents</i>	73
<i>Requête 2 : Liste des paires (bâtiment, parcelle) qui se superposent</i>	73
<i>Requête 3 : Association des bâtis avec les routes</i>	73
<i>Requête 4 : Bâtiments adjacents et sur une même parcelle</i>	74
<i>Requête 5 : Bâtiments adjacents, mais sur différentes parcelles</i>	75
<i>Requête 6 : Bâtiments non adjacents, mais sur la même parcelle</i>	76
<i>Requête 7 : Bâtiments non adjacents, mais sur des parcelles adjacentes</i>	77
<i>Requête 8 : Bâtiments non adjacents, mais sur des parcelles connectées à une même route</i>	78
<i>Requête 9 : Mesure de connectivité locale du bâti</i>	79
<i>Requête 10 : Mesure de densité de bâti relative à une route</i>	80
<i>Requête 11 : Mesure de Metric Compactness</i>	81

Introduction

Contexte

Trouver le moyen de caractériser la façon dont est structuré l'espace pour décrire l'état d'un arrangement géographique ou d'un phénomène est un questionnement scientifique courant pour de nombreuses disciplines, mais aussi un sujet en perpétuelle réflexion et renouvellement. Si dans certains cas ce besoin se limite à une description et un inventaire des éléments qui composent l'espace, il est fréquent que l'ambition et les objectifs soient bien supérieurs et concernent en particulier l'étude de l'évolution de l'espace à des fins d'anticipation et de simulation des phénomènes. Dans ce cadre, en plus d'un inventaire des éléments structurants, il est reconnu qu'une telle démarche doit reposer également sur l'observation et l'analyse des interactions entre les éléments, que ces derniers soient statiques (leur état ne change pas au cours du temps) ou dynamiques (leur état change au cours du temps) (Thériault & Claramunt, 1999). Ainsi, pour citer un exemple local, c'est l'analyse du territoire dans ses multiples composantes (incluant la typologie et l'organisation du relief, l'impact des phénomènes de courants et marées et les capacités et habitudes de manœuvre des bâtiments militaires) qui permet à Vauban d'identifier la position stratégique et la vulnérabilité de l'anse de Camaret, et ainsi de définir la localisation optimale de la « tour Vauban » qui le 18 juin 1694 permettra la défense maritime du goulet de Brest.

Quel que soit le domaine d'application concerné tel que l'urbanisme, l'écologie, la géographie, l'archéologie ou toute autre thématique impliquant la prise en compte de la dimension spatiale de l'environnement, deux catégories de verrous scientifiques se distinguent : d'une part les verrous liés à la *structuration* des territoires et d'autre part les verrous liés à la *dynamique* des territoires. C'est l'implication des actions humaines dans la structuration de l'espace qui est ici abordée (cf. abduction (Couclelis, 2009)). Appréhender l'organisation spatiale d'un territoire génère un vaste champ d'interrogations relatif aux aménagements (distribution et propriétés du bâti, du parcellaire et des voies de circulation) et à la pertinence de ces aménagements en matière d'usages. Premiers jalons nécessaires à la caractérisation, à l'étude et à l'analyse des usages, ce sont ces

interrogations sur la structuration de l'espace qui focalisent l'attention dans cette recherche. Il nous appartiendra de ne pas oublier qu'un lien direct existe entre structuration, usage et dynamique, les trois étant fortement liés en particulier quand la question des flux et des déplacements dans l'espace est évoquée.

Verrous liés à la structuration des territoires

Notre problématique porte en premier lieu sur la représentation de l'espace géographique. Elle inclut plusieurs questionnements sur l'inventaire et la modélisation des éléments à prendre en compte, la définition des relations entre eux, le choix d'un mode de représentation de la structure spatiale et la prise en compte des flux et des déplacements.

- **Quels éléments ?**

Afin de pouvoir représenter un territoire, il est nécessaire de savoir quels sont les éléments géographiques à prendre en compte. On entend ici par éléments géographiques, l'ensemble des composantes géolocalisables dans l'espace et pouvant être représentées et intégrées dans un système d'information avec leurs coordonnées spatiales (le « quoi », définit par (Peuquet, 1984)). On utilisera fréquemment le terme « objet » pour désigner ces éléments en référence à la terminologie usuelle dans le domaine des systèmes d'information. Au-delà du choix de ces objets, c'est aussi la question de leur modélisation qui est posée, avec d'une part les modèles conceptuels, tels que les modèles à base d'ontologies (Noy & McGuinness, 2000), et d'autre part les modèles logiques basés sur la représentation géométrique et thématique des objets (Peuquet, 1984; Goodchild, 1989). Une des problématiques consiste à préciser si l'on traite d'informations discrètes pour lesquelles il est possible d'identifier les limites géographiques, ou d'informations continues pour lesquelles les frontières sont indéfinies (l'élévation du sol, la température ou le niveau de bruit sont des exemples de structures spatiales qui varient dans l'espace continu). Enfin, le dernier aspect à inclure concerne la qualité des données utilisées pour modéliser les objets. La qualité (qui regroupe des notions différentes incluant par exemple la précision géométrique, l'exhaustivité, la précision sémantique, la cohérence logique ou l'actualité) ayant une influence directe sur la fiabilité des analyses spatiales produites et des conclusions qui en découlent.

- **Quelles relations entre les objets ?**

La typologie des relations entre les objets spatiaux se regroupe essentiellement autour de trois notions distinctes relatives aux relations dans l'espace et dans le temps : la relation spatio-temporelle, la relation de filiation et la relation spatiale. Les relations dans le temps, qui ne seront

pas abordées ici, incluent d'une part les relations spatio-temporelles, qui concernent les relations existantes entre des objets présents au même endroit mais à des temps différents (Del Mondo, et al., 2010) et d'autre part les relations de filiation qui décrivent la façon dont les entités appartiennent à la même famille (terme défini par le contexte applicatif) et sont liées par des liens de descendance et/ou identitaires (Hornsby & Egenhofer, 2000). Les relations spatiales proprement dites font référence aux relations topologiques et aux relations algébriques. Les premières s'appuient sur l'observation des liens de connexion, de superposition et d'inclusions entre les objets (Randell & Cohn, 1989), alors que les secondes traitent de relations quantitatives (distances métriques, distances de déplacement...) (Berretti, et al., 2003) ou qualitatives (projectives) (Frank, 1991; Clementini & Laurini, 2008).

- **Quels liens avec les flux ?**

Dans le cas d'un environnement bâti, nous pouvons avoir l'intuition que l'organisation de l'espace et l'usage qui en est fait sont intimement liés (Anderson, et al., 1996; Puissant, et al., 2010). La caractérisation des déplacements des humains dans l'espace est-elle alors une conséquence de cette organisation ou en est-elle affranchie ? C'est la typologie des déplacements qui permet de contraindre l'interrogation en précisant en particulier la place des réseaux de circulation dans ces déplacements. Selon que le déplacement sera piétonnier ou pas, que le véhicule utilisé (si c'est le cas) devra utiliser ou non des voies dédiées (chemin de fer, voies réservées) ou que des déplacements multimodaux (piéton + bus + métro) seront autorisés ou pas, la caractérisation de l'espace sera différente.

- **Quel modèle de représentation ?**

En fonction des objets et des relations entre objets que l'on décide de prendre en compte, il est nécessaire de choisir une structure de données adaptée (Galton, 2000; Peuquet, 2002). Les structures les plus usuelles sont les structures dites vectorielles et rasters. Dans le modèle *raster* l'espace géographique est segmenté en mailles géométriques élémentaires (le plus souvent carrées), et dans le modèle *vecteur*, l'espace géographique est représenté par des primitives géométriques dessinant l'empreinte spatiale des objets. S'appuyant sur ces structures, la caractérisation des espaces s'exprime au travers de deux modèles de représentation : modèle agrégé ou modèle de graphe spatial. Dans le premier cas, l'espace est décomposé en une somme d'unités spatiales (un maillage régulier ou irrégulier par exemple) et les propriétés de l'espace sont décrites à partir d'attributs dérivés décrivant une ou plusieurs de ces propriétés (p. ex. densité d'objets, distance entre les objets...) et agrégés par unité spatiale (Borrego, et al., 2006). Dans le second cas, les objets et les relations entre eux définissent les nœuds et les arcs des graphes dont

les positions correspondent aux coordonnées géographiques des objets qu'ils représentent. Les propriétés géométriques et topologiques de ces graphes alimentent des opérateurs et des fonctions permettant de décrire les propriétés structurales de l'espace (Freeman, 1979). Dans le contexte des espaces bâtis, les réseaux de transports et plus généralement les réseaux viaires sont les éléments les plus représentatifs et sont le support de nombreuses études cherchant à extraire des typologies (Marshall, 2005; Porta, et al., 2006; Badariotti, et al., 2007b; Lagesse, 2015). Dans le même esprit, la recherche en syntaxe spatiale propose également d'extraire les propriétés structurales à partir de graphes représentant non pas des objets de l'espace mais des cartes axiales dessinées par ces objets (Hillier & Hanson, 1984).

Les objectifs et l'organisation du manuscrit

À partir des verrous identifiés autour de la problématique de la caractérisation de la structure spatiale, notre problématique porte sur la conceptualisation et la mise en œuvre d'un modèle de graphe, et le développement d'opérateurs permettant d'analyser les propriétés du graphe et d'en extraire des informations sur la structure spatiale. Dans la continuité des travaux en analyse de graphes spatiaux et en syntaxe spatiale, l'ambition est d'étendre les modèles actuels vers un modèle qui devrait permettre une représentation plus complète de la structure des territoires géographiques. Nous expérimenterons le potentiel de cette modélisation pour une meilleure compréhension de l'organisation spatiale de territoires ruraux, périurbains et urbains. Alors que la majorité des modèles spatiaux actuels traitent préférentiellement de contextes urbains, caractérisés par une forte compacité et densité des objets bâtis et du réseau viaire, l'objectif est ici de proposer un modèle de graphe spatial et des outils de mesure adaptés à des structures plus ouvertes composées de zones faiblement et non habitées comme en milieux ruraux. Pour cela, nous nous intéresserons :

- à la place du parcellaire dans la structuration de l'espace
- à la différenciation entre des réseaux de circulation fortement connectés ou pas.

Concernant la problématique du lien entre la structuration du bâti et celle du réseau de circulation, notre recherche tendra au développement d'éléments de caractérisation de la structure spatiale permettant de qualifier cette adéquation.

La thèse est organisée autour des trois domaines fondamentaux (Figure 1) suivants :

- Les Sciences de l'Information Géographique (SIGs)

Bien plus qu'un outil de visualisation de l'information spatiale, les SIGs sont impliquées dans ce travail comme le support de gestion de l'information spatiale et de production de représentations de l'organisation spatiale. L'implémentation et la mise en œuvre du modèle développé reposeront sur ces capacités et plus spécifiquement sur le couplage entre l'interface de visualisation et les outils de mesure et d'analyse liés aux systèmes de gestion de bases de données spatiales.

- L'analyse spatiale

L'analyse spatiale désigne l'ensemble des méthodes de traitement et de mesure applicables aux données géoréférencées, et fournissant une information sur des propriétés géométriques et/ou géographiques variant en fonction de la position de ces données (De Smith, et al., 2007). Nous nous intéressons particulièrement à la partie de l'analyse spatiale qui tend à décrire l'organisation de l'espace, par exemple via la répartition des objets et leurs connexions.

- La théorie des graphes

Approche clé dans la représentation et l'intégration des relations entre les objets, la théorie des graphes permet de structurer efficacement les données et de disposer des bases de raisonnement et des opérateurs adaptés à notre problématique pour la caractérisation des propriétés topologiques des réseaux (centralité, interconnectivité, etc.) et pour la mesure des propriétés géographiques (distance de parcours, plus court chemin, etc.) (Freeman, 1979).

Ces trois domaines sont fortement complémentaires dans le cadre de cette thèse. La théorie des graphes apporte une structure de données sur laquelle nous pouvons construire notre modèle et proposer des opérateurs d'analyse spatiale. Ces derniers permettent de produire des représentations afin de mieux comprendre l'organisation spatiale des territoires dans les SIGs à partir des résultats obtenus.

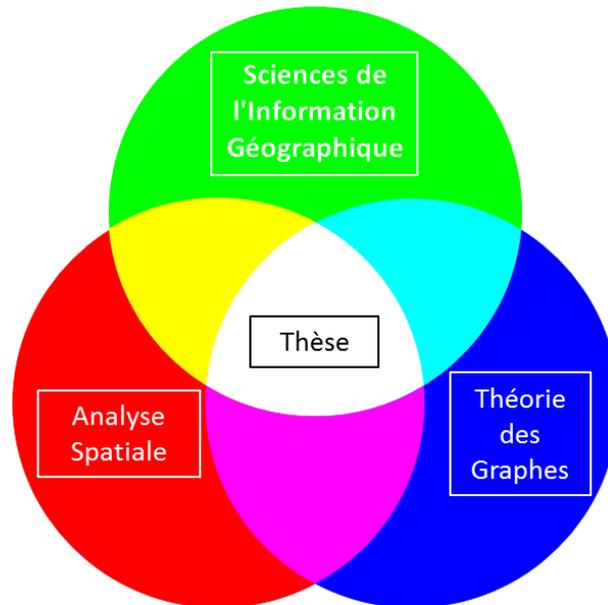


Figure 1 : Contextualisation de la thèse suivant les différents domaines scientifiques abordés

La thèse est organisée en 3 chapitres. Le premier chapitre présente l'état de l'art de l'analyse spatiale et de la théorie des graphes en se focalisant sur les éléments nécessaires à la description de l'organisation spatiale d'un territoire. Ce chapitre traite de la modélisation et de la représentation d'un espace géographique habité au sein des SIGs. Dans un premier temps, il décrit les concepts relatifs à la structuration de l'espace, aux objets présents dans ces espaces, et à la caractérisation des relations spatiales entre ces objets. Dans un second temps, il présente un certain nombre de mesures de la structure spatiale, en particulier les mesures basées sur les graphes. Enfin, il décrit les modèles et les outils d'analyse d'espaces urbains déjà proposés par la littérature, dont notamment ceux issus de la syntaxe spatiale.

Le chapitre 2 introduit la construction de notre modèle basée sur une combinaison de graphes élémentaires ainsi que les opérateurs spatiaux qui en découlent. Dans un premier temps, il décrit les différentes propriétés des espaces habités considérés. Dans un second temps, il introduit le modèle proposé. Enfin, la dernière partie du chapitre présente des opérateurs basés sur le modèle de graphe permettant une approche quantitative des propriétés de l'espace.

Le chapitre 3 propose une application du modèle et des opérateurs sur un cas d'étude : *la commune de Guisseny*, territoire rural situé au nord-ouest de la France. Ce chapitre décrit dans un premier temps l'intégration du modèle de graphe dans une base de données géographiques et les modalités d'implémentation des opérateurs proposés. Dans un second temps, il se focalise sur l'application à des données réelles et l'analyse des résultats qui en découlent.

En conclusion de la thèse se trouve un résumé des contributions relatif au modèle de graphe, à l'introduction d'opérateurs de graphes et à l'analyse de la structuration spatiale d'un territoire habité. Ensuite sont présentées les perspectives de recherches concernant l'enrichissement du modèle, l'exploration de nouveaux opérateurs, l'analyse de la robustesse de l'implémentation et le passage à l'échelle, l'étude de nouvelles pistes de recherches comme le temporel, et la confrontation à des thématiciens.

Chapitre I

État de l'art

I.1	Eléments de modélisation et de représentation	10
I.1.1	Les objets géographiques	10
I.1.2	Structuration des espaces habités	16
I.2	Eléments de caractérisation de la structure spatiale	21
I.2.1	Mesure de Densité.....	21
I.2.2	Mesure de distance.....	22
I.2.3	Mesure de centralité	23
I.2.4	Mesure d'inter-connectivité	25
I.2.5	Extraction de sous-ensembles	26
I.3	Modèles d'analyse d'espaces habités	30
I.3.1	Approche primaire ou duale.....	30
I.3.2	Syntaxe Spatiale	31
I.3.3	Modélisation basée sur l'approche directe	35
I.3.4	Modèle basé sur le bâti.....	36
I.4	Synthèse	37

Ce chapitre traite de la modélisation et de la représentation d'un espace géographique bâti au sein des SIGs. Dans un premier temps, il décrit les concepts relatifs à la structuration de l'espace, aux objets présents dans ces espaces, et à la caractérisation des relations spatiales entre ces objets. Dans un second temps, il présente un certain nombre de mesures de la structure spatiale, en particulier les mesures basées sur les graphes. Enfin, il décrit les modèles et les outils d'analyse

d'espaces urbains déjà proposés par la littérature, dont notamment ceux issus de la syntaxe spatiale.

1.1 Éléments de modélisation et de représentation

1.1.1 Les objets géographiques

1.1.1.1 Modélisation spatiale des objets géographiques

La représentation et la modélisation de l'espace géographique au sein des SIGs reposent souvent sur une discrétisation d'un environnement à l'aide de la notion d'objet. Cette discrétisation amène à rapprocher une vision conceptuelle de la notion d'objet (Goodchild & Sucharita, 1989; Couclelis, 1992; Debardieux, 2004). Dans le domaine des SIGs, les conceptualisations actuellement développées font cohabiter d'une part, la vision *raster* pour laquelle l'environnement est modélisé par la juxtaposition d'éléments géométriques réguliers (le plus souvent une grille régulière composée de pixels) et d'autre part la vision *vecteur* pour laquelle l'environnement est modélisé par la combinaison d'éléments géométriques de type points, polygones et polygones (Peuquet, 2002). La différence conceptuelle fondamentale est que dans le cas d'un modèle *vecteur* l'élément géométrique est défini comme étant l'empreinte spatiale d'un objet géographique identifiable, alors que dans la vision *raster*, la référence à un objet géographique est secondaire, voire absente. Dans ce dernier cas, l'empreinte spatiale représentée (un carré dans le cas d'une grille de pixels) n'est pas équivalente à un objet géographique identifiable (sauf cas fortuit), mais fait référence à une propriété locale de l'espace, le plus souvent un phénomène de nature continue tel que par exemple la température ou l'élévation. Cette distinction entre objets vectorisés et champs de pixels peut également être étendue aux modèles 3D (Peuquet, 2002).

Alors que le choix d'une modélisation *raster* ou *vecteur* est souvent vu comme la conséquence d'un choix technique (quelles sont les données à ma disposition), une approche qui consiste à s'interroger en premier sur les objectifs d'une modélisation permet d'aborder ce point de vue différemment. Ce questionnement, décrit par (Mathian & Sanders, 2014), doit permettre de préciser l'étendue de la notion d'objet (s'agit-il d'objets identifiables, d'éléments géométriques empiriques, ou autre ?) et l'identification et l'analyse des relations entre ces objets. Dans le cas de cette thèse, l'objectif est de traiter des relations spatiales entre des objets géographiques participant à l'émergence d'une structure spatiale du territoire c'est-à-dire ayant un ancrage géométrique clair. Cette définition est à rapporter à la nomenclature de (Smith & Varzi, 2000) qui fait la distinction entre d'une part des objets *bonna fide* dont la géométrie utilisée dans le système

d'information est associée à une empreinte physique sur le terrain (un bâtiment par exemple), et d'autre part des objets *fiat*, dont la géométrie n'apparaît pas sur le terrain (limite administrative par exemple). En fait, indépendamment de cette nomenclature, c'est le fait de pouvoir extraire et manipuler des objets géométriques et des frontières qui est recherché, et c'est l'approche *vecteur* qui est donc privilégiée.

I.1.1.2 Modélisation géométrique des objets

La modélisation géométrique des objets dépend de la *représentation conceptuelle* qui leur est associée. On distingue d'une part la localisation de l'objet, et d'autre part son emprise (son extension et/ou sa forme). Un même objet peut aussi bien être modélisé par un polygone pour représenter son empreinte spatiale que par un point géolocalisé, lorsque seul son emplacement est important et que sa forme importe peu. C'est le cas par exemple dans les modèles de localisation d'activités commerciales qui regroupent la représentation polygonale du bâti et la localisation ponctuelle de l'activité liée à un bâtiment (Codescu, et al., 2011).

Le choix d'une représentation dépend du niveau de granularité spatiale auquel on souhaite prendre en compte l'objet, mais également des concepts de relation spatiale auxquels on fait appel pour les caractériser dans l'espace. Dans les travaux impliquant l'étude du bâti (Badariotti, et al., 2007b; Curie, et al., 2010) et l'analyse cadastrale (Spéry, et al., 2001), on se focalise sur des relations de connectivité et d'inclusion pour pouvoir repérer des ruptures et des tendances dans la structure urbaine. Il est alors nécessaire de traiter des objets géométriques possédant une extension spatiale non nulle (Tarski, 1982), et l'on utilisera classiquement l'empreinte spatiale des bâtiments ou des parcelles, représentée par l'aire du polygone associé. Concernant les réseaux de communication, les travaux relatifs à la description des réseaux et des flux utilisent communément les nœuds de circulation et les segments qui les relient à partir d'une représentation de graphe (Ford & Fulkerson, 1962). Cette approche est particulièrement utile pour des applications en transport ou gestion de trafic, ou pour des analyses de la structure spatiale de ces réseaux (Bhourri, et al., 2008). Néanmoins, l'usage d'une telle représentation de graphe ne convient pas complètement à une prise en compte des espaces à proximité de ce réseau, et en particulier le bâti et le parcellaire qui forment des éléments fondamentaux de la structure urbaine. Comme évoqué plus haut, afin de prendre en compte d'une part la connectivité entre les voies de circulation et d'autre part les relations entre les voies de circulation et l'espace environnant (bâti ou parcelle), il sera nécessaire de représenter cet espace sous une forme appropriée qui reste à construire. Au niveau des carrefours et autres nœuds de circulation, une attention particulière devra cependant être apportée pour permettre de bien différencier les différents tronçons de

communication et leurs connexions, et d'extraire les relations spatiales permettant de rendre compte de la connectivité entre les tronçons et de celle avec l'espace environnant.

I.1.1.3 Les objets structurants

Afin d'éviter une surcharge cognitive, un surcoût d'acquisition des données et des difficultés de manipulation de l'information, il est nécessaire de choisir les objets pertinents, nécessaires et suffisants pour aborder une problématique donnée et éliminer tous les objets superflus. L'identification des objets composant un espace géographique repose sur un inventaire détaillé de ces objets et des relations qui les associent. Regroupées sous le terme d'ontologie, il existe plusieurs façons de mener cette identification, selon que l'on cherche à extraire directement les objets pertinents à considérer, ou bien que l'on s'attache à relever les relations régissant le système en cours d'étude (Bera, 2004). Les SIGs se positionnent dans le premier cas avec le principe d'utilisation des *couches* d'informations et des objets géométriques identifiables et sélectionnables. Ainsi une couche correspond à un ensemble d'objets géographiques regroupés de manière logique (même type d'objets par exemple). Le géoréférencement des objets géographiques permet la superposition et l'association des objets géographiques de différentes couches. Un inventaire des travaux réalisés par la communauté de l'information géographique depuis plusieurs années permet de produire des outils de raisonnement pouvant définir les objets et les opérateurs de relations spatiales à partir d'ontologies (Batsakis & Petrakis, 2010; Jitkajornwanich, et al., 2011).

L'organisation des territoires est fréquemment associée à la répartition et la définition de l'activité humaine (Anderson, et al., 1996). Concernant l'étude des espaces habités, la réflexion porte le plus souvent sur l'influence des interventions humaines dans l'arrangement et la structuration de l'espace (Couclelis, 2009). Cette problématique peut être abordée à partir du rôle des objets introduits ou modifiés par l'homme et de leurs relations. Ce questionnement permet de différencier les objets, et donc de pouvoir les catégoriser, suivant les concepts de propriété ou d'usage. Le premier fait référence à la décomposition de l'espace en aires de propriété associées (juridiquement ou pas) à des individus, des institutions, ou tout équivalent. On fait ici référence au découpage cadastral et/ou à l'identification de parcelles du territoire. Le second fait référence aux objets qui ont été construits, qu'il s'agisse de bâtiments (le bâti) et/ou d'autres types d'infrastructures (terrain de sport, etc.). À la suite de cet inventaire et de cette différenciation, il est possible de diviser l'espace en différents domaines suivant la typologie des usages (p. ex. zone d'habitations versus zone industrielle) ou suivant son arrangement (fonction de la densité ou de la position relative des bâtiments) (Puissant, et al., 2010). Une place à part restera occupée par la

prise en compte des voies de circulation qui focalisent à elles seules une grande partie des questionnements relatifs à la structure des espaces habités.

Les nombreuses études menées dans le domaine de l'analyse des espaces habités identifient trois types d'objets de base comme objets géographiques structurants (Spéry, et al., 2001; Jiang & Claramunt, 2004; Badariotti, et al., 2007b; Curie, et al., 2010; Guyet, 2013) :

- Le bâti, qui correspond à l'ensemble des bâtiments construits par l'homme.
- Les parcelles, qui sont des portions du territoire appartenant à des individus ou des institutions.
- Le réseau de communication, qui est constitué de l'ensemble des voies de circulation.

I.1.1.4 Relations spatiales entre les objets

L'étude de la structuration des espaces habités impose de prendre en compte les relations qui associent les éléments structurants qui les composent. Ces éléments étant caractérisés par leur empreinte dans l'espace, une telle approche est construite à partir des relations spatiales entre ces empreintes. Les deux catégories de relations spatiales sont :

- Les relations quantitatives (c.-à-d., métrique), par exemple les distances entre les objets (Berretti, et al., 2003)
- Les relations qualitatives:
 - o Projectives : orientation, relations cardinales (Frank, 1991; Clementini & Laurini, 2008)
 - o Topologiques : voisinage (Randell & Cohn, 1989)

Les relations quantitatives et projectives sont généralement utiles dans les cas où les empreintes spatiales sont déconnectées. Pour les relations topologiques, le modèle le plus utilisé est le RCC8 (*Region Connection Calculus*) décrit dans (Randell, et al., 1992) qui est une modification de la théorie présentée dans (Randell & Cohn, 1989) basée elle-même sur les travaux de (Clarke, 1981). Ce modèle décrit les huit relations topologiques possibles entre deux régions d'un espace (Figure 2).

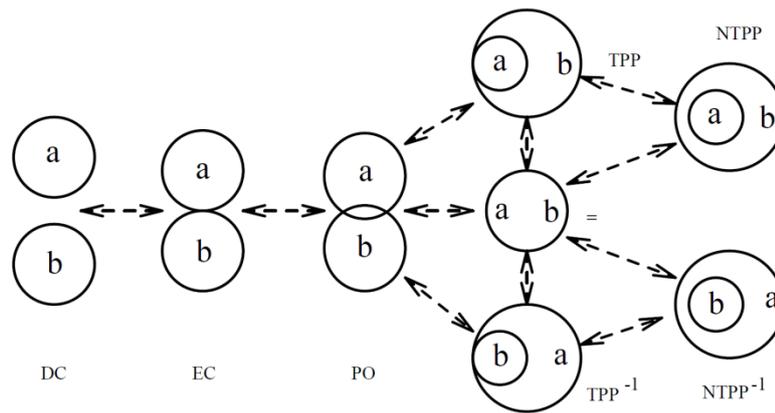


Figure 2 : Les huit relations du RCC8 avec les transitions topologiques possibles

Une autre approche ayant un pouvoir d'expression équivalent pour modéliser les liens entre objets de type région est le modèle 4IM (*Intersection Model*) défini dans (Egenhofer & Franzosa, 1991). Dans ce modèle, une région A est décrite à partir de son intérieur A° et de sa frontière ∂A . Les relations topologiques existantes entre deux régions A et B sont ensuite décrites à l'aide d'un tableau présentant les quatre intersections de leurs frontières et intérieurs (Tableau 1). Ce tableau décrit pour chaque cas si l'intersection est vide ou non vide, ce qui aboutit sur $2^4=16$ cas. Huit cas ne sont pas possibles entre deux régions spatiales, et les huit cas restants correspondent aux huit cas décrits par le RCC8. Le Tableau 2 représente les différentes équivalences de relations entre le RCC8 et le modèle 4IM.

	∂B	B°
∂A	$\partial A \cap \partial B$	$\partial A \cap B^\circ$
A°	$A^\circ \cap \partial B$	$A^\circ \cap B^\circ$

Tableau 1 : Modèle 4IM

	$\partial A \cap \partial B$	\wedge	$\partial A \cap B^\circ$	\wedge	$A^\circ \cap \partial B$	\wedge	$A^\circ \cap B^\circ$
DC (A, B)	\emptyset	\wedge	\emptyset	\wedge	\emptyset	\wedge	\emptyset
EC (A, B)	$\neg\emptyset$	\wedge	\emptyset	\wedge	\emptyset	\wedge	\emptyset
PO (A, B)	$\neg\emptyset$	\wedge	$\neg\emptyset$	\wedge	$\neg\emptyset$	\wedge	$\neg\emptyset$
A=B	$\neg\emptyset$	\wedge	\emptyset	\wedge	\emptyset	\wedge	$\neg\emptyset$
TPP (A, B)	$\neg\emptyset$	\wedge	$\neg\emptyset$	\wedge	\emptyset	\wedge	$\neg\emptyset$
TPP ⁻¹ (A, B)	$\neg\emptyset$	\wedge	\emptyset	\wedge	$\neg\emptyset$	\wedge	$\neg\emptyset$
NTPP (A, B)	\emptyset	\wedge	$\neg\emptyset$	\wedge	\emptyset	\wedge	$\neg\emptyset$
NTPP ⁻¹ (A, B)	\emptyset	\wedge	\emptyset	\wedge	$\neg\emptyset$	\wedge	$\neg\emptyset$

Tableau 2 : Correspondance entre RCC8 et Modèle 4IM

Le modèle 4IM a ensuite été élargi au modèle 9IM (Egenhofer & Herring, 1991), une région A étant désormais définie par son intérieur A° , sa frontière ∂A ainsi que son extérieur A^- . En plus d'être capable de décrire les huit cas du modèle 4IM (ou du RCC8) qui existent entre deux régions, le modèle 9IM peut aussi décrire des relations entre une région et une ligne, et entre deux lignes, à l'aide d'une matrice 3*3 (Tableau 3).

	∂B	B°	B^-
∂A	$\partial A \cap \partial B$	$\partial A \cap B^\circ$	$\partial A \cap B^-$
A°	$A^\circ \cap \partial B$	$A^\circ \cap B^\circ$	$A^\circ \cap B^-$
A^-	$A^- \cap \partial B$	$A^- \cap B^\circ$	$A^- \cap B^-$

Tableau 3 : Modèle 9IM

Une autre extension du modèle 4IM est présentée dans (Clementini, et al., 1993) : le modèle DE-4IM (*Dimension Extended-4IM*) dans lequel les intersections ne sont plus considérées seulement de manière binaire par vide ou non vide, mais via leur dimension (l'intersection est vide, l'intersection contient au moins un point mais pas de ligne ni de zone, l'intersection contient au moins une ligne mais pas de zone, l'intersection contient au moins une zone). Ce modèle aboutit à 52 relations possibles entre deux éléments de type région, ligne ou point. Il y propose aussi un CBM (*Calculus-Based Method*) où cinq noms de relations topologiques (*touch, in, cross, overlap, disjoint*) sont capables de décrire plus simplement les 52 relations possibles que via les matrices.

Bien que le plus souvent les objets géographiques aient été représentés par commodité avec des formes élémentaires bien définies, les relations spatiales de base sont extensibles pour tous types

d'objets géographiques dont les contours peuvent être flous (c.-à-d., localisation probable) (Alboody, et al., 2009) ou vagues (c.-à-d. localisation ambiguë, car liée à des sources de données contradictoires) (Bejaoui, et al., 2009).

I.1.2 Structuration des espaces habités

La différenciation et l'analyse de la structure des espaces habités, et plus particulièrement urbains, sont couramment focalisées sur la variabilité du type d'usage du territoire (Puissant, et al., 2010) et sur les propriétés du réseau de circulation (Buhl, et al., 2006).

I.1.2.1 Influence de l'usage des territoires

En contexte fortement urbanisé, la structure de l'espace est fortement conditionnée par l'activité humaine, et en particulier les mobilités, et peut être caractérisée au travers de trois concepts : la *densité*, la *diversité* et l'*organisation spatiale* (Cervero & Kockelman, 1997). La densité mesure le nombre d'objets (bâties, routes, etc.) présents par unité de surface et caractérise un certain degré d'activité humaine. La diversité fait référence au degré d'interaction spatiale entre différents types d'usages (Claramunt, 2012). Elle exprime le degré de proximité et de distribution spatiale relative entre les différents usages (c.-à-d., distance entre habitations, commerces, industries, services, loisirs, etc.). L'organisation spatiale définit notamment la forme globale de l'espace entre des formes monocentriques versus polycentriques, continues versus discontinues et centralisées versus décentralisées (Tsai, 2005). Cette dernière notion exprime la présence et les dimensions d'ilots et de blocs de bâtiments, et celles d'espaces ouverts (parkings, parcs...). À partir de ces définitions, (Tsai, 2005) propose un coefficient (*Moran Coefficient*) qui permet de décrire la structure des espaces autour des notions de compacité et d'étalement urbain. Ce coefficient, qui prend en compte à la fois le nombre d'habitants, la densité de population, la distribution de l'activité, et la dispersion de l'activité, est fort pour les structures monocentriques et denses, et est faible pour les structures urbaines étalées et décentralisées. De même, (Puissant, et al., 2010) proposent une classification en neuf classes fonction de la densité et du type de bâti, et prenant en compte la présence des réseaux de transport et l'hydrologie (Figure 3).

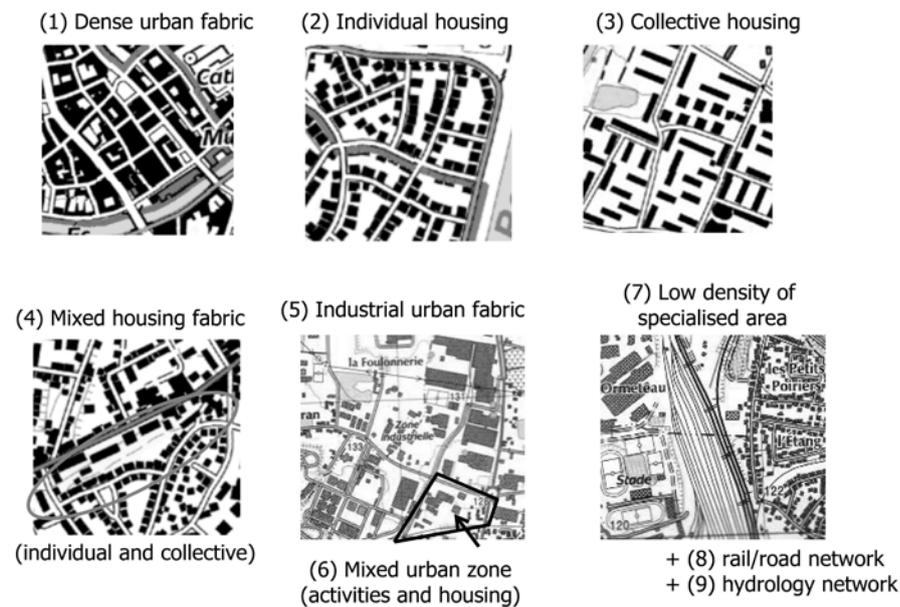


Figure 3 : Exemple de typologie du tissu urbain (Puissant, et al., 2010)

De façon similaire, mais à partir de concepts sensiblement différents, (Borrego, et al., 2006) mettent également en évidence l'influence de l'activité humaine sur les typologies émergentes des structures urbaines. En considérant quatre types de zones (zone verte inhabitée, zone rurale, zone périurbaine et zone urbaine) et en s'interrogeant sur l'organisation des déplacements à l'intérieur de l'espace habité, ils en déduisent trois types de villes possibles :

- La *ville dispersée* (Figure 4a), définie par une zone urbaine étendue mais de faible densité, cernée par une zone périurbaine également étendue et de faible densité. Un tel arrangement impose des déplacements en voiture.
- La *ville couloir* (Figure 4b) structurée le long d'axes de déplacement principaux pouvant accueillir un fort trafic. Elle est constituée d'un centre urbain à la croisée des axes, et de plusieurs zones urbanisées réparties le long des axes de déplacement et pouvant entraîner un étalement important de la ville.
- La *ville compacte* (Figure 4c) où la densité de la population est concentrée au centre. L'objectif est de réduire les temps de trajets et/ou de mutualiser les moyens de transport (transport individuel et collectif).

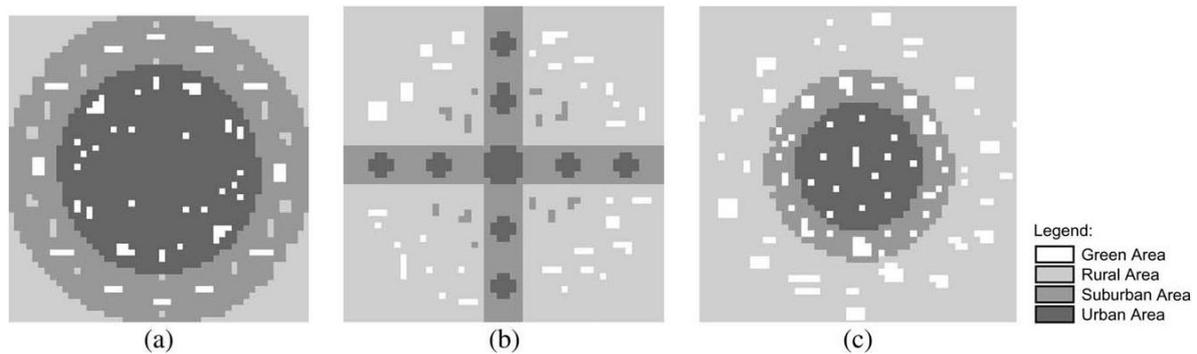


Figure 4 : (a) Ville dispersée, (b) Ville couloir et (c) Ville compacte. (Borrego, et al., 2006)

I.1.2.2 Morphologie des réseaux de déplacement

Si l'on souhaite s'intéresser à la structure des villes, on ne peut se limiter à la seule analyse de la forme globale de la ville, mais on doit aussi s'intéresser dans le détail à l'organisation de son réseau de déplacement (Marshall, 2005).

