



HAL
open science

Analyse de la distribution spatiale des implantations humaines : apports et limites d'indicateurs multi-échelles et trans-échelles

François Sémécurbe

► To cite this version:

François Sémécurbe. Analyse de la distribution spatiale des implantations humaines : apports et limites d'indicateurs multi-échelles et trans-échelles. Géographie. Université Bourgogne Franche-Comté, 2020. Français. NNT : 2020UBFCC013 . tel-03125388

HAL Id: tel-03125388

<https://theses.hal.science/tel-03125388v1>

Submitted on 29 Jan 2021

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.

THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITE BOURGOGNE FRANCHE-COMTE
PREPAREE AU LABORATOIRE ThéMA : Théoriser et Modéliser pour Aménager

Ecole doctorale n°6049
Sociétés, Espaces, Pratiques, Temps

Doctorat de Géographie

Par
Monsieur François Sémécurbe

Analyse de la distribution spatiale des implantations humaines

Thèse présentée et soutenue à Besançon, le 25 septembre 2020

Composition du Jury :

Monsieur Didier Josselin - Directeur de recherche CNRS, laboratoire ESPACE, Avignon - Président du jury
Madame Elsa Arcaute - Associate professor, CASA, University College London - Rapporteur
Monsieur Giovanni Fusco - Chargé de recherche CNRS HDR, laboratoire ESPACE, Nice - Rapporteur
Monsieur Julien Perret - Directeur de Recherche du Développement Durable - Examineur externe
Monsieur Pierre Frankhauser - Professeur émérite, Université de Franche-Comté, ThéMA - Examineur interne
Madame Cécile Tannier - Directrice de Recherche, Université de Franche-Comté, CNRS, ThéMA - Directrice de thèse
Monsieur Stéphane Gaétan Roux - Maître de Conférence HDR, ENS de Lyon, laboratoire de Physique - Codirecteur de thèse

Titre : Analyse de la distribution spatiale des implantations humaines : Apports et limites d'indicateurs multi-échelles et trans-échelles

Mots clés : Analyse fractale, analyse trans-échelle, peuplement, échelle

Résumé :

En tant qu'être humain, il nous est aisé de juger visuellement du caractère dispersé ou concentré d'une distribution. Pour autant, la formalisation quantitative de nos impressions est problématique. Elle est tributaire des échelles d'analyse choisies. Cette dépendance des indicateurs aux échelles a changé de statut. Initialement considérée comme un frein à la connaissance, elle témoigne à présent de l'organisation multi-échelle des distributions étudiées. L'objectif central de cette thèse est d'approfondir les limites et l'apport des indicateurs multi-échelles et trans-échelles à l'étude des distributions spatiales des implantations humaines.

L'analyse spatiale vise à comparer les distributions spatiales à une répartition uniforme. La manière dont on s'éloigne de cette référence est utilisée pour caractériser l'organisation multi-échelle des distributions analysées. L'application de ces méthodes aux implantations humaines n'a pas été

satisfaisante. Le recours à une référence exogène n'est pas adapté à des distributions très inégalement concentrées dans l'espace.

L'analyse fractale, fréquemment utilisée en géographie urbaine, considère que les distributions analysées sont leur propre étalon de mesure. Les dimensions fractales mesurent la façon dont l'espace occupé par celles-ci évolue à travers les échelles. Ce type d'analyse requiert une régularité entre les échelles, l'invariance d'échelle dont l'existence n'est pas vérifiée sur l'ensemble des territoires. L'analyse trans-échelle généralise les principes de l'analyse fractale à toutes les distributions et permet de caractériser l'inégale concentration des implantations humaines dans les territoires ruraux et urbains.

Title : Analysis of the spatial distribution of human settlements : Contributions and limitations of multi-scale and trans-scale indicators

Keywords : Fractal analysis, trans-scale analysis, settlement, scale

Abstract : As human beings, it is easy for us to judge visually whether a distribution is dispersed or concentrated. However, the quantitative formalization of our impressions is problematic. It depends on the scales of the chosen analysis. This dependence of indicators on scales has changed. It is initially considered as a barrier to knowledge, it now reflects the multi-scale organisation of the distributions studied. The central objective of this thesis is to investigate the limits and contribution of multi-scale and trans-scale indicators to the study of the spatial distributions of human settlements.

Spatial analysis aims at comparing spatial distributions to a uniform distribution. The way in which spatial distributions move away from this reference is used to characterize the multi-scale organization of the analyzed distributions. The application of these methods to human settlements has not been satisfactory.

The use of an exogenous reference is not adapted to distributions that are very unevenly concentrated in space.

Fractal analysis used in urban geography considers that the analysed distributions are their own measurement standard. Fractal dimensions measure how the space occupied by them evolves across scales. This type of analysis requires a regularity between scales, the invariance of scale whose existence is not verified on all territories. Trans-scale analysis generalises the principles of fractal analysis to all distributions and makes it possible to characterise the unequal concentration of human settlements in rural and urban territories.

Analyse de la distribution spatiale des implantations humaines

Apports et limites d'indicateurs multi-échelles et trans-échelles

François Sémécurbe

Thèse de doctorat en Géographie

Composition du Jury :

Elsa Arcaute - Associate professor, CASA, University College London - Rapporteur

Giovanni Fusco - Chargé de recherche CNRS HDR, laboratoire ESPACE, Nice - Rapporteur

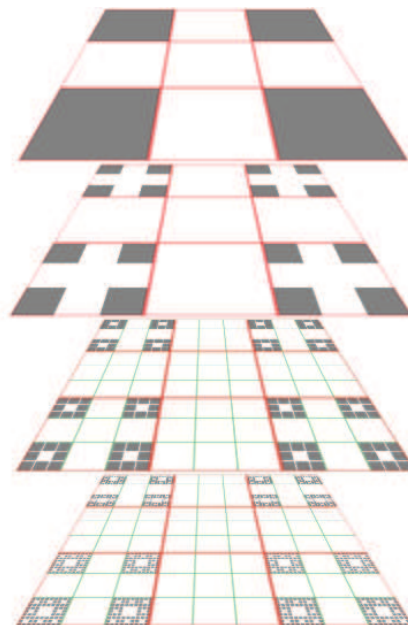
Didier Josselin - Directeur de recherche CNRS, laboratoire ESPACE, Avignon - Président du jury

Julien Perret - Directeur de Recherche du Développement Durable - Examineur externe

Pierre Frankhauser - Professeur émérite, Université de Franche-Comté, laboratoire ThéMA - Examineur interne

Cécile Tannier - Directrice de Recherche, Université de Franche-Comté, CNRS, laboratoire ThéMA - Directrice de thèse

Stéphane Gaétan Roux - Maître de Conférence HDR, ENS de Lyon, laboratoire de Physique - Codirecteur de thèse



Théoriser et Modéliser pour Aménager (ThéMA)

CNRS - Université Bourgogne Franche-Comté

2020

Analyse de la distribution spatiale des implantations humaines

Apports et limites d'indicateurs multi-échelles et trans-échelles

Résumé

En tant qu'être humain, il nous est aisé de juger visuellement du caractère dispersé ou concentré d'une distribution. Pour autant, la formalisation quantitative de nos impressions est problématique. Elle est tributaire des échelles d'analyse choisies. Cette dépendance des indicateurs aux échelles a changé de statut. Initialement considérée comme un frein à la connaissance, elle témoigne à présent de l'organisation multi-échelle des distributions étudiées. L'objectif central de cette thèse est d'approfondir les limites et l'apport des indicateurs multi-échelles et trans-échelles à l'étude des distributions spatiales des implantations humaines.

L'analyse spatiale vise à comparer les distributions spatiales à une répartition uniforme. La manière dont on s'éloigne de cette référence est utilisée pour caractériser l'organisation multi-échelle des distributions analysées. L'application de ces méthodes aux implantations humaines n'a pas été satisfaisante. Le recours à une référence exogène n'est pas adapté à des distributions très inégalement concentrées dans l'espace.

L'analyse fractale, fréquemment utilisée en géographie urbaine, considère que les distributions analysées sont leur propre étalon de mesure. Les dimensions fractales mesurent la façon dont l'espace occupé par celles-ci évolue à travers les échelles. Ce type d'analyse requiert une régularité entre les échelles, l'invariance d'échelle dont l'existence n'est pas vérifiée sur l'ensemble des territoires. L'analyse trans-échelle généralise les principes de l'analyse fractale à toutes les distributions et permet de caractériser l'inégale concentration des implantations humaines dans les territoires ruraux et urbains.

Analysis of the spatial distribution of human settlements

Contributions and limitations of multi-scale and trans-scale indicators

Abstract

As human beings, it is easy for us to judge visually whether a distribution is dispersed or concentrated. However, the quantitative formalization of our impressions is problematic. It depends on the scales of the chosen analysis. This dependence of indicators on scales has changed. It is initially considered as a barrier to knowledge, it now reflects the multi-scale organisation of the distributions studied. The central objective of this thesis is to investigate the limits and contribution of multi-scale and trans-scale indicators to the study of the spatial distributions of human settlements.

Spatial analysis aims at comparing spatial distributions to a uniform distribution. The way in which spatial distributions move away from this reference is used to characterize the multi-scale organization of the analyzed distributions. The application of these methods to human settlements has not been satisfactory. The use of an exogenous reference is not adapted to distributions that are very unevenly concentrated in space.

Fractal analysis used in urban geography considers that the analysed distributions are their own measurement standard. Fractal dimensions measure how the space occupied by them evolves across scales. This type of analysis requires a regularity between scales, the invariance of scale whose existence is not verified on all territories. Trans-scale analysis generalises the principles of fractal analysis to all distributions and makes it possible to characterise the unequal concentration of human settlements in rural and urban territories.

Remerciements

Je remercie chaleureusement toutes les personnes qui m'ont aidé pendant l'élaboration de ma thèse et notamment ma directrice de thèse Cécile Tannier qui, par son intérêt et sa pugnacité m'a permis de réaliser un rêve d'adolescent. Il y a bientôt sept ans, elle m'a accordé sa confiance sans me connaître, ce dont je lui serai toujours reconnaissant. Grâce à ses conseils, j'ai rencontré Stéphane Roux qui est devenu mon co-directeur de thèse et que je remercie pour sa bienveillance et ses explications théoriques imaginées. Je remercie également Marie-Hélène de Sède-Marceau sans qui cette thèse n'aurait pas été possible. Plus largement, je remercie l'ensemble du personnel de ThéMA pour leur gentillesse.

Je remercie mes collègues de l'Insee qui ont nourri mon goût pour la géographie, tout spécialement Brigitte Baccaïni et Jean-Michel Floch. Je remercie également Byl, Jerem Star, Milena, Juju, La Stéph et Elise pour leur soutiens sans faille. Ce travail n'aurait pu être mené à bien sans la complicité de mes amis et de mes bipèdes (mes stagiaires). La liste est longue mais je tenais à remercier nominativement, Alexandre, Duke, Arlindo, Popo, Dridric, Vincent, Jacod, Anne-Cécile, Sylvie, Valérie, Bjorn, Mickaël, Benoît et Joachim pour avoir supporté mes nombreux et récurrents états d'âme tout au long de ces six années. Il y a un peu de vous dans mes recherches.

Au terme de ce parcours, je remercie enfin celles et ceux qui me sont chers et que j'ai quelque peu délaissés ces dernières années pour achever cette thèse. Leurs attentions et encouragements m'ont accompagné tout au long de ces années. Je suis redevable à mes parents, particulièrement à ma mère, Paule, pour son soutien moral indéfectible : une vraie bretonne de Brest. Il va s'en dire que je remercie mon frère (avec qui on a ouvert un Hubblo sur le monde) et ma soeur (pour sa patience et son écoute) et leur famille pour avoir partagé mes nombreuses préoccupations scientifiques. J'embrasse un Hibou du Nord ainsi que ma belle famille pour leur accueil toujours chaleureux. Enfin, j'ai une pensée toute particulière pour celle qui partage mes crêpes et ma vie depuis dix ans, Farida. Ma reconnaissance envers elle va bien au delà de son soutien sans faille.

Table des matières

1	Introduction	14
1.1	Intérêt de l'étude de la distribution spatiale des implantations humaines . . .	14
1.1.1	Contexte	14
1.1.2	L'étude de la distribution spatiale des implantations humaines au cœur des préoccupations des géographes	15
1.2	Méthodes de caractérisation de la distribution spatiale des implantations humaines	17
1.2.1	Premiers indices de concentration de l'habitat et de la population	19
1.2.2	Indices de concentration de la population construits à partir de la courbe de Lorentz	19
1.2.3	Les prémisses de la statistique spatiale	20
1.2.4	L'entropie de Shannon	21
1.2.5	Difficultés inhérentes à la mesure des formes (configurations spatiales) .	22
1.3	Objectif et contenu de la thèse	24
1.3.1	A propos des implantations humaines	24
1.3.2	Choix de la zone d'étude et des entités spatiales considérées	26
1.3.2.1	Choix de la zone d'étude	26
1.3.2.2	Choix des entités spatiales élémentaires	26
1.3.3	Plan de la thèse	29
2	Méthodes de statistique spatiale pour l'analyse de la distribution spatiale des bâtiments	33
2.1	Introduction	34
2.2	Analyse de configurations de points (<i>clustering</i>)	35
2.2.1	Processus ponctuel Poissonien	35
2.2.1.1	Point aléatoire	36
2.2.1.2	Processus binomial	37
2.2.1.3	Processus de Poisson (processus spatial totalement aléatoire)	38
2.2.2	La fonction de Ripley	39
2.2.2.1	Définition de la fonction de Ripley	39
2.2.2.2	Apprécier la significativité des résultats	40
2.2.2.3	Interprétation du rayon r en terme d'échelle	40
2.2.2.4	Mise en œuvre de la fonction de Ripley sur les distributions archétypales	41
2.2.2.5	Application à la caractérisation des tissus bâtis	41
2.2.2.6	Critique des résultats et mises en perspective	44
2.2.3	La DBSCAN (<i>cluster detection</i>)	45
2.2.3.1	Présentation méthodologique de la DBSCAN	45
2.2.3.2	Choix des paramètres eps et min_{pts}	46
2.2.3.3	Cartographie des zones agrégées	52

2.2.3.4	La population agglomérée et dispersée	55
2.2.3.5	La dépendance des résultats aux paramètres : les limites de la recherche de <i>clusters</i>	57
2.3	Conclusion	60
3	Méthodes d'analyses fractales et multifractales pour l'analyse de la distribution spatiale des bâtiments, de la population et des formes d'occupation du sol	61
3.1	Introduction	62
3.2	Analyse fractale d'une distribution de points	63
3.2.1	Les ensembles de points fractals	63
3.2.1.1	Les cascades fractales	63
3.2.2	Les dimensions fractales	64
3.2.2.1	La dimension de <i>box-counting</i>	65
3.2.2.2	L'invariance d'échelle et le calcul de la dimension de <i>box-counting</i> pour les objets naturels	67
3.2.2.3	La dimension de corrélation	67
3.2.3	Application : <i>Geographically Weighted Fractal Analysis</i>	69
3.2.3.1	Contraintes sur les paramètres globaux	72
3.3	Analyse multifractale d'une mesure	73
3.3.1	Définition d'une mesure	73
3.3.1.1	Mesures auto-similaires : cascades multifractales	73
3.3.2	Les dimensions généralisées	75
3.3.2.1	Présentation théorique	75
3.3.2.2	Application du calcul des dimensions généralisées à la GWFA	77
3.3.3	Le spectre de singularité	80
3.3.3.1	<i>Coarse multifractal spectrum</i>	80
3.3.3.2	<i>Fine multifractal spectrum</i>	80
3.3.3.3	Le formalisme multifractal	82
3.3.4	Quantification du <i>MAUP</i> au moyen de l'analyse multifractale : application à l'étude de la population de France métropolitaine	83
3.3.4.1	Données utilisées	86
3.3.4.2	Estimation des indicateurs multifractals	86
3.3.4.3	Choix des paramètres de l'analyse	86
3.3.4.4	Résultats	88
3.4	Analyse multifractale d'une fonction (<i>i.e.</i> une image)	95
3.4.1	De l'analyse multifractale d'une mesure à celle d'une fonction	95
3.4.1.1	Exemples de fonctions invariantes d'échelle	96
3.4.2	Présentation théorique de l'analyse multifractale d'une fonction	97
3.4.2.1	Mesure des fluctuations locales	97
3.4.2.2	Première méthode d'analyse multifractale d'une fonction : la transformée de Fourier	99
3.4.2.3	Présentation intuitive de la transformée en ondelettes	99
3.4.2.4	Deuxième méthode d'analyse multifractale d'une fonction : les Maxima du Module de la Transformée en Ondelettes	101
3.4.2.5	Une méthode plus performante et plus simple à mettre en œuvre : les ondelettes dominantes	103
3.4.2.6	Les log-cumulants : résumer le spectre par deux paramètres	105

3.4.3	Application de l'analyse multifractale en ondelettes dominantes aux orthophotographies de France métropolitaine	106
3.4.3.1	Données et choix des paramètres	106
3.4.4	Résultats de l'analyse	106
3.4.5	L'invariance d'échelle : mythe ou réalité?	108
3.4.5.1	L'invariance d'échelle en tant que norme sociale	108
3.4.5.2	Le statut de l'invariance d'échelle dans les sciences	109
3.4.5.3	Les lois d'échelles en géographie	110
3.4.5.4	Quelques éléments de réflexion sur la pertinence de comparer des dimensions fractales	111
3.4.5.5	Le pouvoir symbolique des fractales	113
3.5	Conclusion	115
4	Dépasser le présupposé d'invariance d'échelle via l'analyse des signatures trans-échelles	116
4.1	Introduction	117
4.2	Les signatures trans-échelles	119
4.3	Analyses statistiques des signatures trans-échelles	120
4.3.1	Analyse en composantes principales (ACP)	120
4.3.2	Application aux tissus bâtis de France métropolitaine (GWFA avec $q = 0$)	121
4.3.3	Implications épistémologiques	126
4.3.4	Comparaison des tissus bâtis locaux et de la forme des enveloppes morphologiques de 20 agglomérations urbaines françaises	127
4.4	Décomposition des signatures trans-échelles au moyen de polynômes orthogonaux	129
4.4.1	Présentation théorique	129
4.4.2	Relation entre la décomposition au moyen de polynômes orthogonaux et les signatures trans-échelles	131
4.4.2.1	Mise en œuvre sur le bâti de la France métropolitaine	132
4.4.2.2	Mise en œuvre sur le bâti de la petite couronne parisienne	133
4.5	Simulation de formes bâties rurales à l'aide de l'analyse trans-échelle	139
4.6	Conclusion	142
5	Conclusion	143
5.1	Retour sur la prise en compte de l'espace dans les analyses multi-échelles et trans-échelles	144
5.1.1	Mesurer la concentration et la dispersion d'une population	144
5.1.2	Concentration et dispersion dans le cas de la ségrégation socio-spatiale	148
5.1.2.1	Mesure multi-échelle de la ségrégation	148
5.2	Le mitage : résultat de l'étalement urbain contemporain ou héritage du passé?	151
5.3	Perspectives : <i>deep learning</i> et réseaux convolutifs	154
5.3.1	Exemple d'application : typologie des styles parcellaires	157
5.3.2	Questions autour de l'usage du <i>deep learning</i>	159
5.3.2.1	Les limites théoriques du <i>deep learning</i>	159
5.3.2.2	Comment contourner l'étape de labellisation?	160
5.3.3	Relation entre l'analyse fractale et le <i>deep learning</i>	161
6	Annexes	164

A	Spatial Distribution of Human Population in France : Exploring the Modifiable Areal Unit Problem Using Multifractal Analysis	165
B	Applying two fractal methods to characterise the local and global deviations from scale invariance of built patterns throughout mainland France	188
C	Principes à la base du <i>deep learning</i>	212
C.1	Perceptron multicouche	212
C.2	Étape d'apprentissage : estimation des poids	213
C.3	Algorithme du gradient descendant	214
C.4	Réseaux convolutifs : prise en compte de la spatialité des données	216

Table des figures

1.1	Répartition de la population actuelle et passée en France métropolitaine. Cartes lissées (rayon = 20 km) réalisées par l'auteur à l'aide du package Python <code>btbpy</code> . Source : Insee 2019	15
1.2	Le mitage et la géographie rurale	16
1.3	Deux configurations archétypales de bâtiments. Les points noirs représentent les bâtiments. Les configurations sont inscrites dans des carrés de 32 km de côté. Sources : les contributeurs d'OpenStreetMap et Bd Topo IGN (2012)	18
1.4	Courbe de Lorentz appliquée à l'étude de la concentration de la population. Δ (la longueur du segment rouge) est la valeur de l'indice de Hoover, et la surface en pointillés est l'aire associée à l'indice de Gini. Source : BIRABEN et DUHOURCAU 1974	20
1.5	Comparaison de l'évolution de l'indices de Hoover pour différentes mailles spatiales. Source : DUNCAN 1957	20
1.6	Concentration et dispersion de l'analyse d'un semis de points selon l'échelle d'observation (HAGGETT, CLIFF et FREY 1977)	21
1.7	Évolution de l'entropie spatiale de la population de France métropolitaine. L'entropie est calculée sur des données lissées à partir d'un noyau <code>biweight</code> (section 3.2.3) paramétré par un rayon de lissage (20, 40 et 80 km). Plus le rayon de lissage est grand et plus l'entropie est petite. Les 3 courbes atteignent leur valeur minimale au milieu des années 1970, période où l'exode rural se termine et la périurbanisation devient un phénomène massif. Source : Insee 2019	22
1.8	Paysage de bocage et d'openfield vus du ciel et depuis la route. Source : Google Maps consulté le 18/05/2020	30
1.9	Localisation des configurations de bocage (à l'ouest) et d'openfield (à l'est).	31
2.1	Trois configurations de référence : agrégée, aléatoire et régulière	35
2.2	Exemples de réalisations d'un processus de points aléatoires	36
2.3	Exemples de réalisations d'un processus binomial	37

2.4	Exemples de réalisations d'un processus de Poisson d'intensité égale à 30. Contrairement au processus binomial dont on connaît exactement le nombre de points, le nombre de points d'une réalisation d'un processus de Poisson est lui-même une variable aléatoire. Ainsi, le nombre de points de la configuration de gauche est 22, 27 pour celle du milieu et 35 pour celle de droite.	38
2.5	Mise en œuvre de la fonction de Besag pour les trois distributions de référence de la figure 2.1	41
2.6	Répartition des bâtiments des deux distributions archétypales de la figure 1.3	42
2.7	Fonction de Besag pour les deux distributions archétypales de la figure 2.6	43
2.8	Densité lissée pour les deux distributions archétypales de la figure 2.6 . Les rayons de lissage ont été choisis par validation croisée.	43
2.9	Test DBSCAN sur le territoire de bocage (en bleu les points agrégés et en rouge les non agrégés)	47
2.10	Test DBSCAN sur le territoire d'openfield (en bleu les points agrégés et en rouge les non agrégés)	48
2.11	En haut, la différences de taux en fonction des paramètres eps et min_{pts} . En bas, les résultats de la DBSCAN pour les paramètres optimaux (un rayon de 330m et un nombre de voisins de 90).	50
2.12	Test de sensibilité du jeu de paramètres optimal, en haut les résultats de la DBSCAN pour le test 1, au milieu pour le test 2 et en bas pour le test 3.	51
2.13	Carte des parts lissées de bâtiments appartenant à un agrégat (à gauche légende comparable, à droite légende obtenue par discrétisation de Jenks). Le rayon de lissage est de 8 km.	53
2.14	Carte des parts lissées de bâtiments appartenant à un agrégat (à gauche légende comparable, à droite légende obtenue par discrétisation de Jenks). Le rayon de lissage est de 8 km.	54
2.15	Part de la population agglomérée en 1876. Source : dénombrement de la population française de 1876, Insee.	56
2.16	Analyse de sensibilité (en bleu les points agrégés et en rouge les non agrégés) d'une zone contenue dans le département de l'Eure. La part des bâtiments appartenant à une zone agrégée est très sensible aux paramètres choisis. On constate à partir de cet exemple, la difficulté à définir de façon univoque la notion de village.	58
3.1	Exemples de cascades fractales	64
3.2	Exemples de grilles recouvrant des configurations de bocage (en haut) et d'openfield (en bas)	66
3.3	$\log(N_\delta)$ en fonction de $\log(\delta)$ pour les deux territoires idéaltypiques (fig. 1.3 et 2.6). A gauche, les droites représentent l'ajustement linéaire pour des grilles de tailles δ comprises entre 50 – 800 m.	68
3.4	Fonctions d'échelles de la dimension de corrélation des configurations idéaltypiques de bocage et d'openfield pour des plages de rayons comprises entre 50 et 800 m (les lignes représentent les ajustements linéaires utilisés pour calculer les dimensions fractales).	70
3.5	Les différentes notions mises en œuvre dans la GWFA. Source : SÉMÉCURBE, TANNIER et ROUX 2019	71

3.6	Comparaison GWFA avec la dimension de corrélation et densité du bâti en France métropolitaine (les rayons considérés pour l'analyse fractale sont 50; 100; 200; 400; 800 m, le rayon de lissage pour les deux analyses est de 8 km et la taille des carreaux est de 2 km).	72
3.7	Exemples de cascades multifractales remarquables	74
3.8	Exemples de cascades multifractales	75
3.9	Dimensions généralisées en fonction de q des 3 cascades multifractales de la figure 3.8	76
3.10	Étude de sensibilité aux paramètres de la GWFA ($q = 0$). Source : SÉMÉCURBE, TANNIER et ROUX 2019	78
3.11	Dimensions généralisées de la GWFA pour différentes valeurs de q	79
3.12	Spectres multifractals des cascades multifractales de la figure 3.8	83
3.13	Déformation des histogrammes à travers les échelles. Chaque histogramme décrit pour une échelle donnée la répartition des carreaux selon la population qu'ils contiennent. Dans la colonne de gauche, la distribution de population est issue d'une cascade monofractale ayant pour support un tapis de Sierpinski dont le générateur comporte 5 éléments non nuls. Dans la colonne de droite, elle est issue d'une cascade multifractale dont le générateur comporte également 5 éléments, mais cette fois-ci les poids des carreaux ne sont pas uniformes.	85
3.14	Emboîtement hiérarchique des unités spatiales d'intérêt choisit en vue de l'analyse multifractale de la distribution de la population en France métropolitaine	88
3.15	Densité de population et dimensions généralisées. Source : SÉMÉCURBE, TANNIER et ROUX 2016.	89
3.16	Dendrogramme de la CAH des spectres multifractals de la distribution de la population en France métropolitaine. Source : SÉMÉCURBE, TANNIER et ROUX 2016	91
3.17	Spectres multifractals moyens des différentes classes de la typologie obtenue par CAH	92
3.18	Carte résultant de la classification des spectres multifractals de la distribution de la population en France. 992 unités spatiales de 25 km de large. D'après SÉMÉCURBE, TANNIER et ROUX 2016	93
3.19	Massifs montagneux et forestiers et principales villes de France métropole	94
3.20	Exemples de fonctions multifractales : la première colonne représente le générateur et les colonnes suivantes, les fonctions à différentes étapes du processus de construction	96
3.21	Exemples de fonctions puissances non dérivables en 0	98
3.22	Spectre multifractal des trois fonctions de la figure 3.20. Les fonctions B et C sont multifractales et la fonction A est monofractale car son spectre est réduit à un point. Les spectres permettent également de classer les fonctions selon leur régularité (de la moins lisse à la plus lisse) : la fonction C puis la fonction A et enfin la fonction B	98
3.23	Transformée de Fourier. Source : WARD et GREENWOOD 2007	99
3.24	Exemple d'ondelettes : le chapeau mexicain. Son aire algébrique est nulle : les deux aires (-) en dessous de l'axe des abscisses compensent l'aire centrale (+).	100