De très nombreux travaux classifient les types de réseaux de déplacements (routiers le plus souvent, mais piétons et deux roues peuvent également être pris en compte), (Spreiregen, 1965; Keeble, 1969; Lynch, 1984; Frey, 1999). La définition des règles et des termes choisis par les différents auteurs pour décrire les réseaux, propose une vision exhaustive et permet de prendre en compte de très nombreux cas. Cependant, chaque définition est souvent spécifique à un contexte particulier, ce qui multiplie les termes, et amène souvent à des définitions floues qui génèrent parfois une confusion¹. Selon les cas, on s'intéresse au nombre de croisements, au caractère rectiligne ou courbe des portions de routes, à la forme radiale ou concentrique du réseau, ou encore à la présence de sous-parties dans le réseau. Devant ce constat, (Marshall, 2005) propose de rationaliser l'analyse en se focalisant sur la présence, ou non, de boucles (*cell/circuit*) et d'embranchements (*branching*). Il propose de classifier les structures des réseaux à partir de la définition d'un ensemble de cinq configurations de base (Figure 5) définies par des relations topologiques. La Figure 5 est illustrée avec des exemples de tronçons de routes sinueuses, mais la classification est également applicable à des réseaux de tronçons rectilignes. En particulier, les configurations de *grid* (*manhattan* ou *rectilinéaire*), de *treillis* ou d'*orbitel*, illustrées sur la Figure 6, correspondent à des formes particulières de réseaux de type *cellular*.

¹ Marshall (2005) identifie plus de cent descripteurs utilisés pour les réseaux urbains (p. ex. structure radiale, structure en grille, structure en arbre, structure linéaire, etc.)

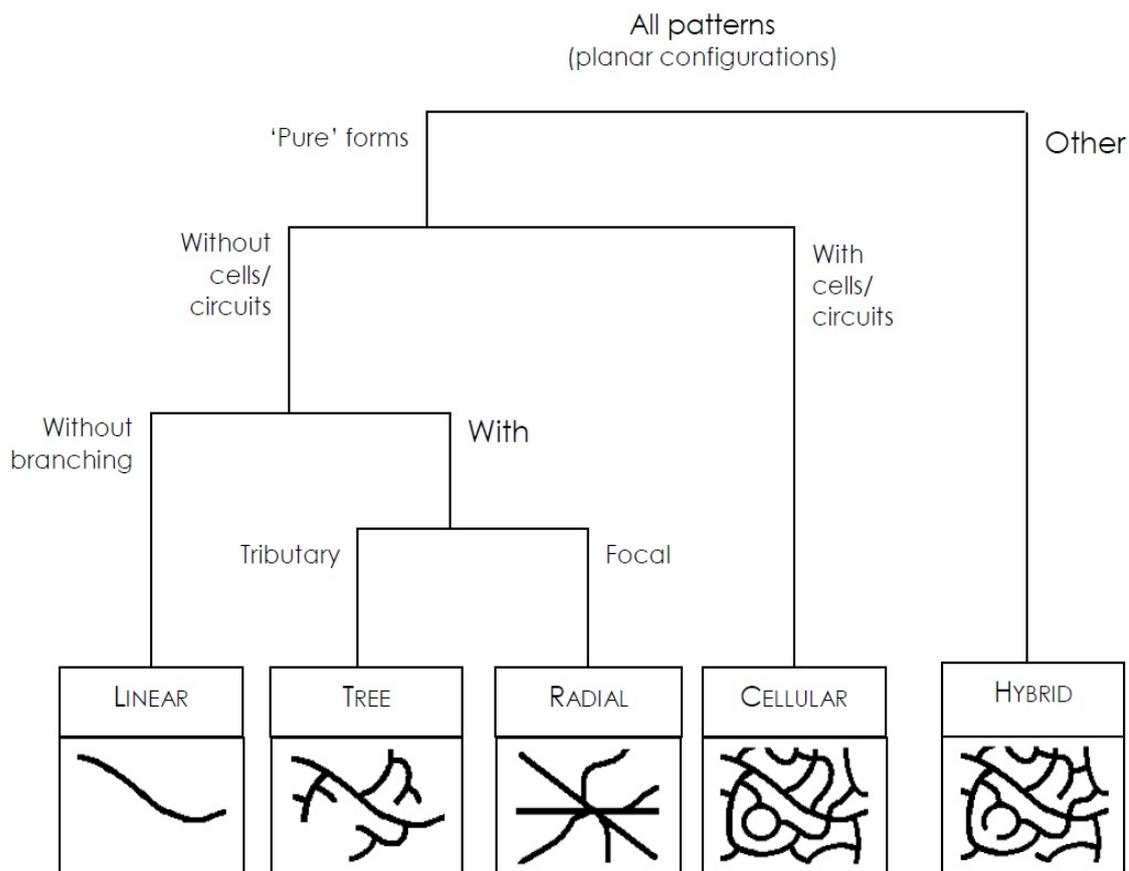


Figure 5 : Taxonomie des formes de réseaux de circulation (Marshall, 2005)

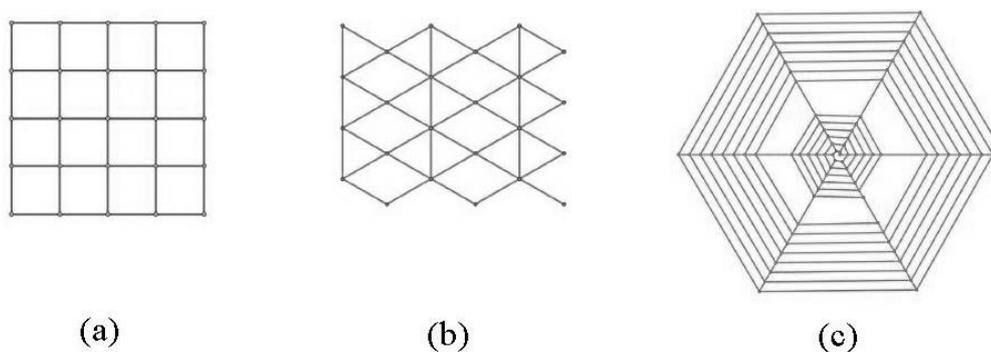


Figure 6 : (a) Réseaux de Manhattan, (b) Réseaux en treillis et (c) Toile orbitèle. (Chekir, et al., 2014)

Au-delà de cette première approche de classification des structures spatiales urbaines, il est possible de compléter celle-ci en prenant en compte la typologie (et/ou la sémantique) des tronçons de route qui composent un réseau de circulation. Ainsi, (Marshall, 2005) propose par exemple de prendre en compte les notions de routes principales et de routes secondaires. Cette

approche permet de décrire le réseau à différents niveaux de granularité (macrostructure et microstructure) et, par exemple, de faire la différence, pour une même macrostructure de type arbre, entre un réseau présentant une microstructure également de type arbre, ou une structure de type grille (Figure 7). Dans l'étude des déplacements, on peut noter que cette différenciation entre macrostructure et microstructure n'a d'importance que si elle est associée à une différenciation des modes et conditions de déplacement entre ces deux ensembles. Enfin, la dernière composante à évaluer est le nombre et/ou la qualité des liens de connectivité. Elle permet d'extraire des sous-ensembles dans le réseau. Cette notion est illustrée par la différenciation entre « structure fortement connectée » et « structure faiblement connectée » illustrée sur la Figure 7. Ce point sera abordé dans l'étude des connectivités.

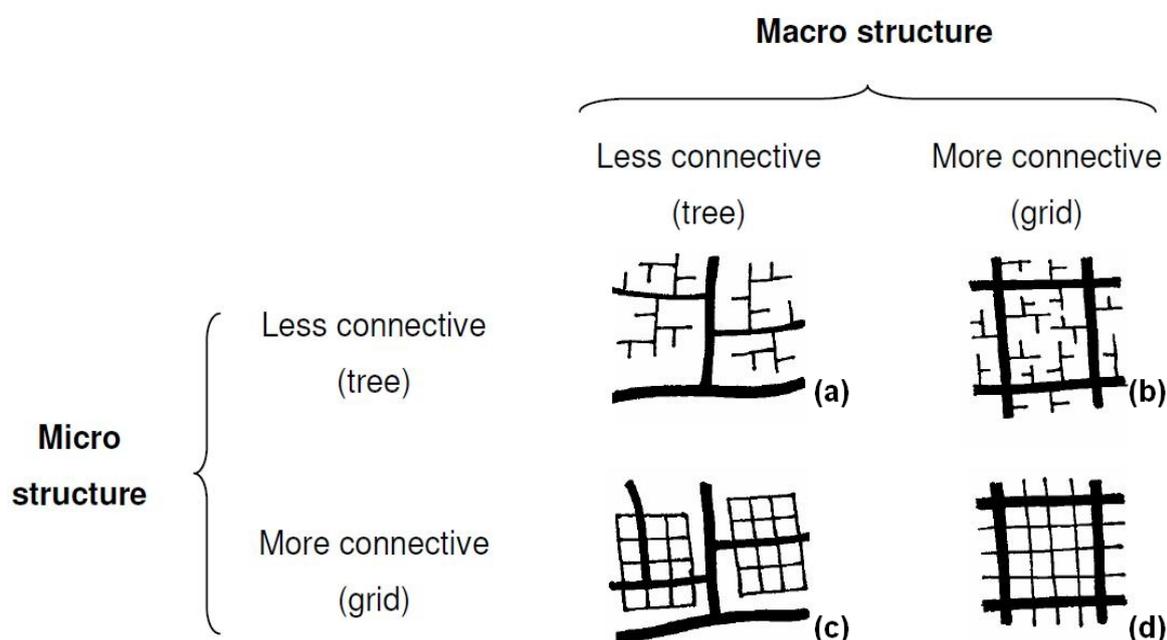


Figure 7 : Exemple de deux structures (*tree, grid*) à différents niveaux de granularité (micro, macro). (Marshall, 2005)

L'étude et l'analyse de la structuration des espaces habités reposent essentiellement sur la caractérisation de relations spatiales entre des éléments remarquables (notamment le bâti, le parcellaire et le réseau de circulation). Les relations spatiales entre ces objets sont de différentes natures (quantitatives et qualitatives). Elles offrent le support conceptuel qui permet de décrire la variabilité des usages du territoire (la densité, la diversité et l'organisation spatiale) et la morphologie du réseau de circulation.

I.2 Éléments de caractérisation de la structure spatiale

L'approche qui consiste à extraire les objets remarquables de l'espace et à identifier les relations entre ces objets nous amène logiquement d'une part à observer la localisation de ces objets pour en extraire une information sur leur distribution spatiale, et d'autre part à utiliser le réseau de connectivité entre ces objets pour caractériser les interactions.

Concernant la modélisation d'un réseau, nous retenons une approche basée sur un modèle de graphe noté $G(S,A)$ tel que S est un ensemble non vide de sommets et A un ensemble d'arêtes dans le cas d'un graphe non orienté ou un ensemble d'arcs pour un graphe orienté. Une arête est une paire de sommets de G . Un arc est une paire orientée de sommets de G , (x,y) représentant un lien du sommet x vers le sommet y . La théorie des graphes associée à ce type de modélisation donne accès à nombre d'opérateurs et de mesures qui permettent de décrire les propriétés du réseau (Freeman, 1979).

I.2.1 Mesure de Densité

Afin de quantifier la « valeur » d'un espace, une approche directe, mais non sans intérêt, consiste à comptabiliser le nombre d'objets présents sur cet espace. On remarque rapidement qu'en fonction de la taille de la zone d'étude, un même nombre d'objets peut refléter une organisation très différente. Cette approche amène à la mesure de densité : le nombre d'objets est divisé par la superficie de la zone prise en compte.

Suivant les objets étudiés et les propriétés spatiales de la zone prise en compte (taille et/ou forme), la densité permet d'obtenir une multitude d'indicateurs globaux et locaux utiles pour comparer différentes zones, ou pour faire un suivi temporel d'une zone. Par exemple dans (Hawbaker, et al., 2006) la densité du réseau routier est mesurée en kilomètres de route par kilomètre carré, sur plusieurs zones d'études et à différentes dates. Dans (Pouyanne, 2004), l'étude s'intéresse à la densité urbaine définie par le rapport entre la taille de la population (c.-à-d. le nombre d'individus) et l'étendue de l'espace qu'elle occupe (habitants/hectare). (Puissant, et al., 2010) définissent une classification des territoires basée sur la densité de bâtiments à l'échelle d'îlots urbains (Figure 3). Enfin, (Long & Kergomard, 2005) proposent une classification à partir de plusieurs mesures de densité surfacique concernant la surface du bâti, de la végétation, des étendues d'eau et des routes.

La mesure de densité permet de décrire la variabilité d'organisation de l'espace à un niveau global. Cependant si l'on souhaite obtenir une mesure à un niveau plus local, il est alors nécessaire

de structurer l'espace en plusieurs sous-espaces, ce qui pose la problématique du choix des paramètres guidant ce découpage. Parmi les différentes solutions proposées par la littérature, (Long & Kergomard, 2005) appliquent un découpage en grille régulière avec des mailles de 50m, 100m et 200m. Cette approche permet d'assurer un cadre efficace pour comparer différentes zones, car leurs dimensions sont les mêmes. L'inconvénient est que ce partitionnement fait abstraction de l'empreinte spatiale des objets qui composent l'espace, et amène à les fractionner. On traite alors des « portions d'objets » ce qui introduit un biais dans l'analyse. Ce constat, identifié par les auteurs, leur a fait écarter la taille de maille de 50m qui découpait trop les bâtiments.

Pour assurer la cohérence entre découpage et empreinte spatiale du bâti, (Hamaina, et al., 2012) proposent de segmenter l'espace à l'aide de diagramme de Voronoï dont les cellules sont centrées sur chaque bâtiment. Pour chaque cellule, les auteurs mesurent une densité surfacique qui correspond à la division de la surface du bâtiment par la surface de la cellule de Voronoï. Cette approche ne prend cependant pas en compte le réseau de circulation. À l'inverse, (Boffet & Serra, 2001; Rousseaux, et al., 2012) proposent de contraindre le découpage de l'espace en utilisant le réseau de circulation. L'espace est alors segmenté en îlots. Cette dernière approche assure le lien entre empreinte spatiale des objets et découpage de l'espace, mais amène à comparer des zones de tailles différentes.

1.2.2 Mesure de distance

Dans le cas générique d'un modèle de graphe, la distance entre deux sommets d'un graphe est donnée par le nombre d'arcs qui les sépare. A contrario, un grand nombre de domaines comme celui de la géographie privilégient l'utilisation de distances métriques. La distance métrique la plus commune est la distance euclidienne (ou distance à vol d'oiseau). Elle correspond à la distance rectiligne entre deux points projetés sur un plan horizontal. Il est possible de définir différemment la distance métrique en prenant en compte le relief et la courbure du planisphère. Il est également possible (et c'est l'usage sur lequel nous nous arrêtons dans notre approche) de définir cette distance comme le parcours du réseau de circulation (route ou chemin) permettant à un individu de rejoindre deux points de l'espace (c.-à-d., distance de parcours). Sauf cas fortuit, cette distance est supérieure à la distance rectiligne, car elle est contrainte par la présence d'obstacles ou de règles de circulation (sens interdit par exemple). Avec la notion de déplacement, la distance peut également être exprimée en temps, référence à la durée de ce déplacement.

Appliquée aux graphes de connectivité et aux réseaux de déplacement, la mesure de distance entre deux sommets est à rapprocher à la mesure du plus court chemin entre deux sommets. Elle fait

l'objet de plusieurs algorithmes d'optimisation de son calcul (Dijkstra, 1959), A* (Hart, et al., 1968), (Floyd, 1962), etc. Soit x et y deux sommets d'un graphe G , $d_{min}(x, y)$ retourne la distance du plus court chemin entre les sommets x et y . Dans le cas où aucun poids n'est appliqué sur les arêtes (arcs), la valeur retournée par $d_{min}(x, y)$ correspond au nombre minimum d'arêtes (arcs) qu'il faut parcourir pour aller du sommet x au sommet y . Dans le cas où c'est la distance métrique qui est prise en compte (distance à vol d'oiseau ou de parcours), une valeur de distance est alors appliquée aux arêtes (arcs), et le plus court chemin entre deux sommets est celui pour lequel la somme des arêtes (arcs) parcourues est la plus petite. Ce plus court chemin ne correspond plus nécessairement au chemin parcourant le plus petit nombre d'arêtes (arcs). Si le graphe n'est pas entièrement connecté, il est possible qu'il n'existe pas de chemin entre le sommet x et le sommet y . Dans ce cas la distance $d_{min}(x, y)$ est considérée comme infinie.

I.2.3 Mesure de centralité

I.2.3.1 Degré

La mesure de graphe la plus simple conceptuellement est le degré (Freeman, 1979). Elle mesure, pour chaque sommet, le nombre d'arêtes qui s'y connectent. Plus un sommet a un degré élevé, plus ce sommet est considéré comme important dans le réseau. Le degré est une valeur locale qui ne dépend pas de la valeur mesurée sur les autres sommets du graphe. Dans le cas d'un graphe orienté, il est possible de distinguer le degré entrant, représentant le nombre d'arcs qui pointent sur le sommet, et le degré sortant représentant le nombre d'arcs qui partent du sommet.

Dans le cas d'un graphe $G(S, A)$ non orienté, pour un sommet $x \in S$, le degré est défini par :

$$Deg(x) = |\{y \in S \mid (x, y) \in A\}|$$

Dans le cas d'un graphe G orienté, pour un sommet $x \in S$, le degré entrant est défini par :

$$Deg^{in}(x) = |\{y \in S \mid (y, x) \in A\}|$$

Et le degré sortant est défini par :

$$Deg^{out}(x) = |\{y \in S \mid (x, y) \in A\}|$$

1.2.3.2 Centralité de proximité (*Closeness centrality*)

L'idée de la centralité de proximité (Bavelas, 1950; Sabidussi, 1966; Freeman, 1979) est qu'un sommet est central s'il est proche de tous les autres sommets, et donc que la somme des distances vers les autres sommets est petite. Pour un sommet x , la centralité de proximité est définie par :

$$C_c(x) = \frac{1}{\sum_{y \neq x} d_{min}(x, y)}$$

Dans le cas où le graphe n'est pas entièrement connecté, si le sommet x ne peut pas atteindre le sommet y , la distance $d_{min}(x, y)$ infinie provoque un résultat de la division toujours égal à 0. Une solution parfois appliquée est de ne pas prendre en compte cette distance infinie.

Dans le cas où le graphe est orienté, il est possible de différencier la centralité de proximité entrante $C_c^{in}(x) = \frac{1}{\sum_{y \neq x} d_{min}(y, x)}$ et sortante $C_c^{out}(x) = \frac{1}{\sum_{y \neq x} d_{min}(x, y)}$

1.2.3.3 Centralité harmonique

La centralité harmonique (Rochat, 2009) est une variante de la centralité de proximité, mais plus performante pour les réseaux qui ne sont pas entièrement connectés. La centralité harmonique d'un sommet x est définie par :

$$H_c(x) = \sum_{y \neq x} \frac{1}{d_{min}(x, y)}$$

Ainsi lorsque la distance $d_{min}(x, y)$ est infinie, la somme totale ne devient pas nulle.

De la même manière que pour la centralité de proximité, lorsque le graphe est orienté il est possible de distinguer la centralité harmonique entrante ($H_c^{in}(x) = \sum_{y \neq x} \frac{1}{d_{min}(y, x)}$) de la centralité harmonique sortante ($H_c^{out}(x) = \sum_{y \neq x} \frac{1}{d_{min}(x, y)}$).

1.2.3.4 Centralité d'intermédiarité (*Betweenness centrality*)

La centralité d'intermédiarité (Freeman, 1979) d'un sommet x est définie par :

$$B_c(x) = \sum_{y, z \neq x, \sigma_{yz} \neq 0} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

où σ_{yz} représente le nombre de plus courts chemins du sommet y au sommet z et $\sigma_{yz}(x)$ le nombre de plus courts chemins du sommet y au sommet z passant par le sommet x . Cette mesure globale évalue l'importance d'un sommet dans l'ensemble des plus courts chemins, l'importance étant d'autant plus grande que le nombre de chemins qui passent par ce sommet est élevé.

1.2.3.5 Centralité spectrale ou Centralité de vecteur propre (*Eigenvector*)

Le principe de la centralité spectrale (Bonacich, 1972) est relativement proche de la mesure du degré. Elle diffère du fait que tous les sommets voisins n'ont pas un poids de 1, mais un poids qui varie en fonction de leur importance. Ainsi un sommet est d'autant plus important qu'il est connecté à des sommets qui sont eux-mêmes importants.

Soit $M = (m_{y,z})$ la matrice d'adjacence du graphe G , $m_{y,z} = 1$ si le sommet y est relié au sommet z et $m_{y,z} = 0$ sinon. Le score de centralité spectrale du sommet y est le nombre x_y vérifiant l'équation suivante :

$$x_y = \frac{1}{\lambda} \sum_{z \in G} m_{y,z} x_z$$

avec λ la plus grande valeur propre de la matrice d'adjacence du graphe.

Lorsque le graphe est orienté, il est possible de différencier la centralité spectrale entrante de la centralité spectrale sortante en utilisant respectivement la matrice d'adjacence et la transposé de la matrice d'adjacence. Ainsi, la littérature décrit de nombreuses variantes basées sur la centralité spectrale, telles que la centralité de Seeley (Seeley, 1949), la centralité de Katz (Katz, 1953), le Pagerank (Page, et al., 1998), l'HITS (*Hyperlink-Induced Topic Search*) (Kleinberg, 1999) ou le SALSA (*Stochastic Approach for Link-Structure Analysis*) (Lempel & Moran, 2001).

1.2.4 Mesure d'interconnectivité

1.2.4.1 Coefficient de *clustering*

Le coefficient de *clustering* (Watts & Strogatz, 1998) est, pour un sommet, une mesure permettant de définir à quel point les voisins de ce sommet sont connectés entre eux. Le coefficient de *clustering* d'un sommet s_i est défini par :

$$C(s_i) = \frac{2l_i}{m_i * (m_i - 1)}$$

où l_i est le nombre d'arêtes qui existent entre les voisins du sommet s_i et m_i le nombre de voisins de s_i . Lorsqu'un sommet a un coefficient de *clustering* de 1, ce sommet et ses voisins forment une clique.

Une généralisation du coefficient de *clustering* (Jiang & Claramunt, 2004) est définie par :

$$C^{(k)}(s_i) = \frac{2l_i^{(k)}}{m_i^{(k)} * (m_i^{(k)} - 1)}$$

où $l_i^{(k)}$ est le nombre d'arêtes qui existent entre les $k^{\text{ième}}$ voisins du sommet s_i et $m_i^{(k)}$ le nombre de $k^{\text{ième}}$ voisin de s_i . Par exemple, pour $k = 2$, cela permet de mesurer à quel point les voisins de voisins du sommet s_i sont interconnectés.

Une variante du coefficient de *clustering* est le coefficient de *clustering* d'arêtes (Radicchi, et al., 2004) et est définie par :

$$C_{ij}^{(3)} = \frac{z_{ij}^{(3)}}{\min[(k_i - 1), (k_j - 1)]}$$

où $z_{ij}^{(3)}$ est le nombre de triangles auxquels appartient l'arête ij , k_i est le degré du sommet i , k_j le degré du sommet j . Ainsi $\min[(k_i - 1), (k_j - 1)]$ représente le nombre maximum de triangles auxquels peut appartenir l'arête ij . L'idée est qu'une arête faisant un lien entre deux communautés appartiendra à peu (ou pas) de triangle. Ce coefficient de *clustering* d'arêtes est aussi généralisé pour fonctionner avec d'autres cycles que des triangles :

$$C_{ij}^{(g)} = \frac{z_{ij}^{(g)}}{s_{ij}^{(g)}}$$

où $z_{ij}^{(g)}$ est le nombre de structures cycliques d'ordre g et $s_{ij}^{(g)}$ représente le nombre maximum de structures cycliques d'ordre g auxquelles peut appartenir l'arête ij .

1.2.5 Extraction de sous-ensembles

1.2.5.1 Utilisation des liens de connectivité

La structuration de l'espace ne peut s'exprimer à partir des seules propriétés individuelles des objets présents dans cet espace, mais également à partir de celles formées par les ensembles d'objets formant des conglomérats à l'intérieur de cette structure. Dans le cadre de représentations

dérivées des graphes de connectivité, les termes *groupes*, *clusters*, *partitions* (MacQueen, 1967) et *communautés* (Flake, et al., 2002) sont utilisés indifféremment pour décrire des groupes de sommets. Ces termes définissent un sous-ensemble de sommets du graphe tel que la connexion entre ces sommets est plus dense qu'avec le reste du réseau (Radicchi, et al., 2004). Ils définissent des sommets fortement connectés entre eux et faiblement connectés aux sommets n'appartenant pas au groupe. Cette notion est à rapprocher de la nomenclature des réseaux de circulation proposée par (Marshall, 2005) avec une différenciation entre « structure fortement connectée » et « structure faiblement connectée » (Figure 7). Cette notion est également à rapprocher de celle des petits mondes (*Small World*) qui, bien que décrit comme étant associée à un réseau faiblement connecté, laisse apparaître des sous-ensembles fortement connectés reliés entre eux par quelques liaisons particulières (Watts & Strogatz, 1998).

Les cliques

Le cas particulier des groupes dont tous les sommets membres sont voisins deux à deux est appelé *clique* (Luce & D., 1949). Les *cliques* les plus simples sont composées de trois sommets. (Palla, et al., 2005) présentent une méthode de percolation de *cliques* pour détecter des communautés. Pour cela, plusieurs termes sont définis :

- Deux *cliques* de k sommets sont dites *adjacentes* si elles partagent $k - 1$ sommets
- L'union de *cliques* adjacentes à k sommets forme des *chaines de cliques* à k sommets.
- Deux *cliques* de k sommets sont *connectées* si elles appartiennent à la même chaîne de *cliques*.

Les *n-cliques*² définies dans (Luce, 1950; Alba, 1973) sont des sous graphes dont les sommets sont, au maximum, éloignés de n arêtes dans le graphe initial. Lorsque $n = 1$, tous les sommets sont éloignés d'une seule arête, ce qui est équivalent à la définition de clique. Par construction, il est possible que des sommets du sous-graphe obtenu soient à une distance supérieure à n des autres sommets du sous-graphe. En effet dans certains cas, certains sommets sont intégrés au sous-graphe sans que le chemin le plus court qui les relie dans le graphe initial ne soit conservé. Afin d'éviter cet inconvénient, (Mokken, 1979) propose les structures de *n-clan*. Les *n-clans* sont des *n-cliques* dont le diamètre est inférieur ou égal à n , ce qui signifie que les sommets sont, au maximum, éloignés de n arêtes dans le *n-clan*.

² Dans certains articles, une 5-clique par exemple ne fait pas référence à cette définition et représente une clique de 5 sommets.

K-plex et k-core

Si l'intérêt de définir des *cliques* est pertinent pour aider à la caractérisation d'un graphe, elles restent des structures rares. Afin de travailler avec des structures moins contraignantes qui n'imposent pas que tous les sommets soient reliés deux à deux, (Seidman, 1983) propose deux autres structures dont les définitions se résument à l'idée que les sommets doivent être (ou peuvent ne pas être) adjacents à un certain nombre des autres sommets. Un *k-core* (Seidman, 1983) est défini comme un sous graphe maximal où chaque sommet est connecté à au moins k autres sommets du *k-core*. À l'inverse, un *k-plex* (Seidman & Foster, 1978) est défini comme un sous graphe maximal où chaque sommet doit, au pire, ne pas être relié avec k sommet du *k-plex*.

La définition des *k-cores* est assez proche de la définition de communauté faible (Radicchi, et al., 2004). Une communauté faible est définie comme un sous graphe où la somme des degrés des sommets vers les sommets de la communauté faible est plus grande que la somme des degrés des sommets vers les sommets n'appartenant pas à la communauté faible. À l'inverse, une communauté forte est définie comme un sous-graphe où chaque sommet a plus de liens vers la communauté forte que vers l'extérieur de la communauté.

1.2.5.2 Approche statistique

Bien que l'utilisation des liens de connectivité soit intuitivement l'approche la plus évidente pour extraire des sous-graphes, l'utilisation d'outils statistiques pour associer des éléments du graphe présentant des propriétés similaires permet également d'extraire des sous-graphes. Les deux approches les plus communes sont le partitionnement en groupe et la classification hiérarchique (MacQueen, 1967; Lesot, 2005).

Partitionnement en groupe

La méthode de partitionnement en groupe nécessite de fixer arbitrairement le nombre de classes d'objets à obtenir, noté k . Le principe est de partir d'un partitionnement aléatoire de K groupes, puis d'améliorer itérativement les groupes.

L'algorithme des *k-moyennes* (*k-means*) (MacQueen, 1967) est la méthode de partitionnement en groupe la plus commune. La première étape consiste à créer k groupes à partir de k sommets sélectionnés aléatoirement (les groupes n'étant composés que d'un sommet au départ). L'étape suivante consiste à sélectionner un sommet qui n'appartient à aucun groupe et à l'ajouter au groupe dont le centroïde est le plus proche de ce sommet. Ce centroïde étant recalculé à chaque fois, l'action est répétée jusqu'à traiter tous les sommets.

Cet algorithme a l'avantage d'être rapide, mais ne fournit pas un résultat stable et n'est donc pas nécessairement optimal. En effet, en fonction des k sommets sélectionnés aléatoirement, et de l'ordre dans lequel les autres sommets sont ajoutés aux groupes, le résultat peut être différent. Une variante (plus couteuse) consiste dans un premier temps à ajouter tous les sommets aux différents groupes sans modifier la valeur du centroïde à chaque ajout, et à en extraire cette valeur seulement à la fin. Cette opération est reproduite plusieurs fois, avec des groupes différents jusqu'à obtenir des centroïdes stables.

Classification hiérarchique

Dans le cas de la classification hiérarchique, le regroupement de sommets est basé sur une mesure de similarité (ou de distance) entre chaque paire de sommets. Une multitude de mesures de similarité existe dans la littérature à mettre en rapport avec la multitude de méthodes et de domaines où une mesure de similarité intervient (Lesot, 2005; Pandit & Gupta, 2011).

Soit $X = (x_1, x_2, \dots, x_n)$ et $Y = (y_1, y_2, \dots, y_n)$ deux éléments ayant chacun n attributs³.

- La distance de Minkowski (ou norme p) est définie par :

$$d_p^{XY} = \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p}$$

Les valeurs de p les plus utilisées sont $p = 1$ (norme 1 ou distance de Manhattan), $p = 2$ (norme 2 ou distance Euclidienne) et $p = \infty$ (norme infinie ou distance de Tchebychev).

Dans le domaine de l'information géographique, la similarité peut être à la fois attributaire et spatiale. En effet, il est possible d'utiliser les coordonnées spatiales des éléments x et y pour calculer les distances métriques afin de regrouper des objets proches (dans le réseau ou dans l'espace).

La classification hiérarchique consiste à construire une hiérarchie à partir de la similarité des sommets (Steinbach, et al., 2000). Elle se présente sous forme de dendrogramme (arbre binaire). Il est possible de distinguer l'approche descendante de l'approche ascendante. Cette dernière considère initialement que chaque sommet est un groupe de taille 1, puis regroupe itérativement les deux groupes les plus proches jusqu'à obtenir un groupe contenant tous les sommets. À

³ On notera qu'avant de calculer de la distance, il est préférable de normaliser les valeurs des attributs sur un intervalle commun pour ne pas favoriser les attributs ayant les plus grandes valeurs absolues.

l'inverse, l'approche descendante consiste à partir de l'ensemble des sommets dans un même groupe, de le diviser en deux groupes les plus distants, puis de diviser un des groupes en deux sous-groupes jusqu'à aboutir à une situation où chaque sommet est dans un groupe distinct. Pour des groupes d'objets (au moins deux éléments), il est nécessaire de préciser comment est définie la distance de similarité : soit vers l'élément le plus proche, soit vers l'élément le plus éloigné, soit vers une valeur (une position) moyenne.

La classification hiérarchique présente l'avantage de ne pas avoir à définir de façon arbitraire ni le nombre de sous-ensembles ni la taille des sous-ensembles. Par contre, elle présente une complexité de calcul assez importante.

La caractérisation de la structure spatiale d'un environnement urbain peut s'appuyer sur des mesures qui quantifient la répartition des objets dans cet espace, et les liens de connectivité entre ces objets. Particulièrement adaptées à l'utilisation de modèles de graphe, ces mesures évaluent les positions relatives des sommets dans le graphe, la centralité et l'interconnectivité de ces sommets, et dérivent des sous-parties du graphe en utilisant les liens de connectivité ou des outils statistiques.

I.3 Modèles d'analyse d'espaces habités

I.3.1 Approche primaire ou duale

L'utilisation d'un graphe donne une représentation efficace et pratique facilitant l'analyse de la structure d'un espace habité composé de bâtis, de parcelles et d'un réseau de circulation. Deux approches sont utilisées pour définir un graphe, impliquant respectivement soit les relations spatiales (qualitatives, quantitatives, ou topologiques), soit le réseau de circulation lui-même. Dans le premier cas, deux objets sont liés dans le graphe s'ils sont associés à une relation spatiale, alors que dans le deuxième cas, deux objets sont liés dans le graphe s'ils sont connectés directement ou indirectement via le réseau de circulation.

Il est fréquent de préférer l'usage des relations spatiales en contexte cadastral (et/ou bâti). Dans ce cas, les relations de voisinage entre différentes parcelles sont représentées par des arêtes, et les parcelles par des nœuds (Le Couédic, et al., 2013). Concernant l'utilisation du réseau de circulation, deux approches sont utilisées. Dans la première approche, chaque intersection est représentée par un sommet du graphe et chaque route par une arête. À l'inverse, la seconde

approche propose que chaque route soit représentée par un sommet du graphe et chaque intersection par une arête. (Porta, et al., 2006) définissent la première approche comme *primaire* (on parle également d'approche directe) et la seconde comme *duale*. Ces appellations sous-entendent une hiérarchie entre les deux qui n'a pas lieu d'être comme le souligne (Marshall, 2016). Ce dernier, propose une nomenclature plus neutre avec les appellations G' et G'' pour distinguer les deux approches.

L'intérêt principal de l'approche G' est son lien intuitif avec la réalité terrain. Elle assure la manipulation des distances métriques (ou temporelles) mesurées sur le réseau de circulation, paramètre clé pour nombre d'études (suivi de mobilité, gestion de trafic, planification de transport...). C'est actuellement un standard en matière de construction et de diffusion de bases de données géographiques (p. ex. BD-CARTO (bd-carto-IGN, 2015) ou TIGER (TIGERweb, 2010) (Topologically Integrated Geographic Encoding and Referencing)). Cette approche centrée sur les intersections génère cependant un découpage des routes en tronçons, entraînant la perte de continuité du réseau au-delà des intersections. À l'inverse, l'approche G'' centrée sur la présence de voies de circulation (routes, chemins et équivalents), intègre naturellement la voie dans sa globalité (une route nommée ou un alignement dans le réseau). Ce dernier choix permet ainsi d'éviter de s'interroger sur la continuité des tronçons de route. Cependant, dans ce dernier cas il n'est pas possible de localiser la position des intersections le long d'une voie. De même, comme le soulignent (Porta, et al., 2006), plus la voie est longue plus elle a de chance de croiser d'autres voies et apparaitre comme centrale (fortement connectée) dans le graphe, alors que dans des études sur l'accessibilité du réseau c'est souvent les distances parcourues qui sont intéressantes plus que la connectivité.

I.3.2 Syntaxe Spatiale

Une communauté scientifique s'est progressivement constituée depuis une trentaine d'années autour de la syntaxe spatiale et les travaux fondateurs de Hillier et Hanson (1984) avec de nombreuses applications et expérimentations menées en milieu urbain. L'objectif principal de la syntaxe spatiale consiste à explorer la relation entre les configurations spatiales et les processus sociaux et économiques d'un environnement urbain.

Une telle approche est basée sur une idée clé : la structuration spatiale doit être confrontée et mise en perspective au regard des activités des individus qui s'y déplacent. En découle une segmentation du réseau reposant sur différentes définitions de la *continuité* basées sur la prise en compte du déplacement des individus.

I.3.2.1 Cartes Axiales

L'approche de segmentation la plus commune en syntaxe spatiale repose sur l'idée que les humains se déplacent en ligne droite dans des espaces convexes guidés par la profondeur de leurs champs de vision (Turner & Penn, 1999)).

Cette prise en compte des contraintes visuelles de mobilité repose principalement sur l'évolution du champ visuel lors du déplacement. Cette analyse est basée sur la construction de cartes axiales⁴ telle qu'illustrée sur la Figure 8. Le principe de construction de ces cartes consiste à garder le plus petit ensemble des lignes axiales les plus longues couvrant l'espace entre les bâtiments.



Figure 8 : Empreinte du bâti et Carte axiale résultante

La syntaxe spatiale adopte les outils et méthodes de la théorie des graphes pour modéliser l'espace à partir de la carte axiale. Chaque ligne axiale est représentée par un sommet dans le graphe, et chaque intersection entre deux lignes axiales est représentée par une arête⁵.

Les différentes mesures de la syntaxe spatiale peuvent être classées en trois catégories en fonction de leurs portées : locales, globales ou contextuelles (Béra, 2004). Les mesures locales correspondent à des mesures prenant en compte le voisinage direct d'un sommet. À l'inverse, les

⁴ Des cartes de 'lignes de vues', une ligne axiale représentant une visibilité continue d'un bout à l'autre de la ligne.

⁵ Bien que les arêtes ne représentent pas les voies de circulation proprement dites, on se rapproche ici d'un modèle de type G'' (Dual)

mesures globales sont basées sur l'ensemble des sommets. Les mesures contextuelles visent à tenir compte de la structure du graphe de manière plus fine en considérant le nombre de possibilités pour aller d'un nœud à un autre.

La syntaxe spatiale utilise plusieurs mesures provenant de la théorie des graphes que nous avons présentées dans la section I.1.2. Les mesures locales reprennent les notions d'interconnectivité (p. ex. *le clustering coefficient*). Les mesures globales présentent les variations de la moyenne des plus courts chemins vers tous les autres nœuds du graphe. On retrouve ici et en particulier la mesure de *closeness centrality*. Une des mesures principales de la syntaxe spatiale est la mesure d'intégration globale (ou d'accessibilité) illustrée par la Figure 9. L'intégration cherche à illustrer le fait que des rues fortement intégrées attirent généralement plus de monde que les rues peu intégrées (Hillier, et al., 1993). Elle est définie pour chaque ligne axiale par :

$$\frac{2(md - 1)}{n - 2}$$

Où md est la profondeur moyenne vers les autres lignes axiales et n est le nombre de lignes axiales. La profondeur moyenne md d'une ligne axiale x est définie par :

$$md = \frac{\sum_{y=1}^n d_{min}(x,y)}{n-1}$$

avec $d_{min}(x, y)$ la distance topologique entre les lignes x et y .