3.25	Analyse MMTO de deux signaux. En haut à gauche est représentée par une courbe mathématique une portion d'ADN. L'ADN est composée de séquences de nucléotides (A, C, G et T). Aux deux premiers nucléotides (A,C) ont associée la valeur 1 et -1 aux deux derniers (G, T). En sommant cumulativement ces valeurs le long d'une séquence, on obtient une fonction caractérisant l'organisation de celle-ci. En abscisse $n(bp)$ est le numéro d'ordre des nucléotides. En haut à droite est représentée la composante longitudinale d'un champs de vitesse le long d'une direction. L'abscisse x est une position dans l'espace et l'ordonnée la valeur de la vitesse au niveau de celle-ci. Cette courbe reflète la variation des vitesses au sein d'un fluide turbulent. Au milieu, est représentée la transformée en ondelettes continues des deux signaux (la couleur rouge est associée aux coefficients de forte amplitude), en bas, la squelettisation de leur transformée d'ondelettes. Source : ARNEODO, AUDIT et al. 2008.	102
3.26	Différentes étapes de l'estimation du spectre multifractal pour les deux signaux précédents. Les ronds rouges et vides sont associées à l'ADN et les ronds bleus et pleins à l'expérience de turbulence. En haut, représentation des moments d'ordre q à travers les échelles. L'alignement des points permet de calculer des coefficients directeurs : les fonctions de partition (c). On déduit les spectres (d) en appliquant la transformée de Legendre aux fonctions de partition. Source : ARNEODO, AUDIT et al. 2008.	103
3.27	Exemples d'ondelettes de Daubechies	104
3.28	Schéma de calcul des ondelettes dominantes. Source : WENDT 2008. En bas de la figure se trouve les échelles les plus fines (2^{j-2}). Chaque point bleu représente un coefficient d'ondelettes. Le cercle rouge est un coefficient d'ondelettes dominante. Il est construit en prenant le maximum des coefficients d'ondelettes (en valeur absolue) contenus dans le rectangle noir.	104
3.29	Deux exemples d'orthophotographies. A gauche un paysage d'openfield et à droite de bocage. La valeur du c_1 est de 0,4 à gauche et de 0,06 à droite. La valeur du c_2 est de -0,4 à gauche contre -0.1 à droite.	105
3.30	Carte des coefficients c_1 et c_2	107
3.31	109
3.32	Analyse multifractale de Kyoto en 1912. Source : SALAT et al. 2018.	112
3.33	Juxtaposition de deux fractales et courbe de comptage associée	112
4.1	Résultat de la GWFA pour $q = 0$. Source : SÉMÉCURBE, TANNIER et ROUX 2019	117
4.2	Dimension fractale des tissus bâtis de France métropolitaine obtenue avec la GWFA pour $q=0$: (a) entre 50 et 200 m; (b) entre 200 et 800 m	118
4.3	Signature trans-échelle de deux configurations bâties idéaltypiques	120
4.4	Représentation des taux de variation $Scr_a - r_b$ calculés à partir de l'analyse faite avec la GWFA sur le bâti de la France métropolitaine pour des carreaux de 2 km de côté	121
4.5	ACP d'un nuage de points décrits par deux variables. Sur cette figure les deux axes de l'ACP sont représentés par les deux flèches noires.	122

4.6	Résultats de l'analyse en correspondante principale des signatures trans-échelles des tissus bâtis de France métropolitaine. Source : SÉMÉCURBE, TANNIER et ROUX 2019	122
4.7	Cartographie des deux premiers axes factoriels. Source : SÉMÉCURBE, TANNIER et ROUX 2019	123
4.8	Algorithme de calcul des nuées dynamiques (K-moyenne). Source : Wikipédia (fr.wikipedia.org/wiki/K-moyennes , consultée le 21/08/2019)	124
4.9	Scree plot de la K-moyenne appliquée aux résultats de l'ACP présentée en figure 4.7. Source : SÉMÉCURBE, TANNIER et ROUX 2019	125
4.10	Représentant idéaltypique de chaque classe issue de la typologie des signatures trans-échelles des espaces bâtis de France métropolitaine. Source : SÉMÉCURBE, TANNIER et ROUX 2019	125
4.11	Typologie de 145 178 tissus bâtis de France métropolitaine basée sur la caractérisation de la façon dont la distribution spatiale des bâtiments dévie localement d'une distribution invariante d'échelle. Source : SÉMÉCURBE, TANNIER et ROUX 2019 (séparer en deux pour faire deux figures)	126
4.12	Caractérisation des tissus bâtis de deux aires urbaines françaises : Rennes et de Reims. Le trait noir représente la délimitation de l'agglomération morphologique identifié au moyen de MorphoLim. Source : SÉMÉCURBE, TANNIER et ROUX 2019	127
4.13	Modélisation des fonctions d'échelles à l'aide d'un polynôme d'ordre 1 et d'un polynôme d'ordre 2	129
4.14	Polynômes orthogonaux calculés sur la plage d'échelle 50 – 800 m	131
4.15	Modélisation des signatures trans-échelles de deux configurations bâties idéaltypiques à l'aide de polynômes orthogonaux	132
4.16	Décomposition en polynômes orthogonaux des signatures trans-échelles calculées avec la GWFA sur la plage d'échelles 50-800 m pour $q = 0$: valeur du coefficient a_2	133
4.17	Analyse trans-échelle (GWFA) du bâti de la petite couronne parisienne. Une discrétisation par quantiles a été appliquée à chaque carte.	134
4.18	Part des bâtiments de surface au sol inférieur à $140 m^2$	135
4.19	Deux quartiers de banlieue parisienne (commune de Colombes) et leur signature trans-échelle. Chaque quartier est inclus dans un carré de 500 m de côté. Source : BD Topo 2017	137
4.20	Calcul de la signature trans-échelle d'un pixel appartenant à une maison individuelle du quartier pavillonnaire de Colombes. La dimension locale est de 1.31 et $a_2 = 0.75$. Source : BD Topo 2017	137
4.21	Analyse trans-échelle des deux quartiers de Colombes. En haut les dimensions locales et en bas les écarts à l'invariance d'échelle. Légende pour les dimensions locales : en bleu, les dimensions locales inférieures à 1.5, en jaune, comprises entre 1.5 et 2, en orange et rouge, les dimensions supérieures à 2. Légende pour le coefficient a_2 : en bleu les valeurs inférieures à -0.5 , en vert comprises entre -0.5 et 0, en orange entre 0 et 0.5 et en rouge supérieures à 0.5. Source : BD Topo 2017	138
4.22	Comparaison entre les configurations simulées et observées	140
4.23	Modélisation trans-échelle dans MobiSim-DR. Source : TANNIER 2020 .	141

5.1	Concentration et dispersion. Les deux configurations sont construites par désagréations spatiales successives (cf section 4.5) à l'aide de deux générateurs. Le premier est associé à la dimension fractale $\frac{\log(4)}{\log 3}$ et le deuxième, à la dimension $\frac{\log(8)}{\log(3)}$. L'échelle la plus grossière est en haut. Le changement de régime a lieu entre la deuxième et la troisième étape : on change de générateur.	145
5.2	Analyse multi-échelle de la ségrégation. Source : REARDON, MATTHEWS et al. 2008	149
5.3	Exemple de spectre bivarié théorique et estimé. Source : JAFFARD et al. 2019	150
5.4	Mitage. Construction isolée à plus de 300 mètres d'une zone bâtie en 2005. Source : ALBIZZATI, POULHES et SULTAN PARRAUD 2017	151
5.5	Carte de Cassini des deux configurations archétypales. Les carrés bleus représentent l'emprise spatiale des deux territoires. Source : geoportail.gouv.fr	152
5.6	Réseau viaire des deux configurations archétypales. Source : OpenStreetMap	153
5.7	Convertisseur Haute basse Tension. Les contours rouges représentent les départements. Source : Open Data Enedis	154
5.8	L'intelligence artificielle, le <i>machine learning</i> et le <i>deep learning</i> . Source : Microsoft.	155
5.9	Exemples de styles parcellaires. Source : Géoportail	158
5.10	Carte des styles parcellaires. En vert : carreaux dont l'agencement des parcelles est similaires à celles du département de la Manche, en orange à celles du département de la Marne. Le rouge correspond aux parcelles urbaines (HIMPENS, SÉMÉCURBE et POULHES 2019). Source : Etalab 2019	158
5.11	Exemple d'auto-encodeur. Source : GUO et al. 2017	161
5.12	Dispersion de l'habitat rural dans le bassin parisien. Source : COLAS 1945	162
C.1	Schéma de fonctionnement d'un neurone artificiel.	213
C.2	Perceptron multicouche. Source : Wikipédia (Perceptron multicouche) consultée le 07/03/2020.	214
C.3	Sur-apprentissage. En vert, les points de l'ensemble d'apprentissage et une régression linéaire sur ces points. En rouge, les points de l'ensemble de test. Source : Wikipédia (Surapprentissage) consultée le 07/03/2020. . . .	215
C.4	Algorithme du gradient descendant	215
C.5	Convolution	217
C.6	Pooling	217
C.7	LeNet est un réseau proposé en 1998 par l'équipe de Yann LeCun pour lire les chiffres contenus dans une image. Il se compose de deux étapes de convolution permettant de passer d'une image de taille 32×32 pixels à une image de 5×5	218

Liste des tableaux

3.1	Générateur et première étape de construction d'une cascade multifractale	74
3.2	Probabilités associées à chaque cellule des générateurs (dans l'ordre, cas 1, cas 2 et cas 3)	75
3.3	Estimation de la qualité des analyses à l'aide des coefficients de détermination	88
4.1	Caractéristiques morphologiques de 20 agglomérations urbaines. Source : SÉMÉCURBE, TANNIER et ROUX 2019	128
4.2	Corrélations statistiques entre le seuil de distance caractérisant chaque agglomération morphologique et les types de tissus bâtis qu'elles contiennent, pour 20 aires urbaines françaises de taille moyenne. p-values *** $p < 0,001$ ** $p < 0,01$ * $p < 0,05$. Source : SÉMÉCURBE, TANNIER et ROUX 2019 . . .	128
4.3	Estimation de la dimension fractale selon différents modèles statistiques .	130

Chapitre 1

Introduction

1.1 Intérêt de l'étude de la distribution spatiale des implantations humaines

1.1.1 Contexte

L'impact des modes d'occupation de l'espace sur la durabilité des modes de vie est au coeur des préoccupations des acteurs publics et ce à toutes les échelles de gouvernance. En 2018, l'OCDE a publié un rapport intitulé *Rethinking Urban Sprawl : Moving Towards Sustainable Cities* sur l'étalement urbain. Il y est abordé sa mesure, les externalités négatives qui lui sont associées et les leviers que peuvent mobiliser les décideurs pour renverser la situation actuelle. Concomitamment, l'ONU a introduit dans ses objectifs de développement durable à l'horizon de 2030 une thématique « Faire en sorte que les villes et les établissements humains soient ouverts à tous, sûrs, résilients et durables »¹. L'enjeu est de garantir la capacité des villes à générer des emplois et de la prospérité, sans grever les sols et les ressources naturelles. La cible 7 de cette thématique, « D'ici à 2030, assurer l'accès de tous, en particulier des femmes et des enfants, des personnes âgées et des personnes handicapées, à des espaces verts et des espaces publics sûrs », témoigne de l'importance que les sociétés modernes accordent à l'organisation de l'espace sur le vivre ensemble et l'émancipation des individus. En France, la loi pour l'accès au logement et un urbanisme rénové (ALUR) promulguée en 2014 et la loi portant sur l'évolution du logement, de l'aménagement et du numérique (ELAN) promulguée en 2018, sont les derniers outils produits par les cadres institutionnels visant à limiter l'étalement urbain et adapter la société française au changement climatique. L'aménagement des centres-villes et des banlieues complète les prérogatives de l'État pour limiter les effets indésirés induits par des configurations spatiales spécifiques. Les quartiers de la politique de la ville² jugés enclavés sont réinsérés dans leur trame urbaine grâce à l'extension des réseaux de transport en commun existants et la destruction des éléments architecturaux faisant office de barrières géographiques. Le bâti collectif dégradé de ces quartiers est rénové et parfois détruit afin de transformer le cadre de vie des habitants. Souvent en coordination avec ces politiques, le plan action Cœur de ville annoncé par Jacques Mézard, ministre de la Cohésion des territoires le 27 mars 2018 se fixe pour objectif de redynamiser les centres-villes en déclin³.

1. www.un.org/sustainabledevelopment/fr/cities/ consulté 18/05/2020

2. www.cohesion-territoires.gouv.fr/acteurs-de-la-politique-de-la-ville consulté le 18/05/2020.

3. www.cget.gouv.fr/dossiers/action-coeur-de-ville consulté le 18/05/2020

Le Commissariat générale à l'égalité des territoires (CGET) coordonne ce dispositif sur les 222 villes sélectionnées et « assure l'implication de tous les acteurs nécessaires dans des projets intégrant toutes les dimensions de revitalisation des centres-villes (habitat, commerces, transports, etc.) ».

Pour comparer les avantages et les inconvénients de telle ou telle organisation spatiale, la caractérisation spatiale des implantations humaines est une étape préalable indispensable, sachant toutefois qu'il ne saurait exister d'organisation spatiale bonne ou mauvaise dans l'absolu, toute analyse d'une implantation humaine devant être contextualisée à partir de l'usage qu'en font les sociétés. Par exemple, l'étalement urbain est devenu un problème social et environnemental uniquement parce qu'il est associé à un mode de vie reposant principalement sur l'utilisation de la voiture individuelle. Avant la révolution industrielle et l'exode rural, les espaces ruraux français étaient nettement plus peuplés (*fig. 1.1*) sans que cela n'induisse des conséquences sociales et environnementales comparables à celles observées actuellement.

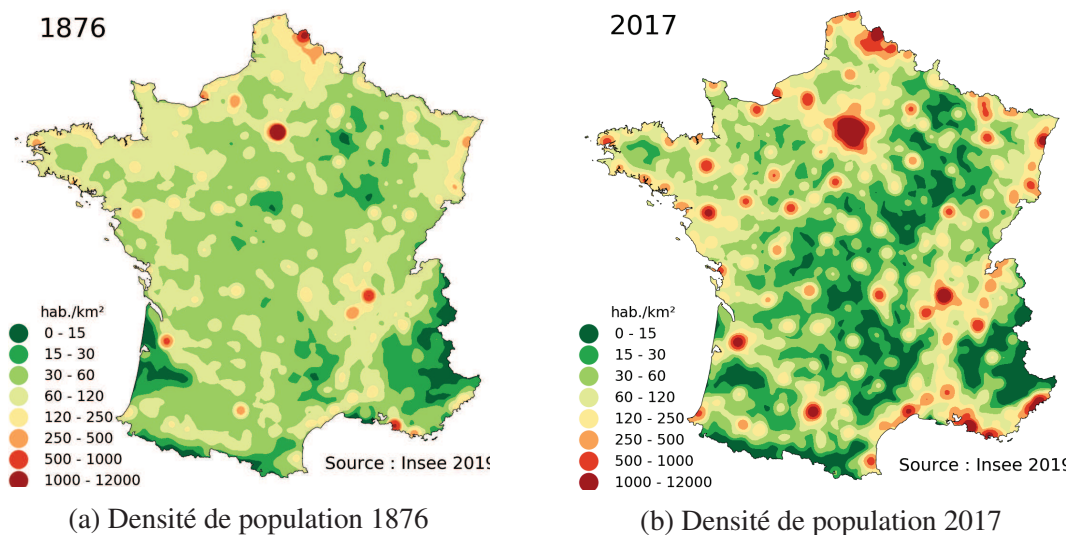


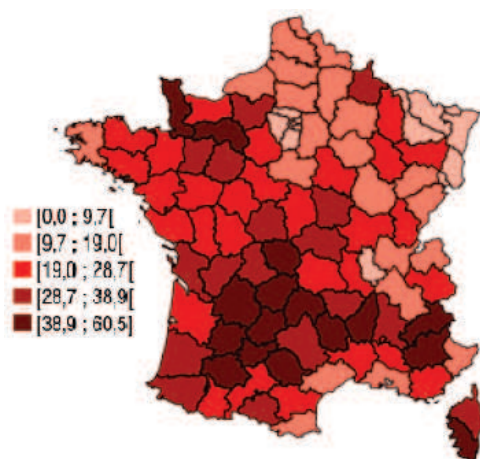
FIGURE 1.1 – Répartition de la population actuelle et passée en France métropolitaine. Cartes lissées (rayon = 20 km) réalisées par l'auteur à l'aide du package Python btbpy. Source : Insee 2019

1.1.2 L'étude de la distribution spatiale des implantations humaines au cœur des préoccupations des géographes

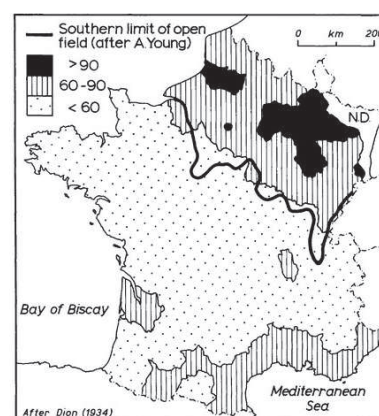
Décrire et caractériser la répartition des implantations humaines dans l'espace est une étape préliminaire pour comprendre les relations qu'entretiennent les hommes avec leur environnement. Au début du 20^e siècle, les analyses quantitatives étaient principalement réservées à l'étude de l'habitat rural. La concentration et la dispersion de la population rurale alimentaient les réflexions sur l'origine des deux systèmes agraires prédominant en Europe : le bocage et l'openfield. Les travaux de CHRISTALLER 1933 sur les lieux centraux orientèrent après guerre les travaux méthodologiques des géographes qui se mirent à tester la validité empirique de sa théorie. Les analyses géographiques quantitatives devinrent dans ce cadre un outil pour penser l'aménagement du territoire.

Depuis les années 1990, une grande part des études géographiques portent sur le phénomène d'étalement urbain. Dans ce cadre, les mesures quantitatives des espaces urbains sont considérées comme des outils de pilotage de l'action publique pour l'aménagement

de villes durables et résilientes. Le mitage des terres agricoles en périphérie des villes est une des manifestations les plus visibles de l'étalement urbain et est l'objet d'une lutte de la part des pouvoirs publics, qui s'est traduite notamment par la promulgation en 2000 de la loi relative à la solidarité et au renouvellement urbains (SRU). Le centre commercial et la maison pavillonnaire, marqueurs physiques et symboliques de l'artificialisation anarchique des terres agricoles, sont accusés d'avoir défigurés et homogénéisés les paysages (POLONY 2014; RAZEMON 2017). Pire, ils impacteraient également le mode de vie urbain car le déclin commercial et résidentiel des centres-villes s'expliquerait par la concurrence déloyale des centralités périphériques fréquentés par les périurbains. La grande majorité de ces discours essentiellement normatifs, verse souvent dans le catastrophisme : si rien n'est fait, la ville aura colonisé tout l'espace avant que l'on ne puisse réagir. Or cette hypothèse où les territoires régionaux et nationaux sont pensés à l'aune de la ville ne résiste pas à la confrontation des chiffres : « même au rythme d'un département tous les 7 ans, il faudra près de trois siècles avant que la moitié du territoire français soit artificialisée » (CHARMES 2013). Par ailleurs, la géographie du mitage n'épouse pas le contour des villes et de leurs périphéries mais se superpose approximativement aux vieux découpages de la géographie rurale entre terres de bocage de population dispersée et terres d'openfield de population concentrée (ALBIZZATI, POULHES et SULTAN PARRAUD 2017) (fig. 4.6). L'inertie des formes de peuplement qui subsistent aux contextes sociaux qui les ont vu émerger participe à limiter ou favoriser l'étalement urbain, de sorte qu'il ne saurait être compris comme un phénomène purement quantitatif. Les formes d'étalement urbain s'adaptent aux différentes configurations paysagères qu'il rencontre. C'est pourquoi l'analyse de la distribution spatiale des implantations humaines selon un prisme qui ne soit ni uniquement urbain, ni uniquement rural, est une nécessité pour comprendre les facteurs géographiques qui expliquent les formes des structures urbaines actuelles. Autrement dit, pour comprendre la forme des espaces urbains, nous devons comprendre également celle des espaces ruraux.



(a) Proportion de constructions neuves construites en dehors des espaces villageois entre 2005 et 2013 (ALBIZZATI, POULHES et SULTAN PARRAUD 2017)



(b) Carte du bocage et de l'openfield selon les relevés d'Arthur Young de 1787. Le dégradé représente la part de la population rurale vivant dans les chefs-lieux en 1891. Le trait noir symbolise les limites de l'openfield observées par Arthur Young. (CLOUT 2013)

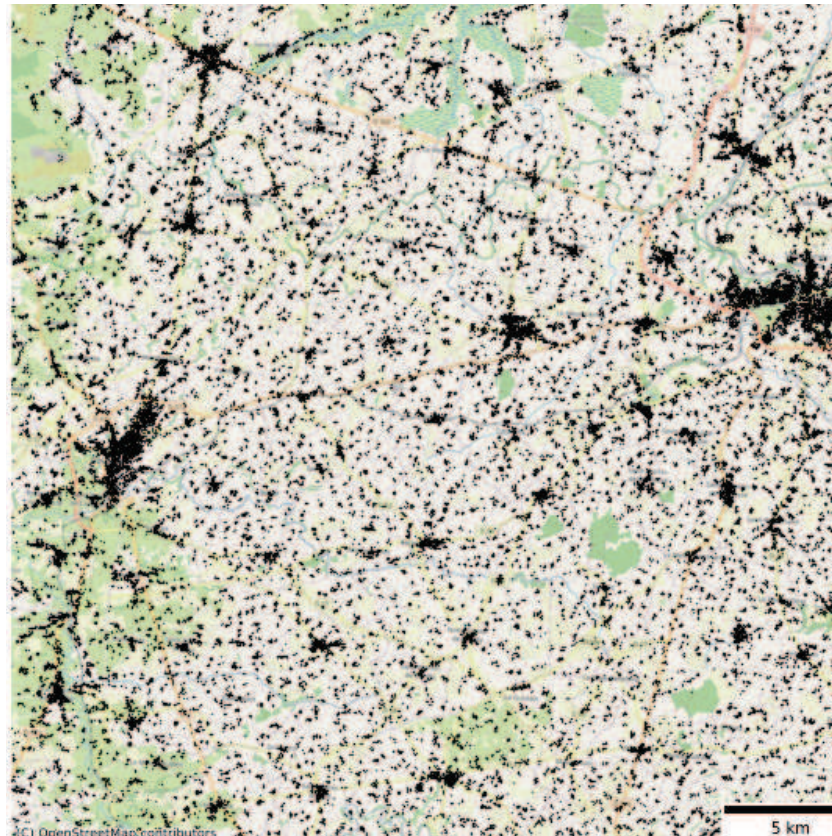
FIGURE 1.2 – Le mitage et la géographie rurale

1.2 Méthodes de caractérisation de la distribution spatiale des implantations humaines

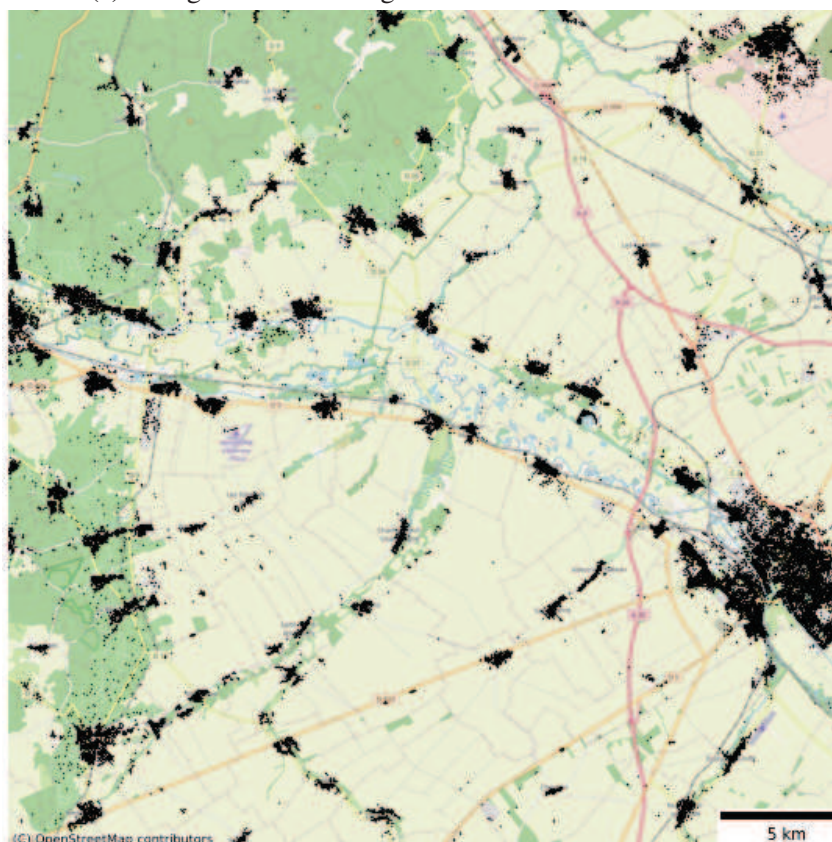
La description de la forme des implantations humaines présente un aspect déroutant : elle est à la fois très difficile à formaliser théoriquement dans un cadre général, et en même temps n'importe quel enfant est capable de catégoriser des formes spatiales complexes. Par exemple, les deux images de bocage et d'openfield (*fig. 1.3*) issues d'une exploitation de Google Maps et d'une couche du bâti (IGN BD Topo) présentent des configurations bâties totalement différentes : dans l'openfield, les bâtiments sont principalement concentrés dans des villages, contrairement au bocage où des bâtiments isolés forment un mouchetis dense et homogène. Ceci a incité certains géographes à réaliser des cartes de concentration de l'habitat rural à l'aide de leur sens plutôt que de recourir à des formules mathématiques (COLAS 1945 ; LEFÈVRE 1926). L'information contenue dans des cartes à grande échelle était synthétisée dans des cartes de paysages en recourant uniquement à la perception visuelle et à la dextérité manuelle. À présent, les services de cartographie en ligne (Google Maps et OpenStreetMap) permettent à tout à chacun d'explorer les paysages et de constater leur grande diversité. Pour autant, la facilité d'accès à ces informations et les capacités de traitement informatique encore inimaginables il y a quelques années n'ont pas transformé notre connaissance de la distribution spatiale des implantations humaines.

Mimer le comportement du cerveau humain afin de réaliser des tâches complexes ou difficiles à réaliser avec les approches classiques est un champ de recherche en plein développement. MARR et HILDRETH 1980 proposent une théorie de la vision dont les implications irriguent encore les travaux les plus récents en apprentissage profond. Selon cette théorie, l'analyse des images par le système visuel humain repose sur une structure hiérarchique compliquée qui implique plusieurs couches de traitement. À chaque couche, la rétine fournit une représentation visuelle qui évolue progressivement de manière géométrique (GRAPS 1995). Au fur et à mesure du traitement, alors que l'échelle d'analyse devient de plus en plus grossière, la représentation du phénomène observé devient de plus en plus abstraite et le sens donné au contenu d'une image émerge à travers les échelles.

En géographie, différentes approches ont été proposées pour mesurer l'organisation spatiale des implantations humaines. Un panorama des approches classiques utilisées peut notamment être trouvé dans DUNCAN 1957 ; BIRABEN et DUHOURCAU 1974.



(a) configuration de bocage : entre Coutances et Saint-Lô



(b) configuration d'openfield : à l'ouest de Châlons-en-Champagne

FIGURE 1.3 – Deux configurations archétypales de bâtiments. Les points noirs représentent les bâtiments. Les configurations sont inscrites dans des carrés de 32 km de côté. Sources : les contributeurs d'OpenStreetMap et Bd Topo IGN (2012)

1.2.1 Premiers indices de concentration de l’habitat et de la population

Très tôt, les géographes se sont aperçus qu’utiliser la densité était insuffisant pour répondre à la question « où se situe la population ? » et pour décrire une configuration spatiale : il était également nécessaire de connaître la façon dont les implantations humaines se positionnent les unes par rapport aux autres (BIRABEN et DUHOURCAU 1974). Au cours de la première moitié du 20^e siècle, la répartition de l’habitat rural a intéressé les géographes qui ont proposé des indicateurs basés sur des statistiques agrégées par communes issues des recensements de la population : la population concentrée autour de la mairie et la population dispersée dans les hameaux villageois et les lieux-dits. DEMANGEON 1939 est le premier à proposer une formule synthétique : $K = \frac{E \times N}{T}$ où N est le nombre d’écart, E la population globale des écarts, T la population totale de la commune et K l’indice de dispersion. La formule de Demangeon a essuyé plusieurs critiques (COLAS 1945) :

- des valeurs très différentes du triplet (E, N, T) peuvent produire des valeurs de K similaires ;
- elle ne tient pas compte de l’écartement des écarts au chef-lieu qui est un facteur essentiel de la dispersion ;
- elle est difficile à interpréter.

Beaucoup d’autres formules ont été proposées afin de dépasser les limites de celle de Demangeon. Néanmoins, comme le précise COLAS 1945 « sommes convaincus qu’il est illusoire de chercher à représenter par un nombre la dispersion rurale, de telle façon qu’à ce nombre corresponde un type de dispersion et un seul ». Finalement, seule la formule la plus simple, à savoir la part de population concentrée autour de la mairie par commune, est encore utilisée de nos jours (BRAS et TODD 2013).

1.2.2 Indices de concentration de la population construits à partir de la courbe de Lorenz

En 1905, Lorenz, un économiste américain, invente la courbe dite de Lorenz pour décrire les inégalités de revenu. La courbe de Lorenz est la représentation graphique de la fonction qui, à la part x des détenteurs d’une part d’une grandeur, associe la part y de la grandeur détenue⁴. Gini propose de synthétiser l’information contenue dans la courbe de Lorenz par une statistique, l’indice de Gini, dont la valeur est le double de l’aire comprise entre la diagonale du carré et la courbe de Lorenz. Un autre indice, celui de Hoover, le plus simple de tous les indices de mesure d’inégalité de revenu, est la portion des revenus totaux qui doit être prise à la partie la plus riche de la population et donnée à la partie la plus pauvre pour qu’il y ait une égalité parfaite⁵. Ces indices sont utilisés par les géographes pour mesurer l’inégale répartition de la population dans l’espace (DUNCAN 1957). La grandeur x est dans ce cas la population et la grandeur y la surface occupée par celle-ci (fig. 1.4).

Bien qu’encore très utilisés pour des comparaisons internationales, l’application de ces indices à l’étude de la concentration spatiale pose question. Comme le soulignait déjà DUNCAN 1957, les indices de Hoover et de Gini sont très sensibles au choix de la maille spatiale utilisée (fig. 1.5).

4. fr.wikipedia.org/wiki/Courbe_de_Lorenz consultée le 27/02/2020

5. fr.wikipedia.org/wiki/Indice_de_Hoover consultée le 27/02/2020

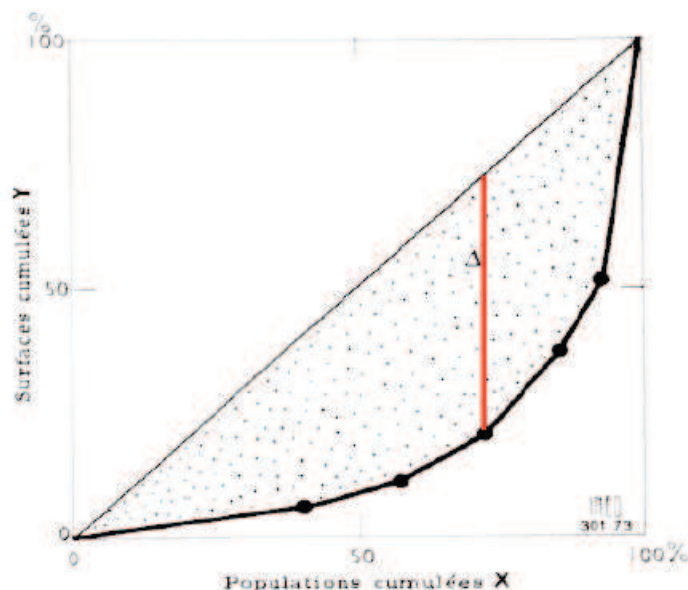


FIGURE 1.4 – Courbe de Lorenz appliquée à l'étude de la concentration de la population. Δ (la longueur du segment rouge) est la valeur de l'indice de Hoover, et la surface en pointillés est l'aire associée à l'indice de Gini. Source : BIRABEN et DUHOURCAU 1974

Year	County basis	Economic Sub-region basis	Geographic division basis
1950	58.9	50.1	39.2
1940	59.1	49.2	40.5
1930	55.9	49.1	40.8
1920	53.8	48.3	41.2
1910	52.8	48.7	41.6
1900	— *	51.6	43.6

* Not computed.
Source : United States Census publications.

FIGURE 1.5 – Comparaison de l'évolution de l'indices de Hoover pour différentes mailles spatiales. Source : DUNCAN 1957

1.2.3 Les prémisses de la statistique spatiale

CLARK et EVANS 1954 ont proposé de calculer la distance moyenne au plus proche voisin de chaque individu d'une population pour caractériser son organisation spatiale. Une comparaison de cette distance moyenne à celle obtenue si la population était répartie aléatoirement permet d'évaluer son état de concentration. En pratique, cette méthode est peu utilisée en géographie car très arbitraire (BIRABEN et DUHOURCAU 1974). Par exemple, si les individus vivent par paires mais que celles-ci soient très dispersées dans l'espace, la distance au plus proche voisin est faible, ce qui suggère que la population est concentrée. Mais si au lieu de considérer la distance au plus proche voisin, on s'intéresse à celle du deuxième voisin le plus proche, la distance devient grande et la population apparaît comme dispersée. Le nombre de voisins joue le rôle d'une échelle d'analyse, mais contrairement à des méthodes plus modernes de statistique spatiale, on ne maîtrise pas directement sa signification. Ceci fait dire à BIRABEN et DUHOURCAU 1974 que « si son application au peuplement humain permet de discerner les tendances à l'uniformité du peuplement, ce qui

est utile pour certaines théories, comme celle de Lösch, l'interprétation des tendances au groupement reste trop incertaine pour en faire un indice de concentration. ». Des méthodes basées sur les quadrats (un découpage de l'espace selon une grille carroyée homogène), ont été utilisées dans les années 1970, mais la variation des résultats obtenus selon l'échelle d'observation rend très difficile leur utilisation (HAGGETT, CLIFF et FREY 1977) (fig. 1.6).

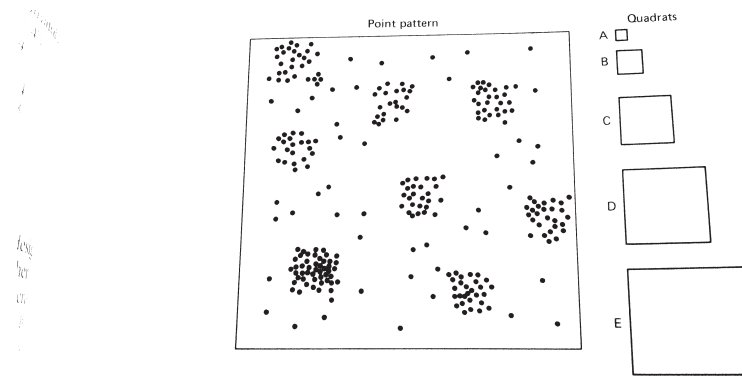


Figure 8.8 Relations between the size of quadrats and spatial pattern. Sampling with smaller quadrats (A and B) will suggest slight clustering, intermediate quadrat (C) strong clustering, and large quadrats (D and E) regularity. The reason why quadrat C will suggest strong clustering is because its size and that of the point clusters are approximately the same. In contrast, E is sufficiently large to include more than one cluster. See Section 13.4. Source: Kershaw, 1964, p. 104.

FIGURE 1.6 – Concentration et dispersion de l'analyse d'un semis de points selon l'échelle d'observation (HAGGETT, CLIFF et FREY 1977)

1.2.4 L'entropie de Shannon

MEDVEDKOV 1967 introduit l'entropie de Shannon en géographie afin de mesurer et détecter les régularités au sein d'une distribution spatiale. Pour ce faire, il déduit de la population pop_i contenue dans la maille i une probabilité $P_i = \frac{pop_i}{\sum pop_i}$ sur laquelle, il applique la formule de Shannon $H(P) = -\sum P_i \log(P_i)$. Très intéressant d'un point de vue théorique, cet indicateur présente néanmoins l'inconvénient d'être aussi particulièrement sensible au découpage spatial utilisé (BATTY 1976). Par exemple, la valeur de l'entropie augmente mécaniquement si l'on sépare en deux parties de densité identique une unité spatiale. BATTY 1976 propose de corriger cet inconvénient en calculant la divergence de Kullback–Leibler entre P et A , où A est l'aire des unités spatiales : $S = -\sum P_i \log(P_i/A_i)$. L'intérêt de ce nouvel indicateur appelé « entropie spatiale » est de permettre des comparaisons de l'organisation spatiale d'une population à différents niveaux d'agrégation et de comparer la quantité d'information apportée entre les niveaux (fig. 1.7). En cela, on retrouve des idées très proches de celles de l'analyse fractale, ce qui n'est pas un hasard quant on sait que Batty a été l'un des chercheurs ayant le premier introduit ce type d'approche en géographie (BATTY et LONGLEY 1994). Néanmoins aussi stimulante qu'elle soit, l'utilisation de l'entropie en géographie est restée principalement théorique. Ceci s'explique probablement par le fait que l'entropie ne mesure pas directement une propriété spatiale mais plutôt la complexité inhérente à une distribution spatiale (BATTY, MORPHET et al. 2014).

A l'issue de ce panorama des méthodes classiquement utilisées pour mesurer l'organisation spatiale des implantations humaines, deux points sont à retenir, que nous allons approfondir dans la section suivante :

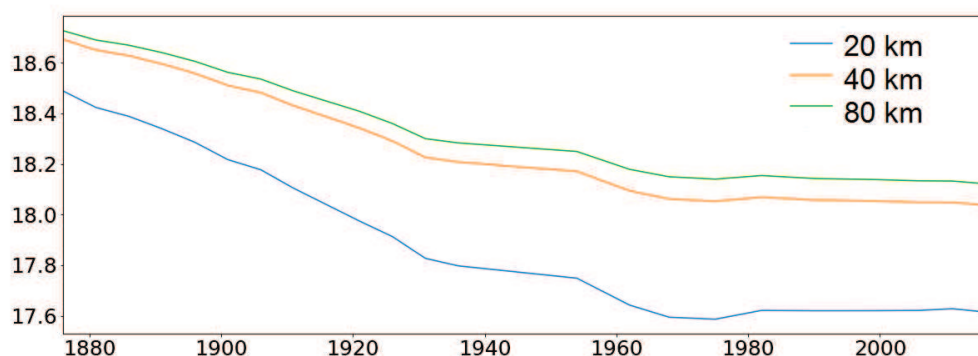


FIGURE 1.7 – Évolution de l'entropie spatiale de la population de France métropolitaine. L'entropie est calculée sur des données lissées à partir d'un noyau biweight (section 3.2.3) paramétré par un rayon de lissage (20, 40 et 80 km). Plus le rayon de lissage est grand et plus l'entropie est petite. Les 3 courbes atteignent leur valeur minimale au milieu des années 1970, période où l'exode rural se termine et la périurbanisation devient un phénomène massif. Source : Insee 2019

- la prise en compte de l'espace avec les méthodes classiques est soit inexistante, soit très rudimentaire ;
- le rôle de l'échelle d'observation est rarement pensé en tant que tel.

1.2.5 Difficultés inhérentes à la mesure des formes (configurations spatiales)

L'emploi du terme « mesurer peut induire en erreur, en laissant dans l'ombre les conventions indispensables à la quantification » (DESROSIÈRES 2014). Mesurer nécessite d'avoir au préalable défini le point de vue méthodologique que l'on porte sur le monde. Comme le précise ORAIN 2004 : « Un consensus s'est établi entre "nouveaux géographes" pour estimer que le "réel" est inconnaissable en soi et que l'on dispose toujours de schémas préalables pour l'ordonner ou lui donner du sens, ce qu'à l'époque on appelait, à la suite de Jean Piaget, des "construits" ». Cette première étape de formalisation que (DESROSIÈRES 2014) nomme convenir « implique une série de conventions préalables, de négociations, de compromis, de traductions, d'inscriptions, de codages et de calculs, conduisant à une mise en nombre ». Convenir puis mesurer sont les deux moments d'une opération plus générale : la quantification.

En géographie, convenir consiste principalement à choisir un modèle qui selon HAGGETT, CLIFF et FREY 1977 est « *a simplified version of reality, built in order to demonstrate certain of the properties of reality* ». Quantifier l'organisation spatiale des implantations humaines revient donc en premier lieu à choisir la modélisation qui permettra de relier les implantations à l'espace, soit la propriété qui nous intéresse. Cette relation qui associe à un objet ou un groupe d'objets une portion plus ou moins grande d'espace est une abstraction, une construction géographique, qui n'existe pas en dehors d'un certain consensus social. Ceci explique à la fois la grande diversité des méthodes proposées dans la section précédente pour quantifier les configurations spatiales mais aussi la dépendance des résultats aux

choix retenus. Cette limite inhérente à la plupart des analyses géographiques a été présentée sous l'intitulé de *Modifiable Areal Unit Problem* par OPENSHAW 1984b. TANNIER 2017 résume le *MAUP* en ces termes :

« Les états de concentration ou de dispersion observés dépendent de la taille des unités spatiales élémentaires considérées (scaling problem), de la forme et du positionnement des unités spatiales élémentaires (zoning problem) et de la délimitation de la zone d'étude (« results depend explicitly on the bounds of the study » (Goodchild 2016)) ».

Les effets du *MAUP* sont loin d'être anodins sur les analyses statistiques : les corrélations entre variables peuvent changer de signe selon le découpage utilisé OPENSHAW 1984b.

Encadré

Dans son article fondateur de 1984, Stan OPENSHAW relie le *MAUP* à l'erreur écologique :

« Finally, it is noted that the MAUP is also closely involved in what is known as the ecological fallacy problem . An ecological fallacy occurs when it is inferred that results based on aggregate zonal (or grouped) data can be applied to the individuals who form the zones or groups being studied ».

Un exemple d'erreur écologique devenu célèbre dans les sciences sociales est le rôle que prête Emile DURKHEIM à la confession religieuse (entre les catholiques et les protestants) sur le taux de suicide (COURGEAU 2000). L'erreur écologique se retrouve également en histoire. Dans ce cas, c'est le découpage du temps en périodes homogènes qui interroge la pertinence des interprétations obtenues (LE GOFF 2014).

Le *MAUP* est l'avatar géographique d'un problème épistémologique partagé par l'ensemble des sciences humaines : quelle signification scientifique donner à l'appartenance d'un individu ou d'un groupe à un collectif social où à un environnement spatial ou temporel particulier.

Avec les méthodes modernes d'analyse géographique basées sur des grilles régulières et des voisinages glissants, limite les effets du choix du découpage spatial sur les résultats d'analyse sont limités et le *MAUP* se résume alors à l'arbitraire induit par le choix des échelles d'analyse et de la zone d'étude englobante (BRUNSDON, FOTHERINGHAM et CHARLTON 2002). Nous laissons de côté ici le choix de la zone d'étude que nous traitons plus loin, afin de nous concentrer sur le rôle joué par les échelles d'analyse. L'échelle d'observation conditionne la quantification des formes observées. W. Tobler (2004) cité par TANNIER 2017 précise à ce sujet : « *Arbia's second law of geography reads "Everything is related to everything else, but things observed at a coarse spatial resolution are more related than things observed at a finer resolution." This suggests that aggregation has a smoothing effect, as is well known (see Tobler 1969, 1990)* ». Plus largement, BEVAN et CONOLLY 2006 constatent que « *the scale of analysis can change what appears to be a nucleated or centralized pattern, to one better described as dispersed* ».

La dépendance de la concentration et la dispersion d'un phénomène selon l'échelle d'observation peut être considérée comme une limite inhérente à la science géographique ou comme le témoignage de l'organisation multi-échelle du phénomène étudié. Selon

SANDERS 2011 : « Dans un cas le *MAUP* est vu comme un problème remettant en cause la robustesse des résultats, dans l'autre comme source de connaissance sur le phénomène étudié ». Cette dernière approche nécessite, pour être efficace, de définir explicitement les notions de concentration et de dispersion en lien avec les échelles d'observation, autrement dit de convenir d'une définition de ces termes. Ceci est d'autant plus important que les termes concentration et dispersion ne sont pas univoques. Tout à chacun les comprend sans pour autant qu'il existe une définition commune et partagée, ce qui réduit leur pouvoir heuristique (O'KEEFFE 2000).

TANNIER 2017 propose une définition de la concentration et de la dispersion à travers les échelles :

« pour un type donné d'implantations humaines, l'existence d'un processus de concentration considérant un certain voisinage implique un processus de dispersion considérant un voisinage plus large (ou plus restreint)... En conséquence de la dépendance d'échelles, les états de concentration observés à une échelle (un certain niveau d'agrégation ou une certaine résolution spatiale) impliquent des états de dispersion à une autre échelle, et vice-versa ».

Selon cette définition, une configuration spatiale est concentrée à une échelle donnée que parce qu'elle est dispersée aux échelles plus grossières et plus fines. Dispersion et concentration sont donc indissociables et sont des phénomènes intrinsèquement multi-échelles.

1.3 Objectif et contenu de la thèse

La thèse porte sur la mesure de la distribution spatiale des implantations humaines. L'objectif est de proposer des méthodes et des indicateurs multi- et trans-échelles qui permettent autant que faire se peut de s'affranchir, ou de tirer partie, du *Modifiable Areal Unit Problem*.

1.3.1 A propos des implantations humaines

« Une implantation humaine consiste en la matérialisation dans l'espace, l'inscription physique, d'une ou plusieurs activités humaines. Elle se caractérise par un certain degré de pérennité. Il peut s'agir d'un bâtiment, d'un groupe de quelques bâtiments, ou d'un ensemble plus vaste de milliers de bâtiments. Chaque bâtiment est le siège d'une ou plusieurs fonctions : résidentielles, commerciales, industrielles, artisanales, religieuses, politiques, logistiques... Il s'inscrit dans un environnement non bâti plus ou moins artificialisé (parcs, parkings, vergers, champs...). Il est relié à d'autres bâtiments, proches ou éloignés, de même fonction ou de fonction différente, au moyen de réseaux de transport et de communication. Tout ceci –bâtiments, environnement non bâti, réseaux– constitue le contenant des activités humaines (...) Le contenu associé à ce contenant sont les individus et les groupes effectuant des activités, ayant des pratiques spatiales et des relations sociales » TANNIER 2017.

Une implantation articule plusieurs notions dont le sens commun est ambigu. Selon Wikipedia⁶ un bâtiment « est une construction immobilière, réalisée par intervention humaine, destinée d'une part à servir d'abri, c'est-à-dire à protéger des intempéries des

6. Consulté le 02/05/2020.

personnes, des biens et des activités, d'autre part à manifester leur permanence comme fonction sociale, politique ou culturelle ». OpenstreetMap⁷ admet que définir concrètement un bâtiment n'est pas une chose simple. Une caravane stationnée au même endroit et occupée pendant plusieurs années n'est pas *stricto sensu* un bâtiment mais exerce une fonction comparable. La pérennité d'un bâtiment n'implique pas la continuité temporelle des activités qu'il contient. La transformation d'une ancienne ferme en une résidence secondaire implique un changement d'activité mais également une rupture qui s'inscrit dans le temps. Pendant les saisons creuses, son activité est seulement potentielle. A l'extrême, peut on encore considérer comme faisant partie d'une implantation humaine, les ruines qui ne sont plus associées à une activité humaine.

« L'artificialisation du sol ou d'un milieu, d'un habitat naturel ou semi-naturel, est la perte des qualités qui sont celles d'un milieu naturel : sa naturalité, qualité qui inclut une capacité autoentretenu à abriter une certaine biodiversité, des cycles naturels et ses qualités biogéochimiques »⁸. Cette opposition marquée entre les territoires « naturels » et ceux transformés par l'homme n'est plus totalement pertinente à l'heure de la mondialisation (BECK, LASH et WYNNE 1992). La pollution et les transformations climatiques impactent l'ensemble de la surface terrestre. Il n'existe donc plus à proprement parler d'espace naturel vierge de toute influence humaine. DESCOLA 2005 pousse plus loin le raisonnement. La forêt amazonienne avant la colonisation européenne n'était déjà plus un espace naturel. Les indiens par leur usage de celle-ci à travers les siècles l'ont transformée en modifiant la répartition des plantes en son sein. En Europe, l'ensemble des paysages, même ceux des espaces montagnards, ont été façonnés par les pratiques culturelles des hommes qui les exploitent depuis plusieurs millénaires.

La dichotomie « nature-culture » est une construction de l'Occident qui n'est pas partagée par l'ensemble des sociétés (DESCOLA 2005). Pour certaines ethnies, la différence entre les hommes et les autres entités vivantes ou inertes avec qui ils partagent leur environnement, est moins prononcées. Ceci ne remet absolument pas en cause, la portée heuristique des notions de contenant et de contenu, mais tend à préciser leur limites d'application : la géographie des espaces européens et occidentaux.

En pratique, rares sont les géographes qui collectent directement leur données. Ils exploitent des systèmes d'information géographique dont ils ne maîtrisent pas tous les biais. Par exemple, lorsque nous utilisons dans nos travaux les données du bâti de la BD TOPO de l'IGN, nous excluons implicitement les personnes sans domicile fixe, alors même que leur localisation spatiale peut-être plus stable dans le temps que celles des locataires qui les environnent. Ils n'apparaissent ni dans les contenants ni dans les contenus. L'usage des données de répartition de la population dans l'espace présente également des limites. Généralement, c'est le lieu de résidence qui permet de définir la localisation des individus. Cette information spatiale ne reflète que très imparfaitement leur localisation qui évolue en fonction des mobilités quotidiennes et touristiques.

Dans cette thèse, nous nous limitons à décrire la concentration et la dispersion des contenants.

7. wiki.openstreetmap.org/wiki/Buildings consulté le 02/05/2020.

8. Source : Wikipedia consultée le 02/05/2020.

1.3.2 Choix de la zone d'étude et des entités spatiales considérées

Analyser et comparer les propriétés d'indicateurs morphologiques n'a pas de sens dans l'absolu mais n'a de sens que pour des territoires de référence et des échelles spécifiques. Réciproquement, la valeur des indicateurs calculés est dépendante de la taille des zones d'intérêt. Il nous apparaît donc primordial de justifier les choix que nous avons faits dans cette thèse.

1.3.2.1 Choix de la zone d'étude

Si la ville est le type de territoire le plus souvent analysé actuellement, ce ne fut pas toujours le cas. Jusqu'à la moitié du 20^e siècle, les espaces ruraux représentaient le principal champ d'investigation des géographes, alors même que la population urbaine représentait déjà plus de 50% de la population de France métropolitaine en 1936. Ce décalage a été documenté par CLOUT 2005 dans son article intitulé *Geographers in Their Ivory Tower : Academic Geography and Popular Geography in Paris 1931*. L'auteur montre, à partir d'une analyse des activités du 13^e congrès de géographie de Paris en 1931, l'attrait des géographes français pour les espaces ruraux ou coloniaux. Après la Seconde Guerre mondiale, l'urbanisation de la France s'accélère et devient progressivement le sujet d'étude dominant en géographie.

Les transformations sociales et environnementales actuellement à l'œuvre nous incitent à sortir de cette vision manichéenne. Il devient de plus en plus pertinent d'étudier conjointement les espaces ruraux et urbains : l'étude de la géographie rurale est indispensable pour anticiper les adaptations des systèmes de production agricole à des conditions météorologiques de plus en plus extrêmes, tandis que celle de la géographie urbaine est nécessaire pour comprendre les déterminants de nos modes de vie et les décisions collectives qui participent à rendre le monde de moins en moins durable.

Dans ce cadre, travailler sur la totalité de la France métropolitaine nous permet de situer les analyses au niveau de l'échelle de gouvernance où la grande majorité des décisions d'aménagement de l'espace sont prises, l'État nation, et de prendre en compte l'interdépendance des territoires et des personnes. Calculer des indicateurs pour la France entière n'empêche pas ensuite d'étudier les résultats spécifiquement sur certains secteurs d'intérêt (une région ou une aire urbaine donnée).