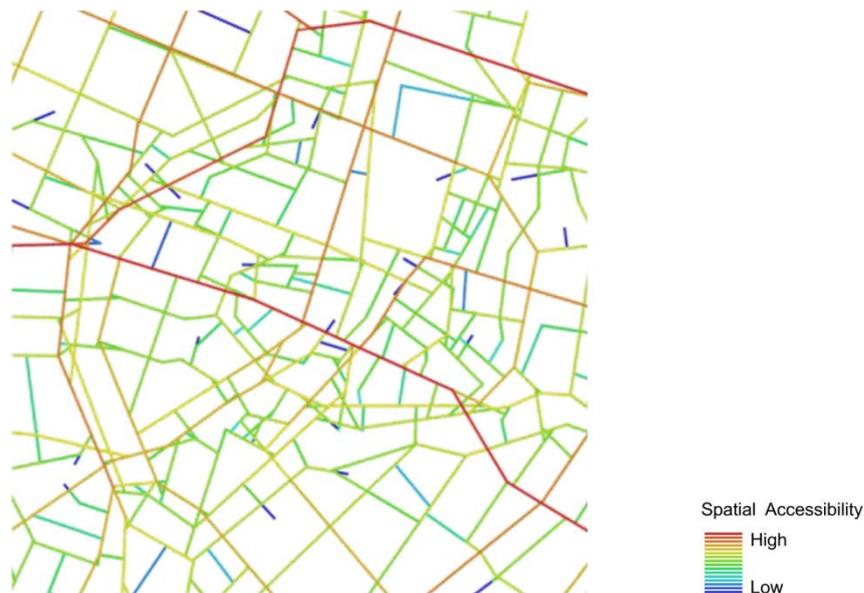


Figure 9 : Mesure d'accessibilité obtenue à partir de la carte axiale de la Figure 8

Les mesures contextuelles « considèrent non seulement les plus courts chemins, mais aussi le nombre de possibilités, c'est-à-dire les différents chemins reliant un nœud à l'autre » (Bera,

2004). C'est le cas par exemple de la *betweenness centrality* (nommée *choice* dans la communauté de la syntaxe spatiale).

En faisant reposer la problématique sur l'importance de prendre en compte l'intégration des lignes axiales dans les études sociales, économiques, et environnementales de l'espace urbain, la syntaxe spatiale s'est avérée être une approche performante dans de nombreux domaines : la criminalité (Jones & Fanek, 1997), la recherche d'itinéraire (*way-finding*) (Peponis, et al., 1990) et la pollution (Penn & Croxford, 1997). Cependant, malgré ce large succès, les critiques les plus souvent évoquées regrettent d'une part la construction d'un graphe dual et la perte d'information sur les distances métriques qui en découle (Porta, et al., 2006), et d'autre part la simplification excessive de la structure spatiale (Ratti, 2004) et en particulier l'absence de prise en compte du parcellaire et de l'organisation du bâti (Buhl, et al., 2006). Aux vues de ces limitations dans la construction des cartes axiales, d'autres approches ont été proposées dans la littérature pour segmenter le réseau routier.

1.3.2.2 Autres approches de segmentation

(Jiang & Claramunt, 2004) proposent par exemple de segmenter un district de la ville de Gävle en Suède en utilisant les noms des rues comme méthode de segmentation. Ce découpage présente l'avantage d'être très performant d'un point de vue sémantique. Ainsi « l'Avenue des Champs-Élysées » par exemple aura bien plus de sens qu'une ligne axiale. Cependant, l'inconvénient de cette méthode est qu'elle introduit une composante nominative qui n'a pas nécessairement de cohérence spatiale :

- comment se décide du nom des rues ?
- à quelle échelle le nom d'une rue est-il significatif ?

(Porta, et al., 2006) proposent un modèle purement spatial qui permet de construire le graphe des rues, nommé *Intersection Continuity Negotiation* (ICN). En partant du graphe primaire, ce modèle permet d'estimer la continuité des tronçons de route à chaque intersection en classant les différentes paires de tronçons de route par mesure d'angle. Plus l'angle obtenu sera proche de 180° plus ils seront alignés et donc susceptible de ne former qu'une seule route. Une fois que tous les tronçons de route sont regroupés, il est possible d'obtenir un graphe dual pour appliquer les mesures de syntaxe spatiale. Après avoir comparé trois méthodes d'association de tronçons de route, (Lagesse, 2015) retient un modèle très similaire incluant un angle seuil de 60 degrés en dessous duquel les tronçons de route ne sont pas associés.

(Marshall, 2015) propose un modèle de continuité de routes basé sur la 'classification des routes' et sur l'alignement des routes. Dans ce cas, les tronçons de route possèdent un 'attribut hiérarchique' (A, B, C ou non classé selon une nomenclature appliquée en Grande-Bretagne) empêchant une route de rang inférieur de continuer au travers une route de rang supérieur. De plus, cet attribut est prédominant sur l'alignement (c.-à-d. une route de rang A continue le long d'un autre tronçon de rang A même si une route de rang B a un meilleur alignement). La question qui se pose (de la même manière que pour les noms de routes précédemment) est de savoir si la classification des routes effectuée par la Grande-Bretagne respecte la notion de continuité.

I.3.3 Modélisation basée sur l'approche directe

L'approche directe qui modélise les intersections par les nœuds et les voies de circulation par les arcs paraît la plus naturelle dans le cas d'applications géographiques dans la mesure où elle préserve la métrique sans abandonner la topologie du système. (Porta, et al., 2006) proposent une analyse approfondie des résultats de mesures de centralité basées sur des distances métriques. Appliqué sur l'ensemble du graphe (Figure 10a), le résultat de la mesure de *centralité de proximité* est dominé par un effet de bord illustrant la concentration des valeurs les plus fortes au centre de l'image. Par contre, les approches locales permettent de distinguer la présence de zones à fortes valeurs de centralité (Figure 10b et c).

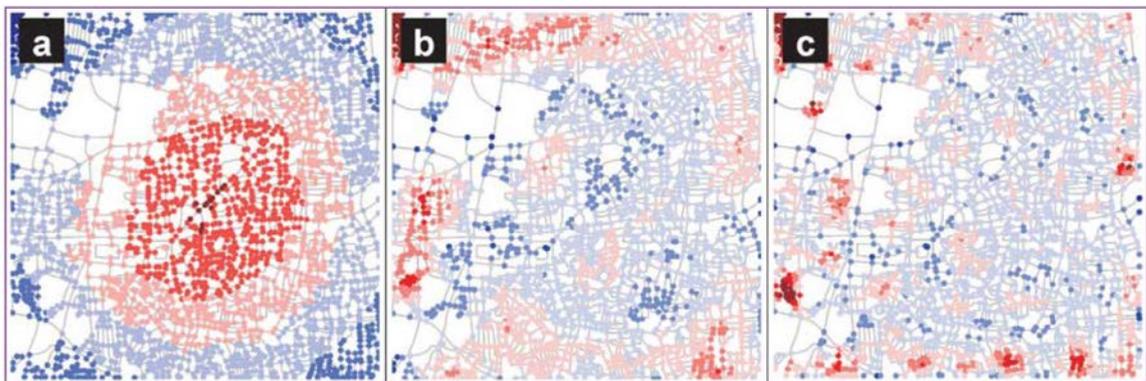


Figure 10 : Mesure de centralité de la ville d'Ahmedabad sur un graphe primaire ; a) centralité globale mesurée sur tout le graphe b) centralité locale avec une distance entre les nœuds inférieure à 400 mètres c) centralité locale avec une distance entre les nœuds inférieure à 200 mètres (Porta, et al., 2006)

Les auteurs montrent que les mesures effectuées avec un modèle direct sur la globalité du graphe sont peu adaptées pour discriminer des propriétés structurelles d'un schéma d'urbanisation. Ils précisent qu'une approche duale a, dans ce cas, plus d'intérêt, mais dépend d'une étape de généralisation du réseau qui modifie le nombre de voies de circulation pris en compte. Afin de

réduire la subjectivité induite par cette généralisation, les auteurs préconisent donc l'usage des distances métriques et le recours à plusieurs indices de centralité.

I.3.4 Modèle basé sur le bâti

Pour mettre en évidence les ruptures dans la proximité et la connectivité de la structure spatiale, une autre approche de modélisation consisterait à se focaliser sur le bâti. Dans ce cas, la problématique devrait s'intéresser aux relations entre les bâtiments, mais également à leurs propriétés géométriques (taille, forme) et à leur localisation. Si dans des contextes très urbanisés les bâtiments sont le plus souvent contigus et regroupés sous forme d'ilots fermés, dans la majorité des cas il est préférable de remplacer cette stricte contiguïté par une distance seuil de voisinage. Dans ce cas, et en première approximation, on peut ainsi considérer que deux bâtiments sont voisins si la distance euclidienne minimale entre eux est inférieure à une distance seuil. Cette distance peut être mesurée entre les frontières de ces bâtiments, ou de leurs centroïdes. Comme précisé en I.2.2, il est également possible de définir cette distance comme le parcours du réseau de circulation qui sépare les deux bâtiments.

À partir de ce constat, (Badariotti, et al., 2007b) décrivent un modèle qui représente, sous la forme d'un graphe, les relations de voisinage au travers des réseaux de circulation (Modèle REMUS – *REticular Model for Urban Simulation*). Ce graphe est construit en deux étapes :

- La première étape est la construction d'un graphe urbain de type primaire : des carrefours sont créés à chaque intersection de tronçons de route ce qui permet de représenter chaque carrefour par un sommet du graphe et chaque tronçon par un arc. Ce graphe urbain est complété par l'ajout des bâtiments. Chaque bâtiment est rattaché au tronçon de route le plus proche de son centroïde (la notion de proximité est définie par une distance seuil qui peut être calibrée). Pour ne pas ajouter de sommets dans le graphe, les bâtiments sont directement connectés aux carrefours identifiés précédemment (Figure 11)
- La seconde étape est le calcul des plus courts chemins entre tous les bâtiments. Chaque bâtiment étant lié à un carrefour du réseau de circulation, les distances affectées aux bâtiments sont celles qui relient les carrefours entre eux le long du réseau de circulation. Pour une distance seuil à calibrer, le résultat est un graphe de voisinage qui relie les bâtiments dont l'éloignement sur le réseau est inférieur ou égal à cette distance seuil.

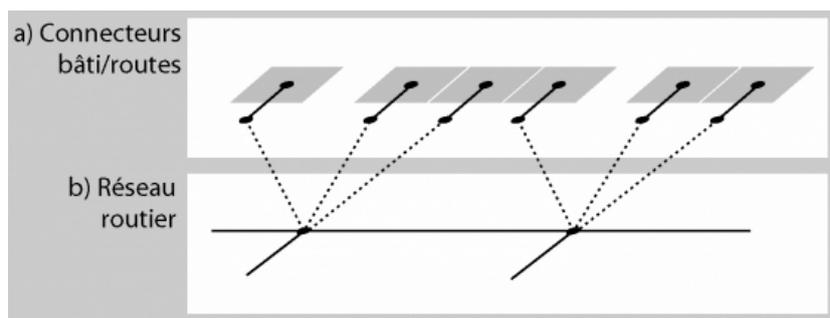


Figure 11 : Schéma de structure du graphe urbain du modèle REMUS. Modifié de (Badariotti, et al., 2007)

Cette approche présente clairement l'intérêt d'inclure la composante « bâti » dans la réflexion sur la structure des espaces habités. L'intérêt de combiner le bâti avec le réseau de circulation permet à (Badariotti, et al., 2007a) de mesurer la tendance de l'habitat à se resserrer autour de pôles urbanisés (illustrée par une mesure du nombre moyen de maisons voisines en fonction d'une distance de seuil de voisinage variable). Ainsi, pour l'agglomération de Pau-Pyrénées, les auteurs déterminent les seuils de voisinage représentatif des centres de villes et villages (entre 5 et 20 m) et des zones périurbaines (> 40 m). Avec pour objectif la construction de graphe de voisinage, le réseau routier est utilisé uniquement pour étudier les continuités ou discontinuités des structures du bâti. En résulte la construction d'un graphe non connexe focalisé uniquement sur la structuration du bâti. De plus, cette approche de modélisation ne permet pas de prendre en compte le cas des bâtiments connectés à plusieurs tronçons de route.

I.4 Synthèse

Les divers aspects concernant la représentation, la modélisation et la mesure de la structure spatiale des espaces habités que nous avons décrits offrent une base conceptuelle et des outils de mesure et d'analyse importants pour modéliser et caractériser un espace géographique bâti. Basé sur la nécessité de caractériser les relations spatiales entre des éléments remarquables de type bâtiments, parcelles et réseaux de circulation, l'usage de modèles et d'opérateurs de graphe s'avère une approche pertinente. La combinaison des différents paramètres qui conditionnent notre réflexion est rappelée dans le Tableau 4 et permet d'évaluer la réponse des trois principales approches de modélisations déjà proposées. Dans le domaine de l'analyse de structures urbaines, la syntaxe spatiale s'est très largement imposée auprès de la communauté d'aménagement du territoire et d'urbanisme, pour traiter des environnements urbains à forte densité. Elle s'appuie sur de nombreux opérateurs de graphe, et permet essentiellement la prise en compte de la continuité dans les déplacements. La seconde approche de modélisation, dite primaire, s'attache à mieux

prendre en compte la métrique des déplacements sans introduire la composante bâtie et parcellaire de l'espace. Enfin, la dernière approche proposée se focalise sur les relations entre les bâtis en introduisant la notion de voisinage réseautique permettant d'étudier les continuités ou discontinuités des structures bâties.

En fait, comme l'illustre le Tableau 4, les solutions proposées n'offrent qu'une prise en compte partielle des besoins que nous souhaitons traiter. La modélisation à base de graphe est une composante essentielle des démarches de représentation des espaces habités avec deux approches respectivement primaire et duale. La syntaxe spatiale amène préférentiellement à proposer des modèles duals peu adaptés à la prise en compte des mesures métriques de distances de déplacement. Centrée sur l'utilisation d'axes de déplacement ou de visibilité, elle n'intègre ni le bâti ni le parcellaire dans le modèle. En ce qui concerne la modélisation directe, l'utilisation de graphes primaires amène à un usage plus intuitif des distances métriques pour évaluer les liens de connectivité dans les réseaux géographiques. Cependant, elle n'intègre ni le bâti ni le parcellaire et ne permet pas de modéliser le concept de continuité de route en se limitant à l'usage de tronçons de route. L'introduction du bâti dans le modèle des voisinages réseaux élargit la capacité d'analyse et de différenciation des typologies de territoires habités. Elle étend la notion de voisinage avec le calcul de distance-temps entre bâtiments par le réseau et ouvre la voie à l'identification de ruptures et de discontinuités morphologiques. Cependant dans ce dernier modèle, le réseau ne joue plus qu'un rôle de médiateur dans la mesure des distances-temps entre bâtiments, et les bâtiments sont rattachés uniquement à une seule route : la plus proche.

Ainsi, même si ces trois approches ont montré leur validité et leur intérêt dans la littérature, il est nécessaire d'étendre ces travaux pour offrir une meilleure prise en compte des spécificités de la structure spatiale des espaces habités.

	Syntaxe spatiale (Hillier, et al., 1993)	Modélisation Directe (Porta, et al., 2006)	Modèle des voisinages réseaux (Badariotti, et al., 2007a)
Approche duale	Oui	Non	Non
Utilisation de relations topologiques	Oui	Oui	Non
Préservation des distances métriques	Non	Oui	Oui
Manipulation du bâti	Non	Non	Oui
Manipulation du parcellaire	Non	Non	Non
Manipulation de tronçons de route	Non	Oui	Oui
Intégration du bâti, du parcellaire et du réseau routier	Non	Non	Non
Manipulation avec des opérateurs de graphes	Oui	Oui	Non
Application aux espaces habités urbains, périurbains et ruraux	Non	Oui	Oui

Tableau 4 : Comparatif des approches de modélisation pour leur application aux espaces habités

La recherche que nous proposons de développer repose sur la mise en place d'un modèle de graphe et de mesures appliquées à un environnement spatial fortement conditionné par l'activité humaine, et en particulier les déplacements, mais pas limité au contexte urbain et pouvant donc être appliqué au contexte rural. L'élément de référence est le tronçon de route (ou autres voies de circulation) qui doit donc apparaître dans le modèle comme le nœud du graphe et non pas simplement comme un lien de connectivité. Cette idée sous-entend une approche de modélisation duale plutôt que directe. L'approche regroupe dans un graphe les parcelles, le bâti et le réseau pour ensuite définir et appliquer des opérateurs. Ceux-ci ont pour objectif l'analyse de la relation entre la distribution spatiale de ces éléments, afin de faire ressortir plusieurs propriétés permettant de distinguer des configurations spatiales différentes.

Chapitre II

Modèle de graphe

II.1 Espace habité	42
II.1.1 Définition.....	42
II.1.2 Hypothèses et propriétés générales.....	43
II.1.3 Propriétés topologiques.....	43
II.2 Modèle de graphe	45
II.2.1 Associations spatiales	46
II.2.2 Graphes élémentaires.....	46
II.2.3 Graphes intermédiaires	48
II.2.4 Graphe intégral	48
II.2.5 Traitement d'un exemple à travers le modèle de graphe proposé.....	49
II.3 Opérateurs.....	50
II.3.1 Mesure de centralité du bâti.....	50
II.3.2 Mesure de connectivité locale du bâti.....	54
II.3.3 Mesure de densité de bâti.....	55
II.3.4 Mesure de compacité du réseau de circulation	56
II.3.5 Impact de la mesure de compacité sur l'analyse de structuration de l'espace	58
II.4 Synthèse.....	61

Le chapitre précédent a montré comment l'usage de graphes pour modéliser et qualifier la structure spatiale d'espaces habités s'avère une approche pertinente. La structuration de ces espaces repose sur la caractérisation de relations spatiales entre des éléments remarquables à savoir, le bâti, le parcellaire et le réseau de circulation. Les modèles existants ne combinant pas

ces éléments, nous proposons un modèle de graphe qui les intègre dans une approche de modélisation intégrée. Dans un premier temps, nous décrivons les différentes propriétés des espaces habités que nous considérons. Dans un second temps, nous introduisons le modèle proposé. Enfin, la dernière partie du chapitre présente des opérateurs basés sur notre modèle de graphe qui permettent de quantifier les propriétés de l'espace.

II.1 Espace habité

II.1.1 Définition

Un espace habité (illustré par la Figure 12) est formellement représenté par un espace E défini dans \mathbb{R}^2 dans lequel sont localisées des régions. L'espace E considéré est défini dans un espace topologique modélisé par un ensemble de points (point set) (Kinsey, 2012). Nous considérons trois types de régions correspondant respectivement au bâti (bâtiments), au parcellaire (parcelles) et au réseau de circulation (routes), modélisés ainsi :

- un ensemble fini B de bâtiments représentés par des régions b_i de E avec $1 \leq i \leq n$ | $B = \{b_1, \dots, b_n\}$ où $n \in \mathbb{N}$
- un ensemble fini P de parcelles représentées par des régions p_j de E avec $1 \leq j \leq m$ | $P = \{p_1, \dots, p_m\}$ où $m \in \mathbb{N}$
- un ensemble fini R de routes représentées par des régions r_k de E avec $1 \leq k \leq o$ | $R = \{r_1, \dots, r_o\}$ où $o \in \mathbb{N}$

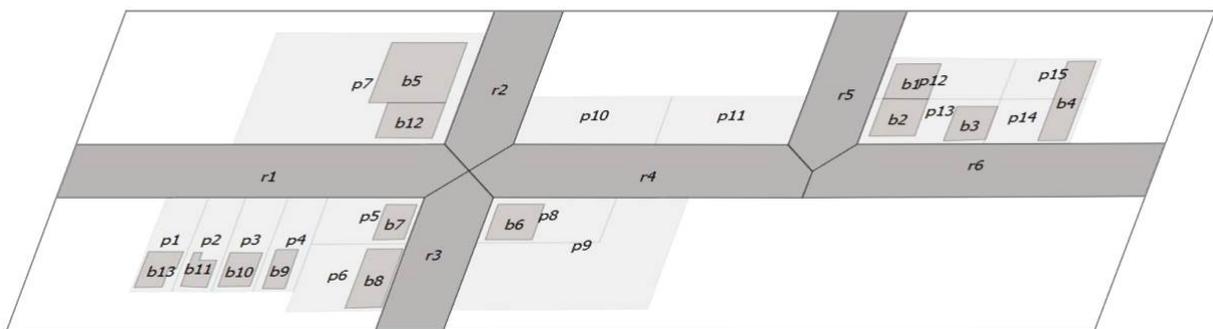


Figure 12 : Représentation schématique d'une partie d'espace habité. Il est formé par la combinaison de bâtiments (b), de parcelles (p) et de routes (r).

II.1.2 Hypothèses et propriétés générales

II.1.2.1 Entre les régions parcelles et routes

La première propriété remarquable que nous assumons par approximation (pour simplifier les cas où nous aurions des rivières par exemple) est que l'ensemble des parcelles P et l'ensemble des routes R forment une partition de E , c'est à dire que les ensembles des parcelles P et R sont spatialement disjoints ($P \cap R = \emptyset$) et que leur union couvre l'ensemble des régions de l'espace E ($P \cup R = E$).

II.1.2.2 Entre les régions bâties et parcelles

Dans un tel espace E , la seconde propriété remarquable est que les bâtiments sont toujours situés sur une ou plusieurs parcelles. L'union des régions bâties de B est donc spatialement incluse dans l'union des régions parcelles de P d'où $B \subseteq P$.

II.1.3 Propriétés topologiques

Afin de ne pas introduire dans le modèle des seuils de distances entraînant des approximations dans les notions de voisinage et de proximité, nous avons choisi de baser les relations entre les éléments de l'espace sur les relations qualitatives. Tous les objets que nous considérons sont de type région et les raisonnements sont basés sur des relations du RCC8 (Randell & Cohn, 1989) (Tableau 5). Ce formalisme est en effet suffisant pour traiter l'ensemble des différents cas possibles entre deux objets de type région.

Soit x, y deux entités de type région :

x DC y	x déconnecté de y
x EC y	x connecté par frontière à y
x TPP y	x contenu dans y et ayant une frontière tangente à y
x NTPP y	x contenu dans y et n'ayant pas de frontière tangente à y
x PO y	x chevauche y
x EQ y	x égal y
x TPP _{i} y	x TPP ⁻¹ y : y contenu dans x et ayant une frontière tangente à x
x NTPP _{i} y	x NTPP ⁻¹ y : y contenu dans x et n'ayant pas de frontière tangente à x

Tableau 5 : Extrait des relations du RCC8 (Randell & Cohn, 1989)

Le graphe spatial que nous souhaitons construire repose sur l'exploitation d'opérateurs permettant d'établir le type de relation entre les régions. La formalisation de ces opérateurs repose sur l'utilisation d'opérateurs génériques permettant de préciser si les régions sont adjacentes, s'intersectent ou sont incluses l'une dans l'autre. L'opérateur d'union est défini comme un opérateur ensembliste sur un espace E représenté par un ensemble de points.

Plus formellement, soit x un point de notre espace, A et B deux ensembles de points de notre espace, chacun définissant une région de l'espace :

–**Union** : $A \cup B = \{x \mid x \in A \vee x \in B\}$

Les opérateurs d'inclusion et d'intersection sont définis comme des opérateurs booléens résultant d'une généralisation des relations du RCC8 :

–**Inclusion** : $A \subset B$ est vrai ssi $A \text{ EQ } B \vee A \text{ TPP } B \vee A \text{ NTPP } B$

–**Intersection** : $A \cap B$ est vrai ssi $\text{PO } B \vee A \text{ EQ } B \vee A \text{ TPP } B \vee A \text{ NTPP } B \vee A \text{ TPP}_i B \vee A \text{ NTPP}_i B$. Ainsi $A \cap B$ est faux si $A \text{ DC } B \vee A \text{ EC } B$

II.1.3.1 Entre deux bâtiments, deux parcelles ou deux routes

Ces propriétés sont immédiatement dérivées de II.1.2 :

- **Propriété 1** : quels que soient deux bâtiments b_i et b_j de B alors, soit b_i est déconnecté de b_j , soit b_i est connecté par frontière à b_j

$\forall b_i, b_j \in B \Rightarrow b_i \text{ DC } b_j \vee b_i \text{ EC } b_j$ avec $1 \leq i \leq n$ et $1 \leq j \leq n$ et $i \neq j$ et $n \in \mathbb{N}$

- **Propriété 2** : quelles que soient deux parcelles p_i et p_j de P alors, soit p_i est déconnectée de p_j , soit p_i est connectée par frontière à p_j

$\forall p_i, p_j \in P \Rightarrow p_i \text{ DC } p_j \vee p_i \text{ EC } p_j$ avec $1 \leq i \leq m$ et $1 \leq j \leq m$ et $i \neq j$ et $m \in \mathbb{N}$

- **Propriété 3** : quelles que soient deux routes r_i et r_j de R alors, soit r_i est déconnectée de r_j , soit r_i est connectée par frontière à r_j

$\forall r_i, r_j \in R \Rightarrow r_i \text{ DC } r_j \vee r_i \text{ EC } r_j$ avec $1 \leq i \leq o$ et $1 \leq j \leq o$ et $i \neq j$ et $o \in \mathbb{N}$

Nous ne considérons pas le cas d'un réseau déconnecté, par exemple une île séparée du continent ce qui nous amène à définir la propriété suivante :

- **Propriété 4** : l'ensemble des routes forme une région connectée : à partir de toute route, il est possible de rejoindre n'importe quelle autre route en ne passant que par des routes.

$$\forall r_i, r_j \in R, \exists (r_k \cup \dots \cup r_m) \in R \mid (r_i \cup (r_k \cup \dots \cup r_m)) \text{ EC } r_j$$

II.1.3.2 Entre un bâtiment et une ou plusieurs parcelles

Cette propriété est immédiatement dérivée de II.1.2.1 :

- **Propriété 5** : tout bâtiment est inclus dans une parcelle ou l'union de plusieurs parcelles :

$$\forall b_i \in B, \exists p_j, \dots, p_o \in P \mid b_i \subseteq (p_j \cup \dots \cup p_o) \text{ avec } 1 \leq i \leq n \text{ et } 1 \leq j \leq o \text{ et } n, o \in \mathbb{N}$$

II.1.3.3 Entre parcelles et routes

- **Propriété 6** : quelles que soient une parcelle p_i de P et une route r_j de R alors soit p_i est déconnectée de r_j soit p_i est connectée par frontière à r_j .

$$\forall p_i \in P, r_j \in R \Rightarrow p_i \text{ DC } r_j \vee p_i \text{ EC } r_j \text{ avec } 1 \leq i \leq m \text{ et } 1 \leq j \leq o \text{ et } m, o \in \mathbb{N}$$

II.2 Modèle de graphe

Après avoir défini les propriétés qui permettent de décrire notre espace de référence, il reste à établir notre modèle en intégrant ces propriétés dans un concept global de graphe spatial.

Un graphe G est un couple (S, A) tel que S est un ensemble non vide de sommets et A un ensemble d'arêtes représentant une relation r entre ces sommets. Une arête de G est une paire de sommets de G .

Nous définissons un modèle de graphe composé de trois graphes élémentaires représentant respectivement l'ensemble des bâtiments, l'ensemble des parcelles et l'ensemble des routes. Les bâtiments, les parcelles et les tronçons de route sont représentés par des sommets et les arêtes représentent les relations topologiques de connexion par frontière entre ces objets. Dans le cadre de la mise en place des graphes, les relations entre les nœuds du graphe sont basées sur l'utilisation de fonctions d'association (définies ci-dessous et nommées : $\text{Asso2B}(b_i, b_j)$),

$\text{Asso2P}(p_i, p_j)$, $\text{Asso2R}(r_i, r_j)$, $\text{AssoBP}(b_i, p_j)$, $\text{AssoPR}(p_i, r_j)$) qui permettent d'identifier si deux régions sont liées.

II.2.1 Associations spatiales

À partir de la définition des propriétés 1 à 6, nous pouvons définir les associations élémentaires suivantes :

- **Asso2B**(b_i, b_j) : Un bâtiment b_i est associé à un bâtiment b_j si et seulement si b_i et b_j sont adjacents.
 - $\text{Asso2B} : B \times B \rightarrow \text{Booléen}$
 - $\text{Asso2B}(b_i, b_j) = \text{vrai}$ ssi $b_i \text{ EC } b_j$
- **Asso2P**(p_i, p_j) : Une parcelle p_i est associée à une parcelle p_j si et seulement si p_i et p_j sont adjacentes.
 - $\text{Asso2P} : P \times P \rightarrow \text{Booléen}$
 - $\text{Asso2P}(p_i, p_j) = \text{vrai}$ ssi $p_i \text{ EC } p_j$
- **Asso2R**(r_i, r_j) : Une route r_i est associée à une route r_j si et seulement si r_i et r_j sont adjacentes.
 - $\text{Asso2R} : R \times R \rightarrow \text{Booléen}$
 - $\text{Asso2R}(r_i, r_j) = \text{vrai}$ ssi $r_i \text{ EC } r_j$
- **AssoBP**(b_i, p_j) : Un bâtiment b_i est associé à une parcelle p_j si et seulement si leur intersection est non vide.
 - $\text{AssoBP} : B \times P \rightarrow \text{Booléen}$
 - $\text{AssoBP}(b_i, p_j) = \text{vrai}$ ssi $b_i \cap p_j \neq \emptyset$
- **AssoPR**(p_i, r_j) : Une parcelle p_i est associée à une route r_j si et seulement si p_i et r_j sont adjacentes.
 - $\text{AssoPR} : P \times R \rightarrow \text{Booléen}$
 - $\text{AssoPR}(p_i, r_j) = \text{vrai}$ ssi $p_i \text{ EC } r_j$

II.2.2 Graphes élémentaires

Les graphes élémentaires des bâtiments adjacents, parcelles adjacentes et des routes adjacentes sont définis respectivement comme suit :

- **Graphe de bâtiments adjacents** $G_B = (S_B, A_B)$ tel que (Figure 13(b)) :
 - $S_B = B$

- $A_B = \{(b_i, b_j) \in B \mid \text{Asso2B}(b_i, b_j) = \text{vrai}\}$ représente les bâtiments adjacents dans E
- **Graphe de parcelles adjacentes** $G_P = (S_P, A_P)$ tel que (Figure 13(c)) :
 - $S_P = P$
 - $A_P = \{(p_i, p_j) \in P \mid \text{Asso2P}(p_i, p_j) = \text{vrai}\}$ représente les parcelles adjacentes dans E
- **Graphe de routes adjacentes** $G_R = (S_R, A_R)$ tel que (Figure 13(d)) :
 - $S_R = R$
 - $A_R = \{(r_i, r_j) \in R \mid \text{Asso2R}(r_i, r_j) = \text{vrai}\}$ représente les routes adjacentes dans E

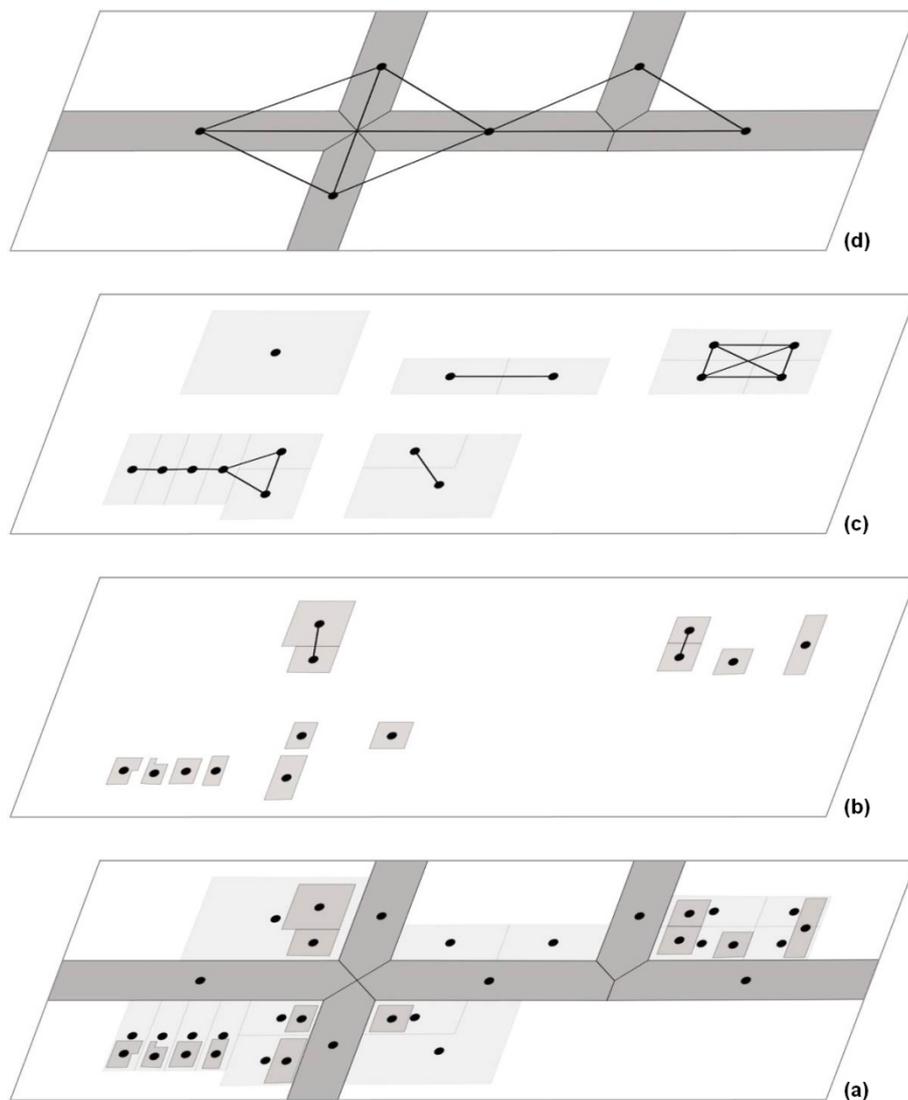


Figure 13 : (a) Représentation schématique d'un espace habité, (b) Graphe de bâtiments adjacents, (c) Graphe de parcelles adjacentes et (d) Graphe de routes adjacentes.

II.2.3 Graphes intermédiaires

Il n'existe pas de lien topologique direct dans notre modèle entre les bâtiments et les routes. Par contre, nous pouvons définir les graphes reliant respectivement les bâtiments et les parcelles, et les parcelles et les routes.

- **Grappe de bâtiments et de parcelles** $G_{BP} = (S_{BP}, A_{BP})$ tel que (Figure 14(b)) :
 - $S_{BP} = B \cup P$
 - $A_{BP} = A_B \cup A_P \cup \{(b_i, p_j) \in B \times P \mid \text{AssoBP}(b_i, p_j) = \text{vrai}\}$ représente les associations des bâtiments avec les parcelles qu'elles intersectent et les associations de bâtiments adjacents et de parcelles adjacentes de E
- **Grappe de parcelles et de routes** $G_{PR} = (S_{PR}, A_{PR})$ tel que (Figure 14(c)) :
 - $S_{PR} = P \cup R$
 - $A_{PR} = A_P \cup A_R \cup \{(p_i, r_j) \in P \times R \mid \text{AssoPR}(p_i, r_j) = \text{vrai}\}$ représente les associations des parcelles avec les routes qui leur sont adjacentes et les associations de parcelles adjacentes et de routes adjacentes de E

II.2.4 Graphe intégral

Le graphe dit intégral est le résultat de la combinaison des graphes précédents et nous permet d'établir un graphe connexe incluant les différents types d'associations qui existent entre les bâtiments, les parcelles et les routes. Le principal intérêt de cette approche est qu'elle assure l'apparition de liens entre bâtiments et routes (Figure 14(d)).

- **Grappe intégral connexe** $G_{BPR} = (S_{BPR}, A_{BPR})$ tel que (Figure 14(d)) :
 - $S_{BPR} = B \cup P \cup R$
 - $A_{BPR} = A_B \cup A_P \cup A_R \cup A_{BP} \cup A_{PR} \cup \{(b_i, r_j) \in B \times R \mid \exists p_k \in P, \text{AssoBP}(b_i, p_k) = \text{vrai} \wedge \text{AssoPR}(p_k, r_j) = \text{vrai}\}$ représente les associations des bâtiments et des routes respectivement associées à une même parcelle en plus des associations de bâtiments adjacents, parcelles adjacentes, routes adjacentes de E ainsi que les associations des bâtiments avec les parcelles qu'elles intersectent et les associations des parcelles avec les routes qui leurs sont adjacentes.