1.3.2.2 Choix des entités spatiales élémentaires

Concernant le choix des entités spatiales sur lesquelles seront calculés les indicateurs et qui serviront de support à l'étude des variabilités spatiales des implantations humaines, le recours à des découpages administratifs ou statistiques posent plusieurs questions que nous allons aborder successivement selon leur ordre d'importance. En premier lieu, les découpages administratifs évoluent constamment et ce, quelle que soit l'échelle géographique considérée. Le nombre de communes a ainsi fortement diminué en 2016 et 2017 suite à des mesures financières incitant celles-ci à fusionner. En parallèle, la réforme territoriale a considérablement transformé, en 2015, l'emprise spatiale des régions françaises introduites à la fin des années 50. Les découpages statistiques évoluent également périodiquement. La statistique publique actualise ses découpages pour tenir compte à la fois de la demande sociale et de l'évolution des phénomènes mesurés à proprement parler. Par exemple, on attend prochainement la livraison de la quatrième version du découpage en aires urbaines fait par l'INSEE.

D'un point de vue strictement heuristique, utiliser des découpages des espaces urbains et espaces ruraux pour étudier l'organisation spatiale des implantations humaines revient à définir *a priori* ce que l'on souhaite découvrir, ce qui représente une forme de raisonnement circulaire. De plus, les espaces périurbains délimités au sein des aires urbaines ne sont pas morphologiquement homogènes. L'aire urbaine de Caen est à ce titre exemplaire et mélange la plaine de Caen, un territoire d'openfield et des espaces de bocages. L'utilisation de l'emprise de l'aire urbaine de Caen comme support de l'analyse moyennerait dans un seul chiffre les différentes formes de peuplement et d'occupation de l'espace.

Par définition, l'étalement urbain et le mitage, qui de façon imagée caractérisent la ville qui déborde de ses frontières, mettent l'accent sur la responsabilité des territoires en périphérie des villes et de leurs habitants dans la crise écologique actuelle. Or l'apparition de la périurbanisation doit autant à des comportements individuels qu'à des politiques globales. L'accès au pavillon est envisagé comme « une machination tactique montée par le conservatisme social afin de réduire les ouvriers à l'obéissance et à la docilité outil de contrôle social » (LAMBERT 2018). Dès 1945, CORBUSIER s'inquiétait des intentions politiques sous-jacentes aux développements des banlieues (cité par LAMBERT 2018) :

« On partit loin des villes, vers des oasis de nature ravissante : des champs, des bocages... Mais lorsque chacun y eut construit sa petite maison, champs et bocages avaient disparu. Au lieu du calme et de la solitude rêvés, ce fut la promiscuité des voisinages immédiats. Ce n'était pourtant qu'une immense déconvenue. D'immenses intérêts privés, cachés derrière une propagande philanthropique, aiguillaient les masses ouvrières vers cette aventure attrayante ; c'était des entreprises de lotissements, de travaux publics, de construction ; c'était encore un profond désir de répandre en poussière d'hommes dispersés aux quatre vents des cieux, les masses ouvrières dangereuses rassemblées dans les centres de découvrir trop vite les voies réelles du travail machinistes... »

L'avènement de ce que les géographes allaient nommer périurbanisation a été rendu possible par la voiture, le supermarché, le pavillon et la télévision selon GUY 1967, pour qui cette nouvelle répartition de la population détruit l'organisation sociale des espaces urbains :

« L'urbanisme qui détruit les villes reconstitue une pseudo-campagne, dans laquelle sont perdus aussi bien les rapports naturels de la campagne ancienne que les rapports sociaux directs et directement mis en question de la ville historique ».

Plus récemment, Claude Parent cité par RAMADE 2007, donne à cette répartition spatiale de la population une dimension explicitement politique de contrôle social et de pacification de la population qui n'est pas uniquement symbolique :

« Une fois tout son monde bien parqué, les vieux dans les asiles pour vieillards, les étudiants dans leurs campus, les industriels dans les zones industrielles, les vacanciers dans les clubs, les employés et les ouvriers dans leurs pavillons, la population est extrêmement facile à contrôler. Il suffit de quelques CRS bien placés, aux carrefours des autoroutes, aux entrées des villes qui seront intégralement réservées aux nantis épris de culture, aux musées et aux commerces de luxe ».

La dimension politique de la périurbanisation, loin d'être une critique portée uniquement par des intellectuels, est un choix totalement assumé par le pouvoir comme en témoigne le discours d'Orléans en 1977 du président Giscard d'Estaing (DEPRAZ 2019) :

« Il faut encourager la construction de logements individuels qui contribuent à la tranquillité sociale et aider tous les Français à accéder à la propriété, car celle-ci favorise chez chacun le sens de la responsabilité ».

Aux États-Unis, cette stratégie de pacification des classes laborieuses est plus ancienne et a débuté juste après la révolution russe de 1917. L'étalement urbain est également motivé pour ce pays par la peur des « élites blanches » de la « promiscuité raciale » (ROTHSTEIN 2017). Le pavillon devient un outil pour garantir la séparation des « races » dans l'espace et explique encore aujourd'hui largement la ségrégation observée dans les villes américaines. En France, la ségrégation raciale n'explique pas directement l'étalement urbain, le périurbain est en effet également un territoire privilégié d'accès à la propriété pour les familles issues de l'immigration (LAMBERT 2015). En revanche, le mouvement des gilets jaunes témoigne de l'existence d'un double processus qui exclut les classes populaires des villes et de leurs périphéries proches (CHARMES 2019) :

- la gentrification favorisée par la concurrence que se livrent les villes à l'échelle nationale, et internationale pour les métropoles les plus importantes, réduit considérablement la capacité des ménages les plus modestes à habiter les centres-villes,
- les banlieues défavorisées des principales grandes villes accueillent par défaut des populations paupérisées victimes de racisme et de discrimination à l'embauche et sont de ce fait socialement peu valorisées.

Le pavillon périurbain représente donc une échappatoire pour les ménages les plus modestes pour accéder à un lieu de résidence décent. Néanmoins, les frais de déplacements et d'entretien inhérents à ce type d'habitat fragilisent le mode de vie des populations périurbaines les plus vulnérables. Le recours aux découpages territoriaux institutionnels tend ainsi à invisibiliser les causes multiples de l'étalement urbain et la souffrance des ménages dont le pavillon est tout autant un choix qu'une nécessité.

Enfin, contrairement à d'autres pays européens, les espaces de très faibles densités sont extrêmement rares en France. Ceci s'explique principalement par l'existence d'un territoire de gouvernance local, la commune qui favorise la promotion de l'attractivité territoriale comme stratégie privilégiée de gouvernance (ESTÈBE 2016). Les 35000 maires et leurs conseillers municipaux représentent le premier niveau de la démocratie française. L'arrivée de nouveaux habitants est généralement perçue dans les communes en déclin comme une aubaine pour garantir l'existence des principaux services publics et des commerces de proximité. Inversement, dans les communes périurbaines les plus prospères, la « clubbisation » de la vie locale tend à exclure les nouveaux arrivants qui sont contraints de se rabattre sur des communes périphériques plus éloignées, ce qui repousse d'autant plus le front pionnier de la périurbanisation (CHARMES 2013). Les découpages territoriaux en France participent donc à la dispersion de la population. Des analyses qui reposeraient sur eux risqueraient alors de légitimer une représentation de l'espace supportant l'idée d'étalement urbain.

Ces considérations nous incitent à nous émanciper de tout découpage préalable de l'espace et de travailler sur l'intégralité de celui-ci sans privilégier aucune de ses parties. Autant que faire se peut, nous recourons à des partitions de l'espace, des carroyages réguliers comme support de nos analyses géographiques. La taille des carreaux sera choisi suffisamment petite pour garantir la robustesse des analyses

et pour restituer la variabilité des phénomènes géographiques. Ces découpages géographiques sont simples à manipuler et peuvent être superposés à d'autres couches d'information géographique, tels que des fonds administratifs afin de contextualiser les résultats obtenus.

Les méthodes d'analyse de la distribution spatiale des implantations humaines présentées dans cette thèse ont été appliquées aux deux configurations idéaltypiques de bocage et d'openfield de la figure 1.3 dont on trouve la localisation sur la carte 1.9. Elles s'inscrivent dans deux environnements ruraux très différents. Les champs de la configuration d'openfield (*fig.* 1.8.a et b), comme son nom l'indique, sont ouverts, seuls quelques rares espaces boisés découpent l'horizon. Dans la configuration de bocage, des haies et des clôtures entourent les champs et les prés (*fig.* 1.8.c et d).

1.3.3 Plan de la thèse

La thèse comporte trois chapitres correspondant à trois manières différentes de considérer le rôle des échelles dans la description et la modélisation des phénomènes analysés. Les deux premiers sont des mises en application et des approfondissements de méthodes modernes d'analyse géographique : l'analyse spatiale et l'analyse fractale. Le troisième chapitre généralise les principes de l'analyse fractale aux structures non fractales.

Dans le premier chapitre de la thèse, nous analysons la distribution spatiale des bâtiments à l'aide de la statistique spatiale. MATHIAN et SANDERS 2004 donnent la définition suivante de la statistique spatiale : « Pris dans son sens méthodologique le plus large, ce terme désigne toute analyse utilisant l'outil statistique et ayant une dimension spatiale, que cette dimension concerne l'outil proprement dit, l'objet analysé ou les variables utilisées comme descripteur de cet objet. Différentes combinaisons sont effectivement possibles ». *A priori* tous les phénomènes, dès qu'ils comportent une composante spatiale, sont susceptibles d'être étudiés à l'aide de la statistique spatiale. La nature des méthodes mises en œuvre et leur objectifs dépendent principalement de la structuration spatiale des données mobilisées (CRESSIE 1992 ; LOONIS et DE BELLEFON 2018) : continues, surfaciques ou ponctuelles.

Les données continues décrivent des phénomènes dont on connaît les propriétés en tout point du territoire d'intérêt. L'analyse de ce type de données est appelée géostatistique et « cherche à prédire la valeur d'une variable en un point où elle n'a pas été échantillonnée, ainsi que la fiabilité de cette prédiction » (LOONIS et DE BELLEFON 2018).

Un jeu de données surfaciques découpe l'espace en polygones disjoints fixes. Les valeurs quantitatives ou qualitatives associées à chaque polygone sont considérées comme aléatoire. Les données territoriales (communales, départementales...) sont des exemples de données surfaciques. L'enjeu pour ces données est de comprendre la façon dont une valeur obtenue pour un polygone influence celles des polygones voisins. Ceci requiert d'avoir défini au préalable une matrice de voisinage qui détermine une relation de voisinage entre les polygones. Elle peut être basée sur un critère de contiguïté ou de distance.

Les données ponctuelles se distinguent des deux précédentes par la nature de l'aléa. On considère cette fois que le caractère aléatoire est directement porté par les localisations. L'emplacement de plantes appartenant à une sous-espèce dans une forêt est souvent donné en exemple dans les manuels d'analyse statistique de données ponctuelles (BADDELEY, RUBAK et TURNER 2015, DIGGLE 2013). L'analyse de ces données a pour objectif de quantifier la façon dont les localisations mesurées se distinguent de celles que l'on aurait



(a) Paysage d' openfield vu du ciel.



(b) Paysage d' openfield



(c) Paysage de bocage vu du ciel.



(d) Paysage de bocage.

FIGURE 1.8 – Paysage de bocage et d' openfield vus du ciel et depuis la route. Source : Google Maps consulté le 18/05/2020



FIGURE 1.9 – Localisation des configurations de bocage (à l'ouest) et d'openfield (à l'est).

observé si la dimension spatiale du phénomènes sous-jacent était totalement aléatoire. On s'intéresse en particulier à discerner si les localisations se regroupent davantage en amas (*clusters* en anglais) ou si elles sont le résultat d'un processus générateur aléatoire. En économie géographique, l'étude de la localisation des entreprises et de magasins à l'aide des méthodes ponctuelles est un domaine de recherche florissant (LANG, MARCON et PUECH 2015). De nombreux raffinements ont été proposés pour adapter ces méthodes à la comparaison d'une distribution spécifique, par exemple les boulangeries, à celle d'une distribution de référence telle que la population sous-jacente. Dans cette thèse, nous appliquons l'analyse des données ponctuelles à l'étude de l'organisation des tissus bâtis. Cette approche se révèle au final décevante, ce qui est paradoxal car elle semble de prime abord plus simple que la comparaison spatiale de deux répartitions. Nous nous sommes efforcés de comprendre les raisons de ce paradoxe ce qui constitue l'apport essentiel de ce premier chapitre.

Dans le second chapitre, nous avons mobilisé les méthodes issues de la géométrie fractale pour décrire aussi bien la localisation des populations que des bâtiments dans l'espace. Les premières applications de cette géométrie à l'étude des tissus bâtis urbains remontent au début des années 1990 avec les travaux notamment de BATTY et LONGLEY 1994 et de FRANKHAUSER 1994. Depuis la géométrie fractale est devenue un outil standard pour mesurer et simuler les espaces urbains (TANNIER et PUMAIN 2005). Ce succès s'explique principalement par le fait que comme les villes, les structures fractales sont, par essence, très hétérogènes et hiérarchiques. Si les villes ne sont pas fractales en tant que telles, en adoptant un regard fractal sur celles-ci, on arrive à prendre en compte leur organisation singulière qui reste hors de portée avec la géométrie euclidienne qui repose sur l'homogénéité et la continuité des objets étudiés.

De nombreux travaux ont été mené sur le calcul des dimensions fractales pour à la fois améliorer la précision des estimations et étendre le champs d'application à l'ensemble des composantes du paysages urbains tels que le réseau viaire ou les lignes de transports. L'originalité de ce chapitre est d'avoir réalisées les analyses fractales sur l'intégralité du territoire métropolitain et de ne plus se cantonner aux régions urbaines. L'application de la géométrie fractale aux répartitions de la population nous a permis de transformer ce qui est considéré parfois comme une limite dans la connaissance, le *MAUP* en un outil

de description à travers les échelles. Ceci suppose pour fonctionner que les conditions d'applications de la géométrie fractale, l'invariance d'échelle soient vérifiées. Or cette propriété que l'on a observé sur les échelles mesoscopiques de notre étude de la population n'est pas vérifiée dans les espaces ruraux aux échelles les plus fines.

L'objectif du troisième chapitre est de dépasser l'hypothèse d'invariance d'échelle est de pouvoir ainsi décrire à travers les échelles toutes les configurations spatiales. L'analyse trans-échelle qui généralise l'analyse fractale est aussi ancienne que cette dernière. On la retrouve dans les travaux de BATTY et LONGLEY 1994 qui évoquent les signatures « fractales » pour mesurer la « variabilité » des dimensions fractales à travers les échelles. Plus récemment, THOMAS et al. 2010 utilisent ce concept pour caractériser la forme de quartiers urbains. Notre apport est de systématiser son utilisation en nous aidant de modélisations statistiques et mathématiques afin de d'extraire des signatures « fractales » l'information pertinente.

Chapitre 2

Méthodes de statistique spatiale pour l'analyse de la distribution spatiale des bâtiments

Résumé

Dans ce chapitre, nous avons appliqué les outils de la statistique spatiale à l'étude des tissus bâtis de France métropolitaine. La fonction de Ripley utilisée dans les sciences naturelles pour décrire la structure spatiale d'un semis de points s'est avérée inadaptée : tous les tissus bâtis étudiés apparaissent comme agrégés.

Une approche plus exploratoire basée sur la recherche d'agrégats nous a permis de révéler la grande variété des formes au sein des espaces ruraux. En revanche, l'arbitraire du choix des paramètres rend difficile l'interprétation des résultats.

2.1 Introduction

La statistique spatiale propose deux types d'analyses multi-échelles basées sur des principes similaires pour étudier et caractériser des semis de points (BESAG et NEWELL 1991) : le *clustering* et la *cluster detection*. La première a vocation à apprécier si une configuration présente des agrégats (*clusters*). Elle est basée sur une comparaison avec un objet de référence, une distribution de points aléatoire, généralement un processus de Poisson. La seconde, plus empirique, vise à circonscrire et délimiter des agrégats dans un ensemble de points. Avec cette approche, c'est le scientifique qui, par le choix ad hoc des valeurs de paramètres, définit ce qu'il considère comme étant des formes agrégées.

Les premières méthodes de statistique spatiale reposent sur des découpages géographiques (des grilles régulières voire des mailles administratives) pour appréhender le positionnement relatif des points les uns par rapport aux autres. Selon cette conception de l'espace, deux points sont considérés comme proches s'ils appartiennent à la même unité spatiale. Cette façon de mesurer la proximité souffre de deux défauts majeurs. Tout d'abord, elle ne restitue que très médiocrement le positionnement mutuel des points dans l'espace. Ensuite, elle est particulièrement sensible à l'arbitraire du découpage géographique, autrement dit au *Modifiable Areal Unit Problem*.

Les travaux de Brian D RIPLEY et Peter DIGGLE au cours des années 1970 ont introduit et banalisé en statistique spatiale l'usage de la notion de voisinage pour décrire localement des distributions spatiales. Le voisinage est, dans la majorité des techniques modernes de la statistique spatiale, un disque centré sur chaque point de l'ensemble étudié. Ce disque joue le rôle d'unité spatiale mobile et sert à dénombrer les points contenus à proximité du point d'intérêt. En faisant varier la taille des disques, différentes échelles d'observation sont introduites dans les analyses. Ainsi, contrairement aux découpages géographiques classiques, l'utilisation de voisinages limite les effets du *MAUP*.

Dans ce chapitre, nous allons nous intéresser principalement à deux méthodes basées sur les voisinages qui occupent une position centrale dans leur champ de savoir respectif :

- la fonction de RIPLEY 1977 qui caractérise globalement le caractère plus ou moins agrégé d'une distribution de points (*clustering*) ;
- la méthode DBSCAN proposée par ESTER et al. 1996 pour délimiter et trouver dans un ensemble de points des agrégats (*cluster detection*).

Nous appliquerons ces méthodes à l'étude de la répartition des bâtiments en France en nous concentrant, en particulier, sur nos deux configurations idéaltypiques présentées dans le chapitre introductif. Nous terminerons ce chapitre en discutant les limites de la statistique spatiale.

Nous ne pouvons pas appliquer directement les méthodes d'analyse spatiale à une couche cartographique de bâtiments. Il est nécessaire d'avoir au préalable transformé les polygones qui représentent les bâtiments en une couche ponctuelle. Ceci a été réalisé en assimilant chaque bâtiment à son centroïde. Les couches de centroïdes constituent les ensembles de points que nous analysons.

2.2 Analyse de configurations de points (*clustering*)

Les techniques d'analyse de configurations de points ont pour objectif d'apprécier l'indépendance spatiale des ensembles de points. Par indépendance, on entend que les points n'interagissent pas entre eux, autrement dit, que la connaissance de la localisation d'un point n'apporte pas d'information sur celle des autres points de l'ensemble étudié. Lorsque cette propriété n'est pas vérifiée, les points peuvent soit avoir tendance à se regrouper dans des amas (*clusters*) ou au contraire à se repousser mutuellement et former des configurations régulières. La figure 2.1 illustre ces trois configurations de référence.

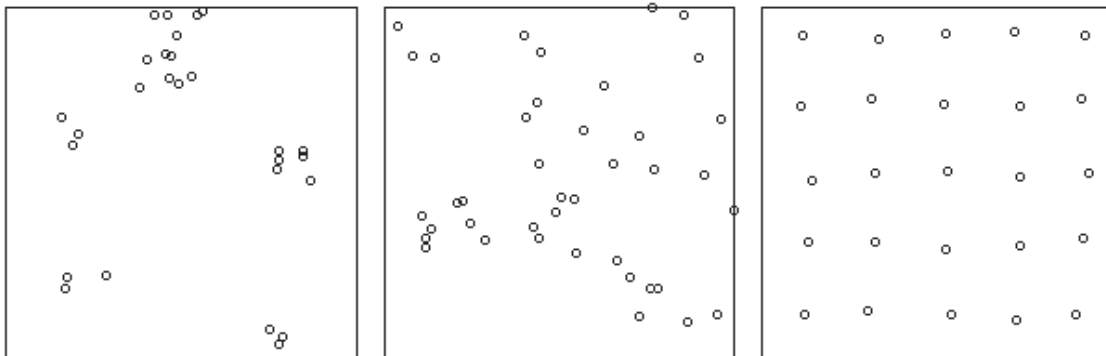


FIGURE 2.1 – Trois configurations de référence : agrégée, aléatoire et régulière

La fonction de Ripley est devenue depuis son introduction par Brian D RIPLEY au milieu des années 1970, l'outil de référence pour décrire la structure spatiale d'un ensemble de points dans de nombreux champs scientifiques (BOLIBOK 2008 ; MARCON et PUECH 2017). Elle est à la fois l'outil le plus utilisé par les chercheurs dans les études appliquées, mais aussi un cadre conceptuel central sur lequel des recherches plus théoriques s'appuient pour proposer de nouvelles méthodes d'analyse.

Comme dans le cas de nombreuses méthodes statistiques, elle est construite selon une logique de test. L'ensemble de points étudié est comparé à une distribution spatiale totalement aléatoire (*Complete Spatial Randomness : CSR*), obtenue au moyen d'un processus de Poisson. La façon dont les ensembles de points s'écartent de ce modèle nous renseigne sur leur caractère plus ou moins agrégé. Les processus spatiaux de Poisson sont paramétrés par une valeur de densité appelée intensité en statistique spatiale. Pour que la comparaison soit pertinente, la fonction de Ripley impose que l'intensité du processus de Poisson corresponde exactement à celui de l'ensemble de points étudié. Avant d'entrer dans les détails de la fonction de Ripley, nous allons d'abord introduire les processus ponctuels et plus spécifiquement ceux de Poisson.

2.2.1 Processus ponctuel Poissonien

Une variable aléatoire est une quantité numérique dont la valeur dépend d'une expérience aléatoire. Par exemple, le lancer d'une pièce équilibrée définit une expérience dont le résultat est soit pile, soit face. Une variable aléatoire construite sur cette expérience consiste à affecter 1 à pile et 0 à face, chacune des valeurs ayant une probabilité de 1/2. Le

résultat d'une variable aléatoire est appelé réalisation. D'un point de vue théorique, les expériences peuvent être répétées une infinité de fois et donc engendrer autant de réalisations. Les processus ponctuels peuvent se comprendre comme une généralisation des variables aléatoires classiques ¹. Ce sont des modélisations qui engendrent des distributions spatiales de points aléatoires. À chaque réalisation d'un processus, un nouvel ensemble de points est produit.

Le processus de Poisson est complexe et contre-intuitif de prime abord. Nous allons le définir progressivement à partir de la généralisation de processus plus élémentaires. Nous reprenons ici les idées contenues dans l'ouvrage de BADDELEY, RUBAK et TURNER 2015, *Spatial Point Patterns : Methodology and Applications with R*.

2.2.1.1 Point aléatoire

Le processus ponctuel le plus simple est le point aléatoire, *uniformly random point* (fig. 2.2). À chaque réalisation de ce processus, deux nombres aléatoires sont tirés et déterminent les coordonnées du point dans le plan usuel. Un unique point est obtenu à chaque tirage. Ce processus est défini sur une sous-zone bornée W du plan ², qui représente pour le géographe sa zone d'intérêt. Au sein de cette zone, aucune partie de l'espace n'est davantage privilégiée qu'une autre. En mathématique, on dit que l'intensité du processus est homogène ou uniforme.

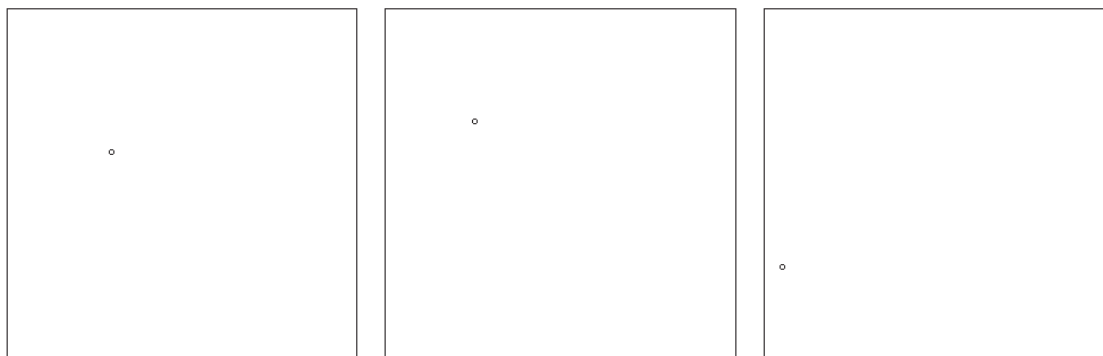


FIGURE 2.2 – Exemples de réalisations d'un processus de points aléatoires

L'intensité en statistique spatiale est une formalisation mathématique de la notion de densité utilisée en géographie. La densité d'une sous-zone de la zone d'intérêt pour une réalisation donnée est le nombre de points qu'elle contient rapporté à sa surface. L'intensité de cette sous zone est la moyenne des densités calculées sur l'ensemble des réalisations. L'intensité locale $\lambda(x,y)$ en un point (x,y) est la valeur de l'intensité d'un disque centré sur celui-ci dont le rayon tend vers 0.

Dans le cas d'un point aléatoire, les intensités locales $\lambda(x,y)$ sont toutes égales à la valeur λ directement déduite de la surface de la zone d'intérêt :

1. Les variables aléatoires et les processus ponctuels se différencient uniquement par la nature des objets engendrés. Chaque réalisation d'une variable aléatoire est un nombre. En revanche, la réalisation d'un processus ponctuel est un ensemble de couples de nombres représentant les coordonnées de points dans l'espace. Les mathématiciens ont unifié ces deux notions sous l'intitulé d'élément aléatoire.

2. Dans la suite de ce chapitre, nous considérons que W est le carré unité, de sorte que les deux nombres aléatoires du processus de points aléatoires ont pour loi, la loi uniforme sur $[0, 1]$.

$$\lambda = \frac{1}{|W|} \quad (2.1)$$

où $|W|$ est l'aire de W .

On peut calculer facilement la probabilité que le point tiré tombe dans une sous-partie B de W :

$$P(B \text{ contient le point}) = \frac{|B|}{|W|} = |B| \times \lambda \quad (2.2)$$

Cette probabilité est le rapport entre les aires de B et W .

2.2.1.2 Processus binomial

Le processus binomial est une extension du processus de points aléatoires : une réalisation de ce processus est le regroupement de n processus de points aléatoires (fig. 2.3). Ainsi, la connaissance de la localisation d'un point ne permet pas d'en inférer une information sur la localisation dans W des $n - 1$ points restants.