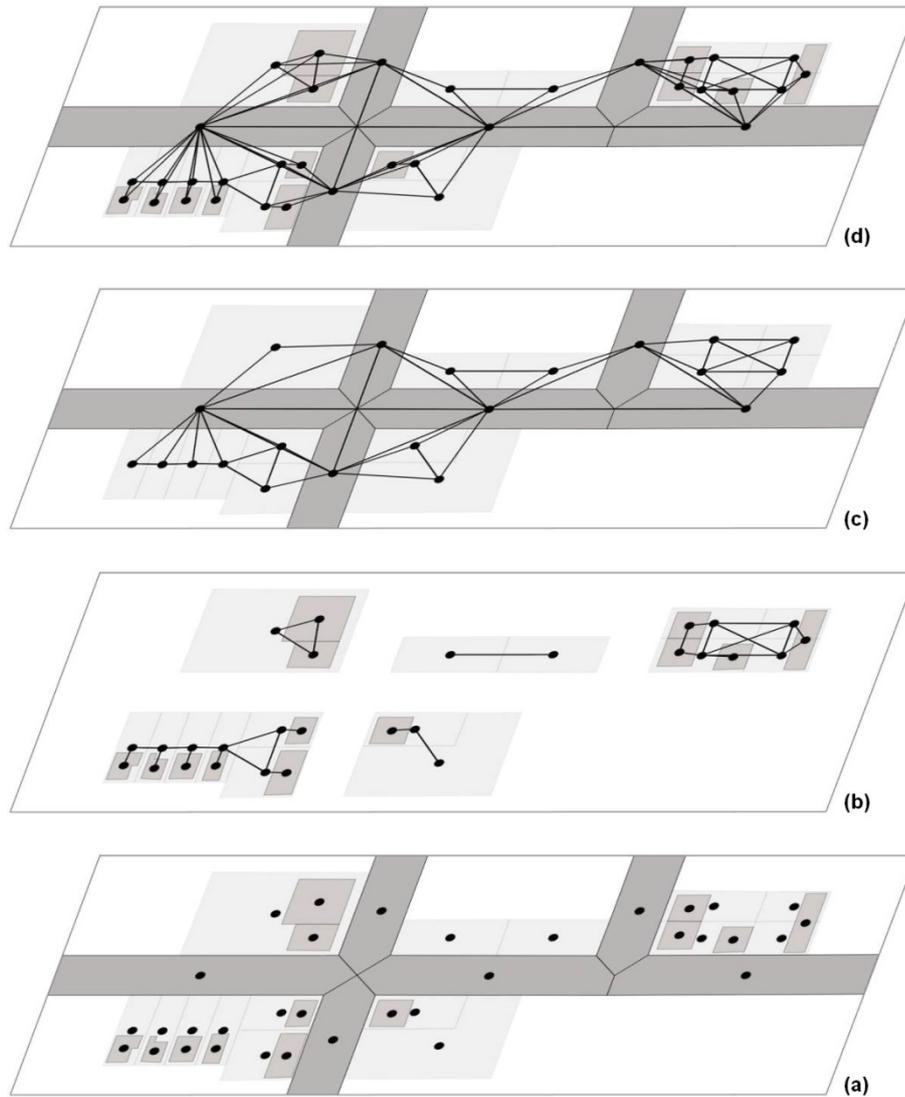


Figure 14 : (a) Représentation schématique d'un espace habité, (b) Graphe de bâtiments et de parcelles, (c) Graphe de parcelles et de routes et (d) Graphe intégral combinant bâtiments, parcelles, et routes.

II.2.5 Traitement d'un exemple à travers le modèle de graphe proposé

Afin d'illustrer la mise en œuvre du graphe intégral, nous utilisons la configuration simplifiée du territoire de la Figure 12. Celui-ci regroupe 13 bâtiments, 15 parcelles et 6 routes.

Les différents sous-graphes et le graphe intégral représentés sur la Figure 13 et la Figure 14 sont construits comme suit :

- Graphe de bâtiments adjacents : $G_B = (S_B, A_B)$ avec $S_B = \{b_1; b_2; b_3; b_4; b_5; b_6; b_7; b_8; b_9; b_{10}; b_{11}; b_{12}; b_{13}\}$ et $A_B = \{(b_1, b_2); (b_5, b_{12})\}$

- Graphe de parcelles adjacentes : $G_P = (S_P, A_P)$ avec $S_P = \{p_1 ; p_2 ; p_3 ; p_4 ; p_5 ; p_6 ; p_7 ; p_8 ; p_9 ; p_{10} ; p_{11} ; p_{12} ; p_{13} ; p_{14} ; p_{15}\}$ et $A_P = \{(p_1, p_2) ; (p_2, p_3) ; (p_4, p_5) ; (p_4, p_6) ; (p_5, p_6) ; (p_8, p_9) ; (p_{10}, p_{11}) ; (p_{12}, p_{13}) ; (p_{12}, p_{14}) ; (p_{12}, p_{15}) ; (p_{13}, p_{14}) ; (p_{13}, p_{15}) ; (p_{14}, p_{15})\}$
- Graphe de routes adjacentes : $G_R = (S_R, A_R)$ avec $S_R = \{r_1 ; r_2 ; r_3 ; r_4 ; r_5 ; r_6\}$ et $A_R = \{(r_1, r_2) ; (r_1, r_3) ; (r_1, r_4) ; (r_2, r_3) ; (r_2, r_4) ; (r_3, r_4) ; (r_4, r_5) ; (r_4, r_6) ; (r_5, r_6)\}$
- Graphe de bâtiments et parcelles : $G_{BP} = (S_{BP}, A_{BP})$ avec $S_{BP} = S_B \cup S_P$ et $A_{BP} = A_B \cup A_P \cup \{(b_1, p_{12}) ; (b_2, p_{13}) ; (b_3, p_{13}) ; (b_4, p_{14}) ; (b_4, p_{15}) ; (b_5, p_7) ; (b_6, p_8) ; (b_7, p_5) ; (b_8, p_6) ; (b_9, p_4) ; (b_{10}, p_3) ; (b_{11}, p_2) ; (b_{12}, p_7) ; (b_{13}, p_1)\}$
- Graphe de parcelles et de routes : $G_{PR} = (S_{PR}, A_{PR})$ avec $S_{PR} = S_P \cup S_R$ et $A_{PR} = A_P \cup A_R \cup \{(p_1, r_1) ; (p_2, r_1) ; (p_3, r_1) ; (p_4, r_1) ; (p_5, r_1) ; (p_5, r_3) ; (p_6, r_3) ; (p_7, r_1) ; (p_7, r_2) ; (p_8, r_3) ; (p_8, r_4) ; (p_9, r_3) ; (p_9, r_4) ; (p_{10}, r_2) ; (p_{10}, r_4) ; (p_{11}, r_4) ; (p_{11}, r_5) ; (p_{12}, r_5) ; (p_{13}, r_5) ; (p_{13}, r_6) ; (p_{14}, r_6) ; \}$
- Graphe intégral : $G_{BPR} = (S_{BPR}, A_{BPR})$ avec $S_{BPR} = S_B \cup S_P \cup S_R$ et $A_{BPR} = A_B \cup A_P \cup A_R \cup A_{BP} \cup A_{PR} \cup \{(b_1, r_5) ; (b_2, r_5) ; (b_2, r_6) ; (b_3, r_5) ; (b_3, r_6) ; (b_4, r_6) ; (b_5, r_1) ; (b_5, r_2) ; (b_6, r_3) ; (b_6, r_4) ; (b_7, r_1) ; (b_7, r_3) ; (b_8, r_3) ; (b_9, r_1) ; (b_{10}, r_1) ; (b_{11}, r_1) ; (b_{12}, r_1) ; (b_{12}, r_2) ; (b_{13}, r_1) ; \}$

II.3 Opérateurs

Le graphe intégral offrant un cadre pour analyser la structure d'un espace habité, nous proposons de définir de nouveaux opérateurs permettant de caractériser les propriétés des bâtiments et celles des tronçons de route.

Pour les bâtiments, on définit une mesure de centralité, et pour les routes des mesures de connectivité, de densité et de compacité. Ces différentes mesures ont pour intérêt de prendre en compte à la fois des notions habituelles (p. ex. la longueur des déplacements sur le réseau, nombre de sauts d'arc sur le graphe), et des notions plus sémantiques comme le type d'élément qui compose le chemin.

II.3.1 Mesure de centralité du bâti

Le premier opérateur que nous proposons concerne l'analyse de la centralité du bâti. Elle dépend du nombre et de la qualité des relations qui existent entre les bâtiments. Les mesures classiques de centralité (cf. I.2.3) sont essentiellement adaptées aux graphes fortement connexes. Si l'on se

limite au graphe de bâtis (c.-à-d. graphe G_B sur la Figure 13b) qui est loin d'être connexe, tous les bâtiments non adjacents à d'autres bâtiments (c.-à-d. les bâtiments isolés) obtiennent une valeur de centralité nulle. L'approche de modélisation que nous proposons permet de répondre à cette lacune. En effet, le modèle de graphe proposé est lui connexe, et contient un ou plusieurs chemins qui permettent de 'connecter' tous les bâtiments entre eux y compris les bâtiments isolés. Il est cependant important de noter que les chemins qui relient les bâtiments sont de différents types selon que les bâtiments soient adjacents ou non, ou que les chemins incluent des parcelles et/ou des routes. En conséquence, nous choisissons de pouvoir distinguer ces différentes configurations dans la mesure de centralité.

Afin de préciser les différentes configurations, nous proposons de qualifier cinq types de chemins possibles qui présentent un intérêt entre deux bâtiments, à savoir :

- adjacents et sur une même parcelle
- adjacents, mais dans différentes parcelles
- non adjacents, mais situés sur une même parcelle
- non adjacents, mais situés sur des parcelles adjacentes
- non adjacents, mais sur des parcelles connectées à une même route

Cette approche qualitative peut alors être évaluée à l'aide d'une valeur α_i avec $1 \leq i \leq 5$. Ces valeurs sont définies par l'utilisateur en fonction de l'importance de chaque configuration vis-à-vis de la problématique. Considérant une fonction $\text{Poids}(b_i, b_j)$ qui quantifie la relation entre deux bâtiments b_i et b_j , on a :

- $\text{Poids}(b_i, b_j) = \alpha_1$, si b_i et b_j sont adjacents et sur la même parcelle
 - $(b_i, b_j) \in A_B$
 - $\wedge \exists p_k \in P \mid (b_i, p_k) \in A_{BP} \wedge (b_j, p_k) \in A_{BP}$
- $\text{Poids}(b_i, b_j) = \alpha_2$, si b_i et b_j sont adjacents mais sur différentes parcelles
 - $(b_i, b_j) \in A_B$
 - $\wedge \nexists p_k \in P \mid (b_i, p_k) \in A_{BP} \wedge (b_j, p_k) \in A_{BP}$
- $\text{Poids}(b_i, b_j) = \alpha_3$, si b_i et b_j ne sont pas adjacents mais sur la même parcelle :
 - $(b_i, b_j) \notin A_B$
 - $\wedge \exists p_k \in P \mid (b_i, p_k) \in A_{BP} \wedge (b_j, p_k) \in A_{BP}$

- $Poids(b_i, b_j) = \alpha_4$, si b_i et b_j ne sont pas adjacents ni sur la même parcelle, mais sur des parcelles adjacentes :
 - $(b_i, b_j) \notin A_B$
 - $\wedge \nexists p_k \in P \mid (b_i, p_k) \in A_{BP} \wedge (b_j, p_k) \in A_{BP}$
 - $\wedge \exists p_k, p_l \in P \mid (b_i, p_k) \in A_{BP} \wedge (b_j, p_l) \in A_{BP} \wedge (p_k, p_l) \in A_P$
- $Poids(b_i, b_j) = \alpha_5$, si b_i et b_j ne sont pas adjacents ni sur la même parcelle ni sur des parcelles adjacentes, mais sont seulement connectés à la même route :
 - $(b_i, b_j) \notin A_B$
 - $\wedge \nexists p_k \in P \mid (b_i, p_k) \in A_{BP} \wedge (b_j, p_k) \in A_{BP}$
 - $\wedge \nexists p_k, p_l \in P \mid (b_i, p_k) \in A_{BP} \wedge (b_j, p_l) \in A_{BP} \wedge (p_k, p_l) \in A_P$
 - $\wedge \exists r_k \in R \mid (b_i, r_k) \in A_{BPR} \wedge (b_j, r_k) \in A_{BPR}$
- $Poids(b_i, b_j) = 0$, dans les autres cas

Nous obtenons pour chaque bâtiment une mesure pondérée de connexion locale qui est définie comme étant la somme des différentes connexions possibles entre un bâtiment b_i et tous les autres bâtiments :

$$C_{BPR}(b_i) = \sum_{j \neq i} Poids(b_i, b_j)$$

Considérant n_1 le nombre de connexions de $Poids(b_i, b_j) = \alpha_1$, n_2 le nombre de connexions de $Poids(b_i, b_j) = \alpha_2$, n_3 le nombre de connexions de $Poids(b_i, b_j) = \alpha_3$, n_4 le nombre de connexions de $Poids(b_i, b_j) = \alpha_4$ et n_5 le nombre de connexions de $Poids(b_i, b_j) = \alpha_5$, on obtient :

$$C_{BPR}(b_i) = n_1 * \alpha_1 + n_2 * \alpha_2 + n_3 * \alpha_3 + n_4 * \alpha_4 + n_5 * \alpha_5$$

À titre d'exemple, en considérant la configuration illustrée sur la Figure 12 et en se focalisant sur les bâtiments b_6 et b_2 , on obtient :

- Exemple pour le Bâtiment b_6 :

Le bâtiment b_6 n'est adjacent à aucun autre bâtiment et l'on obtient donc $n_1 = 0$ et $n_2 = 0$.

Le bâtiment b_6 repose sur la parcelle p_8 qui ne contient aucun autre bâtiment et l'on obtient donc $n_3 = 0$.

La parcelle p_8 est adjacente à la parcelle p_9 . Aucun bâtiment ne repose sur la parcelle p_9 , on obtient donc $n_4 = 0$.

La parcelle p_8 est adjacente aux routes r_3 et r_4 . Les routes r_3 et r_4 sont connectés aux parcelles p_5 , p_6 , p_8 , p_9 , p_{10} et p_{11} . Seules les parcelles p_5 et p_6 contiennent chacune un bâtiment (différent de b_6) : les bâtiments b_7 et b_8 . On obtient donc $n_5 = 2$.

Ainsi on obtient $C_{BPR}(b_6) = 0 * \alpha_1 + 0 * \alpha_2 + 0 * \alpha_3 + 0 * \alpha_4 + 2 * \alpha_5$

Soit $C_{BPR}(b_6) = 2\alpha_5$

- Exemple pour le Bâtiment b_2 :

Le bâtiment b_2 est adjacent au bâtiment b_1 . Les bâtiments b_1 et b_2 ne reposant pas sur une parcelle en commun, on obtient donc $n_1 = 0$ et $n_2 = 1$.

Le bâtiment b_2 repose sur la parcelle p_{13} qui contient un autre bâtiment (b_3) et l'on obtient donc $n_3 = 0$.

La parcelle p_{13} est adjacente aux parcelles p_{12} , p_{14} et p_{15} qui contiennent les bâtiments b_1 et b_4 . Le bâtiment b_1 étant adjacent au bâtiment b_2 , il est déjà compté pour n_2 et n'est donc pas éligible pour le calcul de n_4 ce qui aboutit à ne considérer que le bâtiment b_4 . On obtient donc $n_4 = 1$.

La parcelle p_{13} est adjacente aux routes r_5 et r_6 . Les routes r_5 et r_6 sont connectées aux parcelles p_{11} , p_{12} , p_{13} et p_{14} . Seule la parcelle p_{11} n'a pas encore été traitée et ne contenant pas de bâtiment, on obtient donc $n_5 = 0$.

Ainsi on obtient $C_{BPR}(b_2) = 0 * \alpha_1 + 1 * \alpha_2 + 1 * \alpha_3 + 1 * \alpha_4 + 0 * \alpha_5$

Soit $C_{BPR}(b_2) = \alpha_2 + \alpha_3 + \alpha_4$

Sans avoir défini de valeurs pour α_i , il est ainsi possible d'établir un premier niveau de comparaison en fonction des n_i obtenus par chaque bâtiment. Par exemple, $n_2 + n_4$ est supérieur pour le bâtiment b_2 , il y a donc plus de bâtiment sur les parcelles adjacentes du bâtiment b_2 que sur les parcelles adjacentes du bâtiment b_6 .

Afin de rendre compte qu'une notion de hiérarchie existe dans les différentes configurations, nous suggérons que les liens entre les bâtiments connectés sont plus forts que les liens entre bâtiments

déconnectés (distants). De même, l'association de deux bâtiments situés sur des parcelles connexes est plus importante que le lien entre deux bâtiments localisés sur des parcelles distantes. Afin de refléter cette notion de hiérarchie, on suggère que $\alpha_1 > \alpha_2 > \alpha_3 > \alpha_4 > \alpha_5$.

En prenant en compte une telle hiérarchie, le bâtiment b_2 obtient une valeur de centralité supérieure au bâtiment b_6 . En effet, on a :

$$2\alpha_5 < \alpha_2 + \alpha_3 + \alpha_4$$

$$\text{Soit : } C_{BPR}(b_6) < C_{BPR}(b_2)$$

II.3.2 Mesure de connectivité locale du bâti

La notion de connectivité fait référence ici à la qualité de la relation entre le bâti et le réseau de circulation. Nous introduisons une mesure locale qui permet d'apprécier, pour un tronçon de route, comment le bâti qui y est associé est lié à d'autres tronçons : si le bâti est associé à plusieurs tronçons la mesure de connectivité est plus forte.

Nous définissons la fonction $NbC(b_i)$ qui retourne le nombre de routes connectées à la (ou les) parcelle sur laquelle est situé le bâtiment b_i :

$$NbC(b_i) = |\{r_j \mid \exists p_k \in P \mid (b_i, p_k) \in A_{BP} \wedge (p_k, r_j) \in A_{PR}\}|$$

Nous utilisons ensuite cette fonction pour évaluer le nombre moyen de routes auxquelles sont connectés les bâtiments d'une route. Ainsi pour une route r_j connectée à n bâtiments b_i , la connectivité locale du bâti est donnée par la relation suivante :

$$C_R(r_j) = \frac{\sum_{i=1}^n NbC(b_i)}{n}$$

Soit m le nombre de routes dans le réseau, on a : $1 \leq C_R(r_j) \leq m$.

En effet :

- Si tous les bâtiments de la route r_j ne sont connectés qu'à la route r_j , alors $NbC(b_i) = 1$ et $\sum_{i=1}^n NbC(b_i) = n$. Dans ce cas, $C_R(r_j) = 1$
- Si tous les bâtiments de la route r_j sont connectés à toutes les routes, alors $NbC(b_i) = m$ et $\sum_{i=1}^n NbC(b_i) = n * m$. Dans ce cas $C_R(r_j) = m$.

Notons que cette mesure reflète, pour une route, uniquement une valeur moyenne de « partage » des bâtiments qui lui sont connectés. En conséquence, la mesure n'est pas corrélée avec le nombre de bâtiments concernés.

II.3.3 Mesure de densité de bâti

La densité du bâti est une mesure locale qui, pour un tronçon de route de longueur donnée, évalue le nombre de bâtiments qui lui sont associés. La densité est relative à la longueur du tronçon de route pris en compte. Dans notre modèle de graphe, où les parcelles sont les éléments qui permettent d'associer bâtis et route, cette association est définie par :

$$A_{BR} = \{(b_i, r_j) \in B \times R \mid \exists p_k \in P, \text{AssoBP}(b_i, p_k) = \text{vrai} \wedge \text{AssoBP}(p_k, r_j) = \text{vrai}\}$$

Dans notre approche, la valeur affectée à un bâtiment est exprimée par un poids $P(b_i)$ de valeur égale à 1. Dans le cas où un bâtiment est associé à plusieurs routes (c.-à-d. localisé sur une parcelle connectée à plusieurs routes), ce poids $P(b_i)$ est distribué de façon égale sur chacune de ces routes afin d'éviter une surévaluation du bâti. Ainsi, ce poids est défini par :

$$P(b_i) = \frac{1}{|\{r_k \in R \mid (b_i, r_k) \in A_{BR}\}|}$$

La Densité de Bâti (BRD) d'une route r_j est définie par :

$$BRD(r_j) = \frac{\sum_{b_i \in B \mid (b_i, r_j) \in A_{BR}} P(b_i)}{\text{Longueur}(r_j)}$$

où $\text{Longueur}(r_j)$ représente la longueur de la route r_j .

Soit m le nombre de routes dans le réseau et n le nombre de bâtiments connectés à la route r_j , on a :

$$\frac{n}{\text{Longueur}(r_j) * m} \leq BRD(r_j) \leq \frac{n}{\text{Longueur}(r_j)}$$

En effet :

- Si tous les bâtiments de la route r_j ne sont connectés qu'à la route r_j , alors $P(b_i) = 1$ et $\sum_{i=1}^n P(b_i) = n$. Dans ce cas, $BRD(r_j) = \frac{n}{\text{Longueur}(r_j)}$
- Si tous les bâtiments de la route r_j sont connectés à toutes les routes, alors $P(b_i) = \frac{1}{m}$ et $\sum_{i=1}^n P(b_i) = n * \frac{1}{m}$. Dans ce cas, $BRD(r_j) = \frac{n}{\text{Longueur}(r_j) * m}$

Pour un même nombre de bâtiments et pour une longueur de tronçon de route égale, la valeur de *BRD* sera d'autant plus faible que les bâtiments seront localisés sur des parcelles connectées à plusieurs autres tronçons de route.

À titre d'exemple, en considérant la configuration illustrée sur la Figure 12 et en se focalisant sur la Route r_1 , on obtient :

La route r_1 est connectée aux parcelles p_1, p_2, p_3, p_4, p_5 et p_7 qui contiennent un ou plusieurs bâtiments.

Les parcelles p_1 à p_4 ne sont connectées qu'à la route r_1 et contiennent un seul bâtiment chacune : $b_{13}, b_{11}, b_{10}, b_9$. On obtient donc $P(b_{13}) = P(b_{11}) = P(b_{10}) = P(b_9) = 1$

La parcelle p_5 est aussi connectée à la route r_3 et contient le bâtiment b_7 . Ce dernier étant connecté à deux tronçons de route, on obtient donc : $P(b_7) = \frac{1}{2}$

La parcelle p_7 est aussi connectée à la route r_2 et contient les bâtiments b_5 et b_{12} . Ces derniers étant connectés à deux tronçons de route, on obtient donc : $P(b_5) = P(b_{12}) = \frac{1}{2}$

Ainsi on obtient $BRD(r_1) = \frac{P(b_{13})+P(b_{11})+P(b_{10})+P(b_9)+P(b_7)+P(b_5)+P(b_{12})}{Longueur(r_j)}$

$$\text{Soit : } BRD(r_1) = \frac{5,5}{Longueur(r_j)}$$

Bien que la route r_1 soit connectée à 7 bâtiments via les parcelles adjacentes à r_1 , le partage du poids des bâtiments avec les autres routes diminue le nombre de bâtiments qui lui sont associés de 7 à 5,5 bâtiments adjacents.

II.3.4 Mesure de compacité du réseau de circulation

Nous introduisons et nous définissons la notion de compacité comme une valeur permettant d'évaluer la qualité des relations entre un tronçon de route et le reste du réseau. Dans le domaine de l'information spatiale, nous exprimons cette qualité par un nombre de connexions et une mesure de distance entre les tronçons de route. Ainsi, le niveau de compacité d'un tronçon de route reflète le nombre de tronçons qui lui sont connectés et la distance à parcourir pour les rejoindre. Cet aspect est à rapprocher de la classification de (Marshall, 2005) illustrée sur la Figure 7. Dans ce cadre, notre objectif est de distinguer les tronçons de route fortement connectés (*More Connective*) et ceux faiblement connectés (*Less connective*).

Nous proposons une mesure de compacité, nommée Metric Compactness. Elle représente la capacité d'une route à être reliée aux autres parties du réseau routier. La Metric Compactness est introduite progressivement au travers de deux métriques : *Ratio Minimum Distance* et *Minimum Ratio Minimum Distance*.

II.3.4.1 Ratio Minimum Distance

La *Ratio Minimum Distance* est définie comme la plus petite distance notée $D_{min}^X(r_i)$ qui permet à un tronçon de route $r_i \in R$ (avec $1 \leq i \leq |R|$), d'atteindre X pourcent du nombre de tronçons du réseau routier R .

Elle est définie par :

$$D_{min}^X(r_i) = \text{Min}\{d \mid \left(\frac{|\{r_j \mid d \geq d_{r_i r_j}\}|}{|R|} \geq \frac{X}{100} \right)\}$$

avec d une distance entre deux tronçons et $\text{Min}\{\}$ l'opérateur qui renvoie la valeur minimale d'un ensemble. $d_{r_i r_j}$ correspond à la distance entre le milieu des routes r_i et r_j . La distance est une distance métrique mesurée le long du parcours du réseau de routes.

II.3.4.2 Minimum Ratio Minimum Distance

La *Minimum Ratio Minimum Distance*, notée D_{min}^X , d'un réseau de routes R est la plus petite valeur de $D_{min}^X(r_i)$ avec $1 \leq i \leq |R|$.

Il est défini comme :

$$D_{min}^X = \text{Min}(\{D_{min}^X(r_i)\})$$

II.3.4.3 Metric Compactness

Pour une route r_i , la Metric Compactness représente le nombre de routes atteintes en parcourant la *Minimum Ratio Minimum Distance* D_{min}^X à partir de r_i .

Elle est définie par :

$$MC(r_i) = |\{r_j \mid d_{r_i r_j} \leq D_{min}^X\}|$$

où $d_{r_i r_j}$ représente la distance entre les routes r_i et r_j . Comme pour la *Ratio Minimum Distance*, $d_{r_i r_j}$ correspond à la distance entre le milieu des routes r_i et r_j . La distance est une distance métrique mesurée le long du parcours du réseau de routes.

Une particularité de la mesure de compacité est sa flexibilité donnée par le choix de la valeur X définie par l'utilisateur. Cette valeur est comprise entre 0 et 100 %. Pour $X = 100$, on parlera d'approche globale avec une mesure qui implique l'ensemble du réseau et permet d'extraire les tronçons de route qui sont les plus centraux (les plus proches de l'ensemble des autres tronçons). Pour des valeurs beaucoup plus faibles de X , on parlera d'approches locales avec une mesure qui implique seulement une fraction du réseau et qui permet d'extraire des tronçons de route centraux à l'échelle de cette fraction du réseau. En raison de cette capacité à différencier l'aspect global et l'aspect local, la mesure de compacité offre un réel potentiel pour l'analyse de la structuration de l'espace à plusieurs échelles à l'instar des travaux de (Porta, et al., 2006) (Figure 10).

II.3.5 Impact de la mesure de compacité sur l'analyse de structuration de l'espace

II.3.5.1 Réseaux homogènes ou hétérogènes

La classification de (Marshall, 2005) illustrée sur la Figure 7 permet de distinguer des réseaux très connectés et des réseaux peu connectés. En référence à la mesure de compacité décrite ci-dessus, nous proposons de préciser cette classification en introduisant une distinction entre réseaux homogènes et réseaux hétérogènes. Les réseaux homogènes présentent des valeurs de compacité équivalentes sur tous les tronçons de route, alors que les réseaux hétérogènes présentent, sur le même espace, des tronçons de route fortement connectés et des tronçons de route faiblement connectés. La notion de réseau hétérogène correspond à la présence de clusters (ensemble de tronçons fortement connectés) faiblement connectés aux autres tronçons du réseau. Sur la Figure 7 les schémas a, b et d sont homogènes alors que le schéma c est lui hétérogène. Ainsi nous proposons d'utiliser la mesure de compacité pour faire la distinction entre les deux configurations.

La Figure 15 permet par exemple d'observer la différence de comportement de la mesure de compacité entre un réseau homogène (Figure 15 N_a) et un réseau hétérogène composé de 3 clusters (Figure 15 N_b). Pour un même nombre de tronçons de route et pour des distances de déplacement équivalentes, le rapport $\frac{\Delta D_{min}^X}{\Delta X}$ est constant pour un réseau homogène, et variable pour un réseau hétérogène. Cette illustration souligne l'importance des valeurs de D_{min}^X pour caractériser le réseau.

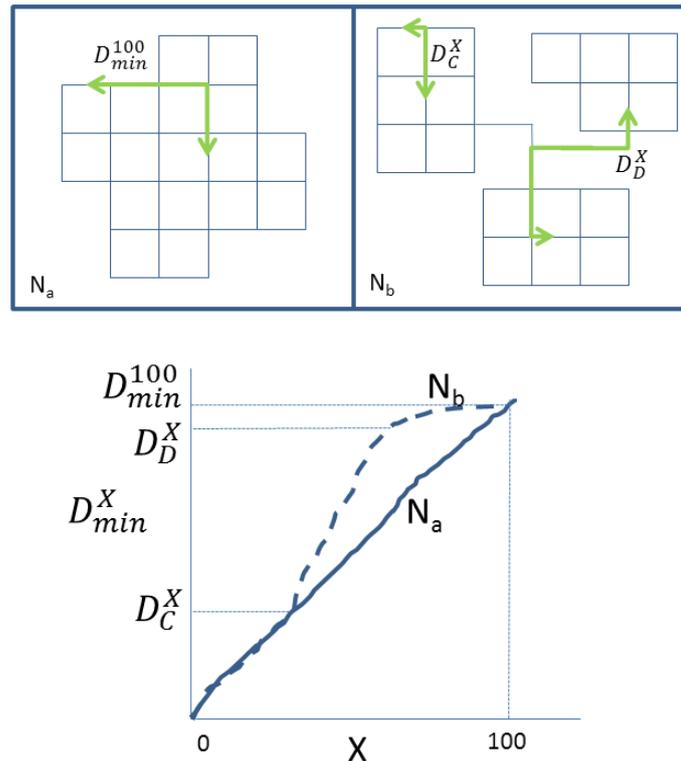


Figure 15 : Influence de l'homogénéité du réseau sur l'évolution de la *Minimum Ratio Minimum Distance* (D_{min}^X)

Dans le cas des réseaux hétérogènes, les variations de D_{min}^X permettent d'extraire certaines propriétés du réseau :

- La valeur de D_C^X caractérise la taille des clusters locaux (c.-à-d. distance du centre à la périphérie du plus petit cluster).
- La valeur de D_D^X caractérise la distance qui sépare les centres des clusters.
- La distance $D_D^X - D_C^X$ représente la distance nécessaire pour connecter les 2 clusters.
- La distance D_{Min}^{100} (c.-à-d. pour $X = 100\%$) est la plus petite distance qui permet à un tronçon de route d'atteindre toutes les autres routes⁶.

⁶ Concernant l'exemple générique de la Figure 15, le nombre de tronçons et les distances à parcourir étant équivalents, les valeurs de D_{Min}^{100} sont équivalentes pour les deux configurations.

II.3.5.2 Variation de densité du réseau

La *Minimum Ratio Minimum Distance* D_{min}^X étant mesurée sur un pourcentage du nombre de tronçons de route d'un réseau, le rapport $\frac{\Delta D_{min}^X}{\Delta X}$ est indépendant du nombre de routes, et est équivalent entre deux réseaux si la distance de déplacement sur ces réseaux est la même. Cette propriété est illustrée sur la Figure 16 avec deux réseaux N_a et N_b , structurellement équivalents (ici homogène) et ayant la même extension spatiale.

Cependant, D_{min}^X variant en fonction du nombre de routes K pris en compte (c.-à-d. avec $K = \text{ceil}(\frac{(n-1)*X}{100})$), le rapport $\frac{\Delta D_{min}^X}{\Delta K}$ est lui variable. Ainsi, comme le montre la Figure 16, la pente de D_{min}^X vs K est plus faible lorsque la densité du réseau est plus élevée, c'est-à-dire lorsque le nombre de routes à parcourir est plus important pour une même distance de déplacement.

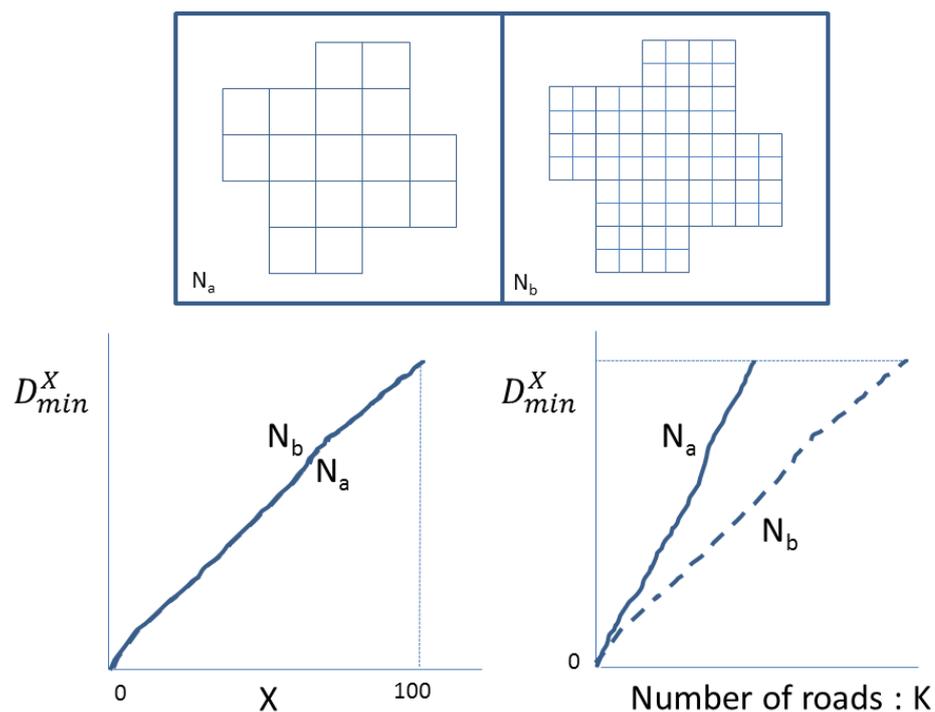


Figure 16 : Impact de la densité sur la Minimum Ratio Minimum Distance

II.4 Synthèse

Ce chapitre décrit le modèle que nous avons mis en place pour représenter, décrire et analyser la structure spatiale d'espaces habités. Nous avons défini formellement le graphe support comme la combinaison de graphes élémentaires représentant respectivement l'ensemble des bâtiments, l'ensemble des parcelles et l'ensemble des routes. Les bâtiments, les parcelles et les tronçons de route sont représentés par des sommets et les arêtes représentent les relations topologiques. Nous avons également introduit plusieurs opérateurs permettant de refléter différentes propriétés de l'espace à savoir la centralité, la connectivité et la densité du bâti, et la compacité du réseau de circulation. Nous avons montré que l'usage de ces opérateurs permettait une analyse de la structuration de l'espace et devait permettre de différencier des organisations spatiales homogènes et hétérogènes et d'identifier la présence de clusters dans le réseau. Dans le chapitre suivant, nous détaillerons la mise en place du modèle sur un cas d'étude et présenterons les résultats des expérimentations des opérateurs que nous venons de présenter.

Chapitre III

Implémentation et Résultats

III.1 Schéma fonctionnel	64
III.2 Le modèle de données	66
III.2.1 Modélisation des objets et des relations dans l'espace.....	66
III.2.2 Modèle conceptuel et modèle relationnel.....	66
III.3 Construction du graphe.....	68
III.3.1 Prétraitement.....	68
III.3.2 Construction des tables relationnelles de relations.....	72
III.4 Implémentation des opérateurs.....	74
III.4.1 Centralité	74
III.4.2 Mesure de connectivité locale du bâti	78
III.4.3 Mesure de densité de bâti relative à une route.....	79
III.4.4 Metric Compactness	80
III.5 Application et résultats	82
III.5.1 Présentation des données géographiques.....	82
III.5.2 Mesure de centralité du bâti.....	83
III.5.3 Mesure de connectivité locale du bâti	98
III.5.4 Mesure de densité de bâti	100
III.5.5 Mesure de compacité.....	101
III.5.6 Corrélation entre la mesure de densité de bâti et la mesure de compacité	109
III.6 Synthèse.....	113

Afin de mettre en application sur des données géoréférencées le modèle de graphe et les opérateurs définis précédemment, il est nécessaire d'une part de proposer un modèle de données permettant de les sauvegarder et de les exploiter, et d'autre part d'utiliser un système d'information et un ensemble de fonctions permettant d'implémenter ces opérateurs. Le choix s'est porté sur l'utilisation d'un système de gestion de bases de données géographiques qui permet de sauvegarder des données géoréférencées et propose de nombreux opérateurs de gestion, de manipulation et d'analyse spatiale. Partant de ce choix nous décrivons dans ce chapitre, dans un premier temps l'intégration du modèle de graphe dans une base de données géographiques et les modalités d'implémentation des opérateurs que nous proposons, et dans un second temps l'application à des données réelles et l'analyse des résultats qui en découlent.

III.1 Schéma fonctionnel

La mise en application du modèle et des opérateurs précédemment définis est décrite en quatre étapes présentées par la Figure 17.

La première étape consiste à intégrer les données géographiques du bâti, du parcellaire et du réseau routier dans une base de données géographiques. Nous avons choisi d'utiliser le système de gestion de base de données relationnelle PostgreSQL avec l'extension PostGIS (2.0), solution générique qui intègre les fonctions spatiales définies par l'OGC (Herring, 2011), ce qui permet de manipuler les données géographiques sous une forme de géométries.

La seconde étape comprend les prétraitements des données géographiques nécessaires pour assurer la cohérence des données en vue de la construction des relations spatiales entre les objets à modéliser. Dans notre modèle, le bâti, les parcelles et le réseau de circulation constituent les régions de l'espace à représenter. Dans le cadre de la modélisation spatiale et structurale de ces régions, la primitive géométrique que nous avons choisi d'utiliser est logiquement le polygone, car c'est la primitive la plus adaptée pour représenter les surfaces et donc les régions de l'espace. Les réseaux routiers étant plus naturellement modélisés sous forme de polygones, il sera nécessaire au niveau de l'implémentation de développer une solution permettant de transformer ces polygones en polygones.

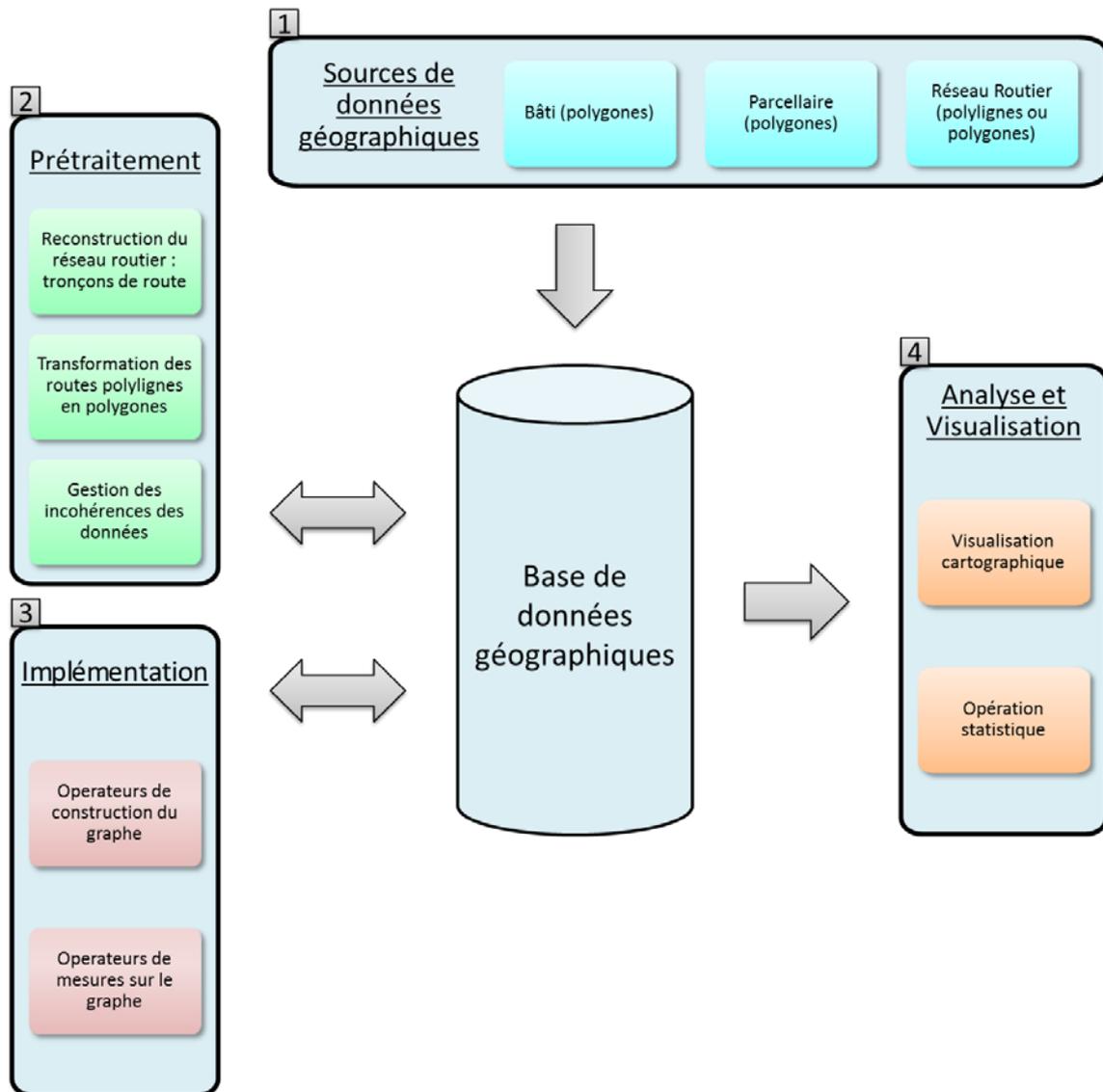


Figure 17 : Schéma fonctionnel

La troisième étape correspond à la phase de construction des relations spatiales et leur enregistrement dans la base de données. Dans un premier temps, les opérateurs de construction des relations topologiques utilisées pour construire le modèle de graphe introduit dans le chapitre II.2 sont implémentés en SQL. Dans un second temps, nous présentons l'implémentation en SQL des opérateurs de mesure de centralité, de mesure de connectivité locale du bâti, de mesure de densité de bâti relative à une route et de mesure de compacité que nous avons présentés dans le chapitre II.3.

Enfin, la dernière étape correspond à la phase d'analyse et de visualisation des résultats.

III.2 Le modèle de données

Dans l'objectif d'implémenter le modèle de graphe décrit précédemment, nous définissons un modèle conceptuel permettant de décrire les objets et la façon dont les relations spatiales qui s'établissent entre ces différents objets s'expriment dans un espace habité.

III.2.1 Modélisation des objets et des relations dans l'espace

Le modèle de graphe proposé considère trois types de régions correspondant respectivement au bâti (bâtiments), au parcellaire (parcelles) et au réseau de circulation (routes). Les relations entre ces objets sont définies à partir des notions d'adjacence, d'intersection et d'inclusion, et ce pour tous les objets. Ce choix impose une modélisation conceptuelle à base de polygones pour les trois types d'objets.

- Le bâti est représenté par un ensemble de polygones définis par l'emprise au sol de l'ensemble des bâtiments. La notion de bâtiment fait référence ici à sa définition cadastrale.
- Le parcellaire est représenté par un ensemble de polygones définis par l'emprise spatiale des délimitations cadastrales
- Le réseau de circulation correspond à l'espace entre les parcelles à l'exception en particulier des étendues d'eau. Il est représenté par un ensemble de polygones correspondant chacun à un tronçon de circulation délimité par les intersections du réseau.

Les relations entre ces objets s'expriment différemment selon ces objets :

Entre deux bâtiments, deux routes ou deux parcelles : relation d'adjacence. (II.2.2)

Entre une parcelle et une route : relation d'adjacence. (II.2.3)

Entre un bâtiment et une parcelle : relation d'inclusion ou d'intersection. (II.2.3)

Entre un bâtiment et une route : pas de relation spatiale directe, mais seulement via une parcelle. (II.2.4)

III.2.2 Modèle conceptuel et modèle relationnel

La Figure 18 est une représentation de notre base de données sous une forme de modèle conceptuel de données. Nous avons choisi d'utiliser la méthode Merise, car elle possède un pouvoir d'expression suffisant pour notre démarche et a l'avantage d'être facilement lisible et interprétable. De plus, elle est particulièrement adaptée pour la modélisation des bases de données géorelationnelles.

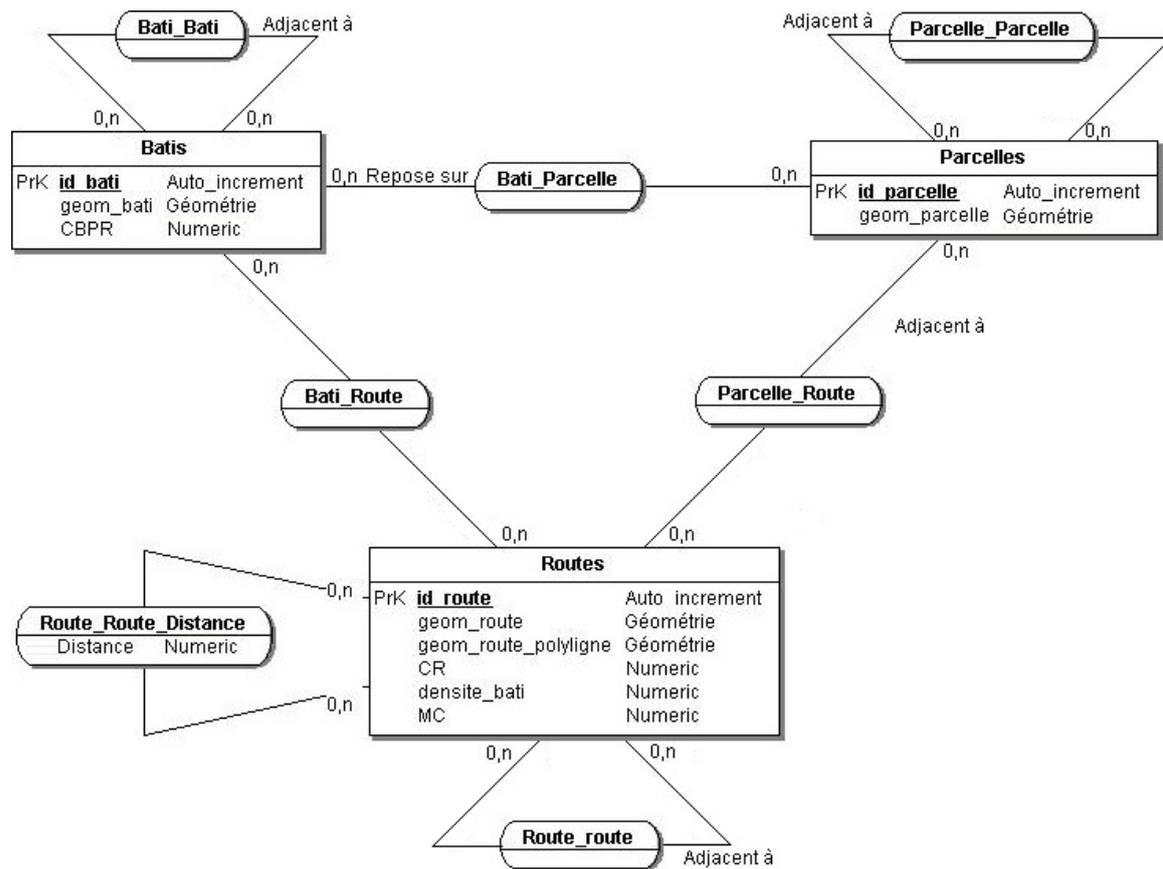


Figure 18 : Modèle conceptuel de données

Le modèle relationnel de données associé au modèle conceptuel de la Figure 18 est le suivant :

Batis (id_bati, geom_bati, CBPR)

Parcelles (id_parcelle, geom_parcelle)

Routes (id_route, geom_route, geom_route_polyligne, CR, densite_bati, MC)

Bati_Bati (id_bati1, id_bati2)

Parcelle_Parcelle (id_parcelle1, id_parcelle2)

Route_Route (id_route1, id_route2)

Bati_Parcelle (id_bati, id_parcelle)

Parcelle_Route (id_parcelle, id_route)

Bati_Route (id_bati, id_route)

Route_Route_Distance (id_route1, id_route2, distance)

Les données géographiques initiales sont stockées dans les trois tables relationnelles nommées Bâties, Parcelles et Routes, chaque géométrie étant associée à un identifiant unique. Le champ

CBPR de la table Bâtis servira à stocker pour chaque bâtiment le résultat de la mesure de centralité proposée au chapitre II.3.1. Les champs CR, densité_bati et MC servent respectivement à stocker les résultats de la mesure de connectivité locale du bâti (II.3.2), de la mesure de densité de bâti (II.3.3) et de la mesure Metric Compactness (II.3.4.3). Dans la table Routes, on différencie une géométrie de type polygone (*geom_route*), et une géométrie de type polyligne (*geom_route_polyligne*) pour un même objet.

Les six tables supplémentaires nommées Bâti_Bati, Parcelle_Parcelle, Route_Route, Bâti_Parcelle, Parcelle_route et Bâti_Route permettent de stocker les relations topologiques existantes dans les données géographiques utiles à la création de notre graphe. Bien que les graphes soient traditionnellement représentés par des matrices d'adjacences, le graphe est ici stocké sous une forme de listes d'adjacences en base de données.

Enfin, la table Route_Route_Distance permet pour chaque tronçon de route de stocker la plus petite distance vers chacun des autres tronçons de route. La distance est calculée de milieu de tronçon de route à milieu de tronçon de route à l'aide de l'algorithme de Dijkstra.

La table Bati_Route étant construite à partir des liens existants entre Bati_Parcelle et Parcelle_Route, elle est redondante et pourrait être supprimée. Cependant, nous avons choisi de la conserver afin d'avoir une correspondance directe avec notre modèle. De plus, cela permet parfois de simplifier les requêtes en évitant de faire la jointure entre les tables Bati_Parcelle et Parcelle_Route.

III.3 Construction du graphe

III.3.1 Prétraitement

Le modèle de graphe reposant sur les relations topologiques entre les entités géographiques, il est nécessaire d'établir des règles de représentation de ces objets. L'application de ces règles repose sur la mise en cohérence des données avec en particulier, l'utilisation exclusive d'objets polygonaux et l'ajustement des frontières entre ces objets.

III.3.1.1 Construction du réseau de circulation

Pour modéliser le réseau routier, l'approche que nous avons choisie consiste à considérer que l'élément de base du réseau est le tronçon. Celui-ci est délimité par une intersection avec au moins deux autres tronçons. Il faut donc s'assurer pour chaque tronçon que :

- Les tronçons ne peuvent pas se croiser. S'il y a un croisement le long d'un tronçon de route, celui-ci doit être divisé en deux tronçons au niveau du point d'intersection (on ne considère pas ici le croisement marqué par un pont ou un tunnel).
- Une intersection implique au moins trois tronçons. S'il existe une limite entre seulement deux tronçons différents, ils doivent être fusionnés.

III.3.1.2 Construction des routes polygones

Le modèle repose sur une représentation des voies de circulation à base de polygones occupant l'espace entre les parcelles (les étendues d'eau ne sont pas considérées). Si ces voies sont enregistrées uniquement par des polygones il est alors nécessaire de les transformer en polygones. À l'intersection des tronçons, une contrainte à respecter est celle de pouvoir dériver et définir les relations spatiales de connexions entre les polygones impliqués. Dans notre cas, la relation de connexion entre deux tronçons étant définie par la relation topologique d'adjacence, ceci implique que les unités spatiales possèdent en commun au moins l'une des arêtes des polygones qui les représentent.

Pour répondre à cette contrainte, nous mettons en place une méthodologie de construction de polygones. Les différentes étapes de la construction de polygone sont illustrées sur la Figure 19. Elles reposent sur l'utilisation d'un opérateur de zone tampon (buffer).

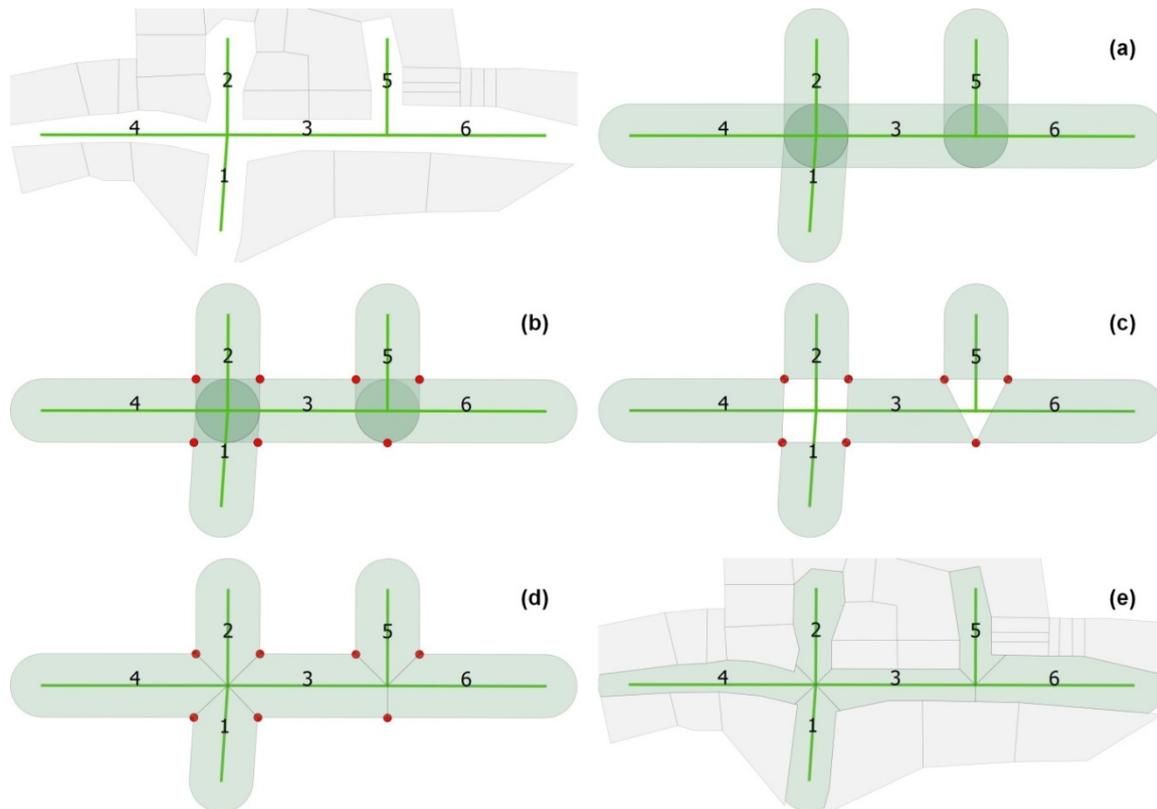


Figure 19 : Création de routes en polygones

Les différentes étapes sont les suivantes :

- Étape 1 : Construction d'une zone tampon autour de chaque tronçon (Figure 19a).

Pour chaque tronçon, application de la méthode *ST_Buffer (geom, dist)* avec *geom* la géométrie de la polyligne associée à un tronçon et *dist* une valeur définissant l'extension de la zone tampon telle que celle-ci soit suffisante pour couvrir l'espace entre les parcelles.

- Étape 2 : Extraction des points d'intersection entre les zones tampons (Figure 19b).

Pour chaque tronçon, application de la méthode *St-Intersection(geom1, geom2)*, avec *geom1* la frontière de la zone tampon de ce tronçon, et *geom2*, la frontière de l'union des zones tampons des autres tronçons.

- Étape 3 : Suppression de la partie des zones tampons qui recouvrent les autres zones tampons (Figure 19c)

Pour chaque tronçon, application de la méthode *ST_Split* (*geom1*, *geom2*), avec *geom1* la géométrie de la zone tampon et *geom2* la géométrie du segment formé par les points d'intersection obtenus à l'étape 2.

- Étape 4 : Modification des zones tampons pour inclure le point au centre de l'intersection (Figure 19d)

Pour chaque tronçon, application de la fonction *ST_MakePolygon* (*geom*), avec *geom* la polyligne fermée regroupant les points de la zone tampon modifiée à l'étape 3 et les points d'intersection de la polyligne initiale associée au tronçon.

- Étape 5 : Modification des zones tampons pour combler l'espace entre les parcelles (Figure 19e)

Pour chaque tronçon, application de la fonction *ST_Difference* (*geom1*, *geom2*) avec *geom1* la géométrie de la zone tampon et *geom2*, la géométrie de l'union de toutes les parcelles.

III.3.1.3 Gestion des incohérences

La représentation géométrique des objets peut introduire un biais entre la réalité terrain et la valeur des données enregistrées. Ainsi, il est possible que deux parcelles voisines et connectées dans le monde réel soient enregistrées comme deux polygones qui se superposent ou sont disjoints. Les approches suivantes sont appliquées pour traiter ces incohérences :

- En ce qui concerne les relations entre deux bâtiments, deux parcelles, deux routes et entre une parcelle et une route, la *superposition* entre les polygones qui les représentent ne doit pas exister. Les données sont systématiquement corrigées en modifiant la position des sommets qui composent les objets, tel que les polygones concernés partagent des frontières sans se superposer. Afin d'estimer la distance d'erreur existante dans le jeu de données, nous conservons la distance moyenne des déplacements (en mètres), nommée *dmp*, et définie comme :

$$dmp = \frac{\sum_{i=1}^n d(i)}{n}$$

avec $d(i)$ la distance euclidienne de déplacement d'un sommet i , et n le nombre de sommets déplacés.

- Pour ces mêmes relations (c.-à-d. entre deux bâtiments, deux parcelles, deux routes et entre une parcelle et une route), le fait qu'ils soient **déconnectés** n'est pas forcément illégitime. Pour différencier les cas conformes (les objets sont réellement déconnectés) de ceux non conformes (les objets devraient se toucher), nous utilisons la distance d'erreur *dmp* comme distance seuil entre les frontières des polygones en dessous de laquelle la déconnexion est considérée comme une erreur dans les données. Les données sont corrigées en modifiant la position des sommets qui composent les objets, pour que les polygones concernés partagent des frontières sans se superposer.
- En ce qui concerne enfin la relation entre bâtiments et parcelles, il est logique d'obtenir des intersections entre ces deux types d'objets. Cependant, quand un bâtiment intersecte très légèrement une parcelle, il est possible que cette intersection soit une erreur de données. Pour différencier les cas conformes (bâtiments et parcelles s'intersectent réellement) et non conformes (bâtiments et parcelles ne s'intersectent pas dans la réalité), nous utilisons à nouveau la distance d'erreur *dmp* comme distance seuil entre les frontières voisines en dessous de laquelle les données enregistrées doivent être modifiées. Les données sont corrigées en modifiant la position des sommets qui composent les objets, pour que le polygone associé au bâtiment se trouve inclus dans le polygone associé à la parcelle.

III.3.2 Construction des tables relationnelles de relations

La construction des tables de relations (bati_bati, parcelle_parcelle, route_route, bati_parcelle, parcelle_route, bati_route) entre les différents éléments (bâtis, parcelles, routes) est basée sur des requêtes associant les objets 2 à 2. À titre d'exemple, les requêtes suivantes illustrent l'approche utilisée :

- Relation d'adjacence entre objets (cf. Requête 1)

Elle concerne les relations bati_bati, parcelle_parcelle, route_route, et parcelle_route,

La requête est basée sur la fonction *ST_Touches* (*geom1*, *geom2*), avec *geom1* et *geom2* les géométries des deux objets pris en compte. Cette fonction retourne VRAI si le contour extérieur du premier objet intersecte le contour extérieur du second objet et que leurs intérieurs respectifs ne s'intersectent pas.

Requête 1 : Liste des paires de bâtiments adjacents

```

insert into Bati_Bati
select a.id_bati id_bati1, b.id_bati id_bati2
from batis a, batis b
where a.id_bati != b.id_bati
and ST_Touches(a.geom_bati, b.geom_bati);

```

- Relation d'intersection entre objets (cf. Requête 2)

Elle concerne la relation bati_parcelle.

La requête est basée sur la fonction *ST_Intersects* (*geom1*, *geom2*) et *ST_Touches* (*geom1*, *geom2*), avec *geom1* et *geom2* les géométries des deux objets pris en compte. La fonction *ST_Intersects* retourne VRAI si les deux objets s'intersectent et la fonction *ST_Touches* retourne VRAI si le contour extérieur du premier objet intersecte le contour extérieur du second objet et que leurs intérieurs respectifs ne s'intersectent pas. Cette dernière permet d'exclure les objets qui n'ont que leurs contours extérieurs qui s'intersectent.

Requête 2 : Liste des paires (bâtiment, parcelle) qui se superposent

```

insert into Bati_Parcelle
select a.id_bati id_bati, b.id_parcelle id_parcelle
from batis a, parcelles b
where ST_Intersects (a.geom_bati, b.geom_parcelle)
and not ST_Touches(a.geom_bati, b.geom_parcelle);

```

- Relation entre bâtiment et route (cf. Requête 3)

Dans notre modèle, il n'existe pas de relation topologique connectée entre bâtiments et routes. La relation entre ces deux objets est identifiée au travers des parcelles. En conséquence, l'insertion des données dans la table Bati_Route s'effectue à partir des tables Bati_Parcelle et Parcelle_Route, le numéro d'identifiant de parcelle servant de jointure entre les bâtis et les routes.

Requête 3 : Association des bâtis avec les routes

```

insert into Bati_Route
select distinct a.id_bati, b.id_route
from bati_parcelle a, parcelle_route b
where a.id_parcelle = b.id_parcelle;

```

III.4 Implémentation des opérateurs

III.4.1 Centralité

La mesure de C_{BPR} que nous avons introduit dans le chapitre II.3.1 est calculée pour chaque bâtiment et est stockée dans le champ CBPR de la table Bâtis. Cette mesure reposant sur la somme de 5 types de chemins possibles entre deux bâtiments, à savoir « adjacents et sur la même parcelle », « adjacents, mais dans différentes parcelles », « non adjacents, mais situés sur une même parcelle », « non adjacents, et situés sur des parcelles adjacentes » et « non adjacents, mais sur des parcelles connectées à une même route », nous présentons séparément les cinq sous requêtes permettant de gérer chaque cas.

- Bâtiments adjacents et sur la même parcelle

La première configuration possible entre deux bâtiments est « adjacents et sur la même parcelle ». La Requête 4 utilise la table Bati_Parcelle pour obtenir, pour un bâtiment donné, premièrement l'ensemble des parcelles sur lesquelles il repose, puis deuxièmement pour obtenir l'ensemble des bâtiments reposant sur ces parcelles. On obtient alors, pour un bâtiment, l'ensemble des bâtiments qui chevauchent les mêmes parcelles que lui. La requête extrait alors de la table Bati_Bati, les bâtiments adjacents qui appartiennent à cet ensemble.

Requête 4 : Bâtiments adjacents et sur une même parcelle

```

update Batis b set CBPR =  $\alpha_1$  (
  select count(*)
  from Bati_Bati bb
  where b.id_bati=bb.id_bati1
  and bb.id_bati2 in (
    select bp.id_bati
    from Bati_Parcelle bp
    where bp.id_parcelle in (
      select bp.id_parcelle
      from Bati_Parcelle bp
      where b.id_bati=bp.id_bati)))

```

- Bâtiments adjacents, mais sur des parcelles différentes

La seconde configuration entre deux bâtiments est « adjacents, mais dans différentes parcelles ». De manière très analogue à la Requête 4, la Requête 5 utilise la table Bati_Parcelle pour obtenir,

pour un bâtiment, l'ensemble des bâtiments qui chevauchent les mêmes parcelles que lui. La requête extrait alors de la table Bati_Bati, les bâtiments adjacents qui n'appartiennent pas à cet ensemble.

Requête 5 : Bâtiments adjacents, mais sur différentes parcelles

```
+  $\alpha_2$  * (  
  select count(*)  
  from Bati_Bati bb  
  where b.id_bati=bb.id_bati1  
  and bb.id_bati2 not in (  
    select bp.id_bati  
    from Bati_Parcelle bp  
    where bp.id_parcelle in (  
      select bp.id_parcelle  
      from Bati_Parcelle bp  
      where b.id_bati=bp.id_bati)))
```

- Bâtiments non adjacents, mais sur la même parcelle

La troisième configuration entre deux bâtiments est « non adjacents, mais situés sur une même parcelle ». Pour un bâtiment, la Requête 6 extrait de la table Bati_Parcelle les bâtiments situés sur la même parcelle que lui, et enlève de cet ensemble les bâtiments qui lui sont adjacents (c.-à-d. apparaissent dans la table Bati_Bati).

Requête 6 : Bâtiments non adjacents mais sur la même parcelle

```
+  $\alpha_3$  * (
  select count(distinct bp.id_bati)
  from Bati_Parcelle bp
  where bp.id_bati <> b.id_bati
  and bp.id_parcelle in (
    select bp.id_parcelle
    from Bati_Parcelle bp
    where b.id_bati = bp.id_bati)
  and bp.id_bati not in (
    select id_bati2
    from Bati_Bati bb
    where b.id_bati = bb.id_bati1))
```

- Bâtiments non adjacents, mais sur des parcelles adjacentes

La quatrième configuration entre deux bâtiments est « non adjacents et situés sur des parcelles adjacentes ». Dans un premier temps, la Requête 7 utilise la table Bati_Parcelle pour obtenir, pour un bâtiment, dans un premier temps l'ensemble des parcelles sur lesquelles il repose, puis dans un second temps l'ensemble des parcelles adjacentes à celles-ci (c.-à-d. qui apparaissent dans la table Parcelle_Parcelle). On obtient alors, pour un bâtiment, l'ensemble des bâtiments qui sont sur la même parcelle que lui ou sur les parcelles adjacentes. Dans un second temps, la requête soustrait de ce premier ensemble tous les bâtiments adjacents (c.-à-d. présents dans la table Bati_Bati) ou qui sont sur la même parcelle que le bâtiment évalué.

Requête 7 : Bâtiments non adjacents, mais sur des parcelles adjacentes

```

+  $\alpha_4$  * (
  select count(*)
  from (
    select distinct bp.id_bati
    from Bati_Parcelle bp
    where bp.id_parcelle in (
      select distinct pp.id_parcelle2
      from Parcelle_Parcelle pp
      where pp.id_parcelle1 in (
        select bp.id_parcelle
        from Bati_Parcelle bp
        where b.id_bati=bp.id_bati )
      )except (
        select distinct bp.id_bati
        from Bati_Parcelle bp
        where bp.id_parcelle in (
          select bp.id_parcelle
          from Bati_Parcelle bp
          where b.id_bati=bp.id_bati)
        )union (
          select bb.id_bati2
          from Bati_Bati bb
          where b.id_bati=bb.id_bati1)
    )temp)

```

- Bâtiments non adjacents, mais sur des parcelles connectées à une même route

La cinquième configuration considérée entre deux bâtiments est « non adjacents, mais sur des parcelles connectées à une même route ». Dans un premier temps, la Requête 8 utilise la table Bati_Route pour obtenir, pour un bâtiment, l'ensemble des bâtiments connectés aux mêmes routes que lui. Dans un second temps, la requête soustrait de cet ensemble les bâtiments qui sont sur les mêmes parcelles que le bâtiment évalué (c.-à-d. présents dans la table Bati_Parcelle) ou sur les parcelles adjacentes (c.-à-d. présents dans la table Parcelle_Parcelle).

Requête 8 : Bâtiments non adjacents, mais sur des parcelles connectées à une même route

```
+  $\alpha_5$  * (
  select count(*)
  from (
    select distinct id_bati
    from Bati_Route
    where id_route in (
      select br.id_route
      from Bati_Route br
      where id_bati=b.id_bati)
    except (
      select distinct bp.id_bati
      from Bati_Parcelle bp
      where bp.id_parcelle in (
        (select distinct pp.id_parcelle2
        from Parcelle_Parcelle pp
        where pp.id_parcelle1 in (
          select bp.id_parcelle
          from Bati_Parcelle bp
          where bp.id_bati=b.id_bati)
        union (
          select id_parcelle
          from Bati_Parcelle bp
          where id_bati=b.id_bati))))))
  )temp)
```

III.4.2 Mesure de connectivité locale du bâti

La mesure de connectivité locale du bâti que nous avons introduit dans le chapitre II.3.2 est calculée pour chaque route et est stockée dans le champ CR de la table Routes. Cette mesure permet d'évaluer pour chaque route connectée à des bâtiments, le nombre moyen de routes auxquels sont connectés ces bâtiments. La Requête 9 utilise la table Bati_Route pour déterminer d'une part, pour une route, l'ensemble des bâtiments qui y est connecté, et d'autre part, pour un bâtiment, le nombre de routes auquel il est connecté. Enfin, un score de 0 est attribué à toutes les routes n'étant connectées à aucun bâtiment.

Requête 9 : Mesure de connectivité locale du bâti

```
update Routes r set CR = (  
  select avg  
  from (  
    select br.id_route , avg(nb_routes)  
    from Bati_Route br, (  
      select id_bati, count(id_route) as nb_routes  
      from Bati_Route  
      group by id_bati ) bati_nb_route  
    where br.id_bati=bati_nb_route.id_bati  
    group by br.id_route ) resultat  
  where r.id_route=resultat.id_route);  
update Routes set CR = 0 where CR is null;
```

III.4.3 Mesure de densité de bâti relative à une route

La mesure de *densité de bâti* introduite dans le chapitre II.3.3 mesure pour chaque route le nombre de bâtiments connectés, relativement à la longueur de la route. Cette valeur est pondérée en répartissant le poids d'un bâtiment sur l'ensemble des routes auxquelles il est connecté.

La Requête 10 utilise la table Bati_Route d'une part pour déterminer, pour une route, l'ensemble des bâtiments qui y est connecté, et d'autre part pour comptabiliser le nombre de routes auquel est connecté un bâtiment. La longueur de la route est issue par la fonction spatiale ST_Length (geom) qui retourne une distance euclidienne, dont l'unité dépend du système de projection utilisé.

Requête 10 : Mesure de densité de bâti relative à une route

```

update Routes set densite_bati = t3.nb_bati/(select
ST_Length(r.geom_route_polyligne))
from (
  select id_route, sum(taux_partage) as nb_bati
  from (
    select br.id_route, br.id_bati, 1/(nb_lien_route) taux_partage
    from Bati_Route br, (
      select id_bati, cast(count(id_route)as double precision) as
nb_lien_route
      from Bati_Route
      group by id_bati) t1
    where br.id_bati=t1.id_bati) t2
    group by id_route) t3
  where routes.id_route=t3.id_route;
update Routes set densite_bati = 0 where densite_bati is null;

```

III.4.4 Metric Compactness

La mesure de Metric Compactness introduite dans le chapitre II.3.4 repose sur la mesure de *Ratio Minimum Distance* qui, à partir d'une route, mesure la plus petite distance permettant d'atteindre X pourcent de routes. Pour une valeur de X , la mesure de *Minimum Ratio Minimum Distance* correspond à la plus petite des valeurs de *Ratio Minimum Distance* mesurées sur l'ensemble des routes. Enfin, la mesure de Metric Compactness, calculée pour chaque route, correspond au nombre de routes atteintes en parcourant cette *Minimum Ratio Minimum Distance*. Les résultats de Metric Compactness sont stockés dans le champ MC de la table Routes.

Nous utilisons la table Route_Route_Distance pour calculer la *Minimum Ratio Minimum Distance* pour un X choisi, puis nous utilisons à nouveau la table Route_Route_Distance pour compter, pour chaque route, le nombre de routes atteintes (i.e. la Metric Compactness) en parcourant la *Minimum Ratio Minimum Distance* calculée.

Requête 11 : Mesure de Metric Compactness

```
update Routes set MC = nb_route_atteinte
from (
  select id_route1, count(distance)as nb_route_atteinte
  from Route_Route_Distance
  where distance <= (
    select distance from (
      select row_number() over(partition by id_route1 order by
id_route1, distance)as ligne, id_route1, distance
      from Route_Route_Distance)t1
      where t1.ligne=cast((((select count(*) from routes)-1)* X /
100.0) as bigint)
      order by distance limit 1)
    group by id_route1
    order by nb_route_atteinte desc)t2
  where id_route=id_route1;
update routes set MC = 0 where MC is null;
```

III.5 Application et résultats

III.5.1 Présentation des données géographiques

Les données géographiques utilisées dans notre cas d'étude sont celles du territoire de la commune de Guisseny (Figure 20), situé au nord-ouest de la France. Territoire à dominante rurale et agricole, cette commune a vu son urbanisation modifiée depuis une trentaine d'années par l'attractivité du littoral. Ainsi, alors que le centre du village (le bourg) et les espaces les plus éloignés de la côte ont conservé leur aspect, la côte s'est fortement modifiée avec la construction d'un grand nombre de maisons (habitations principales ou secondaires) occupées par une population attirée par le lieu. L'intérêt pour notre étude est donc de regrouper sur un même territoire trois physionomies différentes : rurale, pavillonnaire et compacte.

Délimitée au nord par la mer (la Manche), la carte est composée de 1026 bâtiments, 2598 parcelles cadastrales, et 244 tronçons de route reliés au niveau de 213 intersections. La zone est caractérisée par une dispersion du bâti et par une répartition du réseau de circulation sur l'ensemble du territoire. L'urbanisation est plus importante à proximité de la mer (au nord de la région) et le centre du village est clairement associé à une concentration plus importante de bâtiments. Sur le reste du territoire, le bâti se compose de petits hameaux et de bâtiments dispersés.

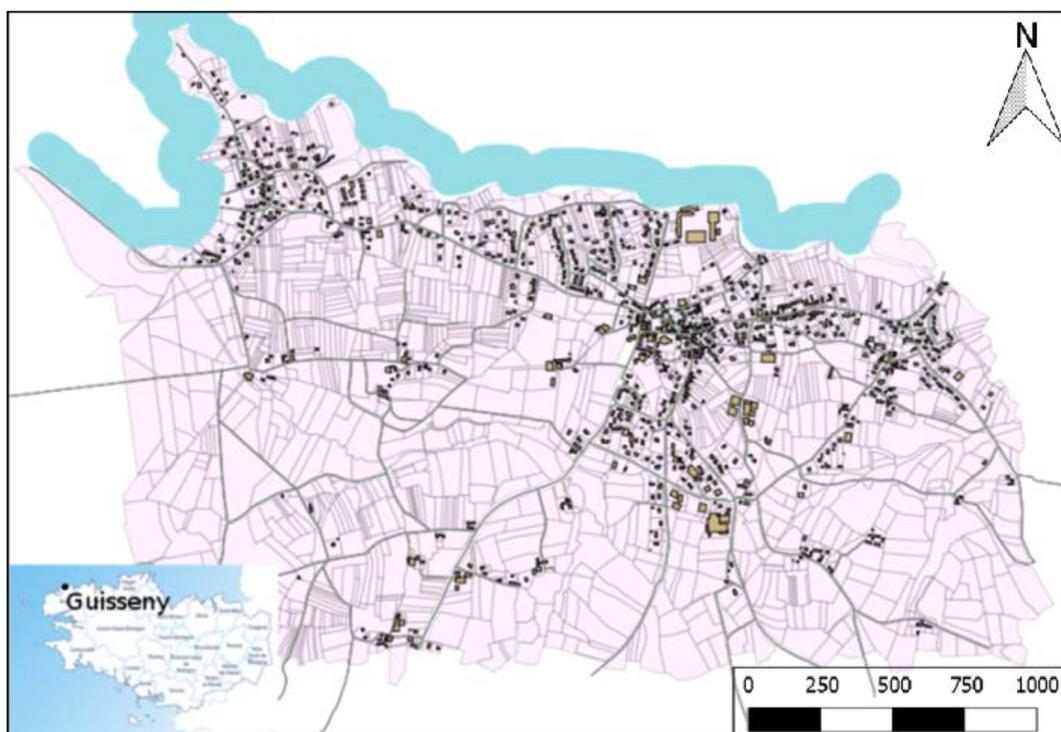


Figure 20 : Territoire de Guisseny

Les données du bâti et du parcellaire sont extraites du Plan Cadastral Informatisé (PCI) et de la base de données MAJIC associée, entretenue par la Direction Générale des Finances Publiques (DGFIP). Le réseau routier quant à lui est extrait de la BD Topo de l'IGN (Institut national de l'information géographique et forestière). L'ensemble de ces données ont été enregistrées dans une base de données PostgreSQL/Postgis dont le modèle correspond à la Figure 18. Les géométries du bâti et du parcellaire sont des polygones tandis que celles du réseau routier sont des polygones converties en polygones.

III.5.2 Mesure de centralité du bâti

Afin d'illustrer l'intérêt de la mesure de centralité, l'analyse des résultats se fait dans un premier temps d'un point de vue global (c.-à-d. sur toute la carte), puis dans un second temps seulement sur des zones plus locales permettant une analyse plus fine et des comparaisons plus détaillées des résultats.

III.5.2.1 Mesures sur toute la carte

- Mesure de Centralité de degré

La mesure de centralité de degré que nous avons présentée dans le chapitre I.2.3.1 mesure pour chaque sommet le nombre d'arêtes qui s'y connecte. Concernant l'analyse du bâti, l'approche la plus simple est d'appliquer cette mesure au graphe de bâti qui reflète les associations entre les bâtiments qui sont physiquement reliés par des murs mitoyens (cf. Figure 21).