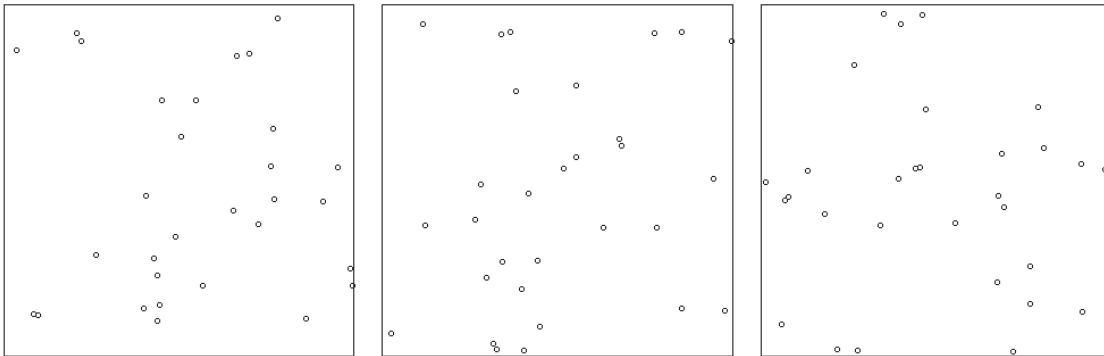


FIGURE 2.3 – Exemples de réalisations d'un processus binomial

La probabilité que k points tombent dans la partie B de $|W|$ se déduit directement de la loi binomiale s'écrivant comme suit :

$$P(B \text{ contient } k \text{ points}) = \binom{k}{n} (\lambda |B|)^k (1 - \lambda |B|)^{(n-k)} \quad (2.3)$$

La localisation aléatoire des n points dans W ne signifie pas pour autant que le processus binomial est totalement aléatoire et que les points soient indépendants les uns des autres. En effet, si nous connaissons k le nombre de points contenus dans B , nous en déduisons mécaniquement $n - k$ le nombre de points contenus dans le complémentaire de B dans W .

Le processus de Poisson, que nous allons aborder à présent, est une généralisation du processus binomial qui garantit la nature totalement aléatoire dans l'espace des distributions générées.

2.2.1.3 Processus de Poisson (processus spatial totalement aléatoire)

Un processus est considéré comme totalement aléatoire (CSR) lorsqu’il vérifie les propriétés suivantes :

- homogénéité : aucune partie de l’espace de la zone d’intérêt n’est privilégiée dans sa capacité à contenir un point ;
- indépendance : la connaissance du nombre de points contenus dans une partie de la zone d’intérêt n’apporte pas d’information sur le nombre de points contenus dans une autre zone (les deux zones ne devant pas se chevaucher).

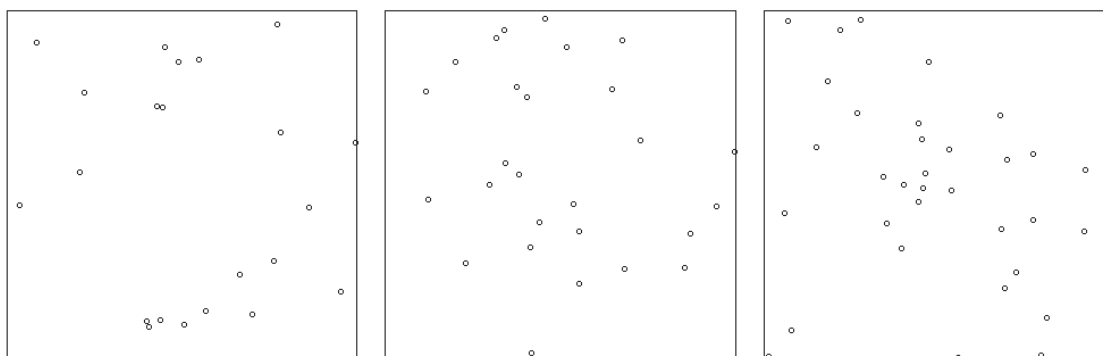


FIGURE 2.4 – Exemples de réalisations d’un processus de Poisson d’intensité égale à 30. Contrairement au processus binomial dont on connaît exactement le nombre de points, le nombre de points d’une réalisation d’un processus de Poisson est lui-même une variable aléatoire. Ainsi, le nombre de points de la configuration de gauche est 22, 27 pour celle du milieu et 35 pour celle de droite.

On peut montrer que le seul processus respectant ces propriétés est le processus de Poisson homogène. Il est caractérisé également par son intensité homogène λ ³. La probabilité qu’une partie B contienne k points suit une loi de Poisson de paramètre $p = \lambda|B|$ (LOONIS et DE BELLEFON 2018) :

$$P(k \text{ points tombent dans la partie } B) = \frac{(\lambda|B|)^k}{k!} e^{-\lambda|B|} \quad (2.4)$$

En pratique, la simulation d’un processus de Poisson d’intensité λ sur une zone d’intérêt W s’obtient en deux temps :

1. Tout d’abord, on tire le nombre total n de points de la distribution selon une loi de Poisson de paramètre $\lambda|W|$: $P(\text{nombre de points dans } W \text{ égal à } n) = \frac{(\lambda|W|)^n}{n!} e^{-\lambda|W|}$.
2. Ensuite, on tire une répartition spatiale de points selon un processus binomial de paramètre n .

3. Pour qu’en moyenne, le nombre de points tirés vaille n , λ doit être égale à $\frac{n}{|W|}$

Encadré

BADDELEY, RUBAK et TURNER 2015 justifient cette formule par passage à la limite. Supposons qu'un processus suive les propriétés CSR. Le nombre de points dans chaque cellule d'une grille de taille 5×5 qui subdivise la zone d'intérêt forme une variable aléatoire indépendante. Le nombre de points dans une cellule est $\frac{\lambda}{5 \times 5}$. Lorsque la taille des cellules devient très petite, la probabilité qu'une cellule contienne un point devient quasi proportionnelle à $\frac{\lambda}{N}$ (où N est le nombre de cellules) et la probabilité qu'elle contienne plus d'un point devient négligeable. En faisant tendre la taille des cellules vers 0, nous nous retrouvons dans le cadre d'application du théorème dit « loi des petits nombres » qui nous indique que la somme infinie d'événements très rares (une partie très petite contenant au moins un point) converge vers une loi de Poisson.

Remarque

BADDELEY, RUBAK et TURNER 2015 soulignent que, par définition, les processus de Poisson ne peuvent ni générer des configurations avec des superpositions des points dans l'espace, ni des configurations fractales.

2.2.2 La fonction de Ripley

2.2.2.1 Définition de la fonction de Ripley

La fonction de Ripley notée \hat{K} est construite à partir du rapport entre une densité globale et des densités locales calculées autour de chaque point de l'ensemble étudié composé de n points :

$$\hat{K}(r) = \frac{|W|}{n(n-1)} \sum_{i=1}^n \sum_{j \neq i} 1_{\{d_{ij} \leq r\}} \quad (2.5)$$

où i et j sont des points de l'ensemble étudié et $\sum_{i=1}^n \sum_{j \neq i} 1_{\{d_{ij} \leq r\}}$ est le nombre de points situés dans le voisinage de taille r centré autour du point i .

Pour un rayon r , la fonction de Ripley compare le nombre moyen de voisins autour de chaque point à la densité calculée sur la zone d'intérêt (BADDELEY, RUBAK et TURNER 2015). Si l'ensemble de points est une réalisation d'un processus de Poisson autour de chaque point dans un rayon r , le nombre de points attendus est $\pi r^2 \lambda = \pi r^2 \frac{|W|}{n}$, de sorte que la valeur de \hat{K} en fonction de l'aléa des configurations tirées oscille autour de πr^2 . BESAG 1977 proposa à partir de ce constat de normaliser l'expression de la fonction de Ripley de la façon suivante :

$$L(r) = \sqrt{\frac{\hat{K}(r)}{\pi}} - r \quad (2.6)$$

La nouvelle fonction ainsi définie est centrée autour de 0. Lorsqu'elle est positive, en moyenne chaque point a plus de voisins dans son environnement de taille r qu'un processus de Poisson de même intensité. L'ensemble de points apparaît, pour cette valeur de r , plus agrégé qu'un processus spatial totalement aléatoire (CSR). Réciproquement, une valeur négative indique que l'ensemble de points est plus régulier que la référence (CSR).

Remarque

Utiliser la fonction de Ripley suppose implicitement que l'intensité λ du processus sous-jacent soit homogène.

2.2.2.2 Apprécier la significativité des résultats

Afin de juger la significativité de l'écart à un processus de Poisson, les statisticiens ont développé plusieurs approches. La première est basée sur des calculs analytiques. Gabriel LANG et Eric MARCON ont publié un article en 2013 explicitant un test statistique classique reposant sur les propriétés mathématiques du processus de Poisson (LOONIS et DE BELLEFON 2018). Néanmoins, dans l'immense majorité des situations concrètes, les chercheurs utilisent plutôt des méthodes de Monte Carlo. L'idée est de simuler un grand nombre de réalisations de processus de Poisson de même intensité que la configuration de points étudiés. Ensuite, pour chaque ensemble de points simulé, on calcule sa fonction de Ripley (ou fonction de Besag si on utilise cette dernière). On extrait ensuite de l'emprise de ces fonctions simulées une enveloppe⁴ qui quantifie la variabilité et la fluctuation naturelles des réalisations d'un processus de Poisson. L'enveloppe ainsi calculée a pour objectif d'aider le praticien à apprécier si sa configuration peut être considérée comme le résultat d'un processus de Poisson, ou si, au contraire celle-ci déroge à l'hypothèse d'indépendance. BADDELEY, RUBAK et TURNER 2015 différencient deux types d'enveloppes :

- Les enveloppes point par point sont calculées comme leur nom l'indique pour chaque valeur de r . Bien que très souvent utilisées de par leur simplicité conceptuelle, elles sont néanmoins sources de mauvaises interprétations. En toute rigueur, nous devons fixer le ou les rayons r dont on souhaite vérifier l'écart avant d'observer les courbes. Sinon, on risque de rejeter à tort l'hypothèse d'indépendance et de fausser les conclusions de l'analyse⁵ ;
- Les méthodes globales ont été initiées par RIPLEY 1997 dès la fin années 70. Dans ce cas, l'enveloppe est calculée sur l'intégralité des rayons simultanément et les risques de mauvaises interprétations sont limités. Dans le reste de ce chapitre, nous utilisons ce type d'enveloppe dans nos applications pratiques.

2.2.2.3 Interprétation du rayon r en terme d'échelle

La fonction de Ripley apporte une information sur la distribution spatiale d'un ensemble de points pour une large gamme d'échelles (BADDELEY, RUBAK et TURNER 2015). Pour autant, la notion d'échelle mobilisée par celle-ci ne correspond pas directement à celle usuelle des géographes. En effet, l'observation pour un rayon donné d'une dérogation à l'hypothèse d'indépendance ne permet pas d'en déduire que l'ensemble de points est agrégé ou régulier de façon univoque à cette échelle d'observation car la fonction de Ripley est par construction cumulative : les voisinages de petite taille sont contenus dans ceux de plus grande taille, de sorte que la valeur de la fonction de Ripley pour un rayon donné est tributaire des valeurs de celle-ci pour tous les rayons inférieurs. Selon GABRIEL et al.

4. L'enveloppe est constituée en prenant pour chaque rayon r le minimum et le maximum des fonctions de Ripley calculées sur les ensembles de points simulés.

5. Ceci est appelé *data dredging* ou *datasnooping* par les statisticiens. Selon AKTAS 2004, le *data dredging* « consiste à ne publier que les compositions d'échantillon et les périodes d'observation favorables à l'hypothèse testée » (cité dans l'article Wikipédia sur le sujet).

2016, le pic de la fonction de Besag (cf. équation 2.6), nous indique le rayon d'agrégation maximal au-delà duquel la nature agrégée de l'ensemble de points doit être interprétée avec précaution. Néanmoins, BADDELEY, RUBAK et TURNER 2015 restent très mesurés quant à la possibilité de définir à partir de ce pic une échelle géographique qui refléterait la nature spatiale des interactions entre les points.

La fonction de corrélation de paires de points est une méthode d'analyse non cumulative. Elle utilise non pas le support d'un disque pour appréhender l'environnement d'un point mais un anneau de rayon r dont l'épaisseur est infiniment petite. Elle peut être déduite de la fonction de Ripley par une opération de dérivation. Si d'un point de vue théorique, elle semble être une alternative pertinente à la fonction de Besag, en pratique son calcul se révèle difficile (BADDELEY, RUBAK et TURNER 2015). En effet, le nombre de points dans les anneaux est souvent minuscule, en particulier pour les plus petites valeurs de r , ce qui réduit considérablement la robustesse de l'estimation de cette fonction.

2.2.2.4 Mise en œuvre de la fonction de Ripley sur les distributions archétypales

Nous avons représenté sur la figure 2.5, les trois fonctions de Besag pour les distributions de référence de la figure 2.1. La ligne en pointillés rouges est la valeur de référence théorique pour un processus de Poisson, à savoir dans le cas de la fonction de Besag, 0. Le rectangle gris autour de cette ligne représente les enveloppes estimées. On constate que pour la distribution aléatoire, la fonction de Besag est contenue dans l'emprise de l'enveloppe. L'hypothèse selon laquelle l'ensemble de points a été créé à partir d'un processus totalement aléatoire ne peut pas être rejetée. Pour les deux autres distributions, l'hypothèse est rejetée. La distribution agrégée a plus de points dans ses voisinages qu'une distribution aléatoire. A contrario, la distribution régulière a moins de points dans ses voisinages que ne laisserait espérer une distribution aléatoire.

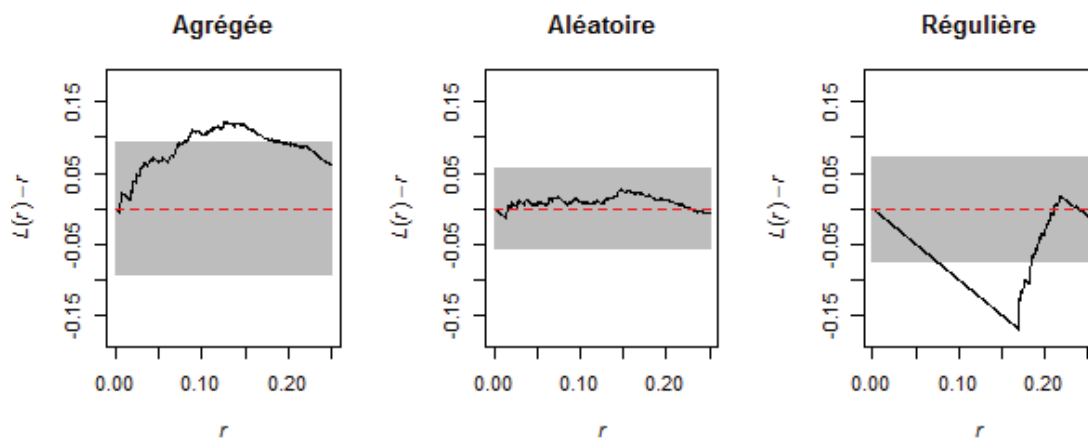


FIGURE 2.5 – Mise en œuvre de la fonction de Besag pour les trois distributions de référence de la figure 2.1

2.2.2.5 Application à la caractérisation des tissus bâtis

Nous avons appliqué la fonction de Besag aux deux distributions archétypales présentées dans l'introduction (fig. 1.3 et 2.6). Visuellement, la répartition des bâtiments nous apparaît dispersée dans le cas de la configuration bocagère et agrégée pour le territoire

d'openfield (*fig. 2.7*). Cette opposition entre les deux configurations ne se retrouve pas dans les analyses statistiques. Toutes les deux présentent, en effet, des courbes très au-dessus de leur enveloppe, à tel point que celles-ci se confondent avec l'axe des origines. Ceci est à nuancer seulement par le fait que la courbe associée au bocage est très en-dessous de la courbe d'openfield : la concentration des bâtiments semble encore plus forte pour l'openfield que le bocage. Si l'on s'intéresse à des valeurs de rayon très petites, les courbes sont en-dessous de leur enveloppe et atteignent leur minimum vers 5 mètres. En soi, ce phénomène n'est pas porteur d'une signification géographique mais lié à un artefact de mesure. En effet, afin de pouvoir mettre en œuvre des méthodes de statistique spatiale, nous avons simplifié la géométrie des bâtiments en assimilant ceux-ci à leur centroïde, de sorte qu'en moyenne deux points associés à deux bâtiments sont séparés mécaniquement par une distance induite par l'emprise spatiale moyenne des bâtiments. Cette distance minimum entre deux bâtiments explique le caractère régulier des configurations observées pour des petits rayons.

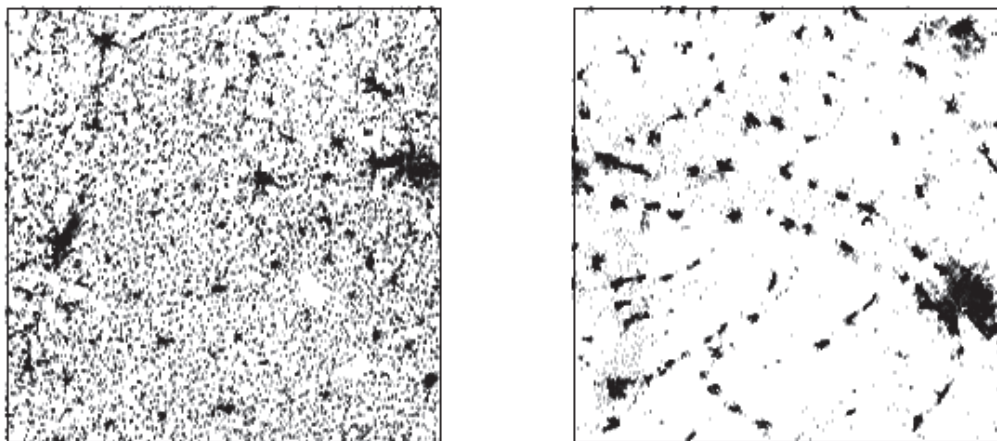


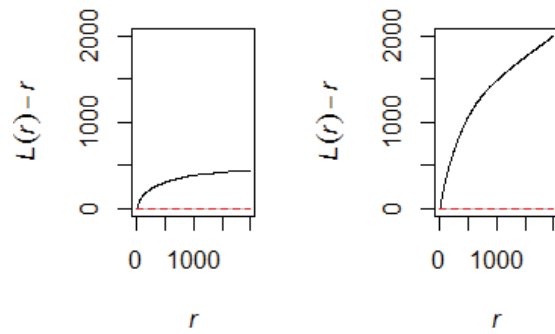
FIGURE 2.6 – Répartition des bâtiments des deux distributions archétypales de la figure 1.3

Si nous examinons de plus près les configurations, nous constatons que l'hypothèse d'homogénéité des intensités n'est pas vérifiée pour celles-ci. Les cartes de densité lissée (*fig. 2.8*) montrent des zones de très forte densité recouvrant approximativement les villages et les hameaux. En toute rigueur, nous ne pouvons donc pas appliquer la fonction de Ripley et ses dérivées pour analyser ces deux distributions de bâtiments.

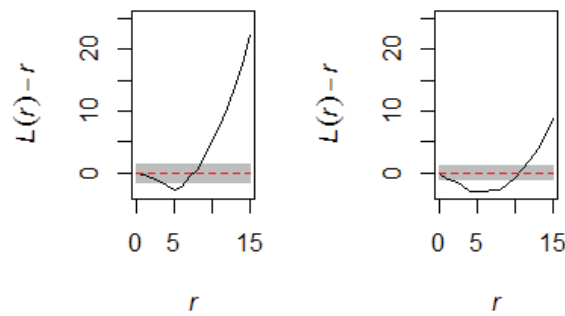
BADDELEY, MØLLER et WAAGEPETERSEN 2000 proposèrent d'étendre le champ d'application de la fonction de Ripley aux distributions d'intensité non homogène. La fonction de Ripley inhomogène est définie ainsi :

$$\hat{K}_{inhom} = \frac{1}{|W|} \sum_i^n \sum_{j \neq i} \frac{1_{\{d_{ij} \leq r\}}}{\hat{\lambda}(i)\hat{\lambda}(j)} \quad (2.7)$$

Où $\hat{\lambda}(i)$ est une estimation de l'intensité au point i obtenue par une méthode de lissage à noyau. L'introduction de deux intensités locales a pour objectif de prendre en compte les effets mécaniques de l'inhomogénéité de l'intensité sur le nombre de voisins : les points contenus dans des zones de forte intensité ont de nombreux voisins et réciproquement,



(a) Entre 0 et 2000m



(b) Entre 0 et 15m

FIGURE 2.7 – Fonction de Besag pour les deux distributions archétypales de la figure 2.6

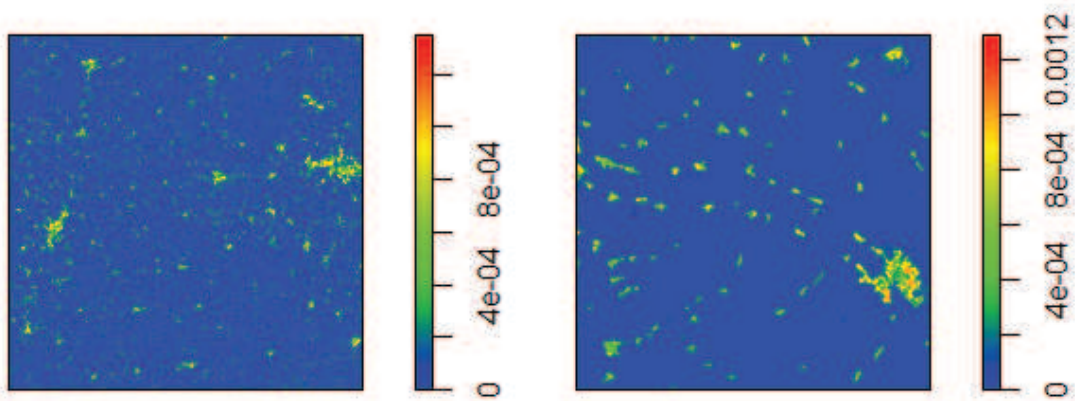


FIGURE 2.8 – Densité lissée pour les deux distributions archétypales de la figure 2.6 . Les rayons de lissage ont été choisis par validation croisée.

dans les zones de faible intensité, les points ont peu de voisins. La formule de Ripley inhomogène permet donc d'apprécier le caractère agrégé d'une distribution de points une fois que l'on a contrôlé les effets mécaniques induits par l'inhomogénéité de l'intensité.

Pour un processus de Poisson d'intensité non homogène la fonction de Ripley \hat{K}_{inhom} a pour espérance πr^2 comme pour un processus de Poisson homogène. Ceci permet en théorie, comme précédemment, de tester la nature agrégée ou régulière des configurations de points analysées. En pratique, nous n'avons pas réussi à appliquer la fonction de Ripley inhomogène à nos deux configurations de bâtiments car le package R spatstat⁶, dédié à la statistique spatiale, ne supporte pas des configurations de points de si grande taille et les résultats que nous obtenions étaient totalement aberrants.

2.2.2.6 Critique des résultats et mises en perspective

En statistique usuelle, le premier ordre représente l'ensemble des indicateurs et des approches qui sont de près ou de loin apparentés à une idée de moyenne. Les analyses de second ordre reposent quant à elles principalement sur la notion de variance et décrivent l'écart à la moyenne. La statistique spatiale a repris à son compte cette distinction entre deux niveaux d'analyse. Le premier ordre est dans ce cas basé sur la notion d'intensité et le second ordre sur l'étude de l'écart à une répartition d'intensité connue.

Néanmoins, alors que dans la statistique non spatiale cette séparation ne pose guère de problèmes conceptuels et pratiques, en statistique spatiale, il en va tout autrement. Le calcul de l'intensité nécessite de définir un territoire de référence. Or ce choix territorial, comme l'ont démontré les travaux d'OPENSHAW 1984a sur le *MAUP* ne va pas de soi. Les statistiques spatiales sont en effet très dépendantes du découpage utilisé et ceci est particulièrement vrai pour le calcul de l'intensité. BADDELEY, RUBAK et TURNER 2015 démontrent ainsi que le choix d'une fenêtre d'analyse trop resserrée sur les points étudiés⁷ tend à surestimer l'intensité et donc mécaniquement à surestimer le caractère agrégé de ceux-ci. Ce faisant, la séparation entre premier ordre et deuxième ordre est beaucoup plus floue.

Dans notre analyse, cette séparation est d'autant plus compliquée à effectuer que l'intensité des points n'est pas homogène. L'estimation de l'intensité locale repose sur des méthodes de lissage à noyau qui requièrent de fixer un rayon de lissage. Celui-ci joue, grossièrement parlant, le rôle d'un territoire de référence pour calculer des densités locales. Ce choix, même dicté par la minimisation d'un critère statistique, est très fragile. En effet, en fonction du critère utilisé, le rayon de lissage peut varier selon un facteur 10 (LOONIS et DE BELLEFON 2018). Or, BADDELEY, RUBAK et TURNER 2015, soulignent la grande variabilité des résultats de la fonction de Ripley inhomogène selon le choix du rayon de lissage.

Les difficultés que nous avons rencontrées à séparer le premier du second ordre ont été théorisées par BARTLETT 1964 qui démontre que la simple réalisation d'un processus ponctuel qui est stationnaire et agrégé peut être identique à la réalisation d'un processus inhomogène non agrégé. Dit autrement, il est quasi-impossible de distinguer de façon univoque les zones de forte densité des espaces agrégés lorsque l'on dispose d'une seule réalisation d'un processus.

6. Selon notre expérience, au-delà de 50 000 points les calculs deviennent très longs et les résultats peu fiables.

7. Par exemple en ayant recours au polygone convexe

Afin de dépasser ces limites liées au caractère global des analyses, GETIS et FRANKLIN 1987 proposèrent des déclinaisons locales de la fonction de Ripley. L'idée de leur analyse est de calculer localement des indicateurs reposant sur une décomposition de la fonction de Ripley selon la contribution de chacun des points analysés. Cette approche revient à chercher localement les points regroupés dans des agrégats (*clusters*). Par rapport à la fonction de Ripley, elle présente également un autre avantage très important : elle permet de comparer l'organisation de différents territoires. Dans la suite de ce chapitre, nous proposons d'exploiter la méthode DBSCAN pour explorer la distribution des bâtiments dans l'espace. Par rapport, aux indicateurs locaux, cette approche présente l'avantage de ne pas dire uniquement si un point est agrégé mais définit également explicitement des agrégats, ce qui n'est pas anodin. Par exemple les villes, qui sont des espaces de forte densité et d'agrégation des bâtiments se différencient également par leur taille. Différencier les agrégats selon leur taille peut donc se révéler potentiellement utile.

2.2.3 La DBSCAN (*cluster detection*)

La *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) est une méthode de délimitation d'agrégats proposée par ESTER et al. 1996. Comme son nom l'indique, elle est basée sur la densité. Soit un ensemble de points donné, elle regroupe les points se situant dans des zones de forte densité et marque comme aberrants les points isolés localisés dans les zones de faible densité.

En 2014⁸, la méthode DBSCAN a été reconnue par la SIGKDD⁹, émanation de *Association for Computing Machinery*, comme étant un des algorithmes qui a le plus contribué à développer le *machine learning*. Le succès de cette méthode de *clustering*, tient à la fois à la simplicité de son algorithme et à sa rapidité. Ainsi pour analyser les 24 millions de points correspondant à l'emplacement des centroïdes des bâtiments en France, l'implémentation R de la méthode fournie dans le package DBSCAN met moins d'une minute sur un ordinateur de bureau standard¹⁰.