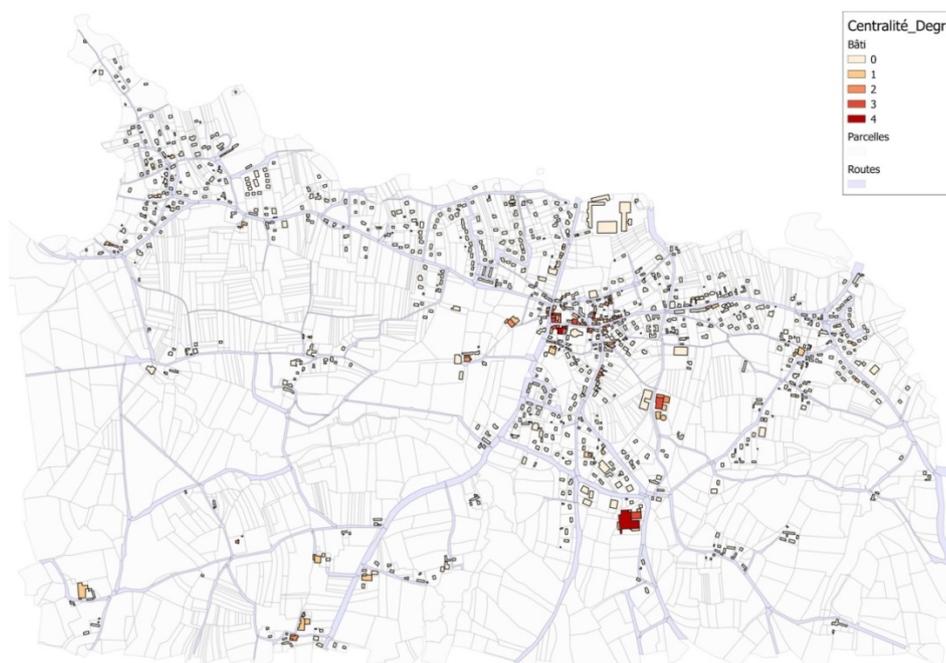


Figure 21 : Centralité de degré appliquée au graphe de bâti

Le résultat montre que 164 bâtiments sont connectés par adjacence avec au maximum 4 autres bâtiments :

- 121 bâtiments sont adjacents à 1 seul bâtiment
- 32 bâtiments sont adjacents à 2 bâtiments
- 9 bâtiments sont adjacents à 3 bâtiments
- 2 bâtiments sont adjacents à 4 bâtiments

En complément, 862 bâtiments ne sont pas connectés et sont considérés comme isolés.

Les bâtiments mitoyens sont localisés essentiellement au centre du village, et en dehors ils correspondent uniquement à des bâtiments de grandes dimensions (bâtiments agricoles en l'occurrence) ou localisés dans quelques hameaux. Par contre, sur la zone côtière, les bâtiments sont essentiellement non mitoyens. Ces premiers éléments renseignent sur l'évolution du type d'urbanisation, entre des constructions plus anciennes regroupées au centre du village ou dans de petits hameaux, et des constructions plus récentes avec des habitants réclamant moins de promiscuité et préférant des habitations non mitoyennes.

Bien que l'analyse reste pertinente pour décrire le bâti, son intérêt est limité concernant la structuration du territoire. En particulier, le graphe utilisé contient trop de nœuds non connectés, et confirme la nécessité d'analyser un graphe connexe plus complet (c.-à-d. le graphe intégral).

- Mesure de Centralité sur le graphe intégral

La mesure de centralité sur le graphe intégral, C_{BPR} , que nous avons introduite dans le chapitre II.3.1 est une mesure reposant sur la distinction de 5 types de chemins possibles entre deux bâtiments (adjacents et sur une même parcelle, adjacents mais dans différentes parcelles, non adjacents mais situés sur une même parcelle, non adjacents mais situés sur des parcelles adjacentes, non adjacents mais sur des parcelles connectées à une même route). Chaque type de chemin est associé à un coefficient α_i défini par l'utilisateur qui pondère l'importance relative des chemins. La mesure de C_{BPR} retourne pour chaque bâtiment la somme de chacun de ces 5 types de chemin.

Dans une première approche, nous considérons que tous les types de chemins sont équivalents et donc que tous les coefficients sont égaux : $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 1$ (cf. Figure 22). Ainsi dans ce cas, que deux bâtiments soient adjacents et sur la même parcelle ou seulement connectés à une même rue, le poids résultant dans leurs scores de C_{BPR} sera le même : 1. Sur les 1026 bâtiments, seuls 75 obtiennent une valeur de C_{BPR} supérieur à 30. Ils sont localisés au sud et à

l'Est du bourg et sont connectés à des tronçons de route plutôt longs (> 200 mètres). Cette observation nous amène à nous interroger sur le lien entre la longueur totale des routes à laquelle un bâtiment est lié et le score de C_{BPR} obtenu. Le résultat illustré sur la Figure 23 ne permet pas d'établir une relation claire entre le score de C_{BPR} et la somme des longueurs des routes. En effet, notre jeu de données correspond à un milieu à dominante rurale composé de maisons individuelles où il existe un nombre important de bâtiments relativement isolés indépendamment de la longueur des routes associées. On peut alors s'interroger pour savoir si le rapport entre centralité et longueur de route serait différent entre les zones rurales et les zones urbanisées. Ce point sera évoqué plus tard (cf. chapitre III.5.4).

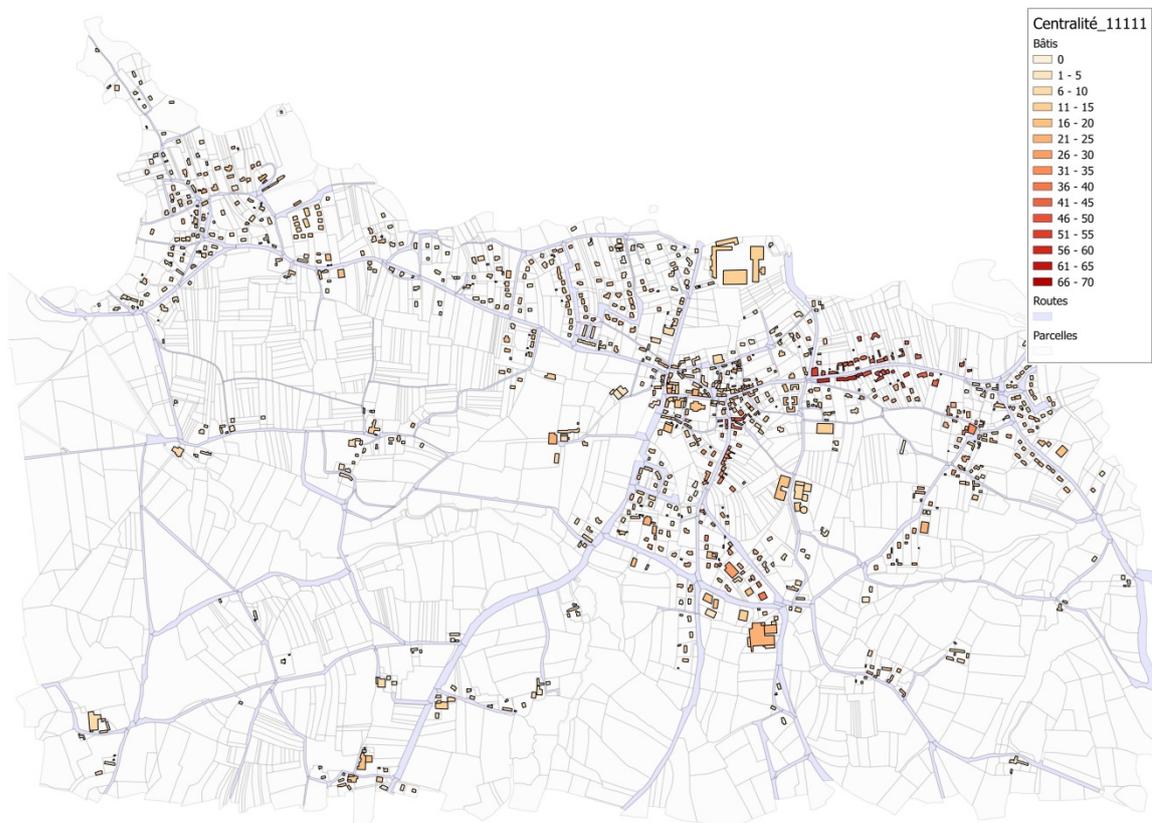


Figure 22 : C_{BPR} avec $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 1$

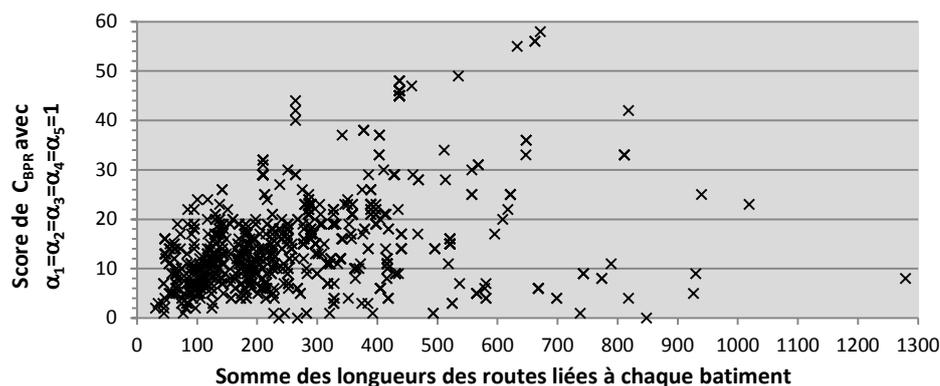


Figure 23 : Rapport entre centralité et longueur de route

La deuxième approche cherche à prendre en compte la notion de hiérarchie des chemins, basée sur l'idée que les liens entre les bâtiments connectés (mitoyens) sont plus forts que les liens entre bâtiments déconnectés (distants) et que l'association de deux bâtiments situés sur des parcelles connexes est plus importante que le lien entre deux bâtiments localisés sur des parcelles distantes. La Figure 24 illustre le résultat de C_{BPR} avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$. Ainsi dans ce cas, lorsque deux bâtiments sont seulement connectés via une rue, le poids résultant dans le score de C_{BPR} sera toujours 1. Par contre pour les autres configurations, au lieu d'un poids de 1, le poids sera plus important : 2 pour deux bâtiments non adjacents mais situés sur des parcelles adjacentes, 3 lorsqu'ils sont non adjacents mais situés sur une même parcelle, 4 s'ils sont adjacents mais dans différentes parcelles, et 5 s'ils sont adjacents et sur une même parcelle. Après avoir effectué la somme des poids sur les 1026 bâtiments, 169 obtiennent une valeur de C_{BPR} supérieur à 30. Comme précédemment, on observe la même concentration de bâtiments à forte centralité au sud et à l'Est du bourg, mais on note en plus des bâtiments à forte centralité dispersés plus aléatoirement sur le reste du territoire sans localisation particulière.

Entre les deux approches, les valeurs de centralité sont sensiblement différentes. La Figure 25 permet d'observer une différence importante du nombre de bâtiments pour les classes 1-5 et 6-10 avec respectivement une diminution de 84 et 112 bâtiments entre $C_{BPR_{11111}}$ et $C_{BPR_{54321}}$. A l'inverse pour les classes 21-25 et 26-30 on note une augmentation de respectivement 59 et 79 bâtiments entre $C_{BPR_{11111}}$ et $C_{BPR_{54321}}$. L'ajout d'une pondération dans les valeurs de α_i amène une diminution du nombre de bâtiments ayant une faible centralité ($C_{BPR_{54321}} < 11$) et à l'inverse une augmentation du nombre de bâtiments avec une forte centralité ($C_{BPR_{54321}} > 21$).

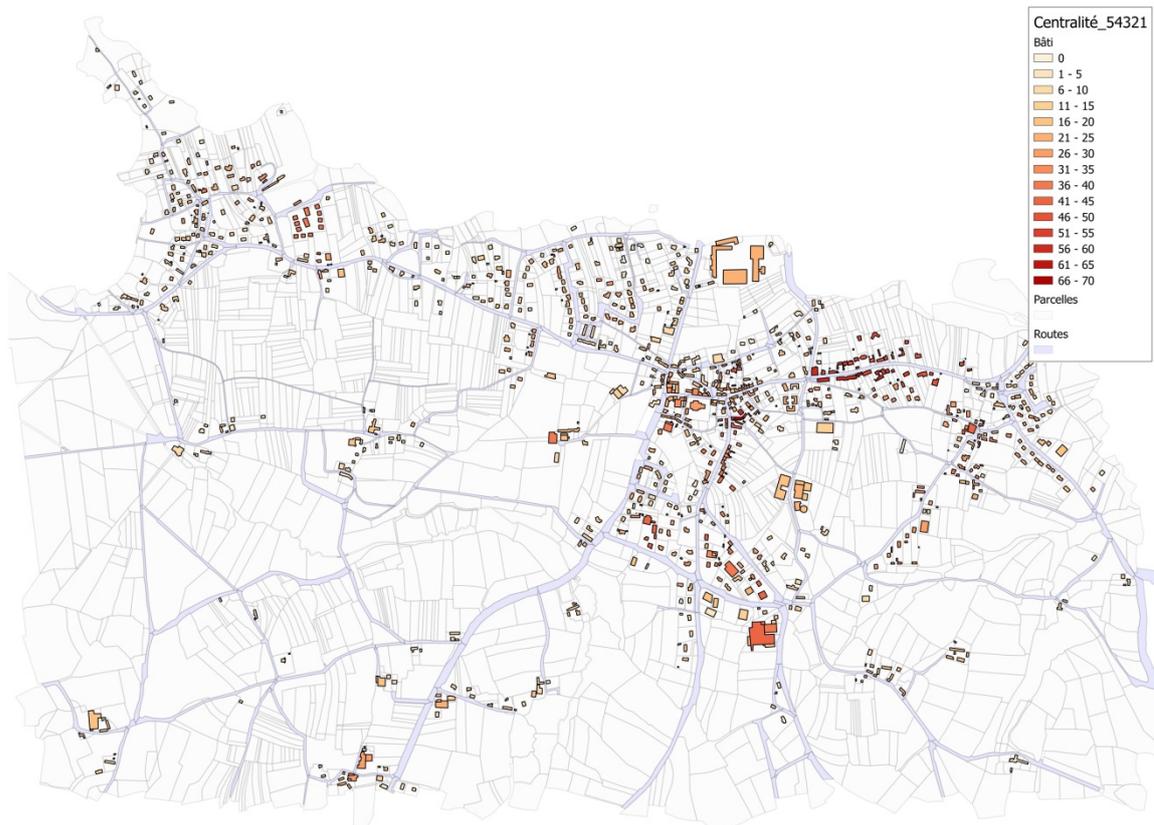


Figure 24 : C_{BPR} avec $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$

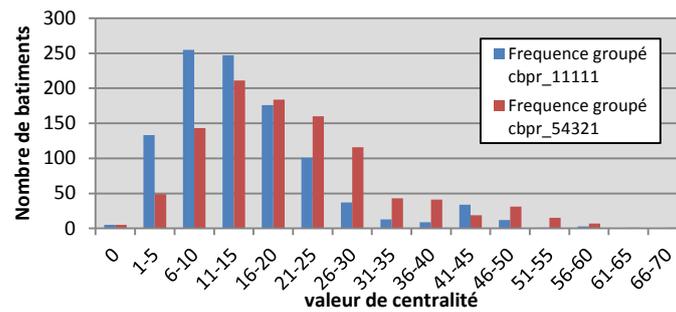


Figure 25 : Histogrammes des valeurs de C_{BPR} pour $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 1$ et $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$

Afin de préciser ces différences, nous nous intéressons à la différence de valeurs de C_{BPR} obtenue par chaque bâtiment (Figure 26 et 27). Dans tous les cas, la valeur obtenue pour $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 1$ est inférieure ou égale à la valeur obtenue pour $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$. Dans 70% des cas, la différence est inférieure à 6. Ainsi, le choix de valoriser les connexions impliquant les parcelles aux dépens des relations via les routes a donc peu d'effet sur la valeur de centralité. Ceci indique que le nombre de bâtiments mitoyens ou voisins par parcelle est faible. Sur la Figure 26, on observe que les bâtiments pour lesquels la différence est la plus importante sont localisés essentiellement dans le centre bourg du village et sont souvent de grands bâtiments (en l'occurrence des bâtiments agricoles type hangar ou serres).

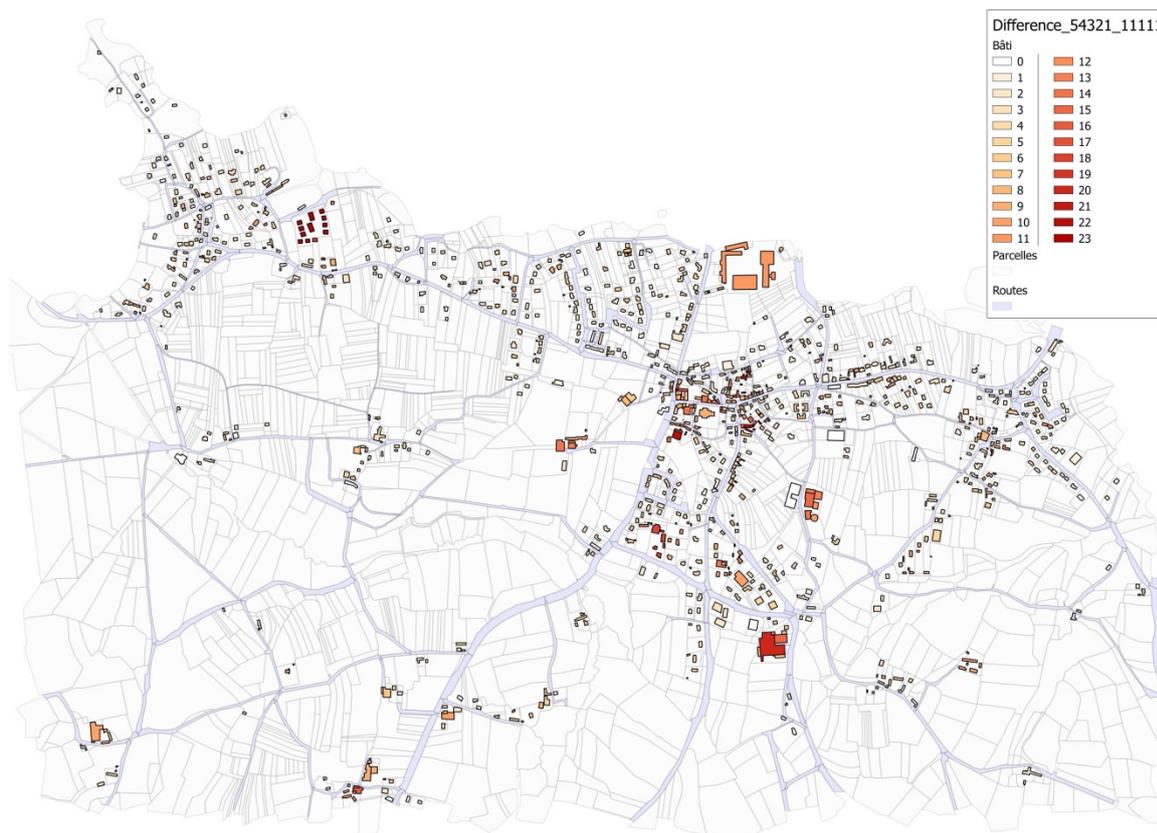


Figure 26 : Différence des valeurs de C_{BPR} obtenus avec $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$ et avec $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 1$

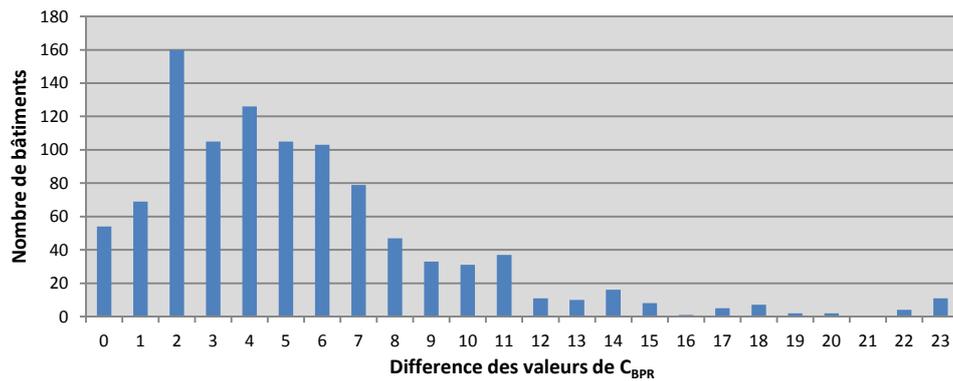


Figure 27 : Histogramme des différences de valeurs de C_{BPR}

III.5.2.2 Analyse de détail

Afin de préciser les résultats obtenus et analyser la différence entre une approche non pondérée (c.-à-d. les valeurs de α_i sont égales) et une approche pondérée (c.-à-d. les valeurs de α_i sont différentes) nous nous focalisons sur trois zones (Figure 28).

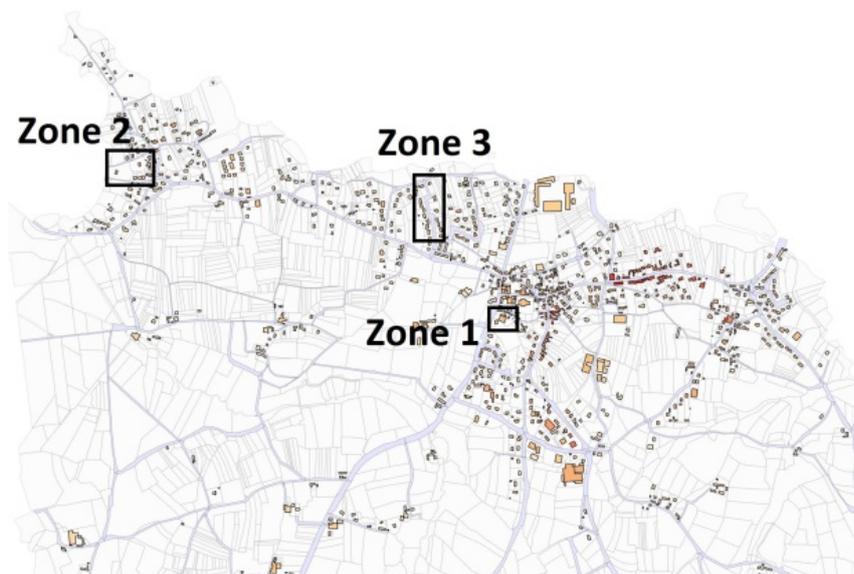


Figure 28 : Localisation des zones 1, 2 et 3

- Analyse de la Zone 1

La zone 1 est localisée au niveau du centre du village de Guisseny. Les scores de C_{BPR} varient entre 10 et 20 pour les coefficients $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 1$, et entre 13 et 40 pour les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$ (Figure 29).

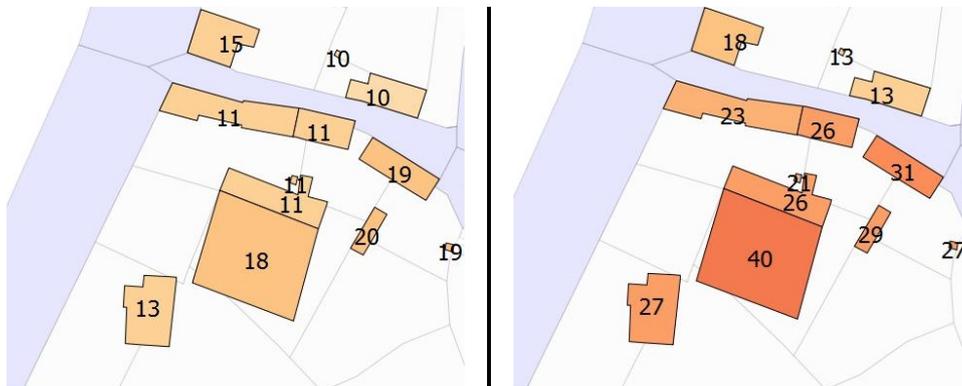


Figure 29 : C_{BPR} avec tous les coefficients à 1 à gauche et C_{BPR} avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$ à droite

Comme le montre la Figure 30, l'intérêt est que cette zone présente à la fois des bâtiments ayant une différence de score faible (3), des bâtiments ayant une différence de score modérée (8 à 15) et un bâtiment ayant une différence de score importante (22).

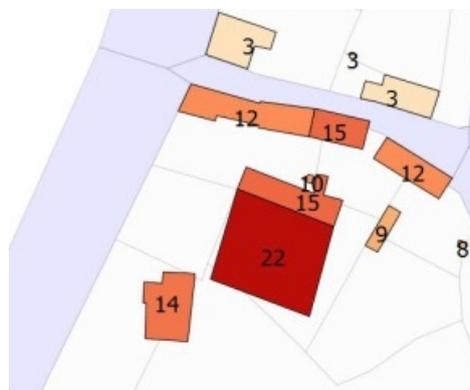


Figure 30 : Différence des scores de C_{BPR}

Les trois bâtiments au nord obtiennent une très faible augmentation, car la majorité de leurs connexions se font par les routes :

- le bâtiment au nord-ouest n'obtient qu'une augmentation de 3 au score de C_{BPR} car il est relié à trois bâtiments via les parcelles adjacentes à la parcelle sur laquelle il repose (les deux bâtiments à l'est et un bâtiment au nord (hors champ de vision)).

- les deux bâtiments au nord-est ont une augmentation de 2, car ils reposent sur la même parcelle et une augmentation de 1, car ils sont sur une parcelle adjacente au bâtiment du nord-ouest.

Les autres bâtiments obtiennent une augmentation modérée ou importante pour différentes raisons :

- Les deux bâtiments au sud (+14 et +22) reposent sur une grande parcelle créant de nombreux liens avec les bâtiments reposant sur les parcelles adjacentes ce qui produit une augmentation de 12 et 10. Étant sur la même parcelle sans être adjacent, cela génère également une augmentation de 2.
- Les bâtiments au centre (+22 et +15) et au nord (+12 et +15) ont un bâtiment adjacent sur la même parcelle permettant une augmentation de 4 au score de C_{BPR} .
- La proximité des bâtiments crée des liens entre les bâtiments de type non adjacents mais sur la même parcelle ou sur des parcelles adjacentes générant respectivement des augmentations de score de C_{BPR} de +6/+8 et +2/+4 selon les bâtiments.

L'augmentation importante du score de C_{BPR} du bâtiment au centre (+22) résulte du cumul de ces liens avec les autres bâtiments et permet de mettre en avant sa centralité vis-à-vis des autres bâtiments (Figure 29).

- Analyse de la Zone 2

La zone 2 est localisée au Nord-Ouest du domaine d'étude (cf. Figure 28). Elle représente un îlot habité, entouré par des routes non adjacentes. La majorité des bâtiments sont disjoints et répartis sur la presque totalité des parcelles. Parmi les 5 types de relations que nous avons différenciés, seule la relation « adjacents et sur une même parcelle » n'est pas présente.

La Figure 31 illustre les résultats de la centralité de degré. La majorité des bâtiments obtiennent un score de 0 et seulement quatre bâtiments obtiennent un score de 1. L'information est laconique et permet de discrétiser uniquement deux configurations : Bâtiments adjacents ou Bâtiments disjoints.



Figure 31 : Valeur de centralité de degré

La mesure de C_{BPR} (Figure 32 et Figure 33) permet d'obtenir un score discriminant pour l'ensemble des bâtiments. Les valeurs obtenues prennent en compte des éléments (bâti, route, parcelle) qui n'apparaissent pas sur la figure. Dans le cas où le poids (α_i) des différents types de relations sont toujours égaux à 1 (Figure 32), la valeur de C_{BPR} est comprise entre 7 et 19. Dans le cas où le poids des différents types de relations est différent (Figure 33), la valeur de C_{BPR} est comprise entre 7 et 27.

Dans les deux cas :

- Les bâtiments connectés à plusieurs routes obtiennent les scores les plus grands.
- Le bâtiment au centre de l'image obtient la valeur la plus grande
- Les deux bâtiments localisés dans le coin Sud-Est obtiennent des valeurs importantes
- Il n'existe pas de lien entre la distance séparant les bâtiments et leur valeur de centralité.

Ces résultats mettent en évidence l'intérêt du graphe intégral pour préciser les relations entre les bâtiments. En effet, l'approche permet de différencier clairement les objets les mieux connectés, indépendamment des distances qui les séparent.



Figure 32 : C_{BPR} avec tous les coefficients α_i égaux à 1



Figure 33 : C_{BPR} avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$

Il est intéressant d'observer que le bâtiment le plus à l'ouest obtient un score important dans les deux cas (12 et 14) alors qu'il semble visuellement être isolé du reste des autres bâtiments. À l'inverse, le bâtiment au nord-ouest obtient en score beaucoup plus faible (7 dans les deux cas) alors qu'il semble mieux intégré à l'ensemble (c.-à-d. plus proche des voies de circulation et des autres bâtiments). Cette différence illustre l'intérêt de l'approche choisie ici pour mettre en valeur l'intégration d'un objet sans se limiter à la proximité au réseau de communication et des autres bâtiments. Ainsi, c'est la connexion via le parcellaire qui prédomine dans ce cas et permet de mieux apprécier la centralité de l'objet dans l'ensemble de l'espace.

Entre les deux configurations (Figure 32 et Figure 33) la valeur de C_{BPR} augmente pour tous les bâtiments, excepté pour un seul dont la valeur reste égale à 7. Celui-ci a pour particularité spatiale de se trouver isolé par des routes des autres bâtiments. Les augmentations les plus importantes correspondent d'une part aux bâtiments mitoyens (+9 au nord, et +6 au sud), et d'autre part au bâtiment le plus connecté via les parcelles, localisé au centre de l'image (+8). Ceci est à associer au choix qui est fait ici d'attribuer une pondération plus importante aux relations d'adjacences entre bâtiments et entre parcelles. À l'inverse, pour les bâtiments reliés essentiellement par des connexions via les routes, la différence entre les deux configurations est faible. Ce résultat montre clairement l'intérêt de valoriser différemment le poids des types de connexion. Cette approche permet de mieux discrétiser les valeurs de centralité et de préciser la qualité des relations.

- Analyse de la Zone 3

La zone 3 est localisée au Nord du territoire de Guisseny dans la frange littorale (cf. Figure 28). Elle représente une organisation linéaire (type lotissement pavillonnaire) avec de part et d'autre de la route une succession régulière d'habitations le plus souvent disjointes.

Dans le cas où tous les coefficients de pondération sont équivalents (Figure 34) les bâtiments ayant les plus fortes valeurs de centralité sont soit localisés à l'intersection des voies de circulation (c.-à-d. au nord de la zone), soit adjacents. De part et d'autre de la route, une partie importante des bâtiments obtiennent la même valeur (i.e. 16), ce qui suppose qu'ils sont dans la même configuration. En fait, ils ont des contextes différents : un ou plusieurs bâtiments par parcelle, un ou plusieurs bâtiments sur les parcelles adjacentes.

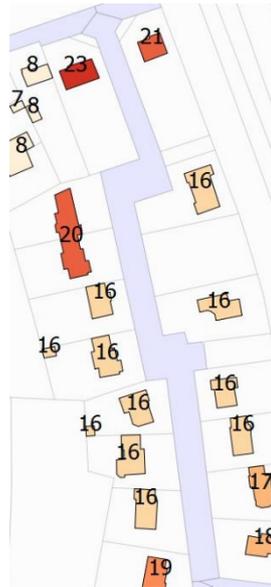


Figure 34 : C_{BPR} avec tous les coefficients α_i égales à 1



Figure 35 : C_{BPR} le long d'une route avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$

Avec une pondération différente pour les types de relation, le résultat est sensiblement différent (Figure 35) : les bâtiments proches des intersections et/ou adjacents ont toujours une très forte centralité ; par contre, pour les bâtiments le long de la route, les valeurs de centralité sont moins homogènes et permettent de différencier plusieurs organisations spatiales. Ce résultat montre clairement l'intérêt de différencier les relations pour décrire la structure spatiale. Bien que les

coefficients α_i aient été choisis empiriquement, l'approche pondérée offre une vision plus précise de la structure de l'espace.

III.5.2.3 Application de la mesure de centralité du bâti à un sous graphe du graphe intégral

La zone 4 englobe le tracé d'une route à la transition entre la zone littorale et la zone rurale (Figure 36). L'extrémité nord de la route nord présente une succession régulière de bâtiments localisés de part et d'autre de la route, alors que l'extrémité sud correspond à l'entrée dans un hameau d'une dizaine de bâtiments. Entre les deux extrémités, la route serpente entre des parcelles sans aucun bâtiment.

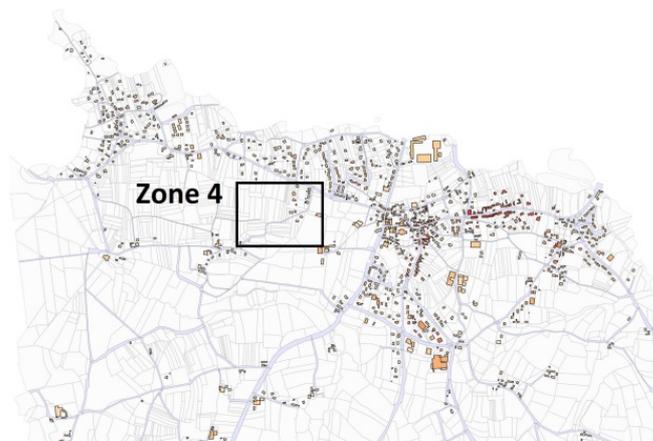


Figure 36 : Localisation de la zone 4

Avec l'approche pondérée (Figure 37), sans surprise, les bâtiments proches des intersections et voisins de parcelles construites présentent les valeurs les plus fortes, et à l'inverse, les bâtiments isolés ont des valeurs de centralité plus faible. Cette règle ne semble pas vraie pour le bâtiment localisé à l'extrémité sud de la route. En effet, bien qu'apparemment isolé, la mesure de centralité de ce bâtiment est importante (22). Ceci s'explique aisément, en précisant qu'il est fortement connecté à des bâtiments localisés en dehors de l'image. Il faut noter que ceux-ci sont localisés sur des parcelles non connectées à la route qui nous intéresse ici. Ainsi, à l'échelle de cette route, la différence de centralité entre ses deux extrémités n'apparaît pas.

Afin de répondre à cette lacune, il est possible d'utiliser l'opérateur de centralité, non pas sur la totalité du graphe, mais seulement sur un sous-graphe correspondant aux bâtiments et parcelles en relation de connexion avec la route étudiée. Un tel exemple, représenté sur la Figure 38, permet de montrer clairement la différence de résultat. Ainsi, dans ce cas la valeur de centralité diminue

progressivement vers le sud, pour atteindre la valeur la plus faible pour le bâtiment isolé localisé à l'extrémité sud. La valeur obtenue (i.e. 15) est alors égale seulement au nombre de bâtiments localisés sur des parcelles connectées à la route (1 bâtiment n'apparaît pas sur l'image).

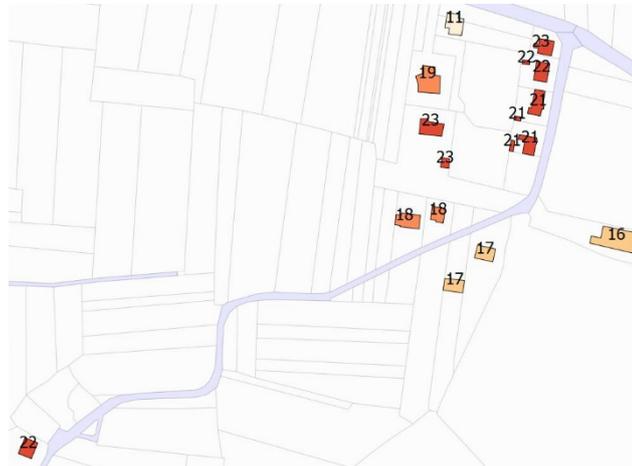


Figure 37 : C_{BPR} avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$ appliqué à l'ensemble du graphe intégral



Figure 38 : C_{BPR} avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$ appliqué sur un sous-graphe

III.5.3 Mesure de connectivité locale du bâti

La mesure de connectivité locale du bâti que nous avons introduit dans le chapitre II.3.2 permet d'évaluer pour chaque route si les bâtiments auxquels elle est connectée sont connectés à d'autres routes.

Une route ne partageant aucun de ses bâtiments avec une autre route obtiendra un score de 1. Si une route obtient un score de 2 par exemple, cela signifie qu'en moyenne l'ensemble des bâtiments qui lui sont connectés sont partagés avec 2 routes. Nous signalons que cette mesure ne dépend pas du nombre de bâtiments connectés à la route.



Figure 39 : Mesure de connectivité locale du bâti

Sur la zone d'étude, les valeurs sont comprises entre 0 et 4.3. Telle qu'illustrée sur la Figure 39, cette mesure permet de localiser dans l'espace les taux de partage du bâti de chaque route. On peut y observer que les scores ne sont pas clairement influencés par le type de territoires (c.-à-d. rurale, pavillonnaire et compacte). En effet, que l'on se situe sur le littoral dans le centre du village ou dans les zones plus rurales ou agricoles, les valeurs sont équivalentes. Ainsi sur la Figure 40, qui montre le détail des valeurs obtenues au sud du bourg du village de Guisseny, on observe une forte variabilité des scores de connectivité locale du bâti, allant de 1,17 à 3 sur la

zone. Les scores les plus élevés correspondent aux routes les plus courtes fortement connectées. Cette dernière tendance se retrouve sur l'ensemble du territoire étudié où l'on observe une relation entre la valeur de connectivité locale du bâti et longueur de route (Figure 41). La valeur de connectivité locale du bâti n'apparaît donc pas comme un marqueur du type de territoire, mais comme une information locale à l'échelle de la route.

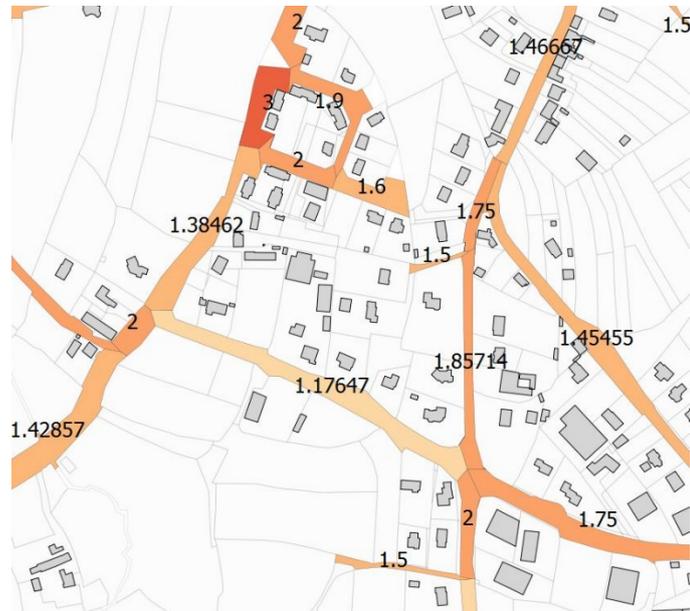


Figure 40 : Mesure de connectivité locale du bâti

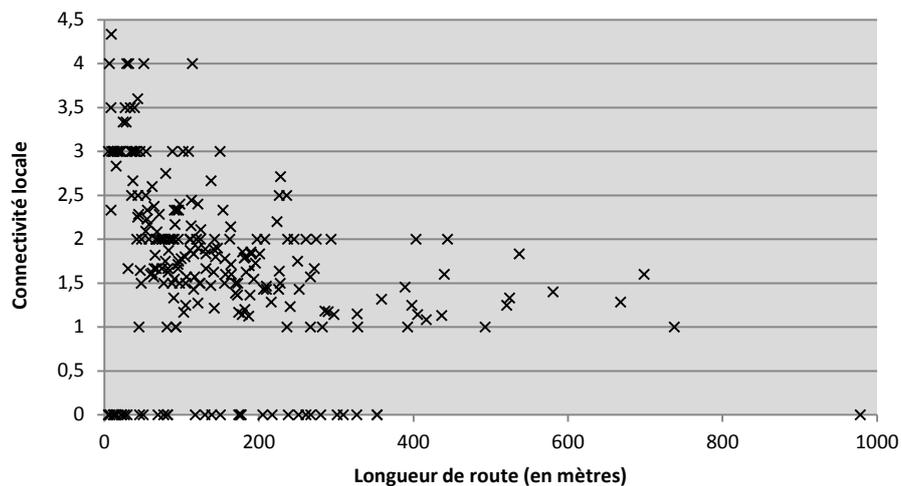


Figure 41 : Mesure de connectivité locale du bâti et longueur de route

III.5.4 Mesure de densité de bâti

Contrairement à la mesure de connectivité locale du bâti (voir ci-dessus), la mesure de densité de bâti introduite dans le chapitre II.3.3 ($BRD(r_j)$) dépend essentiellement du nombre de bâtiments connectés à chaque route. Le poids de chaque bâtiment est reparti de manière équitable entre chaque route à laquelle il est connecté.

La Figure 42 montre la distribution des valeurs de densité sur l'ensemble du réseau. Les valeurs les plus basses sont principalement localisées à l'ouest et au sud de la zone, dans la partie la plus rurale du territoire, tandis que les valeurs les plus hautes font ressortir les zones les plus urbanisées : le bourg du village et la zone littorale.

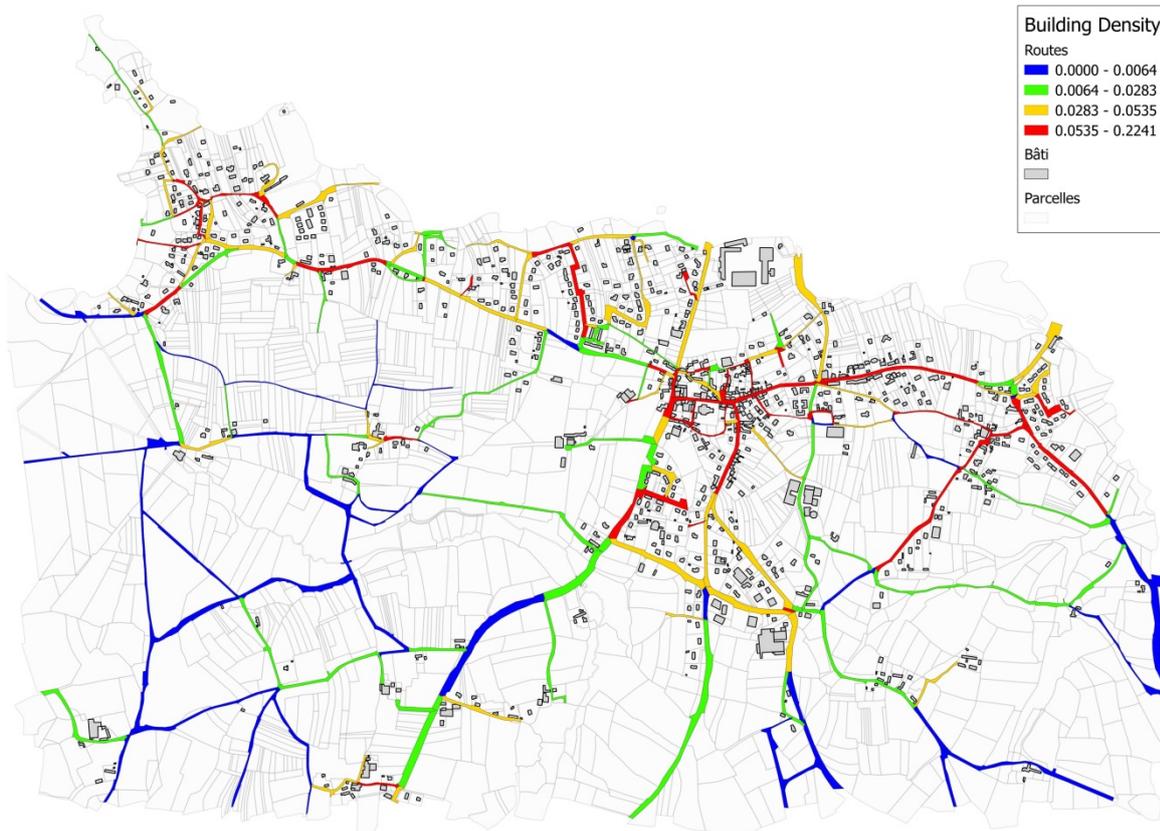


Figure 42 : Building Road Density

À partir de la valeur de densité de bâti, il est possible de préciser la relation qui pourrait exister entre la longueur de route et la valeur de centralité mesurée sur le bâtiment connecté à cette route. En effet, en considérant d'une part les routes à plus forte densité ($BRD(r_j) > 0.0501$ (valeur

médiane)) et d'autre part les routes à plus faible densité ($BRD(r_j) < 0.0501$) on distingue deux comportements sensiblement différents (Figure 43). Ainsi, dans le cas des routes à faible densité (c.-à-d. dans la partie rurale du territoire) il n'existe pas de tendance forte qui associe longueur de route et valeur de centralité ; par contre, pour les zones fortement bâties, la tendance est claire, avec des valeurs de C_{BPR} plus fortes pour les routes les plus longues.

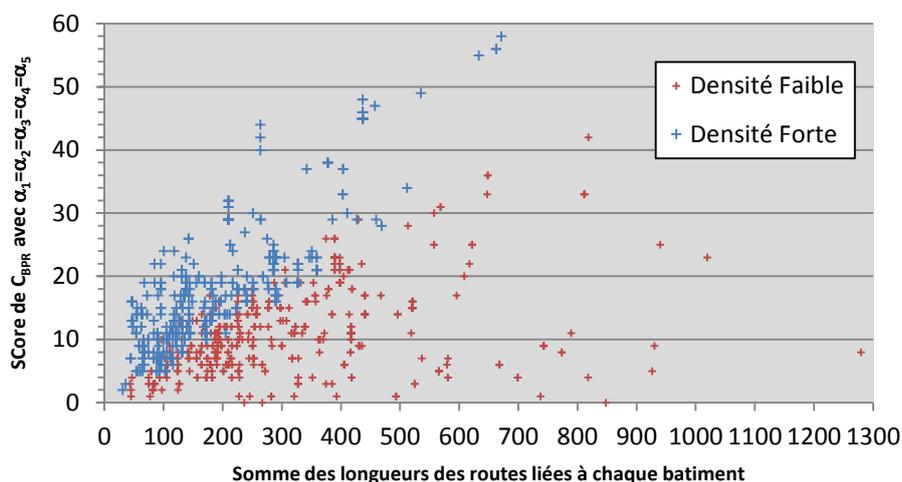


Figure 43 : Score de centralité et longueur de route par score de densité (idem Figure 23)

En comparaison des mesures de densités rappelées au chapitre I.2.1, la mesure de densité de bâti proposée permet d'obtenir une mesure à un niveau local grâce à une approche permettant de la mesurer sur les tronçons de route, et en évitant les problèmes de segmentation de l'espace tel que la taille de la grille ou le découpage d'objets par exemple.

III.5.5 Mesure de compacité

III.5.5.1 Variation de la Metric Compactness

Dans le chapitre II.3.4, nous avons introduit une série de mesures dans l'objectif de distinguer les tronçons de route fortement connectés de ceux faiblement connectés en mesurant la capacité d'une route à être reliée à d'autres parties du réseau routier. Pour rappel, $D_{min}^X(r_i)$ représente la plus petite distance permettant à une route r_i d'atteindre X pourcents du réseau routier et D_{min}^X représente la plus petite de ces distances pour l'ensemble des routes. Enfin $MC(r_i)$ représente le nombre de routes atteintes en parcourant la distance D_{min}^X sur le réseau routier à partir du milieu de r_i .

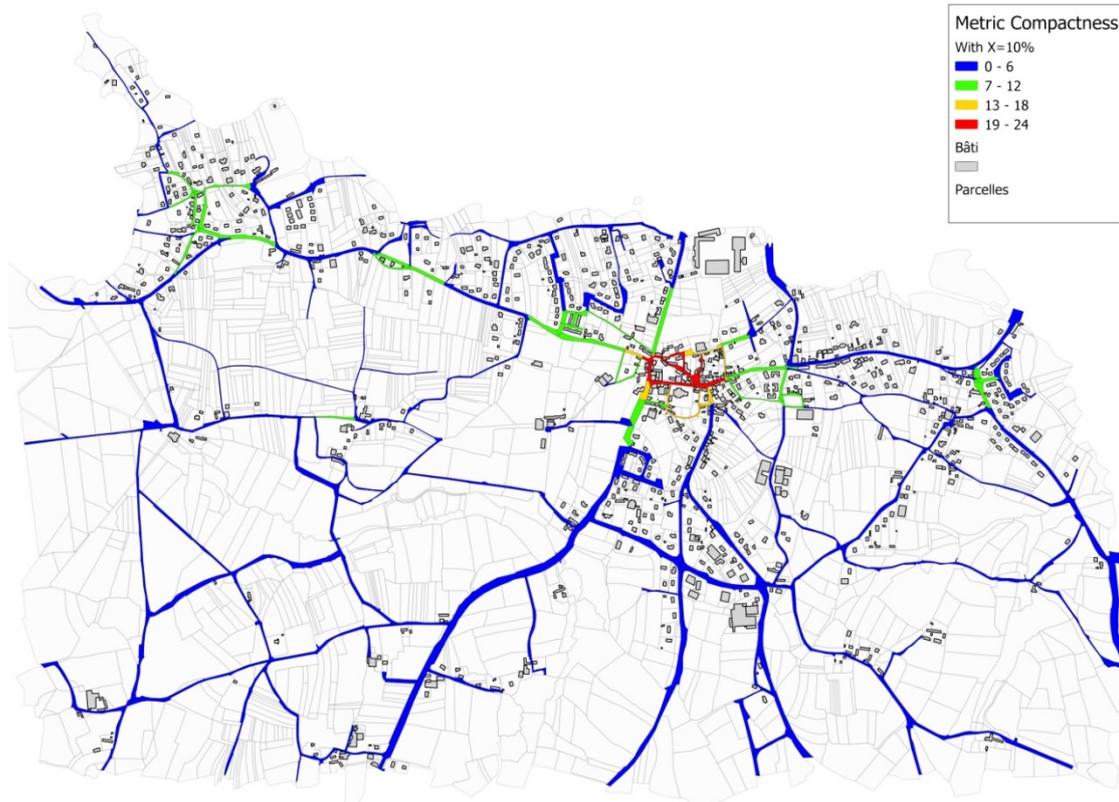
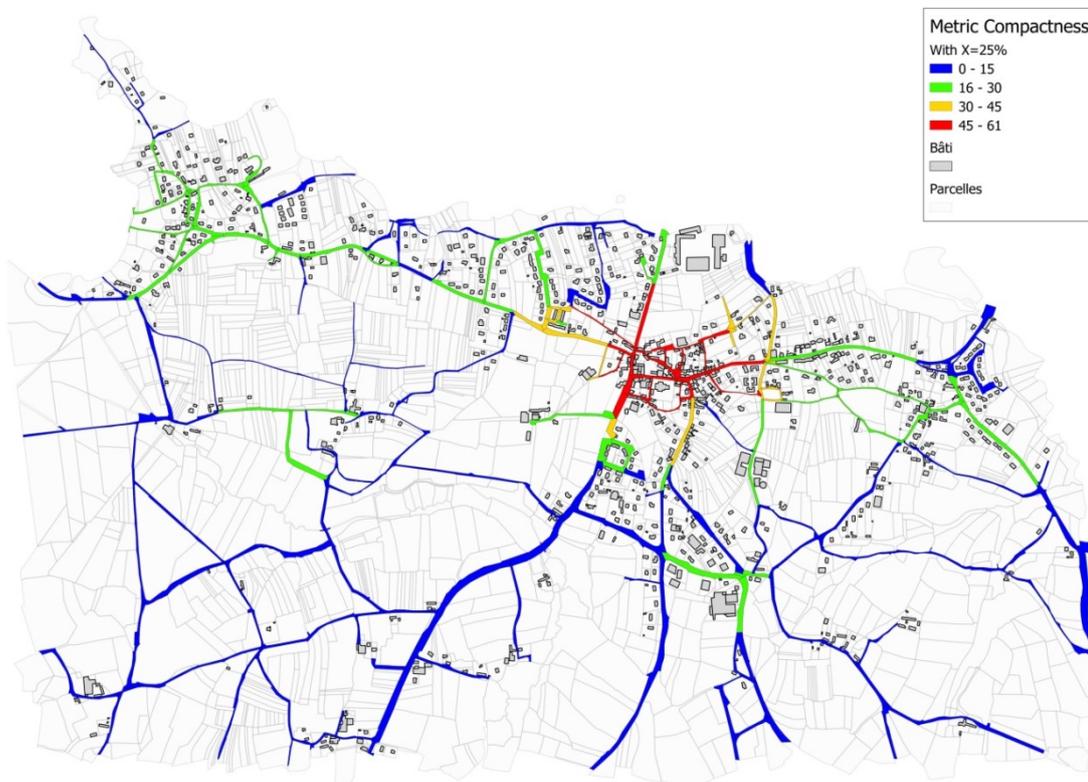
Les figures 44, 45, 46 et 47 présentent les résultats des mesures de la Metric Compactness pour quatre valeurs de X : 10%, 25%, 45% et 100%. On considère dans cette approche qu'un tronçon est fortement connecté s'il atteint plus de 75% du nombre de routes défini par X . Il apparaît en rouge sur les figures. Ainsi, dans l'exemple donné Figure 44, pour $X = 10\%$ (i.e : 24 tronçons), on considère qu'un tronçon est fortement connecté s'il atteint au moins 18 autres tronçons du réseau (c.-à-d. $MC(r_i) > 18$).

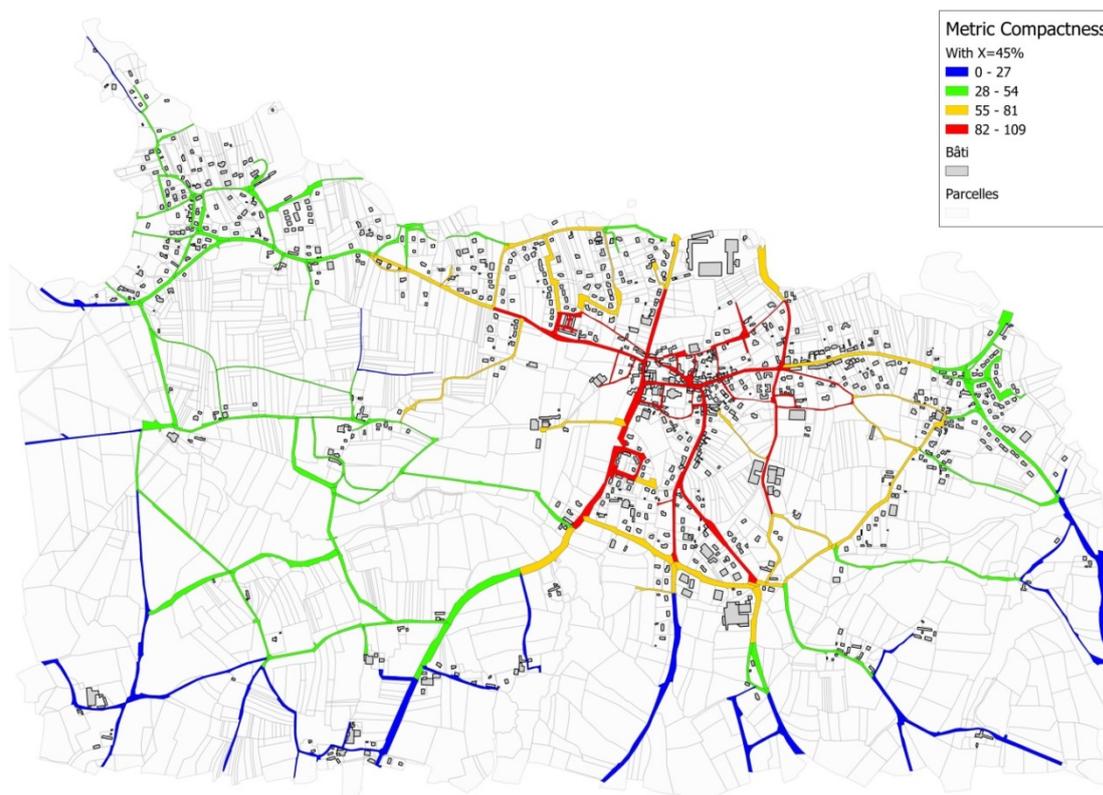
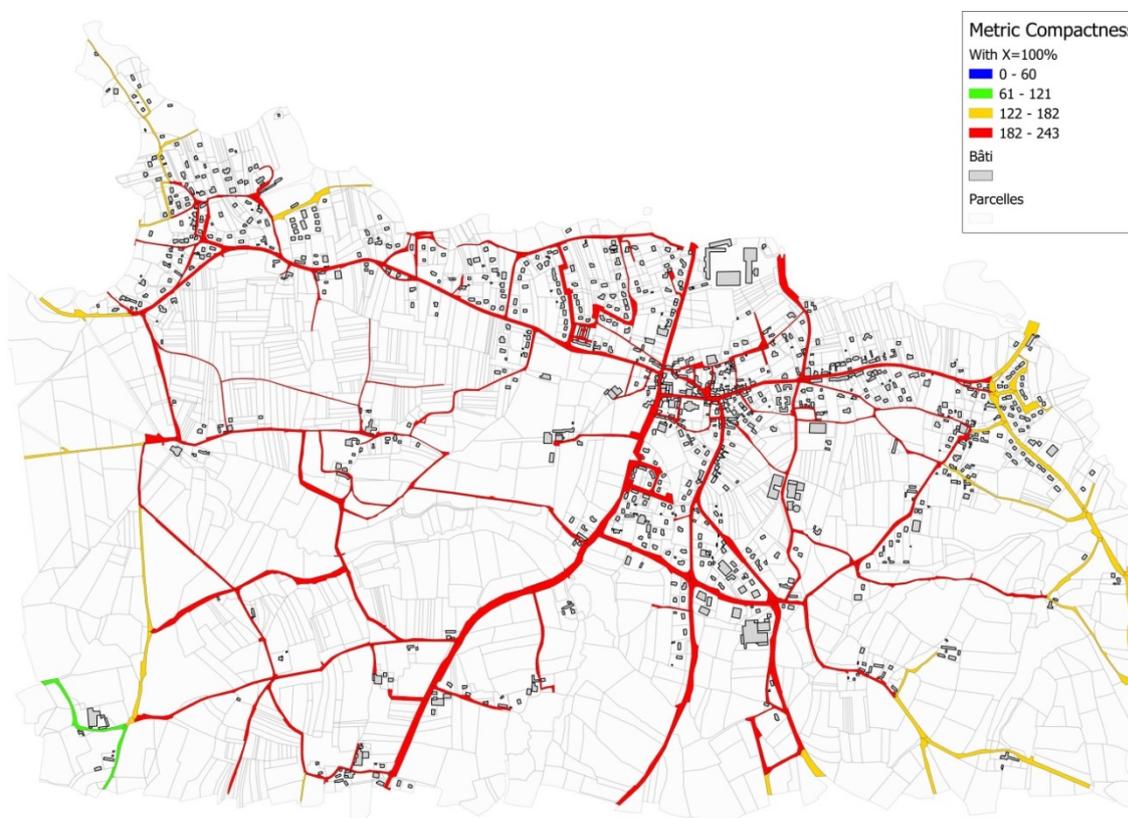
Pour $X < 45\%$ (Figure 46), la méthode permet d'associer aisément les tronçons les mieux connectés au bourg du village. Centrée sur ce bourg, l'extension géographique et le nombre de routes augmentent régulièrement avec X . À l'extrême (i.e. $X=100\%$, Figure 47), c'est l'ensemble du territoire qui apparaît bien connecté, ce qui n'est pas significatif.

La Figure 48 permet de différencier la répartition du nombre de routes sur l'ensemble des valeurs de Metric Compactness possibles. Sur les 244 tronçons :

- pour $X=10\%$, seuls 26 tronçons sont fortement connectés (c.-à-d. $MC(r_i) > 18$).
- pour $X=25\%$, 52 tronçons sont fortement connectés (c.-à-d. $MC(r_i) > 45$),
- pour $X=45\%$, 94 tronçons sont fortement connectés (i.e $MC(r_i) > 82$).

Les figures 44, 45 et 46 montrent que sur le territoire de Guisseny, la localisation des tronçons fortement connectés (en rouge) se concentre uniquement autour du bourg du village et qu'il n'est pas possible de distinguer d'autres lieux ayant les mêmes caractéristiques : l'organisation du territoire est très homogène, et il n'est pas possible de distinguer de sous-parties du graphe. L'analyse est différente si l'on s'intéresse aux tronçons de route qui peuvent atteindre au moins 25% du reste du réseau (en vert). Dans ce cas, pour $X=10\%$ (Figure 44) on peut distinguer 4 zones distinctes en dehors du bourg, réparties le long de la zone littorale. Ces résultats indiquent que la mesure de Metric Compactness permet d'identifier des sous-parties du réseau, mais cette identification est dépendante du choix de la valeur du paramètre X .

Figure 44 : Metric Compactness avec $X = 10\%$ Figure 45 : Metric Compactness avec $X = 25\%$

Figure 46 : Metric Compactness avec $X = 45\%$ Figure 47 : Metric Compactness avec $X = 100\%$

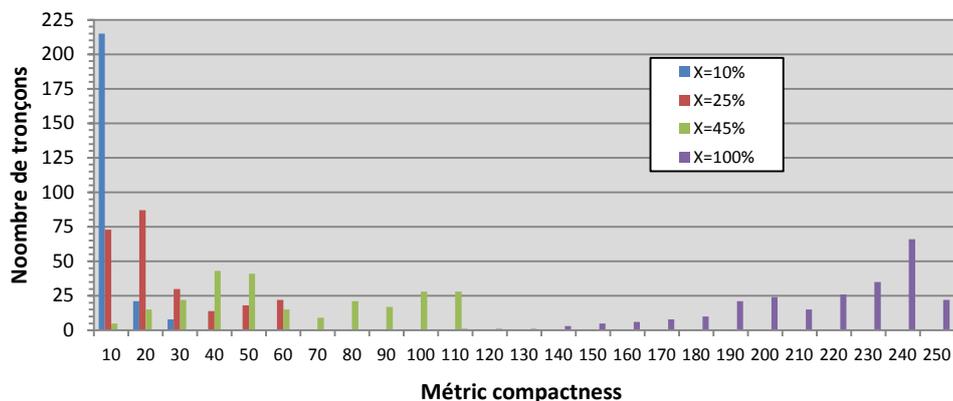


Figure 48 : Distribution des valeurs de Metric Compactness

Ce résultat est proche de la mesure de centralité de proximité décrite au chapitre I.2.3.2. En effet, dans les deux cas nous obtenons des cercles concentriques autour du centre du réseau routier lorsque la mesure est appliquée à l'ensemble du réseau. Une première différence entre les deux approches se situe au niveau de la sémantique du résultat. En effet, elle correspond à un nombre de routes dans notre approche ce qui est plus clair que l'inverse d'une somme de distance pour la centralité de proximité. En ce qui concerne les valeurs obtenues pour la Metric Compactness, on n'observe pas clairement de relation entre la longueur de tronçons et le nombre de tronçons atteints.

L'approche proposée pour introduire la Metric Compactness repose sur deux métriques ce qui nous permet de compléter l'analyse comme nous le montrons dans la section suivante.

III.5.5.2 Variation de D_{min}^X

Sur notre jeu de données, la croissance de D_{min}^X en fonction de X (Figure 49) est constante, présentant seulement une rupture de pente pour $X = 95\%$. La rupture de pente pour $X = 95\%$ est due à un effet de bord de notre jeu de données. On peut en conclure que l'augmentation de la distance à parcourir pour atteindre une proportion de plus en plus importante du nombre de tronçons du réseau est régulière. Selon le principe décrit dans la section II.3.5.1, ce résultat correspond donc à un réseau routier globalement homogène, c'est-à-dire sans sous-parties du réseau isolées (c.-à-d. mal connectées) du reste du réseau.

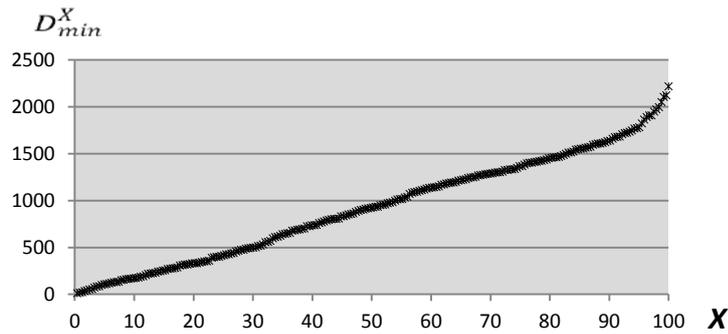


Figure 49 : Variations de D_{min}^X en fonction de X sur le territoire de Guisseny

Afin d'illustrer la capacité de la Metric Compactness à décrire la structure spatiale, nous avons modifié le jeu de données en supprimant 3 routes au centre du territoire afin d'isoler la partie Sud-Ouest de la partie Centre et Est. L'objectif est de créer deux sous-parties dans le réseau. Le résultat de cette modification sur la mesure de Metric Compactness est illustré sur la Figure 50, pour $X=45\%$ (à comparer avec la Figure 46). On observe qu'entre les deux configurations le résultat est très semblable : la localisation géographique et le nombre de routes fortement connectées sont les mêmes.

La Figure 51 permet de comparer la différence entre les deux configurations sur la relation entre D_{min}^X et X . On observe que les valeurs de D_{min}^X sont équivalentes pour $X < 50\%$. Au-delà, la variation de D_{min}^X est plus forte pour le jeu de données modifié. Ce résultat est à rapprocher de l'étude illustrée sur la Figure 15. On en déduit que la modification du réseau de routes crée un réseau hétérogène dont la dimension (le rayon) des sous-parties est de 950 mètres. Ainsi, pour des valeurs de X inférieures à 50%, les valeurs de Metric Compactness sont représentatives de la relation entre les tronçons distants de moins de 950 mètres. Ces relations sont limitées à l'intérieur des sous-parties du réseau et sont indépendantes des modifications que nous avons introduites en modifiant le jeu de données. C'est le cas illustré par la Figure 46 et la Figure 50 pour $X=45\%$. À l'inverse, pour $X > 50\%$, les valeurs de Metric Compactness sont représentatives de la relation entre les tronçons distants de plus de 950 mètres. Elles sont donc influencées par notre modification du jeu de données. Les figures 52 et 53 avec $X = 85\%$, illustrent cette conséquence. Les résultats montrent clairement la différence de structuration du territoire, avec dans un cas, une structure centrée sur le bourg du village, vers lequel convergent les voies de circulation de façon homogène, et dans l'autre cas, une zone côtière très fortement connectée et une zone rurale particulièrement isolée dans sa partie Sud-Ouest créant un déséquilibre très fort dans le territoire.

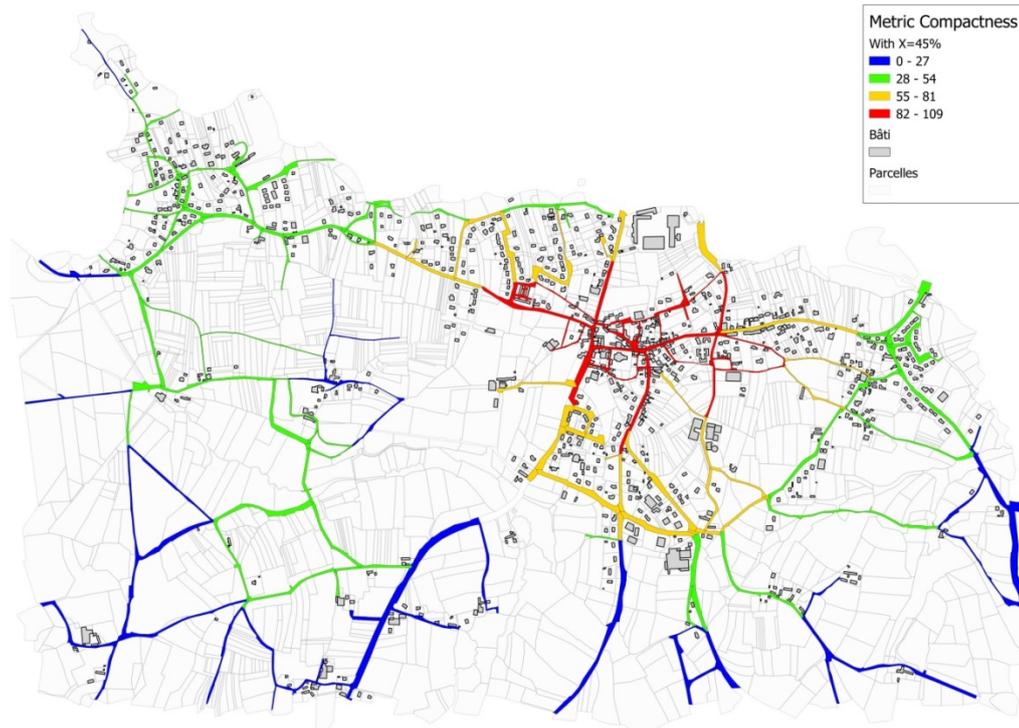


Figure 50 : Metric Compactness avec $X = 45\%$ pour le jeu de données modifié

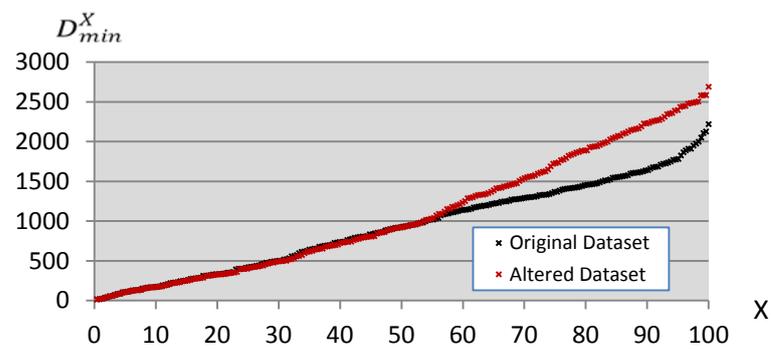


Figure 51 : Variations de D_{min}^X en fonction de X pour le jeu de données d'origine et modifié

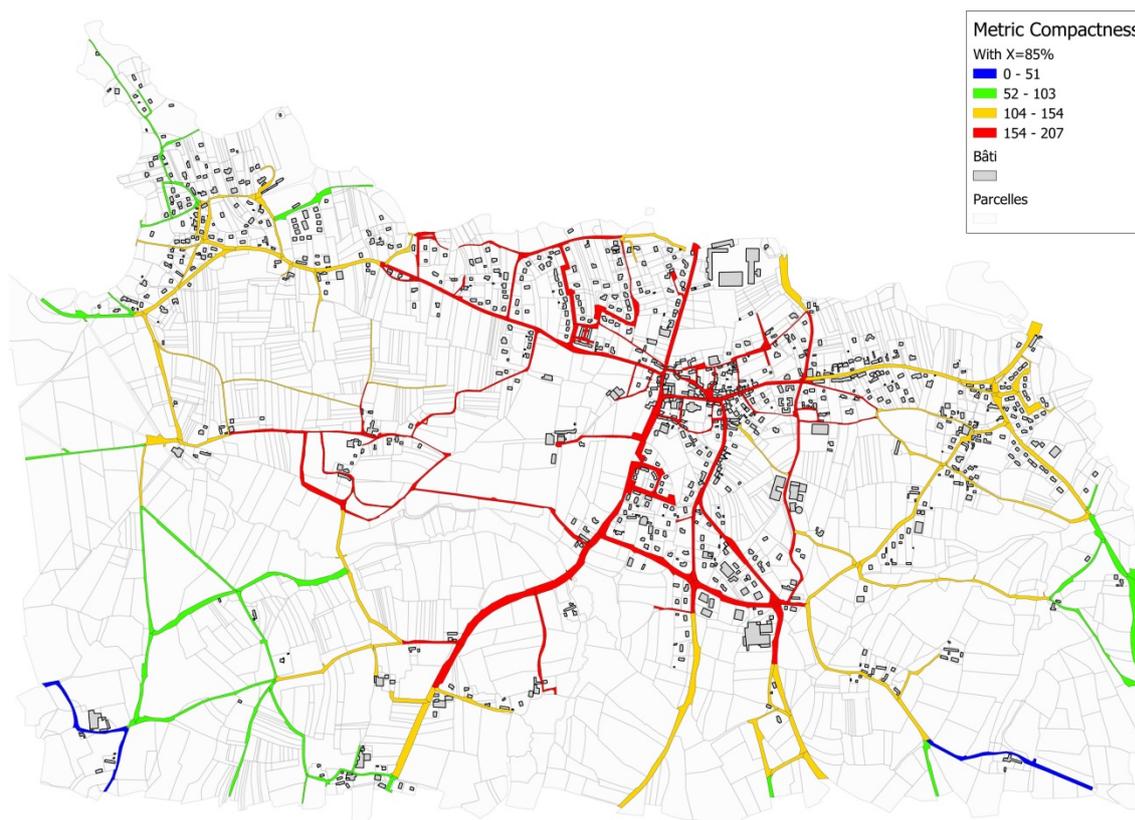


Figure 52 : Metric Compactness avec $X = 85\%$ pour le jeu de données d'origine

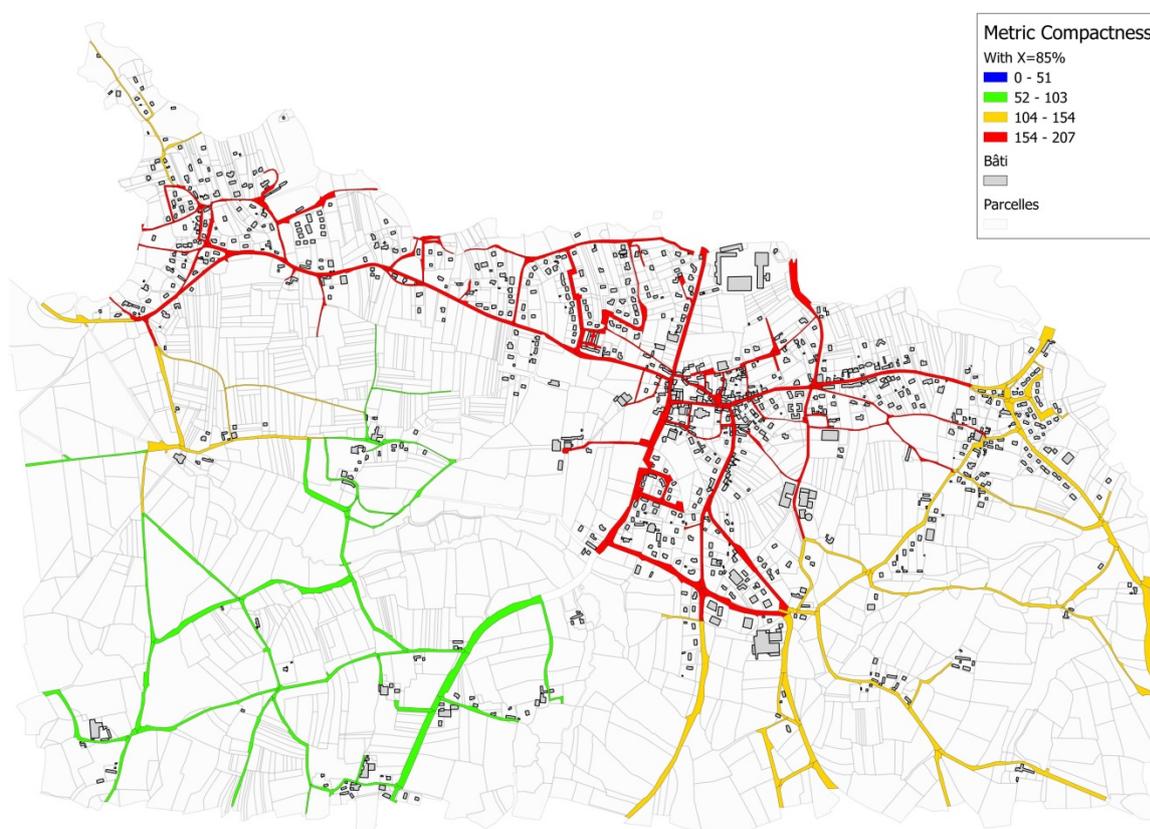


Figure 53 : Metric Compactness avec $X = 85\%$ pour le jeu de données modifié

III.5.6 Corrélation entre la mesure de densité de bâti et la mesure de compacité

Dans les deux sections précédentes, nous avons proposé des opérateurs permettant d'une part d'estimer la capacité des voies de circulation à atteindre d'autres voies de circulation (la Metric Compactness), et d'autre part de mesurer la quantité de bâtis liés à chaque voie de circulation ($BRD(r_i)$). Dans cette section nous cherchons à identifier une corrélation entre ces 2 paramètres afin d'établir les éléments de caractérisation de l'espace regroupant la connaissance sur le réseau de circulation, le parcellaire et le bâti.