Dans l'article originel de ESTER et al. 1996, l'algorithme de la DBSCAN est présenté de façon très théorique. Nous proposons ici de simplifier la présentation de celui-ci sans pour autant trahir l'intuition originale.

2.2.3.1 Présentation méthodologique de la DBSCAN

L'idée de la DBSCAN est de construire des agrégats en regroupant les points dont le voisinage est suffisamment dense. Le voisinage d'un point i de taille eps , $V_{eps}(i)$, est le nombre de points contenus dans le disque de rayon eps centré sur le point i . Un point est considéré comme appartenant à un agrégat lorsque son voisinage est supérieur ou égal à min_{pts} . L'algorithme de la DBSCAN peut être décomposé en trois étapes :

1. On repère tous les points i , tels que $V_{eps}(i) \geq min_{pts}$. Ces points sont appelés points-cœurs (les points non cœurs sont dits isolés). On forme à partir de ces points-cœurs un graphe. Deux points-cœurs sont connectés si leur distance est inférieure à eps .

8. <http://www.kdd.org/News/view/2014-sigkdd-test-of-time-award>

9. *Special Interest Group on Knowledge Discovery and Data Mining*.

10. Pour obtenir de telles performances, les implémentations de la DBSCAN reposent sur l'utilisation d'algorithmes de calcul de distances entre points très rapides tel que le *kd-tree*.

2. On recherche dans le graphe formé à l'étape précédente, toutes les composantes connexes (agrégats de points coeurs connectés). Chaque composante représente le noyau intérieur d'un agrégat.
3. On rajoute pour former les agrégats définitifs aux noyaux intérieurs les points isolés situés à une distance inférieure à eps d'un point cœur. Ces points forment la frontière extérieure des agrégats. Les points isolés trop loin d'une composante connexe pour appartenir à un agrégat sont dits aberrants.

Bien que l'algorithme de la DBSCAN soit déterministe, l'affectation des points frontières aux agrégats peut varier en fonction de l'ordre de traitement des parties connexes. En pratique, ce phénomène très marginal n'altère en rien la pertinence de cette méthode.

2.2.3.2 Choix des paramètres eps et min_{pts}

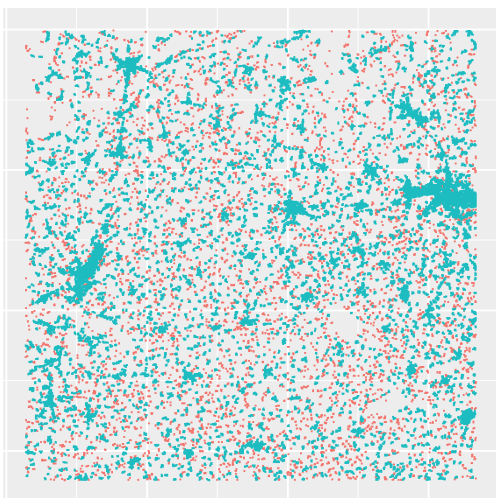
La mise en oeuvre de la DBSCAN requiert de choisir la valeur de deux paramètres : eps et min_{pts} qui ont chacun une grande influence sur les agrégats obtenus (KARAMI et JOHANSSON 2014). Différentes stratégies ont été proposées par les chercheurs (ANKERST et al. 1999, DARONG et PENG 2012, LIU et al. 2015) depuis la publication de l'article de ESTER et al. 1996 pour déterminer des paramètres optimaux et réduire ainsi l'arbitraire des analyses. La grande majorité de celles-ci supposent de fixer un des paramètres, généralement min_{pts} et d'en déduire ensuite selon différents tests, le paramètre manquant. Néanmoins, en l'absence d'a priori sur les phénomènes que nous souhaitons mesurer, ce type d'approche se révèle inopérant.

Afin de pallier cette limite, nous proposons ici d'utiliser les deux distributions de référence pour déterminer un jeu de paramètres optimal. L'idée est de choisir les valeurs des paramètres qui discriminent le plus fortement la distribution de bocage de la distribution d'openfield. Avant de rentrer dans le détail de notre démarche, nous allons tout d'abord observer les résultats de la DBSCAN pour différents couples de valeurs (figures 2.9 et 2.10). On constate la grande sensibilité des agrégats obtenus selon les paramètres choisis, comme la théorie le prévoit. Ainsi, sur ces exemples, la part des bâtiments appartenant à un agrégat (tx_{ag}) s'échelonne entre 0% et 100% pour les deux configurations. Plus précisément, pour un même un rayon eps , plus la valeur de min_{pts} augmente et plus, mécaniquement, la part des points agrégés diminue. De façon symétrique, pour une même valeur de min_{pts} , plus la valeur de eps augmente et plus la part de points agrégés augmente.

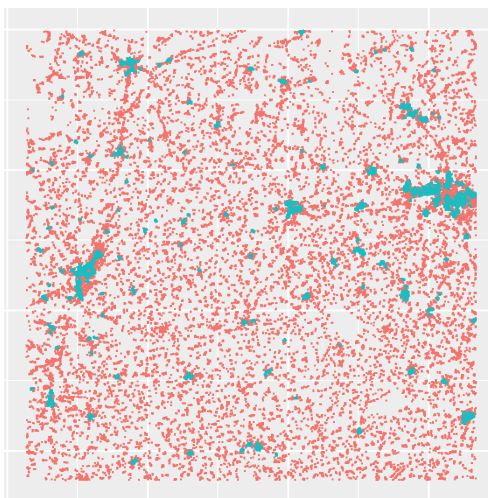
Pour les couples de paramètres (200, 20), (200, 40) et (400, 40), la DBSCAN semble séparer de façon efficace les espaces « villageois » des bâtiments isolés. La différence entre le taux des points agrégés de la configuration d'openfield et de bocage est respectivement de 37%, 53% et 31%. Nous pouvons utiliser cette différence de taux pour tenter de calibrer un jeu de paramètres pertinents. L'idée est de trouver le couple de paramètres qui maximise la différence de parts de bâtiments agrégés entre les deux configurations.

En théorie, cette opération nécessiterait d'explorer tous les couples possibles de paramètres. En pratique, plusieurs remarques permettent de limiter l'étendue de la recherche. Premièrement, l'emprise spatiale des configurations étant limitée, si l'on souhaite minimiser les effets de bord, il est impératif de limiter la taille des rayons admissibles. De plus, en dessous de 50 m, les bâtiments, du fait de leur emprise spatiale, ont très peu de voisins. Il n'est donc pas utile de tester des valeurs de rayon trop petites. Ensuite, pour un rayon donné, les valeurs admissibles du nombre de points sont bornées. En effet, les taux de points agrégés étant décroissants avec le rayon fixé, ceux-ci deviennent nuls à partir d'une certaine valeur et il n'y a donc pas d'intérêt à aller explorer l'espace des

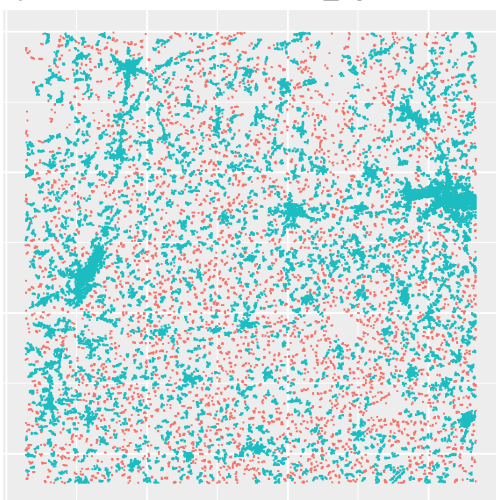
eps = 100, minPts = 5 et tx_ag = 85.1%



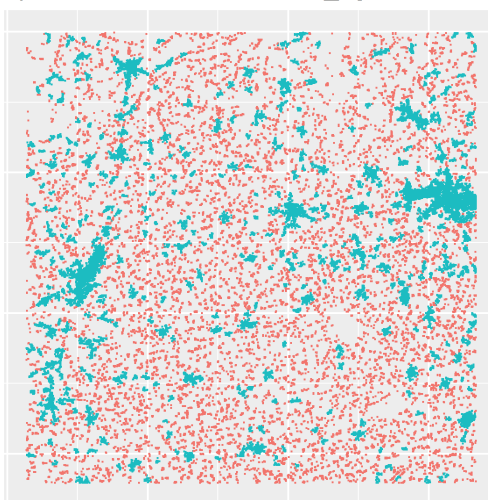
eps = 100, minPts = 20 et tx_ag = 25.6%



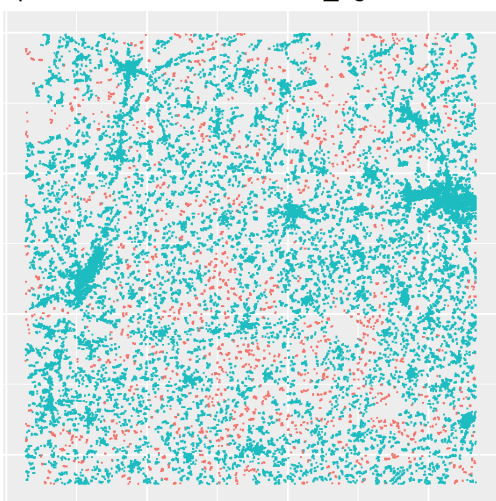
eps = 200, minPts = 10 et tx_ag = 81.8%



eps = 200, minPts = 20 et tx_ag = 55.5%



eps = 400, minPts = 20 et tx_ag = 92.3%



eps = 400, minPts = 40 et tx_ag = 63.3%

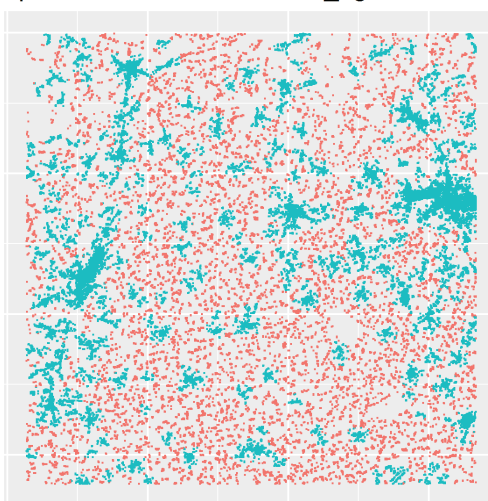


FIGURE 2.9 – Test DBSCAN sur le territoire de bocage (en bleu les points agrégés et en rouge les non agrégés)

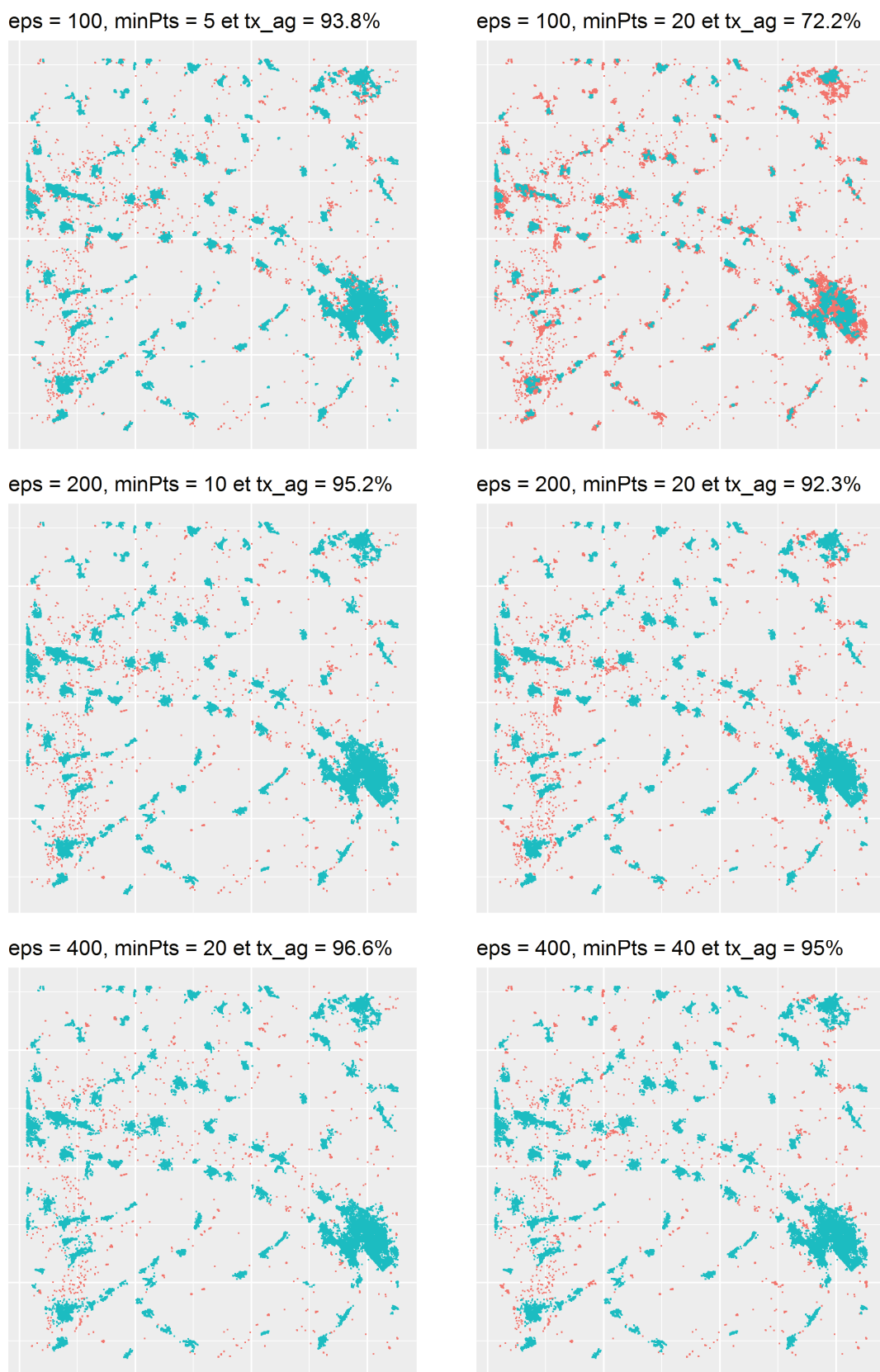


FIGURE 2.10 – Test DBSCAN sur le territoire d’openfield (en bleu les points agrégés et en rouge les non agrégés)

paramètres au-delà. De sorte que nous nous sommes limités après différents tests, aux valeurs de rayon eps comprises entre 50 et 1000 m et aux valeurs de min_{pts} comprises entre 5 et 100 points. Néanmoins, même pour ce sous-espace, le nombre de croisements possibles demeure encore gigantesque quand bien même on se limite à des valeurs entières pour les rayons. Nous avons donc décidé de discrétiser l'espace des paramètres selon un pas de 10 m pour les rayons et de 5 pour le nombre de points. Cette opération ne semble pas préjudiciable car, comme nos expérimentations nous ont permis de le constater, les variations de parts agrégées n'évoluent pas brutalement et semblent au contraire présenter une forme de continuité. Au final, nous calculons la différence de taux tx_{ag} sur une grille comprenant 1820 couples de paramètres.

Sur la figure 2.11, nous avons représenté sous la forme d'une représentation tridimensionnelle les différences de taux entre les deux configurations calculées sur la grille de paramètres. L'écart maximal (55 points de pourcentage) est atteint pour un rayon de 330 m et un nombre de voisins de 90. Le test des paramètres suivants (figure 2.12) permet de mettre à l'épreuve l'optimalité de la configuration correspondant à un rayon de 330m et un nombre de voisins de 90 (figure 2.11) :

- Test 1 : un rayon eps de 330 m, mais avec une densité de points voisins moitié moins que celle de la configuration optimale, soit une valeur de min_{pts} égale à 45 voisins ;
- Test 2 : un rayon eps deux fois plus petit que le rayon optimal (170 m) avec un nombre de points correspondant à une densité comparable soit à une valeur de min_{pts} égale à 25 ;
- Test 3 : des paramètres définis de façon ad-hoc pour que les formes observées se rapprochent de celles de la carte de population dispersée et concentrée de 1876 présentée en figure figure 2.15, à savoir un rayon eps égal à 200 m et une densité de points voisins $minPts$ égale à 20.

Pour ces deux jeux de paramètres, les différences de taux sont respectivement de 45% et de 47% et les zones agrégées apparaissent très proches, ce qui tend à démontrer la robustesse du maximum. Cette robustesse est par ailleurs confirmée par l'organisation de la surface formée par la différence entre les taux tx_{ag} selon le rayon eps et le nombre de points considérés min_{pts} (fig. 2.11). Les résultats observés sont donc très proches les uns des autres, lorsque l'on se situe à proximité du maximum.

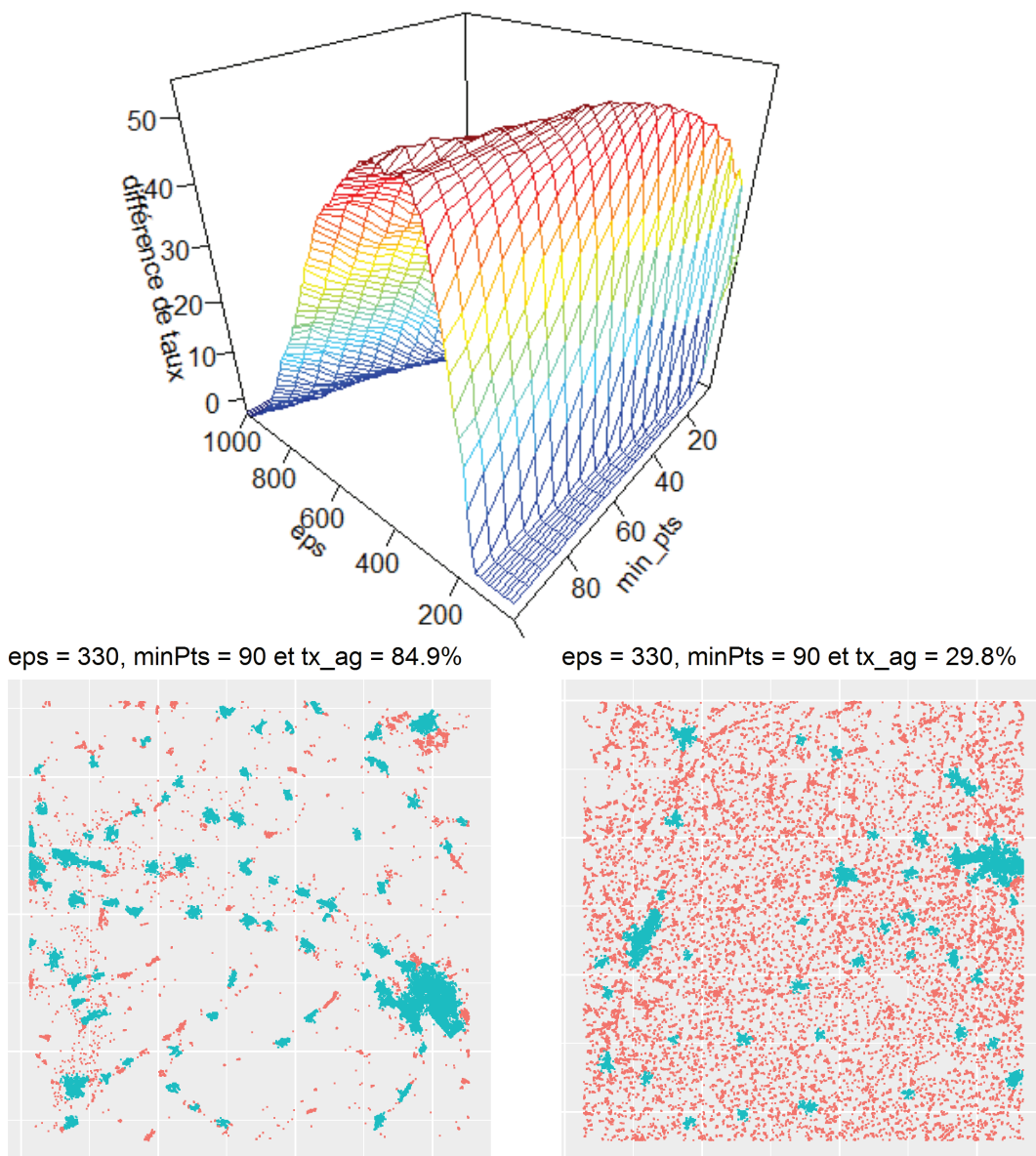


FIGURE 2.11 – En haut, la différences de taux en fonction des paramètres eps et min_{pts} . En bas, les résultats de la DBSCAN pour les paramètres optimaux (un rayon de 330m et un nombre de voisins de 90).

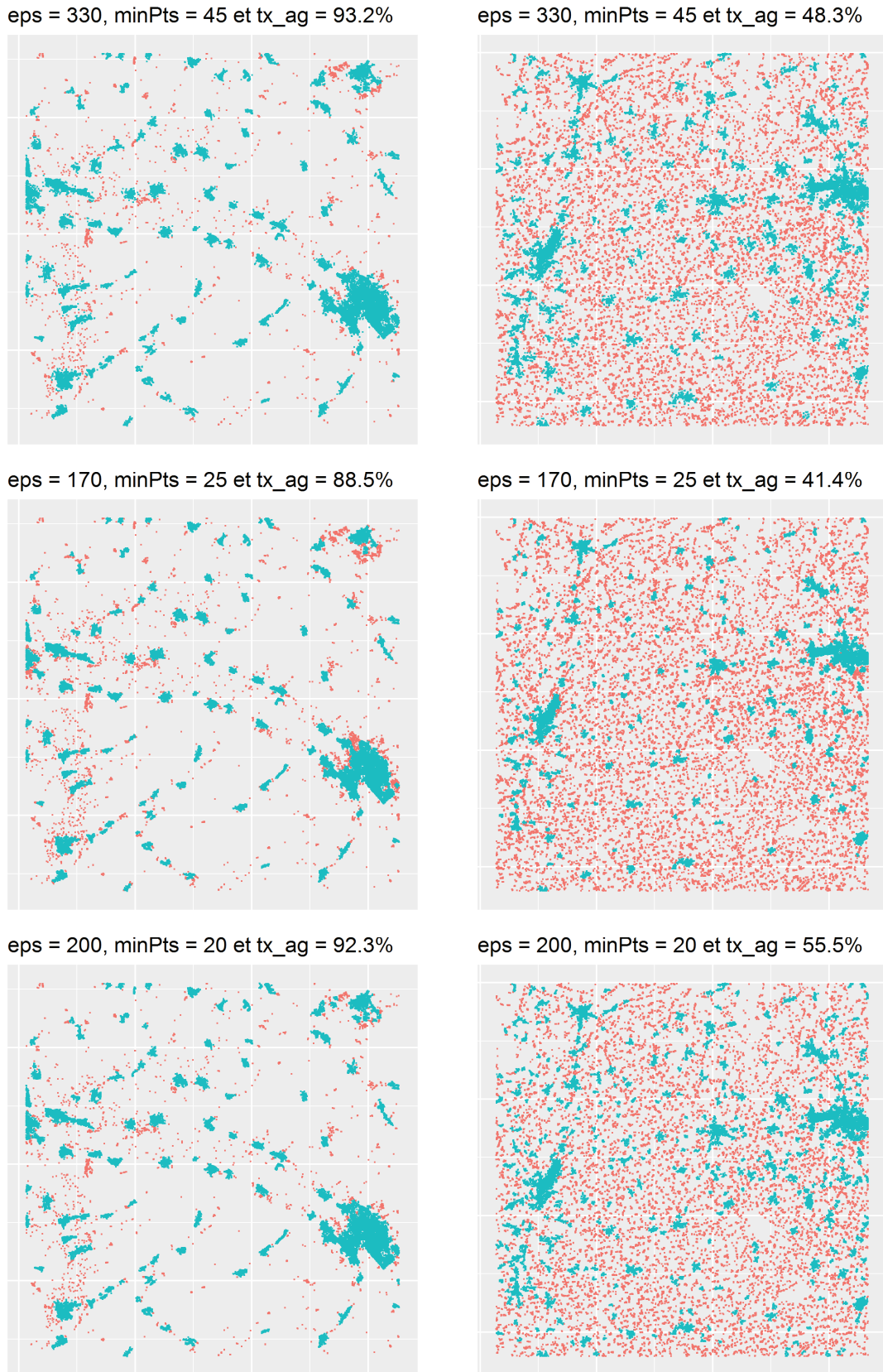


FIGURE 2.12 – Test de sensibilité du jeu de paramètres optimal, en haut les résultats de la DBSCAN pour le test 1, au milieu pour le test 2 et en bas pour le test 3.

2.2.3.3 Cartographie des zones agrégées

Maintenant que nous disposons d'un jeu de paramètres que l'on juge pertinent, nous pouvons l'appliquer à l'ensemble du territoire de la France métropolitaine. Afin d'obtenir une représentation cartographique de l'organisation des bâtiments selon leur nature agrégée ou isolée, nous avons calculé une part lissée des bâtiments agrégés. Pour cette opération, nous avons utilisé un noyau quadratique, *biweight*, (cf. équation 3.12) et un rayon de 8 km qui permet de restituer les détails des différentes configurations spatiales tout en garantissant une lisibilité des aires observées. Sur les figures 2.13 et 2.14, nous avons ainsi représenté le taux d'agrégation tx_{ag} lissé pour le jeu de paramètres optimal et pour les jeux de paramètres qui nous ont permis d'apprécier la sensibilité de notre analyse. Afin de ne pas biaiser les interprétations, nous proposons pour chaque variable, deux légendes. La première (dégradé rouge) est commune aux 3 jeux de paramètres et permet des comparaisons directes entre ceux-ci. Pour les cartes de droite (dégradé bleu), la légende est le résultat de l'application de l'algorithme de Jenks ¹¹ pour chaque variable.

Nous observons globalement des formes spatiales communes entre toutes les cartes. Les territoires du Nord-Est de la France métropolitaine ont des parts de bâtiments agrégés globalement très supérieures à ceux de l'Ouest et du Sud-Ouest de la France. Par ailleurs, dans les grandes villes et ce, quelque soit le jeu de paramètres et la méthode de discrétisation, la part de bâtiments agrégés est proche de 100%. Ce phénomène est mécanique : par définition les espaces urbains concentrent de fortes densités de bâtiments, tel que pratiquement chaque bâtiment a de nombreux voisins et est donc considéré comme agrégé par la DBSCAN.

Pour autant, cette ressemblance générale masque des différences entre les jeux de paramètres. Les espaces agrégés repérés par la carte basée sur les paramètres optimaux sont nettement plus restreints que ceux des deux autres cartes. Ce constat peut être surprenant de prime abord tant les résultats sur les configurations qui nous ont permis de calibrer notre analyse étaient comparables.

11. L'algorithme de Jenks vise à regrouper au mieux les valeurs similaires et à maximiser les différences entre les classes.

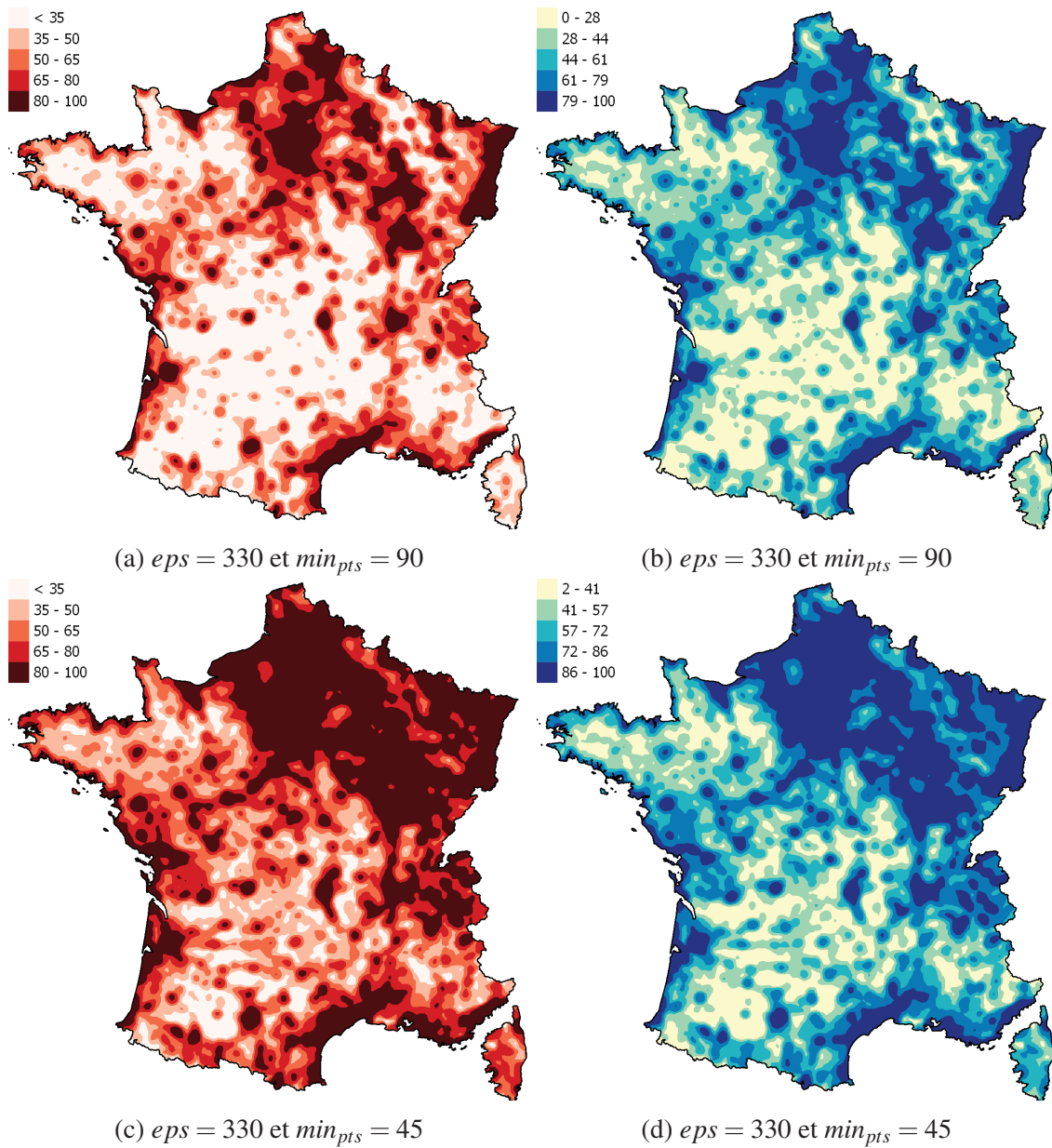


FIGURE 2.13 – Carte des parts lissées de bâtiments appartenant à un agrégat (à gauche légende comparable, à droite légende obtenue par discrétisation de Jenks). Le rayon de lissage est de 8 km.

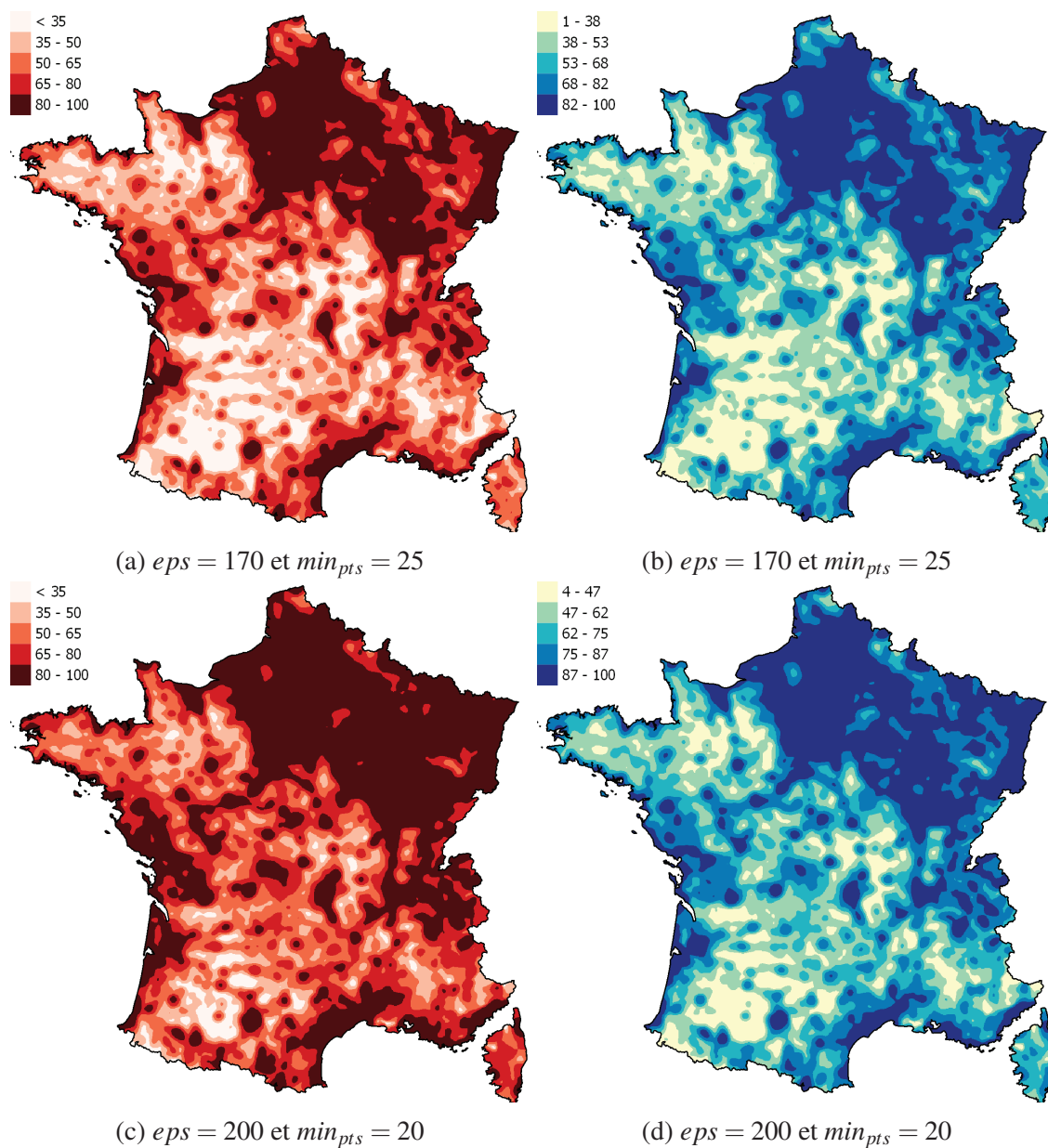


FIGURE 2.14 – Carte des parts lissées de bâtiments appartenant à un agrégat (à gauche légende comparable, à droite légende obtenue par discrétisation de Jenks). Le rayon de lissage est de 8 km.

2.2.3.4 La population agglomérée et dispersée

Afin d'approfondir notre compréhension de l'organisation spatiale observée et de la variabilité des résultats, nous allons mettre en perspective notre analyse avec des méthodes et outils autrefois utilisés par les géographes pour décrire l'organisation de la population rurale et qui présentent de nombreuses similarités avec notre démarche : les populations éparses et agrégées.

Notre utilisation de la DBSCAN revient à donner implicitement une définition des villes et des villages. L'Insee, avec les unités urbaines, procède selon une méthode analogue pour définir l'emprise spatiale des villes. Historiquement, une méthode similaire était utilisée pour décrire la morphologie des espaces ruraux : à partir du début du 19^e siècle (LE MÉE 1971) et jusqu'au recensement de 1982, la statistique publique¹² publiait des populations éparses et agrégées. Initialement, ces chiffres répondaient à de nombreux besoins. Le premier d'entre eux était d'ordre fiscal. Un décret impérial pris en vertu de ce texte, le 21 décembre 1808, cité par LE MÉE 1971, précise en son titre II :

« Article 7. Les droits d'entrée établis par l'article 18 de la loi du 25 novembre ne seront perçus que dans les lieux dont la population agglomérée sera de deux mille âmes au moins, non compris celle éparsée dans les hameaux ou les villages dépendants de la commune ».

Néanmoins, les contraintes fiscales n'excluaient pas une nécessité de connaissances géographiques plus larges, comme en témoigne la lettre envoyée par Coquebert de Montbret, dirigeant alors de la statistique publique, au préfet de la Drôme en 1809, également cité dans LE MÉE 1971 :

« On a bien dit quel était le nombre des habitants du territoire de chaque commune mais on a négligé d'indiquer combien de ces habitants sont réunis au chef-lieu, et combien sont disséminés dans les hameaux, écarts, fermes et habitations isolées qui se trouvent sur le même territoire. Il y a cependant, sous un grand nombre de rapports, une différence très essentielle entre une population agglomérée et une population éparsée, et, dans un grand nombre de cas, il importe beaucoup au gouvernement de ne pas être exposé à se tromper à cet égard ».

Dans son acceptation la plus large, la population agglomérée¹³ est la population rassemblée autour de la mairie en habitations contiguës ou très proches dans un village de type concentré ou dans le chef-lieu d'une localité. Elle s'oppose à la population éparsée contenue dans des hameaux ou des fermes isolées qui est son complémentaire au sein de la population communale.

La carte de la population agglomérée en 1882 (figure 2.15) présente, nonobstant les différences de concepts, une forte ressemblance avec les cartes que nous proposons à partir de nos mises en œuvre de la DBSCAN. Cette opposition entre les territoires ruraux du Nord-Est et de l'Ouest de la France a été signalée dès la fin du 18^e siècle par l'agronome britannique Arthur Young. Au début du 20^e siècle en France, de nombreux chercheurs tentèrent d'expliquer cette séparation de la France en deux parties. Au-delà de leurs divergences, cette génération pionnière s'accorda pour penser l'opposition entre un habitat agrégé et un habitat dispersé dans un cadre d'analyse plus large : les structures agraires.

12. Initialement, la Statistique générale de la France, puis à la sortie de la seconde guerre mondiale, l'Institut National de la Statistique et des Études Économiques.

13. Définition provenant du site du Centre National de Ressources Textuelles et Lexicales.

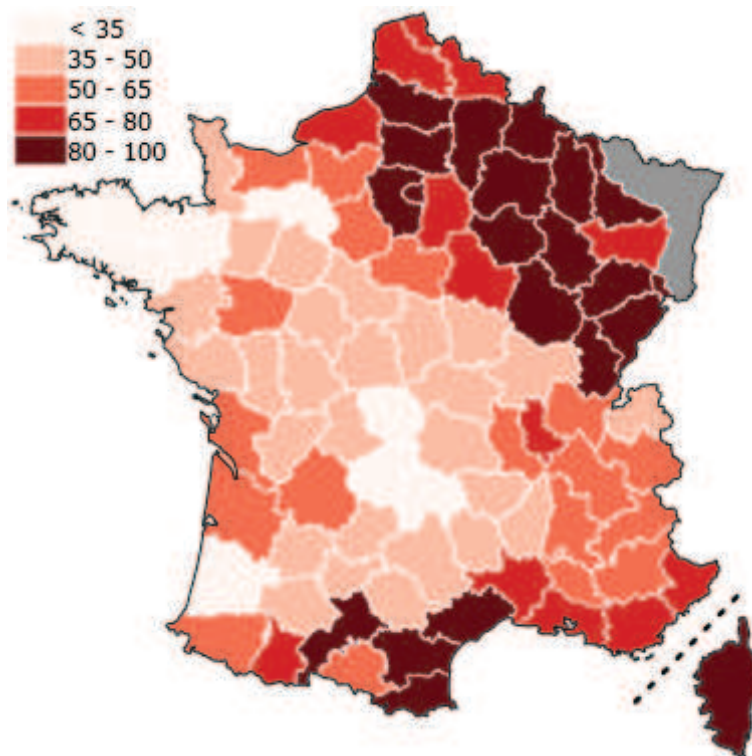


FIGURE 2.15 – Part de la population agglomérée en 1876. Source : dénombrement de la population française de 1876, Insee.

D'un côté, l'habitat dispersé est ainsi une composante du bocage et est associé à des champs fermés par des enclos et des haies, de l'autre côté, l'habitat aggloméré s'intègre à des champs ouverts pour former des paysages dits d'openfield.

Cette manière de représenter l'espace rural suscita aux moins deux grandes questions de recherches interdépendantes :

- Pourquoi trouve-t-on cette répartition des structures agraires dans l'espace métropolitain français ?
- Quels sont les mécanismes physiques et sociaux qui expliquent cette association (corrélation spatiale) entre les formes agricoles et les formes de peuplement ?

DEMANGEON 1927 en géographie et BLOCH 1931 en histoire, forgèrent chacun à partir de leur domaine de compétences des réponses à ces questions qui guident encore aujourd'hui notre compréhension des espaces ruraux. En particulier, ils réfutèrent l'argumentation de RATZEL 1891 selon laquelle ces différences observées résulteraient de différences ethniques de conception et d'organisation de l'espace entre les peuples germaniques et celtes. Plus largement, ils démontrèrent que les explications étaient moins à rechercher du côté de la géographie physique que de la géographie sociale, la première jouant principalement un rôle facilitant ou limitant mais ne serait pas capable à elle seule de contraindre une organisation spatiale.

Au moins trois grandes causes interdépendantes ont été proposées par DEMANGEON 1927, BLOCH 1931 et leurs successeurs pour expliquer cette répartition de l'habitat rural (CURTIS 2013) :

- La première est liée au mode de production agricole. L'association de champs ouverts et d'un habitat groupé trouverait son origine dans une mise en commun et un partage des ressources agricoles induit, entre autres, par l'assolement triennal. A l'inverse, l'habitat dispersé s'expliquerait par une agriculture bocagère qui nécessiterait en permanence la présence des cultivateurs à proximité de leur champ ;
- La seconde est liée à l'organisation du pouvoir. Localement, l'existence d'un pouvoir fort impliquerait un contrôle social fort. Afin de faciliter celui-ci, les populations seraient regroupées par les élites et les groupes dominants ;
- Enfin, le troisième type serait lié à des considérations de défense. L'organisation du peuplement répondrait à des logiques de protection des habitants face à des menaces extérieures : il serait plus aisé de garantir la sécurité de la population dans l'habitat groupé que dispersé. Ceci expliquerait le caractère groupé des populations du nord-est de la France situées à proximité du Saint-Empire romain germanique.

La validité de ces explications repose sur l'hypothèse que les concepts d'habitats dispersés et concentrés soient univoques. Or comme le souligne O'KEEFFE 2000, rares sont les chercheurs qui s'accordent sur la définition de la concentration et de la dispersion. Par ailleurs, le caractère universel de la description des paysages ruraux selon l'opposition entre bocage et openfield a été remise en question par LAVIGNE 2003. De sorte que l'hypothèse d'une corrélation systématique entre forme d'habitat et forme agraire a pu être remise en cause récemment par WATTEAUX 2005.

Au-delà de ces enjeux de recherche qui dépassent largement la portée de cette thèse, on voit la nécessité d'explicitier de façon précise les termes que nous employons pour décrire des configurations géographiques, sans quoi nous risquons de surinterpréter et réifier les formes observées.

Remarque

Le couple bocage-openfield n'est pas de même nature que le couple ville-périphérie (O'KEEFFE 2000). La ville forme avec sa périphérie un système contenant de nombreuses rétroactions : il est difficile d'étudier leur fonctionnement séparément. En revanche, le bocage et l'openfield ne constitue pas un système : l'existence du bocage n'implique pas celle de l'openfield. Les deux systèmes agraires peuvent être analysés indépendamment l'un de l'autre.

2.2.3.5 La dépendance des résultats aux paramètres : les limites de la recherche de *clusters*

A l'époque de la collecte des populations agglomérées, la question de la pertinence des définitions des objets mesurés était déjà posée. Ainsi DION 1934 dans son ouvrage, *Essai sur la formation du paysage rural français*, a pu écrire :

« Une population qui ne se groupe pas autour du clocher n'est pas nécessairement disséminée. Elle peut former, en un ou plusieurs points du même territoire communal, des hameaux assez importants pour être qualifiés d'agglomérations ».

Reformulé avec les concepts du *Machine Learning*, nous dirons aujourd'hui que nous sommes confrontés à un problème de sur-apprentissage. Notre calibration optimale fonctionne très bien, trop bien sur nos deux exemples de configurations, mais n'est pas



FIGURE 2.16 – Analyse de sensibilité (en bleu les points agrégés et en rouge les non agrégés) d’une zone contenue dans le département de l’Eure. La part des bâtiments appartenant à une zone agrégée est très sensible aux paramètres choisis. On constate à partir de cet exemple, la difficulté à définir de façon univoque la notion de village.

adaptée pour décrire des territoires qui s’éloignent morphologiquement trop de celles-ci. En d’autres mots, l’analyse réalisée sur les deux configurations idéaltypiques n’est pas généralisable à tout l’espace français sans une perte de pertinence d’autant plus forte que les configurations locales observées dérogent aux formes qui ont permis de calibrer les paramètres.

Pour préciser ceci, nous avons étudié une zone de France où le taux de bâtiments agrégés est le plus sensible aux valeurs des paramètres (figure 2.16). Nous constatons qu’avec le jeu de paramètres (200,20), nous retrouvons approximativement des noyaux villageois, tels que notre intuition semble nous le suggérer. Néanmoins, ces structures sont invisibles pour les autres jeux de paramètres. Ceci s’explique avec la taille des noyaux villageois dans ce territoire¹⁴. Avec les paramètres 200 et 20, on constate que 50% des

14. La DBSCAN, contrairement aux méthodes basées sur la fonction de Ripley fournit des agrégats, nous pouvons donc étudier les propriétés de ces derniers

agrégats de ce territoire regroupent moins de 53 bâtiments. A titre de comparaison, dans la distribution archétypale d'openfield, avec le même jeu de paramètres, 50% des agrégats ont plus de 145 bâtiments. Ainsi d'avoir calibré les paramètres de l'algorithme DBSCAN sur une zone d'openfield où les villages sont de grandes tailles, fait que l'algorithme est moins pertinent pour décrire les zones d'habitat concentré où les agrégats sont de tailles plus modestes.

L'arbitraire induit par notre façon de définir la concentration et la dispersion est accentué par le choix d'une méthode d'agrégation statistique des données primaires. Nous avons utilisé des ratios de moyennes pour synthétiser la répartition locale des bâtiments selon leur caractère concentré ou dispersé. Un autre indicateur, tel que l'entropie ou la taille des *clusters*, aurait produit des résultats différents. Si bien que, comme nous l'écrivions dans la section précédente, il est très important d'être prudent dans l'interprétation des résultats pour ne pas réifier de façon trop forte des situations floues et amener à des interprétations erronées. Ainsi, en toute rigueur, nous ne pouvons pas déduire de la carte de paramètres ($eps = 200m$ et $min_{pts} = 20$) l'inertie des structures locales de peuplement entre 1876 et 2012. En dehors du fait que la carte de population dispersée de 1876 (figure 2.15) repose sur des données de population, nous commettrions un raisonnement circulaire à s'étonner de leurs ressemblances car l'élaboration des cartes plus récentes (fig. 2.13 et 2.14) a été influencée par la carte de 1876.

Remarque

ZHANG et al. 2014 analysent le peuplement d'une région chinoise à l'aide de la fonction de Ripley. Dans cette publication, la fonction de Ripley contrairement aux résultats que nous obtenons permet de différencier des formes de peuplement. Néanmoins cette analyse contrairement à la nôtre n'est pas menée sur des bâtiments mais sur des centroïdes de villages. Nous retrouvons dans ce cas implicitement la question de la définition des entités primaires utilisées, les villages, et de l'impact de celle-ci sur la sensibilité des résultats. Les auteurs de l'article ne s'attardent pas sur cette question et considèrent comme allant de soi la définition officielle des villages, ce qui limite fortement la généralisation des résultats de cet article.

2.3 Conclusion

D'un point de vue opérationnel, le recours à la statistique spatiale pour étudier l'organisation des tissus bâtis ne s'est pas révélé totalement satisfaisant. Pour autant, cette première étape dans notre cheminement ne doit pas être considérée comme inutile. L'incapacité de la fonction de Ripley et de ses dérivés à distinguer des formes que l'œil humain différencie sans difficulté est en soi un résultat. Outre le fait que nous justifions ainsi le peu d'utilisation de cette méthode pour décrire les tissus bâtis, nous montrons également que toutes les modélisations de l'espace n'ont pas la même capacité à décrire tous les phénomènes observables. Les paysages humains et les formes d'occupation de l'espace résultent d'une série d'héritages et d'actualisations par chaque génération, l'héritage étant à la fois une contrainte qui limite les champs du possible, et le support et la trame des éventuelles transformations liées aux évolutions techniques et sociales. Ainsi, contrairement aux exemples issus des sciences naturelles, tels que l'organisation des cellules dans un organe ou la répartition des arbres dans une forêt, l'hypothèse selon laquelle les bâtiments seraient une réalisation d'un processus ponctuel n'est guère pertinente (VAN ANDEL, RUNNELS et POPE 1986).

La DBSCAN, plus empirique, ne repose pas sur des hypothèses très fortes et donne de prime abord des résultats pertinents : on retrouve les structures agraires héritées du passé dans les formes rurales actuelles. Néanmoins, la forte variabilité de ceux-ci aux choix des paramètres limite leur pouvoir heuristique. Nous retrouvons dans cette limite l'impossibilité de résumer, même avec un ensemble d'outils très larges et conceptuellement très puissants, la variabilité des formes spatiales observée dans les sociétés humaines, déjà formulée par les géographes au début du 20^e siècle.

Bien que très différentes dans leurs présupposés conceptuels, les méthodes multi-échelles de la statistique spatiale ont en commun de comparer les configurations analysées à un objet tiers, étranger à celles-ci. Dans le cas de la fonction de Ripley, la référence est un processus de Poisson et avec la DBSCAN celle-ci est directement donnée par le praticien à travers le choix des paramètres. Ce type d'approches est classique en statistique mais interroge en géographie lorsque la configuration de référence est trop différente ou sur-adaptée aux configurations analysées.

Chapitre 3

Méthodes d'analyses fractales et multifractales pour l'analyse de la distribution spatiale des bâtiments, de la population et des formes d'occupation du sol

Résumé

Dans ce chapitre, nous présentons une exploration de la concentration et de la dispersion des implantations humaines à travers les échelles à l'aune de la géométrie fractale. Nous abordons d'abord la répartition des bâtiments dans l'espace en les modélisant sous la forme d'ensemble de points. L'analyse fractale, dans ce cas, s'intéresse principalement à l'entremêlement des espaces occupés et vides à travers les échelles. Ensuite, la prise en compte de l'intensité du peuplement dans l'espace, nous amène à introduire de nouveaux outils tels que les mesures multifractales. Enfin, l'analyse du paysage à l'aide de photographies demande de mettre en œuvre des méthodes complexes basées sur les ondelettes. La pertinence des résultats obtenus à chaque fois est moins fonction de la complexité des analyses mises en œuvre que de notre capacité à nous poser les bonnes questions sur ce que l'on mesure avec des méthodes trans-échelles.

3.1 Introduction

Entre les premiers écrits de Mandelbrot basés sur des modélisations sous la forme d'objets mathématiques relativement simples, tels que des ensembles de points, et les dernières productions abordant l'étude des trajectoires boursières, le champ d'application de l'analyse fractale s'est considérablement élargi. Cette extension des phénomènes naturels modélisables sous la forme d'objets fractals est le fruit d'une étroite collaboration entre les recherches appliquées et théoriques, les premières apportant de nouveaux défis aux secondes. Si les limites du champ d'application des fractales sont sans cesse repoussées, c'est aussi parce que les mathématiciens ont été capables de donner un sens fractal à des objets mathématiques théoriques tels que des distributions de probabilités ou des fonctions.

Cet élargissement des modélisations est à la fois une aubaine pour le praticien et un danger. Aubaine car il bénéficie de la puissance des fractales pour analyser des objets qui échappaient jusqu'à présent à ce paradigme théorique. Danger car de nombreux phénomènes peuvent être modélisés au moyen de divers objets plus ou moins complexes dont la modélisation impacte les descriptions fractales obtenues.

Dans ce chapitre, après avoir présenté l'analyse fractale classique et son application pour décrire les tissus bâtis, nous nous intéressons à des méthodes plus complexes pour prendre en compte les variations d'intensité des phénomènes dans l'espace et non plus uniquement leur localisation spatiale. Nous étudions dès lors la distribution spatiale de la population et les formes d'occupation du sol à partir d'orthophotographies. Introduire dans l'analyse une information supplémentaire, à savoir une troisième dimension (hauteur), n'est pas simple et nécessite de se poser des questions sur la nature du phénomène étudié. Lorsque celui-ci est discontinu (distributions de population) la modélisation sous la forme d'une mesure (une probabilité) est adaptée et au contraire, lorsque cette hauteur est continue (orthophotographies), la représentation sous la forme d'une fonction est plus appropriée.

Dans ce chapitre, nous nous sommes efforcés tant que faire se peut de montrer les limites des modèles sur les résultats obtenus et de prendre du recul sur l'usage des fractales en géographie. Que le monde soit fractal ou non nous importe peu tant que les outils fractals nous aident à le décrire de façon satisfaisante.