Concernant les valeurs de *densité de bâti*, nous différencions les voies de circulation en trois sous-ensembles :

- L'ensemble B_0 correspond aux routes connectées à aucun bâti :

$$B_0 = \{r_i \in R \mid \nexists b_j \in B, (b_j, r_i) \in A_{BR}\}$$

- L'ensemble B_1 correspond aux routes connectées à un ou plusieurs bâtiments :

$$B_1 = \{r_i \in R \mid \exists b_j \in B, (b_j, r_i) \in A_{BR}\} = B - B_0$$

Nous utilisons la méthode des *k-moyennes* introduite au chapitre I.2.5.2 pour partitionner B_1 . Ainsi, on a $B_1 = \{B-, B+\}$.

- L'ensemble $B-$ correspond aux routes qui ont les plus faibles nombres de bâtiments connectés (faibles scores de BRD) :

$$B- = \{r_i \in R, BDR(r_i) < \text{Min}(B+)\}$$

- L'ensemble $B+$ correspond aux routes liées à un nombre élevé de bâtiments (hauts scores de BRD) :

$$B+ = \{r_i \in R, BDR(r_i) > \text{Max}(B-)\}$$

Concernant les valeurs de *Metric Compactness*, nous utilisons à nouveau la méthode des *k-moyennes* pour différencier les voies de circulation en deux sous-ensembles. Ainsi on a $R = \{MC-, MC+\}$.

- l'ensemble $MC-$ correspond aux routes dont les valeurs de *Metric Compactness* sont les plus faibles :

$$MC- = \{r_i \in R, MC(r_i) < \text{Min}(MC+)\}$$

- l'ensemble $MC+$ correspond aux routes dont les valeurs de *Metric Compactness* sont les plus fortes :

$$MC+ = \{r_i \in R, MC(r_i) > \text{Max}(MC-)\}$$

À partir de ces différents ensembles, nous définissons 6 classes représentatives des propriétés des tronçons de route (Tableau 6). Plusieurs de ces configurations représentent des structurations du territoire cohérentes avec les observations et les nomenclatures décrites par ailleurs :

- $B+MC+$ et $B-MC+$ correspondent à des espaces urbanisés ou périurbains, regroupant une forte compacité des voies de circulation et un nombre important de bâtiments.
- B_0MC- et $B-MC-$ correspondent et à des espaces ruraux, regroupant une faible compacité des voies de circulation et un faible nombre (voir nul) de bâtiments.

À l'inverse, deux configurations peuvent être considérées comme des anomalies dans la structuration de l'espace :

- B_0MC+ correspondent à des lieux pour lesquels la compacité des voies de circulation est importante, mais n'est associée à aucun bâtiment.
- $B+MC-$ correspondent à des lieux pour lesquels la compacité des voies de circulation est faible, mais est associée à un grand nombre de bâtiments.

	<i>MC-</i>	<i>MC+</i>
<i>B₀</i>		
<i>B-</i>		
<i>B+</i>		

Tableau 6 : Classification des tronçons de circulation

La Figure 54 illustre les résultats obtenus avec cette classification pour le territoire de Guisseny. La mesure de densité étant une mesure locale, nous avons choisi d'utiliser $X = 10\%$ pour la mesure de compacité afin que la portée de la mesure de compacité soit locale elle aussi.

On observe que la classification proposée par notre approche permet d'extraire clairement le centre bourg, caractérisé par une forte compacité du réseau et la présence de bâtiments ($B+MC+$ et $B-MC+$). De même, la classification permet d'identifier les parties rurales pour lesquelles la compacité des tronçons de route et le nombre de bâtiments sont faibles (B_0MC- et $B-MC-$).



Figure 54 : Corrélation entre densité de bâti et compacité de route

En périphérie du centre bourg et le long du littoral, la structuration du territoire apparaît moins homogène. Dans ces lieux, pour une partie importante des tronçons, on observe une certaine cohérence avec des valeurs faibles de Metric Compactness et un nombre peu important de bâtiments associés (B-MC-). Par contre, dans de nombreux autres cas, on observe que la valeur de Metric Compactness est également faible, mais le nombre de bâtiments est lui important (B+MC-). On est ici dans le cas qu'on considère comme une anomalie, correspondant à des zones où la connectivité du réseau n'est peut-être pas suffisante pour assurer une accessibilité satisfaisante au regard du nombre de bâtiments présent.

Ce résultat, et la détection de ces anomalies qu'il permet, ouvrent des perspectives vers la catégorisation des aménagements à prévoir. Ainsi, la question qui est posée pour les tronçons de type (B+MC-) est de savoir si les futurs aménagements doivent chercher à augmenter la connectivité des voies de circulation ou à réduire la densité du bâti.

III.6 Synthèse

Ce chapitre décrit l'implémentation du modèle de graphe et des opérateurs que nous avons introduits dans le Chapitre II ainsi que les résultats obtenus lors de l'application de chaque opérateur sur les données du village de Guisseny.

Afin d'implémenter le graphe, nous avons commencé par présenter un modèle de données à partir des objets retenus pour modéliser l'espace (bâtiments, parcelles et routes). Le choix a été fait d'implémenter le graphe dans une base de données relationnelle PostgreSQL avec l'extension PostGIS, car elle présente l'intérêt de posséder toutes les primitives spatiales nécessaires au stockage des objets géographiques et à la construction du graphe. Elle peut également être facilement connectée à d'autres outils tels que QGIS dans notre cas.

À partir de la construction du graphe reposant sur les relations topologiques existantes entre les différents objets géographiques, nous avons pu décrire des prétraitements pouvant être appliqués pour rendre les données d'origine conformes au modèle de graphe, à savoir la transformation du réseau de circulation en tronçons de route, la construction des routes sous forme de polygones à partir des polygones ainsi que la gestion de certaines incohérences au niveau des intersections d'objets. À la suite de ces prétraitements, nous avons présenté l'implémentation de la construction du graphe et des opérateurs.

Au niveau global, la mesure de centralité de bâti permet d'obtenir plus d'informations que la mesure de centralité de degré. Au niveau local, l'approche non pondérée (c.-à-d. avec les coefficients $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 1$) donne des résultats moins discriminants que l'approche pondérée (c.-à-d. avec les coefficients $\alpha_1 = 5, \alpha_2 = 4, \alpha_3 = 3, \alpha_4 = 2, \alpha_5 = 1$), notamment à cause des scores trop similaires le long d'une route, d'où l'intérêt de différencier les relations pour décrire la structure spatiale.

La mesure de connectivité locale a montré qu'il n'y a pas clairement de lien entre les scores de taux de partage du bâti et les types de territoires (c.-à-d. rurale, pavillonnaire et compacte). Ainsi la valeur de connectivité locale du bâti n'apparaît donc pas comme un marqueur du type de territoire, mais seulement comme une information locale à l'échelle de la route.

À l'inverse, la mesure de densité de bâti se comporte plus comme un marqueur du territoire avec des scores faibles dans la partie la plus rurale du territoire, et des valeurs les plus hautes dans les zones les plus urbanisées. La mesure proposée étant mesurée sur les tronçons de route, elle

présente l'avantage d'éviter les problèmes de segmentation de l'espace tels que la taille de la grille ou le découpage d'objets par exemple.

Pour la mesure de compacité, l'analyse des variations de D_{min}^X en fonction de X permet d'estimer l'homogénéité de la répartition des routes sur le territoire. En faisant varier la valeur de X , il est possible d'analyser la structure du réseau routier aussi bien à un niveau local que global. Sur la commune de Guisseny, cette analyse a permis de montrer un réseau globalement homogène contrairement au réseau de Guisseny modifié qui présente une partie du réseau plus faiblement connecté.

La corrélation entre la mesure de densité de bâti et la mesure de compacité cherche à classer les routes en 6 catégories. Elle permet notamment d'extraire clairement le centre bourg et les parties rurales. Cette catégorisation permet de détecter les tronçons présentant un déséquilibre entre la densité du bâti et l'intégration du tronçon au reste du réseau.

Conclusion

L'objectif de cette thèse a consisté à étudier la structure des territoires habités à partir d'une approche basée sur une analyse structurale et spatiale. Notre recherche se démarque des modèles existants principalement focalisés sur l'analyse du réseau routier sans prendre compte du bâti ou du parcellaire.

Contribution

Modèle de graphe

Nous avons présenté un modèle de graphe permettant d'inclure au sein d'une même représentation formelle les données cartographiques du bâti, du parcellaire et du réseau routier en utilisant les relations topologiques existantes entre ces éléments. Pour assembler ces différentes informations, nous avons opté pour une approche modulaire en construisant plusieurs graphes :

- trois graphes pour les relations d'adjacences entre bâtis, entre parcelles et entre tronçons de route
- deux graphes pour les relations d'intersections entre bâtis et parcelles et les relations d'adjacences entre parcelles et routes
- un graphe représentant l'intégralité des liens entre les bâtiments, les parcelles et les routes, les parcelles servant de jointure entre le bâti et les routes.

En ce qui concerne la modélisation du réseau routier, nous avons considéré que l'élément important est le tronçon de route plutôt que l'intersection de route ce qui nous a conduit à adopter l'approche *duale*. Cependant, au lieu de considérer des routes *continues* (telles que les routes nommées, lignes axiales ou autres) nous avons choisi de considérer le tronçon de route comme élément de base pour la représentation du réseau routier. Une telle modélisation nous a permis de conserver la notion de distance métrique qui est perdue dans les autres modélisations *duales* (par exemple si une route continue s'intersecte avec deux autres routes, la distance entre ces intersections est inconnue dans les autres modélisations duales.), mais en contrepartie cela nous

fait perdre la notion de *continuité* de route, notion qui peut être discutable dans sa définition et pourrait être définie dans notre modèle comme un enchaînement de tronçons de route.

Le choix d'utiliser la topologie pour relier les bâtiments aux parcelles puis aux routes nous a permis de ne pas utiliser de distance seuil choisie arbitrairement pour construire les relations. Cependant, un biais est introduit si l'on souhaite mesurer les distances entre les bâtiments en utilisant le réseau routier. En effet dans notre modèle, seule la distance du milieu du tronçon de route lié au premier bâtiment jusqu'au milieu du tronçon de route lié au deuxième bâtiment est connue. Pour mesurer la distance entre bâtiments en évitant ce biais, il faudrait connaître la distance entre chaque bâtiment et le milieu du tronçon de route.

Introduction d'opérateurs de graphe

Nous avons également introduit plusieurs opérateurs de graphe, décrits formellement et illustrés de manière théorique dans un premier temps, puis implémentés et appliqués sur un jeu de données de la commune de Guisseny dans un second temps.

Le premier opérateur proposé est une mesure de centralité ayant pour objectif de mesurer le niveau d'intégration locale du bâti en distinguant cinq configurations locales entre deux bâtiments, à savoir : adjacents et sur une même parcelle, adjacents mais dans différentes parcelles, non adjacents mais situés sur une même parcelle, non adjacents mais situés sur des parcelles adjacentes, non adjacents mais sur des parcelles connectées à une même route. Un système de coefficients permet de pondérer les différents cas pour pouvoir adapter facilement l'opérateur à différentes problématiques.

Le second opérateur est une mesure de connectivité locale du bâti reflétant à quel point les tronçons de route partagent les bâtis connectés via les parcelles avec d'autres tronçons de route.

Le troisième opérateur est une mesure de densité de bâti relative à une route visant, contrairement à l'opérateur précédent, à refléter la quantité de bâtiments connectés à chaque route. Afin de ne pas surévaluer cette quantité, le poids d'un bâtiment connecté à plusieurs routes est distribué de manière égale à chacune de ces routes.

Le quatrième opérateur proposé est construit en trois mesures, nommées *Ratio Minimum Distance*, *Minimum Ratio Minimum Distance* et *Metric Compactness*, et a pour objectif de mesurer la compacité du réseau de circulation, c'est-à-dire de mesurer la capacité pour une rue à atteindre facilement d'autres tronçons de route.

Ces trois derniers opérateurs permettent d’avoir différents angles de vision pour analyser la répartition du bâti et du réseau routier.

Analyse de la structuration spatiale d’un territoire habité

Ces différents opérateurs permettent de faire une analyse de la structuration de l’espace à différents niveaux. Ainsi la mesure de centralité de bâti proposée repose sur la distinction de différents contextes existants entre deux bâtiments au niveau des bâtis, des parcelles, et des routes. La mesure de densité de bâti permet de distinguer les parties les plus rurales du territoire avec des valeurs de densité faibles, et les zones les plus urbanisées avec des valeurs plus hautes. La mesure de compacité permet d’analyser la structure du réseau routier aussi bien à un niveau local que global faisant varier X et d’estimer l’homogénéité de la répartition des routes sur le territoire. La combinaison des deux permet de décomposer les structures spéciales en six classes. Quatre d’entre elles représentent des structurations du territoire cohérentes avec l’attendu, alors que deux autres peuvent être considérées comme des anomalies dans la structuration de l’espace.

Synthèse

En résumé, notre contribution est la création d’un modèle de graphe combinant le bâti, le parcellaire et le réseau routier à partir des relations topologiques, avec, pour la partie réseau routier, une approche duale permettant de conserver la distance métrique. Notre contribution est aussi la définition de plusieurs opérateurs de graphe adapté au modèle et permettant d’analyser la structure de l’espace tout en étant facilement adaptable aux contextes ruraux, périurbains et/ou compacts, et à leurs intermédiaires, grâce à une paramétrisation permettant d’ajuster l’échelle d’observation et d’analyse.

Perspectives

Enrichissement du modèle

Comme nous l’avons dit dans la partie contribution, l’atome de notre réseau routier étant le tronçon de route, nous n’avons pas la notion de *continuité* d’une route. Cependant, dans notre modèle, cette notion peut être très facilement modélisable comme un enchaînement de tronçons de route. Il serait donc intéressant de chercher une façon de construire cette continuité de route, ne reposant pas sur des noms de route (Jiang & Claramunt, 2004) ou des niveaux de classification (Marshall, 2015) qui sont des éléments empiriques ne reflétant pas nécessairement l’idée de continuité de route.

Exploration de nouveaux opérateurs

Bien que nous ayons introduit quelques opérateurs sur notre modèle de graphe permettant d'analyser l'espace, il est nécessaire de renforcer le nombre d'opérateurs pour obtenir une analyse plus complète, notamment au niveau des parcelles. Ainsi il faudra construire des ilots (c.-à-d. un ensemble de parcelles adjacentes délimité par les routes) et en extraire diverses informations telles que leurs superficies, le nombre et la superficie moyenne des bâtiments et des parcelles qui les composent par exemple.

Suite à cela, à partir du modèle et idéalement des opérateurs déjà proposés, la classification de différentes zones introduite lors de la corrélation entre la mesure de densité de bâti et la mesure de compacité (B_0MC- , B_0MC+ , etc.) devra être complétée afin de réussir à identifier différents types de zones, correspondant par exemple à la classification en neuf classes, proposée par (Puissant, et al., 2010) (cf. Figure 3). Il sera ainsi possible d'introduire des bases de notions de sémantiques telles que des ilots agricoles ou des ilots résidentiels par exemple.

Robustesse de l'implémentation et passage à l'échelle

Une fois que nous aurons renforcé le nombre d'opérateurs et amélioré les méthodes de classifications, il faudra obtenir et analyser d'autres jeux de données :

- Des jeux de données plus grands : pour s'assurer que les opérateurs actuels sont efficaces tout en garantissant que les temps de calcul restent raisonnables. Cela sera également l'occasion de créer des opérateurs permettant de comparer plusieurs jeux de données, de même ou de différentes tailles.
- Des jeux de données plus hétérogènes, avec plusieurs centres par exemple. L'espace que nous considérons n'incluant pas les rivières, une analyse de certaines villes où les rivières ont un effet structurant certain, tel que Venise ou Delft par exemple, permettrait de faire ressortir de probables limitations et de voir comment elles pourraient être intégrées dans notre modèle.

Lorsque la robustesse du modèle et des opérateurs aura été établie sur plusieurs jeux de données, il faudra réaliser et mettre à disposition un package de fonctions PostgreSQL permettant au plus grand nombre d'utiliser le modèle décrit dans cette thèse et d'appliquer les différents opérateurs. Les fonctions de visualisation des erreurs seront éventuellement complétées par l'ajout de fonctions de correction (semi)automatique de données, notamment pour les erreurs dans la topologie du réseau et dans les intersections invalides entre éléments de notre espace. Il faudra

aussi voir s'il est possible de coupler les fonctions PostgreSQL des opérateurs dans un plugin Qgis afin de les exécuter directement depuis Qgis, et ainsi simplifier leurs utilisations.

Nouvelles pistes de recherches : le temporel

Dans les données de Guisseny que nous avons utilisées, chaque bâtiment est référencé temporellement par rapport à la date de construction ou à la date de publication du document (cadastre, photographie aérienne) qui le laisse apparaître pour la première fois. Les dates s'étendent sur plus de 50 ans, et permettent d'observer principalement l'extension du bourg et l'urbanisation de la zone littorale. Il est donc possible de construire plusieurs cartes structurales à des temps différents pour décrire la façon dont s'est façonné le territoire (soit une extension à partir du bourg, soit le bourgeonnement de plusieurs hameaux finissant par se rejoindre). C'est également la possibilité de savoir si les anomalies que nous décrivons sont des événements temporaires (la collectivité sait réagir) ou qui perdurent dans le temps (la collectivité s'y adapte). Cette perspective ouvre une opportunité pour faire évoluer le modèle de graphe proposé vers un modèle de graphe spatio-temporel. Cette évolution pourrait se répercuter également sur le modèle conceptuel de la base de données pour tenir compte de la dimension temporelle.

Confrontation à des thématiciens

Dans l'état de nos recherches, l'information sémantique et le lien à l'usage des objets modélisés sont très pauvres. Le modèle ayant une vocation générique, ces points n'ont pas été approfondis. L'implication des thématiciens, spécialistes de l'urbanisme et de l'aménagement est prépondérante pour s'assurer de la pertinence des opérateurs proposés et des informations extraites à des fins d'aide à la décision. En effet, il est intéressant d'observer nos opérateurs en les appliquant seulement sur certains types de bâtiments, de parcelles ou de voies de communication. De même, c'est certainement de leurs expertises que viendra le meilleur choix dans la valeur de X utilisées pour la Metric Compactness. Mieux appréhender ces paramètres pourrait permettre de remplacer les mesures de k-means pour définir les seuils de partitionnement lors de la corrélation entre la mesure de densité de bâti et la mesure de compacité.

Bibliographie

Alba, R. D., 1973. A graph-theoretic definition of a sociometric clique. *Journal of Mathematical Sociology*, 3(1), pp. 113-126.

Alboody, A., Sedes, F. & Inglada, J., 2009. *Fuzzy Intersection and Difference Model for Topological Relations*. Lisbon, J. P. Carvalho, D. Dubois, U. Kaymak and J. M. C. Sousa.

Anderson, W., Kanaroglou, P. & Miller, E., 1996. Urban form, energy and the environment: a review of issues, evidence and policy.. *Urban studies*, 33(1), pp. 7-35.

Badariotti, D., Banos, A. & Moreno, D., 2007a. *Discontinuités, catastrophes et réseaux urbains: le modèle Remus*. s.l., 14èmes journées de Rochebrune.

Badariotti, D., Banos, A. & Moreno, D., 2007b. Conception d'un automate cellulaire non stationnaire à base de graphe pour modéliser la structure spatiale urbaine: le modèle Remus.. *Cybergeo: European Journal of Geography*.

Batsakis, S. & Petrakis, E., 2010. *SOWL: spatio-temporal representation, reasoning and querying over the semantic web*. Graz, Proceedings of the 6th International Conference on Semantic Systems..

Bavelas, A., 1950. Communication patterns in task-oriented groups.. *Journal of the acoustical society of America*, Volume 22, pp. 725-730.

bd-carto-IGN, 2015. <http://professionnels.ign.fr/bdcarto>, s.l.: IGN.

Bejaoui, L., Pinet, F., Y., B. & M., S., 2009. Qualified topological relations between spatial objects with possible vague shape. *International Journal of Geographical Information Science*, 23(7), pp. 877-921.

Bera, C., 2004. *Une Etude Contextuelle de l'Influence de l'Environnement Spatial dans l'Appréhension de la Notion de Proximité*, Univ. Rennes1: Thèse de doctorat.

Béra, R., 2004. *L'adjacence relative. Une étude contextuelle de l'influence de l'environnement spatial dans l'appréhension de la notion de proximité*. Rennes: s.n.

Bera, R., 2004. *Une Etude Contextuelle de l'Influence de l'Environnement Spatial dans l'Appréhension de la Notion de Proximité*, Univ. Rennes1: Thèse de doctorat.

Berretti, S., Del Bimbo, A. & Vicario, E., 2003. Weighted walkthroughs between extended entities for retrieval by spatial arrangement.. *IEEE Transactions on*, 5(1), pp. 52-70.

Boffet, A. & Serra, S. R., 2001. *Identification of spatial structures within urban blocks for town characterization*.. Beijing, Proceedings 20th International Cartographic Conference..

Bonacich, P., 1972. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1), pp. 113-120.

Borrego, C. et al., 2006. How urban structure can affect city sustainability from an air quality perspective.. *Environmental modelling & software*, 21(4), pp. 461-467.

Buhl, J. et al., 2006. Topological patterns in street networks of self-organized urban settlements. *The European Physical Journal B*, Volume 49, p. 513 – 522.

Cervero, C. & Kockelman, K., 1997. Travel demand and the 3Ds: density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2(3), pp. 199-219.

Chekir, S., Josselin, D., Pasquet, A. & al., e., 2014. *Modélisation, simulation et analyse de propriétés de réseaux orbitales*.. Grenoble, Sageo.

Claramunt, C., 2012. Towards a spatio-temporal form of entropy. In: *Advances in Conceptual Modeling*. s.l.:Springer Berlin Heidelberg, pp. 221-230.

Clarke, B. L., 1981. A calculus of individuals based on connection. *Notre Dame Journal of Formal Logic Notre-Dame*, 22(3), pp. 204-219.

Clementini, E., Di Felice, P. & Van Oosterom, P., 1993. *A small set of formal topological relationships suitable for end-user interaction*. Singapore, Springer.

Clementini, E. & Laurini, R., 2008. Un cadre conceptuel pour modéliser les relations spatiales.. *Revue des Nouvelles Technologies de l'Information (RNTI)*, Volume RNTI-E-14, pp. 1-17.

Codescu, M. et al., 2011. *DO-ROAM: Activity-Oriented Search and Navigation with OpenStreetMap*. Brest, GeoSpatial Semantics Proceedings.

Couclelis, H., 1992. People manipulate objects (but cultivate fields): beyond the raster-vector debate in GIS. In: *Theories and methods of spatio-temporal reasoning in geographic space*. Berlin: Springer Berlin Heidelberg, pp. 65-77.

Couclelis, H., 2009. *The abduction of geographic information science: transporting spatial reasoning to the realm of purpose and design*. Aber Wrac'h, France, Springer Berlin Heidelberg, pp. 342-356.

Curie, F., Perret, J. & Ruas, A., 2010. *Simulation of urban blocks densification*. Guimaraes, Portugal, 13th AGILE International Conference on Geographic Information Science.

De Smith, M. J., Goodchild, M. F. & Longley, P., 2007. *Geospatial analysis: a comprehensive guide to principles, techniques and software tools*. 2e ed. s.l.:Troubador Publishing Ltd..

Debardieux, B., 2004. Présentation générale. De l'objet spatial à l'effet géographique.. In: *L'effet géographique. Construction sociale, appréhension cognitive et configuration matérielle des objets géographiques*. Grenoble: Maison des sciences de l'homme-Alpes - CNRS, pp. 11-33.

Del Mondo, G., Stell, J. G., Claramunt, C. & Thibaud, R., 2010. A graph model for spatio-temporal evolution.. *Journal of Universal Computer Science*, 16(11), pp. 1452-1477.

Dijkstra, E. W., 1959. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1), pp. 269-271.

Egenhofer, M. J. & Franzosa, R. D., 1991. Point-set topological spatial relations. *International Journal of Geographical Information System*, 5(2), pp. 161-174.

Egenhofer, M. J. & Herring, J., 1991. *Categorizing binary topological relations between regions, lines, and points in geographic databases*, Departement of Surveying Engineering, University of Maine: Technical Report.

Flake, G. W., Lawrence, S., Giles, C. L. & Coetzee, F. M., 2002. Self-organization and identification of web communities. *IEEE Computer*, 35(3), pp. 66-70.

Floyd, R. W., 1962. Algorithm 97: shortest path. *Communications of the ACM*, 5(6), p. 345.

Ford, J. L. R. & Fulkerson, D. R., 1962. *Flows in networks*. s.l.:Princeton university press.

Frank, A. U., 1991. *Qualitative spatial reasoning with cardinal directions*.. Vienna, Springer Berlin Heidelberg, pp. 157-167.

- Freeman, L. C., 1979. A set of measures of centrality based on betweenness. *Sociometry*, 1(3), pp. 215-239.
- Frey, H., 1999. *Designing the city: Towards a more sustainable urban form..* London: E & FN Spon.
- Galton, A., 2000. *Qualitative Spatial Change*. 1 ed. Oxford: Oxford University Press.
- Goodchild, M. F., 1989. Tiling large geographical databases. *SSD'89*, p. 137-146.
- Goodchild, M. F. & Sucharita, G., 1989. *The accuracy of spatial databases*. London: Taylor and Francis.
- Guyet, T., 2013. *Visualisation de données relationnelles*. Brest, Conférence Internationale de Géomatique et d'Analyse Spatiale (SAGEO).
- Hamaina, R., Leduc, T. & Moreau, G., 2012. *Towards urban fabrics characterization based on buildings footprints*. Avignon, Springer Berlin Heidelberg, pp. 327-346.
- Hart, P. E., Nilsson, N. J. & Raphael, B., 1968. A formal basis for the heuristic determination of minimum cost paths. *Systems Science and Cybernetics, IEEE Transactions on*, 4(2), pp. 100-107.
- Hawbaker, T. J. et al., 2006. Road development, housing growth, and landscape fragmentation in northern Wisconsin: 1937-1999.. *Ecological Applications*, 16(3), pp. 1222-1237.
- Herring, J. R. ed., 2011. *OpenGIS® Implementation Standard for Geographic information - Simple feature access*. OGC 06-103r4 - Version 1.2.1: Open Geospatial Consortium, Inc.
- Hillier, B. & Hanson, J., 1984. *The social logic of space..* Cambridge: Cambridge university press.
- Hillier, B. et al., 1993. Natural movement or configuration and attraction in urban pedestrian movement. *Environment and Planning B*, Volume 20, pp. 29-66.
- Hornsby, k. & Egenhofer, M., 2000. Identity-based change: A foundation for spatio-temporal knowledge representation. *International Journal of Geographical Information Science*, Volume 14, pp. 207-224.
- Jiang, B. & Claramunt, C., 2004. Topological analysis of urban street networks. *Environment and Planning B*, 31(1), pp. 151-162.

- Jitkajornwanich, K., Elmasri, R., Li, C. & McEnery, J., 2011. *Formalization of 2-d spatial ontology and owl/protégé realization*. Athens, Proceedings of the International Workshop on Semantic Web Information Management (p. 9). ACM.
- Jones, M. A. & Fanek, M. F., 1997. *Crime in the urban environment*. University College London, Proceedings of First International Symposium on Space Syntax.
- Katz, L., 1953. A new status index derived from sociometric analysis. *Psychometrika*, 18(1), pp. 39-43.
- Keeble, L. B., 1969. Principles and practice of town and country planning. *Estates Gazette*.
- Kinsey, L. C., 2012. *Topology of surfaces*. s.l., Springer Science & Business Media.
- Kleinberg, J. M., 1999. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5), pp. 604-632.
- Lagesse, C., 2015. *Lire les lignes de la ville : méthodologie de caractérisation des graphes spatiaux*, Paris: s.n.
- Le Couédic, M. et al., 2013. Du cadastre ancien au graphe. Les dynamiques spatiales dans les sources fiscales médiévales et modernes.. *ArchéoSciences*, 36(1), pp. 71-84.
- Lempel, R. & Moran, S., 2001. SALSA: the stochastic approach for link-structure analysis.. *ACM Transactions on Information Systems (TOIS)*, 19(2), pp. 131-160.
- Lesot, M. J., 2005. Similarity, typicality and fuzzy prototypes for numerical data. *6th European Congress on Systems Science, Workshop Similarity and resemblance*, Volume 94, p. 95.
- Long, N. & Kergomard, C., 2005. Classification Morphologique du tissu urbain pour des applications climatologiques. *Revue Internationale de Géomatique*, Volume 15, pp. 487-512.
- Luce, R. D., 1950. Connectivity and generalized cliques in sociometric group structure.. *Psychometrika*, 15(2), pp. 169-190.
- Luce, R. D. & D., P. A., 1949. A method of matrix analysis of group structure. *Psychometrika*, 14(2), pp. 95-116.
- Lynch, K., 1984. *Good city form*. s.l.:MIT press.

- MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations.. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1(14), pp. 281-297.
- Marshall, S., 2005. *Urban pattern specification*. London, Institute of Community Studies.
- Marshall, S., 2015. *Route Structure Analysis Working Paper 3: Route Interpretation Protocol*, London: academia.edu.
- Marshall, S., 2016. Line structure representation for road network analysis.. *Journal of Transport and Land Use*, 9(1), pp. 1-38.
- Mathian, H. & Sanders, L., 2014. *Objets géographiques et processus de changement*. Collection système d'information géographique ed. Londre: ISTE editions.
- Mokken, R. J., 1979. Cliques, clubs and clans.. *Quality & Quantity*, 13(2), pp. 161-173.
- Noy, N. F. & McGuinness, D. L., 2000. *Développement d'une ontologie 101: Guide pour la création de votre première ontologie.*, Stanford: Université de Stanford.
- Page, L., Brin, S., Motwani, R. & al, e., 1998. *The PageRank citation ranking: bringing order to the Web.*, Standford: Technical report, Standford Digital Library Technologies Project.
- Palla, G., Derényi, I., Farkas, I. & Vicsek, T., 2005. Uncovering the overlapping community structure of complex networks in nature and society.. *Nature*, 435(7043), pp. 814-818.
- Pandit, S. & Gupta, S., 2011. A comparative study on distance measuring approaches for clustering. *International Journal of Research in Computer Science*, 2(1), pp. 29-31.
- Penn, A. & Croxford, B., 1997. *Effects of street grid configuration on Kerbside concentrations of vehicular emissions*. University College London, Proceedings of First International Symposium on Space Syntax.
- Peponis, J., Zimring, C. & Choi, Y. K., 1990. Finding the building in wayfinding.. *Environment and Behavior*, Volume 22, pp. 555-590.
- Peuquet, D., 1984. A conceptual framework and comparison of spatial data. *Cartographica The International Journal for Geographic Information and Geovisualization*, 21(4), p. 66.
- Peuquet, D. J., 2002. *Representations of space and time*. New York: Guilford Press.

- Porta, S., Crucitti, P. & Latora, V., 2006. The network analysis of urban streets: a dual approach.. *Physica A: Statistical Mechanics and its Applications*, 369(2), pp. 853-866.
- Porta, S., Crucitti, P. & Latora, V., 2006. The network analysis of urban streets: a primal approach. *Environment and Planning B : Planning and design*, 33(5), pp. 705-725.
- Pouyanne, G., 2004. Des avantages comparatifs de la ville compacte à l'interaction mobilité-forme urbaine. Méthodologie et premiers résultats.. *Cahiers Scientifiques du Transport*, 45(1), pp. 49-82.
- Puissant, A. et al., 2010. Classification des tissus urbains à partir de données vectorielles - application à Strasbourg. *Proceeding du Colloque International de Géomatique et d'Analyse Spatiale SAGEO'10*, pp. 198-210.
- Radicchi, F. et al., 2004. Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(9), pp. 2658-2663.
- Randell, D. A. & Cohn, A. G., 1989. *Modelling Topological and Metrical Properties in Physical Processes*. s.l., s.n., pp. 357-368.
- Randell, D. A., Cui, Z. & Cohn, A. G., 1992. *A spatial logic based on regions and connection*. s.l., s.n., pp. 165-176.
- Ratti, C., 2004. Space syntax: some inconsistencies. *Environment and Planning B: Planning and Design*, 31(4), pp. 501-511.
- Rochat, Y., 2009. *Closeness centrality extended to unconnected graphs: The harmonic centrality index*.. Zürich, ASNA.
- Rousseaux, F., Long, N. & Renouard, A., 2012. ers une simulation de l'évolution des structures urbaines à partir d'une modélisation multi-agents.. *VertigO-la revue électronique en sciences de l'environnement*, 11(3).
- Sabidussi, G., 1966. The centrality index of a graph. *Psychometrika*, 31(4), pp. 581-603.
- Seeley, J. R., 1949. The net of reciprocal influence; a problem in treating sociometric data. *Canadian Journal of Psychology Revue Canadienne de Psychologie*, Issue 3, pp. 234-240.
- Seidman, S. B., 1983. Network structure and minimum degree. *Social networks*, 5(3), pp. 269-287.

- Seidman, S. B. & Foster, B. L., 1978. A graph-theoretic generalization of the clique concept*. *Journal of Mathematical sociology*, 6(1), pp. 139-154.
- Smith, B. & Varzi, A., 2000. Fiat and bona fide boundaries. *Philosophical and Phenomenological Research*, pp. 401-420.
- Spéry, L., Claramunt, C. & Libourel, T., 2001. A spatio-temporal model for the manipulation of lineage metadata.. *Geoinformatica*, 5(1), pp. 51-70.
- Spreiregen, P. D., 1965. *Urban design: The architecture of towns and cities*. s.l.:New York: McGraw-Hill.
- Steinbach, M., Karypis, G. & Kumar, V., 2000. *A Comparison of Document Clustering Techniques*, University of Minnesota Minneapolis: Technical Report 00-034.
- Tarski, A., 1982. *Logic, Semantics, Metamathematics*. édition d'origine : Oxford Clarendon Press, 1956. épuisé . ed. Indianapolis: Hackett Publishing Co..
- Thériault, M. & Claramunt, C., 1999. La représentation du temps et des processus dans les SIG: une nécessité pour la recherche interdisciplinaire. *Revue internationale de géomatique*, 9(1), pp. 67-99.
- TIGERweb, 2010. http://tigerweb.geo.census.gov/tigerwebmain/TIGERweb_main.html, s.l.: s.n.
- Tsai, Y.-H., 2005. Quantifying urban form: compactness versus 'sprawl'. *Urban studies*, 42(1), pp. 141-161.
- Turner, A. & Penn, A., 1999. *Making isovists syntactic: isovist integration analysis*. Brasilia, 2nd International Symposium on Space Syntax.
- Watts, D. J. & Strogatz, S. H., 1998. Collective dynamics of 'small-world' networks. *Nature*, 393(6684), pp. 440-442.

Modélisation de structures habitées à partir d'une approche de graphe

Résumé : Appréhender l'organisation spatiale d'un territoire urbain ou périurbain génère un vaste champ de questionnements concernant la distribution et les propriétés du bâti, du parcellaire et des voies de circulation, et les interrelations de cet ensemble d'objets géographiques avec les usages de la cité. Cette thèse a pour objectif de proposer un modèle de représentation et de quantifier la structuration d'un tel espace afin de mieux en appréhender les usages. Cette démarche de modélisation repose sur la caractérisation des relations spatiales existantes entre les éléments remarquables identifiés. Afin de prendre en compte les entités d'un tel territoire et les relations spatiales qui les caractérisent, nous caractérisons un modèle de graphe où les noeuds représentent respectivement les bâtiments, les parcelles, et les routes, et les arcs correspondent aux relations topologiques de connexion, adjacence ou intersection, de ces éléments. Afin de quantifier les propriétés d'un territoire, une mesure de centralité du bâti, une mesure de connectivité locale du bâti, une mesure de densité du bâti et une mesure de la compacité du réseau de circulation sont proposées et formalisées. Ces mesures sont adaptables aux contextes ruraux, périurbains et urbains grâce aux paramètres permettant d'ajuster l'échelle d'observation et d'analyse. Le modèle et les mesures proposés sont implémentés dans le Système de Gestion de Bases de Données relationnel PostgreSQL avec prise en compte de l'extension spatiale PostGIS pour des applications grand public. L'application sur un cas d'étude, la commune de Guissény, tend à rendre opérationnel le modèle de graphe proposé et permet d'examiner les capacités d'analyse de la structuration spatiale de chaque mesure identifiée.

Mots clés : Modèle de Graphe, Syntaxe Spatiale, SIG, Aménagement du territoire, espace urbain, espace rural.

A graph-based model for the representation of land spaces

Abstract : Understanding the spatial organization of a given territory generates a wide range of urban planning and research issues such as the distribution and properties of buildings, parcels and roads, and the relevance of it in terms of uses. The research developed by this thesis model the structuring of space, this being one of the first steps towards a better characterization, study, interpretation and simulation of the evolution of such objects.

The structuring of space relies on the characterization of the spatial relationships between remarkable elements: buildings, parcels and road networks. Since existing models do not combine these elements, this thesis proposes a graph model that integrates all of them. The nodes represent buildings, parcels, and roads respectively, and the arcs correspond to the topological connection relationships, adjacency or intersection, of these elements.

In order to quantify the properties that emerges from a given territory, a building centrality measure, a building local connectivity measure, a building density measure and a road network compactness measure have been proposed and formalized. These measures adapt to the rural, peri-urban and urban contexts thanks to the parameters that allow to adjust, scale of observation, and analysis. The model and measures proposed are implemented in the relational database management system PostgreSQL with PostGIS extension. The application on a case study, the village of Guisseny, makes the proposed graph model operational and allows to examine the spatial structuring analytical capacities of each measure.

Keywords : Graph model, Space Syntax, GIS, Spatial Planning.