3.2 Analyse fractale d'une distribution de points

3.2.1 Les ensembles de points fractals

Les scientifiques, en particulier les physiciens et les mathématiciens, utilisaient des objets fractals pour décrire leurs phénomènes d'intérêt bien avant que Benoît Mandelbrot n'invente le terme fractal. L'apport fondamental de Mandelbrot est d'avoir regroupé en un seul concept un ensemble d'objets et de phénomènes que l'on considérait auparavant comme étant très différents. Cette unification a néanmoins un inconvénient : il est presque impossible de donner une définition univoque des objets fractals. La première définition d'un ensemble fractal que proposa Benoît Mandelbrot fut la suivante :

« un ensemble de points de dimension de Hausdorff strictement plus grande que sa dimension topologique » (MANDELBROT 1975).

Cette définition repose sur une notion théorique, la dimension de Hausdorff et exclut de ce fait une grande partie des phénomènes naturels pour lesquels cette dimension n'est pas calculable. D'autres définitions ont été proposées par la suite, sans qu'aucune ne fasse consensus. À la manière des biologistes, qui rencontrent de grandes difficultés à donner une définition unique et universelle de la vie, FALCONER 2013 propose une définition à tiroir des objets fractals. Selon celle-ci, un ensemble F est fractal lorsqu'il vérifie une majorité des assertions suivantes :

1. F présente de nombreux détails quelle que soit l'échelle d'observation ;
2. F est trop irrégulier pour être décrit efficacement avec le langage de la géométrie traditionnelle, aussi bien localement que globalement ;
3. F est caractérisé par des propriétés d'auto-similarité (exactes, approchées ou statistiques) ;
4. Le principe de construction de F est très simple et peut être récursif ;
5. La dimension fractale de F (la dimension ici n'est pas nécessairement celle de Hausdorff) est plus grande que sa dimension topologique.

A notre connaissance, en dehors de la cinquième assertion qui est très technique, les autres assertions sont couramment utilisées en géographie pour penser sous un prisme fractal les espaces urbains.

3.2.1.1 Les cascades fractales

Plusieurs méthodes pour construire des fractales théoriques ont été détaillées dans la littérature (FALCONER 2013 ; BARNSELY 2014)¹. Les cascades fractales, comme nous allons le découvrir à présent, sont construites selon des règles très simples mais sont susceptibles de produire une grande variété de formes.

La construction d'une cascade fractale commence tout d'abord par le choix d'un générateur (une matrice composée de 0 pour les espaces vides et de 1 pour les espaces pleins) qui décrit les motifs qui vont être dupliqués à travers les échelles. A la première étape, les espaces pleins du générateur sont remplacés par une copie miniature de celui-ci. La fractale est obtenue en répétant cette opération, c'est-à-dire en remplaçant les

1. FRANKHAUSER 2015 utilise le principe des cascades fractales pour dessiner des villes théoriques qui optimisent l'accessibilité à différents types d'aménités.

espaces pleins obtenus à l'étape précédente par une copie miniature du générateur. D'un point de vue algorithmique cette opération est très simple à réaliser avec les langages de programmation de haut niveau tels que Python (cf code² ci-dessous).

```
def cascade(generateur, nb_iteractions):
    fractal = generateur
    for i in range(nb_iteractions):
        fractal = np.kron(fractal, generateur)
    return(fractal)
```

La commande *kron* est la fonction de Kronecker. Elle permet de remplacer les éléments non nuls de la matrice *fractale* par le générateur.

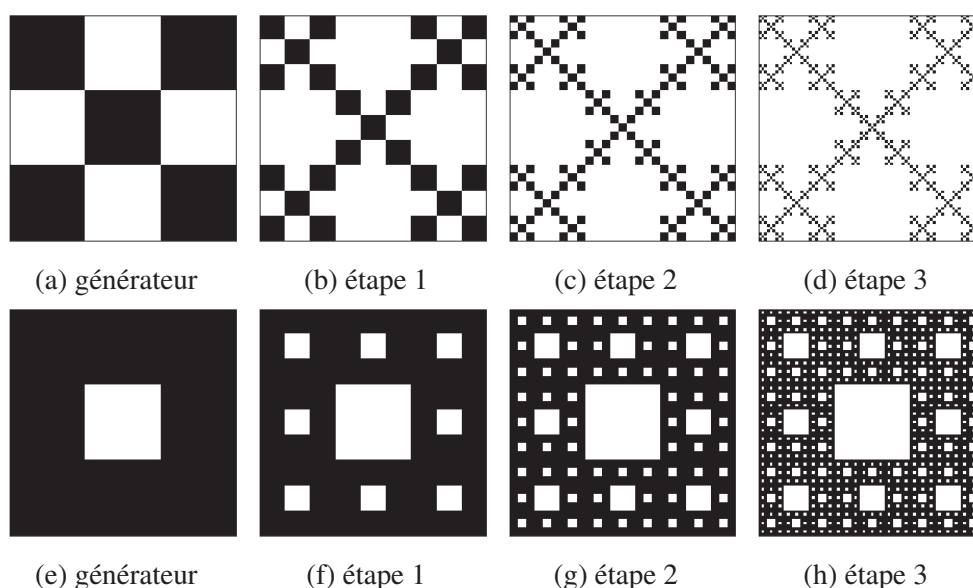


FIGURE 3.1 – Exemples de cascades fractales

Les cascades de la figure 3.1, appelées tapis de Sierpinski, sont construites à l'aide de deux générateurs différents. À chaque étape, la part des surfaces pleines décroît d'un facteur respectivement de $5/9$ pour la première et de $8/9$ pour la deuxième. Asymptotiquement, la surface pleine des deux fractales suit une loi d'échelle et tend vers 0. On constate que la deuxième « remplit » nettement plus l'espace à travers les échelles que la première. Cette intuition est à la base de la notion de dimension fractale.

3.2.2 Les dimensions fractales

Les dimensions fractales mesurent la façon dont les objets remplissent l'espace à travers les échelles. Les scientifiques les utilisent pour décrire et quantifier les organisations trans-échelles des objets fractales. Les mathématiciens et les physiciens ont développé de nouvelles dimensions au gré des besoins rencontrés. La dimension de Hausdorff est la dimension mathématique de référence. Son formalisme est complexe et s'applique uniquement aux objets théoriques. Dans les sciences appliquées, les chercheurs privilégient des dimensions plus simples à mettre en œuvre.

2. Ce code a été utilisé dans l'ensemble de la thèse pour produire différents exemples de cascades fractales et multifractales.

3.2.2.1 La dimension de *box-counting*

La dimension de *box-counting* repose sur la notion de mesure à une échelle donnée δ (FALCONER 2013). Pour chaque échelle δ , on mesure l'objet d'intérêt en négligeant les détails de taille plus petite que δ et l'on s'intéresse à la façon dont cette mesure varie à travers les échelles.

Supposons que l'on souhaite mesurer l'aire A_F d'un ensemble de points F , appartenant à un plan. Une façon de procéder consiste à rechercher le plus petit nombre $N_\delta(F)$ de disques de diamètre δ qui recouvrent F . A partir de cette information, on obtient une estimation de l'aire A_δ à l'aide de la formule suivante :

$$A_\delta = N_\delta(F) \times \delta^2 \quad (3.1)$$

Dans le cas où l'ensemble de points est une surface régulière, lorsque δ devient de plus en plus petit, l'aire approximée A_δ converge vers l'aire A de l'ensemble considéré. On peut alors écrire $N_\delta(F)$ sous la forme d'une loi de puissance :

$$N_\delta(F) \approx A \times \delta^{-2} \quad (3.2)$$

Remarque

$N_\delta(F)$ est appelée fonction d'échelle. Elle caractérise l'organisation de l'ensemble F lorsque l'on néglige les détails de taille inférieure à la résolution δ .

L'exposant de la loi de puissance (égal à 2) correspond à la dimension géométrique d'une surface. Si l'ensemble F est une courbe régulière et non plus une surface, l'aire d'une courbe étant nulle, la relation précédente ne s'applique plus. En revanche, en cherchant à estimer la longueur L , toujours en exploitant le même dispositif, nous obtenons :

$$N_\delta(F) \approx L \times \delta^{-1} \quad (3.3)$$

Cette fois, l'exposant correspond à la dimension d'une ligne. Cette propriété est générale et sert de base à la définition de la dimension de *box-counting*, définie comme l'exposant d de la loi de puissance caractérisant le comportement trans-échelle de $N_\delta(F)$:

$$N_\delta(F) \approx c \times \delta^{-d} \quad (3.4)$$

c étant un paramètre de forme. On obtient d par passage au logarithme :

$$\log(N_\delta(F)) \approx \log(c) - d \log(\delta) \quad (3.5)$$

Soit :

$$d \approx \frac{\log(N_\delta(F))}{-\log(\delta)} + \frac{\log(c)}{\log(\delta)} \quad (3.6)$$

En considérant le comportement asymptotique lorsque $\delta \rightarrow 0$, le second terme disparaît et reste uniquement la dimension de *box-counting* :

$$d = \lim_{\delta \rightarrow 0} \frac{\log(N_\delta(F))}{-\log(\delta)} \quad (3.7)$$

La recherche du plus petit nombre de disques de diamètre δ recouvrant F est une opération impossible à mener sur les objets réels et difficile à réaliser sur les objets théoriques. Les mathématiciens ont montré que l'on pouvait définir la dimension de *box-counting* plus simplement à l'aide d'une grille régulière composée de carreaux de côté δ projetée sur l'ensemble F . La fonction d'échelle $N_\delta(F)$ est alors le nombre de carreaux contenant au moins un point de F (fig. 3.2).

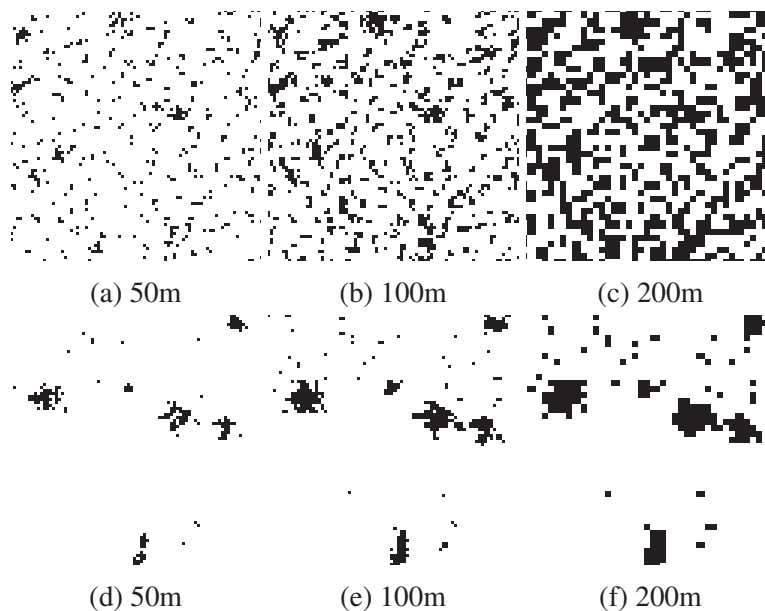


FIGURE 3.2 – Exemples de grilles recouvrant des configurations de bocage (en haut) et d'openfield (en bas)

Dans le cas des cascades fractales, la dimension de *box-counting* est tout simplement déterminée par le nombre d'éléments du générateur et le facteur de réduction, soit respectivement $\log(5)/\log(3) \approx 1.46$ et $\log(8)/\log(3) \approx 1.89$ pour les deux tapis de Sierpinski (fig. 3.1). On confirme en cela l'impression visuelle d'un remplissage plus important du plan par le deuxième tapis que par le premier.

Plus généralement, plus la dimension fractale est proche de 2 est plus l'ensemble de points tend à remplir uniformément le plan. Des valeurs proches de 1 caractérisent, elles, des objets qui ressemblent à des courbes régulières. Enfin, une dimension fractale proche de 0 indique que les points sont isolés les uns des autres.

Remarque

La limite $\lim_{\delta \rightarrow 0} \frac{\log(N_\delta(F))}{-\log(\delta)}$ n'est pas définie pour certains ensembles de points théoriques. Les mathématiciens privilégient pour ces ensembles la dimension de Hausdorff qui repose également sur une logique de recouvrement. Les objets utilisés pour recouvrir l'ensemble étudié n'ont pas tous le même diamètre : on impose seulement que celui-ci soit inférieur à δ . Ce changement permet de garantir l'existence de la dimension de Hausdorff pour n'importe quel ensemble de points et confère à cette dimension des propriétés mathématiques très pertinentes. Dans bon nombre de cas théoriques, les deux dimensions, *box-counting* et Hausdorff, sont identiques.

3.2.2.2 L'invariance d'échelle et le calcul de la dimension de *box-counting* pour les objets naturels

Dans les sciences naturelles, le calcul des dimensions fractales sur des plages d'échelles infinies n'a pas de sens. On recherche, dans ce cas, une plage de résolutions suffisamment large où $N_\delta(F)$ suit une loi d'échelle (équation 3.4). On qualifie d'invariants d'échelle les objets qui respectent une telle propriété. La dimension de *box-counting* est déterminée pour de tels objets en examinant la pente de la courbe représentant $\log(N_\delta(F))$ en fonction de $\log(\delta)$.

En pratique, lorsque le nombre de configurations analysées est relativement restreint, le choix de la plage de calcul est déterminée en recherchant les échelles communes où les courbes $\log(N_\delta(F))$ forment approximativement des droites. Sur la figure 3.3.a, nous avons représenté celles-ci pour les deux figures idéaltypiques de bocage et d'openfield pour des résolutions (longueurs des côtés des carreaux) comprises entre 12.5 m et 6400 m³. On remarque deux échelles de ruptures qui sont caractéristiques des objets naturels (LESNE et LAGUËS 2015). La première située à gauche de la plage d'échelles, lorsque δ est proche de 0, est induite par la précision des données analysées. Les bâtiments ont une emprise spatiale limitée de l'ordre d'une dizaine de mètres de côté de sorte que la distance moyenne entre deux centroïdes voisins est approximativement déterminée par cette valeur caractéristique du phénomène observé. Lorsque δ s'approche de cette valeur, les points semblent isolés les uns des autres et la pente, et donc la dimension fractale, se rapproche de 0. La deuxième rupture située à droite est provoquée par le fait qu'au-delà d'une certaine emprise spatiale, chaque carreau contient au moins un village et donc un bâtiment. Pour ces valeurs de δ , la dimension fractale se rapproche de 2. Au-delà et en-deçà de ces échelles de ruptures, l'information apportée par la dimension fractale est triviale et ne permet plus de distinguer les configurations de les unes des autres (PUMAIN 2010).

Pour l'analyse des deux configurations idéaltypiques de bocage et d'openfield (fig. 1.3 et 2.6), nous avons choisi de restreindre l'analyse aux grilles de carreaux dont les côtés des carreaux sont compris entre 50 et 800 m (fig. 3.3.b). A ce niveau de précision, nous nous intéressons principalement à la façon dont les bâtiments s'organisent en villages et en hameaux villageois. Les droites représentent l'ajustement linéaire aux points. On constate visuellement que la qualité de l'ajustement est meilleure pour la configuration d'openfield que de bocage. Ce fait est confirmé par les coefficients de détermination (R^2), qui sont respectivement de 0.9995 pour l'openfield et de 0.9675 pour le bocage. Pour cette deuxième configuration, l'invariance d'échelle semble moins avérée. La pente des droites, c'est-à-dire les dimensions fractales, sont proches : 1.13 pour l'openfield et 1.07 pour le bocage.

3.2.2.3 La dimension de corrélation

L'application de la dimension de *box-counting* aux objets naturels n'est pas sans poser quelques problèmes pratiques (BOUDA, CAPLAN et SAIERS 2016). Notamment, les résultats obtenus sont très sensibles aux choix de l'orientation et de l'origine des grilles utilisées. La dimension de corrélation (GRASSBERGER et PROCACCIA 1983) mise en œuvre dès 1997 par Pierre Frankhauser en géographie, en s'affranchissant de toute notion de grille, est nettement plus robuste. Sa fonction d'échelle est la moyenne du nombre de

3. Les codes associés sont disponibles sur le dépôt github <https://github.com/fsemecurbe/Fractal-Analysis.git>.

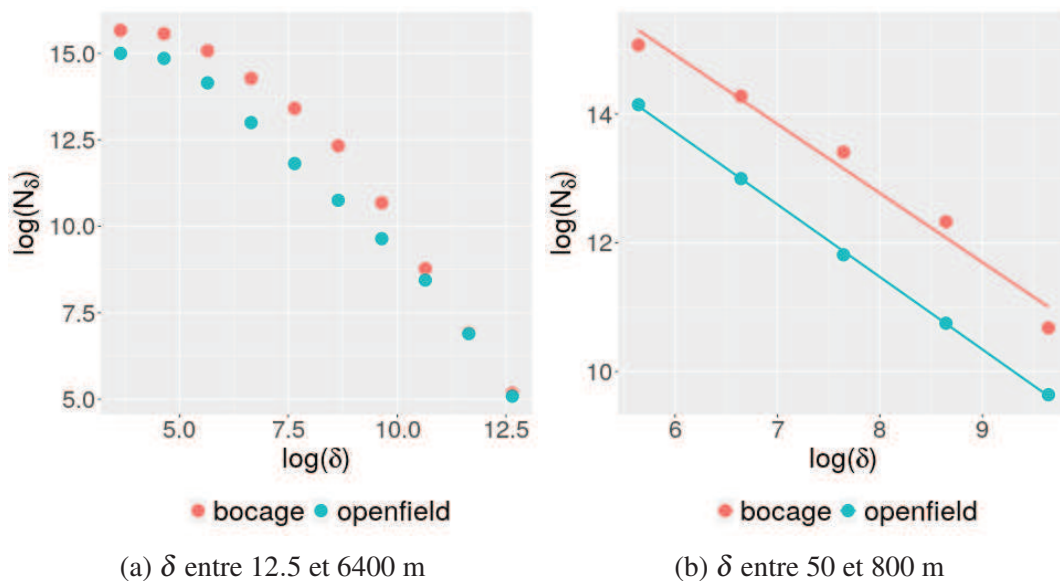


FIGURE 3.3 – $\log(N_\delta)$ en fonction de $\log(\delta)$ pour les deux territoires idéaltypiques (fig. 1.3 et 2.6). A gauche, les droites représentent l’ajustement linéaire pour des grilles de tailles δ comprises entre 50 – 800 m.

voisins que possèdent chaque point de l’ensemble F :

$$M_r(F) = \frac{1}{|F|} \sum_e M_r(e) \quad (3.8)$$

Où $|F|$ est le cardinal de l’ensemble F et $M_r(e)$ est le nombre de points contenus dans le disque centré sur le point e de rayon r .

De façon imagée, les disques forment des grilles « glissantes » qui s’adaptent aux configurations étudiées. Le nom de « corrélation » provient du fait que l’on s’intéresse à la façon dont les points se positionnent les uns par rapport aux autres. Comme pour la dimension de *box-counting*, le calcul de la dimension de corrélation pour les objets naturels consiste à rechercher une plage de rayons où $M_r(F)$ suit une propriété d’invariance d’échelle :

$$M_r(F) \approx r^d \quad (3.9)$$

Sur cette plage de rayons, une estimation de la dimension fractale d est obtenue en calculant la pente de la courbe de $\log(M_r(F))$ en fonction de $\log(r)$ ⁴. Comme le montre la figure 3.4, comparativement aux fonctions d’échelle précédentes, l’alignement des points est aussi bien vérifié pour la configuration de bocage que d’openfield pour les rayons compris entre 50 et 800 m : les coefficients de détermination (R^2) sont réciproquement de 0.9991 pour l’openfield et de 0.9981 pour le bocage. L’invariance d’échelle semble vérifiée pour les deux configurations. Les dimensions fractales sont de 1.57 pour l’openfield et de 1.42 pour le bocage. Ces dimensions de corrélation sont plus élevées que les dimensions de *box-counting* calculées précédemment et elles discriminent mieux les deux tissus bâtis.

4. Les codes associés sont disponibles sur le dépôt github <https://github.com/fsemecurbe/Fractal-Analysis.git>.

Encadré

Dans le sens commun, la notion de dimension renvoie au nombre de paramètres minimum nécessaires pour localiser un point dans un environnement donné sans ambiguïté (ARNEODO et al. 1996). En géographie, ces paramètres sont appelés coordonnées. La position d'un objet situé à la surface du globe terrestre est définie à l'aide de deux coordonnées, la longitude et la latitude. En rajoutant une troisième coordonnée, l'altitude, nous avons accès à l'ensemble de l'espace en trois dimensions.

La dimension fractale est une généralisation de la dimension géométrique, où l'on a remplacé le concept de coordonnée par celui d'information. SHANNON 1948 emprunte à la thermodynamique le concept d'entropie pour quantifier l'information contenue dans une distribution quelconque. L'entropie mesure la taille du code optimal permettant de communiquer à un tiers le résultat d'un tirage aléatoire parmi les éléments de la distribution. Plus le code est long et plus l'information contenue dans la distribution est importante.

Dans notre cas, la situation de communication est la suivante. Un évènement vient de se produire et a pour localisation un point de l'ensemble F . Un expert souhaite localiser celui-ci. Sa seule possibilité est de s'adresser à un témoin qui ne peut répondre que par « oui » ou « non ». $\log(N_\delta)$ est le nombre minimal de questions que devra poser l'expert en procédant par dichotomie pour localiser l'évènement avec une précision δ . $\log(N_\delta)$ est l'entropie dite de Hartley. Elle correspond à l'entropie de Shannon sous l'hypothèse que la distribution analysée suit une répartition uniforme.

Lorsque l'expert souhaite doubler la précision de la localisation (passer de δ à $\delta/2$), il doit poser $\log(N_{\delta/2}) - \log(N_\delta) = d$ questions supplémentaires. Cette quantité d'information d est égale à la dimension de *box-counting*, qui s'interprète donc comme un taux marginal d'information défini à travers les échelles.

3.2.3 Application : *Geographically Weighted Fractal Analysis*

Les dimensions fractales, en plus d'être mobilisées pour comparer des territoires différents, servent également à explorer l'organisation interne d'un unique territoire. DE KEERSMAECKER, FRANKHAUSER et THOMAS 2004 utilisent le découpage communal pour caractériser la variabilité spatiale du bâti au sein de l'agglomération bruxelloise. Pour chaque commune, une dimension fractale est calculée sur leur bâti. DE KEERSMAECKER, FRANKHAUSER et THOMAS 2003 et ENCARNÇÃO et al. 2012 au lieu de recourir à un découpage administratif, exploitent un carroyage régulier pour décrire leur territoire d'analyse. Ce choix permet de lever une partie de l'arbitraire induit par l'hétérogénéité des formes des divisions administratives utilisées.

En analyse spatiale, les méthodes « géographiquement pondérées » (*Geographically Weighted*), inspirées de la statistique non paramétrique, ont été développées dans l'objectif de « territorialiser » des analyses globales. BRUNSDON, FOTHERINGHAM et CHARLTON 1996 inaugureront celles-ci par une déclinaison locale des méthodes de régression statistique (*Geographically Weighted Regression*). A la place de calculer une unique régression, les auteurs proposent d'estimer des régressions en tout point de l'espace, l'objectif étant de décrire la variation des relations dans l'espace entre une variable expliquée et des variables explicatives. Le caractère local de l'analyse est obtenu en affectant un poids aux données

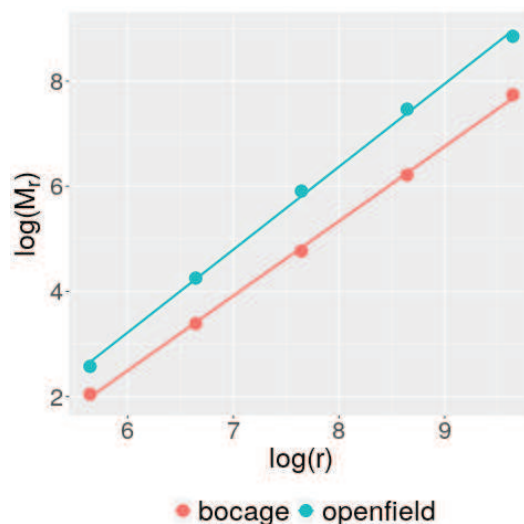


FIGURE 3.4 – Fonctions d'échelles de la dimension de corrélation des configurations idéaltypiques de bocage et d'openfield pour des plages de rayons comprises entre 50 et 800 m (les lignes représentent les ajustements linéaires utilisés pour calculer les dimensions fractales).

Remarque

AGTERBERG 2013 constate l'existence d'une forte ressemblance entre les méthodes de la statistique spatiale et de l'analyse fractale. D'une part, l'écriture de la fonction d'échelles de la dimension de corrélation est très proche de celle de la fonction de Ripley. D'autre part, la dimension de *box-counting* est reliée à la méthode des quadrats car les deux reposent sur l'utilisation d'une grille régulière de carreaux de taille variable.

statistiques selon leur proximité au point d'estimation. Plus les observations statistiques sont éloignées du point d'estimation considéré et moins leur contribution dans le calcul est importante. Les pondérations sont déterminées à l'aide d'une fonction décroissante positive de la distance, appelée noyau. Cette fonction est paramétrée par un rayon de lissage (on parle également de bande passante ou *bandwidth*) qui précise la taille du voisinage pris en compte dans l'analyse. D'un point de vue théorique, cette « territorialisation » de l'analyse est directement inspirée des méthodes de lissage⁵. Plus récemment, BRUNSDON, FOTHERINGHAM et CHARLTON 2002 et HARRIS et al. 2015 ont étendu cette démarche à une large gamme de statistiques telles que les médianes et les analyses en composantes principales (GWPCA).

La Geographically Weighted Fractal Analysis (SÉMÉCURBE, TANNIER et ROUX 2019)⁶ est l'application aux dimensions fractales des méthodes de « territorialisation » par noyau. Cette méthode contrôle les effets de zonages (y compris l'arbitraire engendré par l'orientation et l'origine des unités spatiales) du MAUP qui se réduit alors principalement à la question du choix de l'échelle de restitution de l'information géographique.

5. Les densités lissées peuvent être considérées à cet égard comme une « territorialisation » de la notion de densité globale

6. L'article est en Annexe B. Un package R CRAN.R-project.org/package=gwfa a été développé durant le cadre de la thèse et une application de la GWFA est disponible sur le dépôt github github.com/fsemecurbe/GWFA.git

La dimension de corrélation, contrairement à la dimension de *box-counting* qui repose sur l'usage de grille, se prête aisément à l'introduction d'une pondération par noyau. Soit i un point de l'espace géographique, la fonction d'échelle locale $M_r(i)$ est définie de la sorte (fig. 3.5)⁷ :

$$M_r(i) = \sum_e W_{i,e} M_r(e) \quad (3.10)$$

où i est un point d'observation, e est un point bâti, et $M_r(e)$ est le nombre de points bâtis situés à une distance inférieure ou égale à r du point e . Le poids $W_{i,e}$ est fonction de la distance qui sépare le point d'estimation i du point bâti e : plus celle-ci est importante est plus le poids est faible. Ce poids est calculé en recourant à un noyau *biweight* K :

$$W_{i,e} = K(d_{i,e}/L) \text{ où } L \text{ est le rayon de lissage.} \quad (3.11)$$

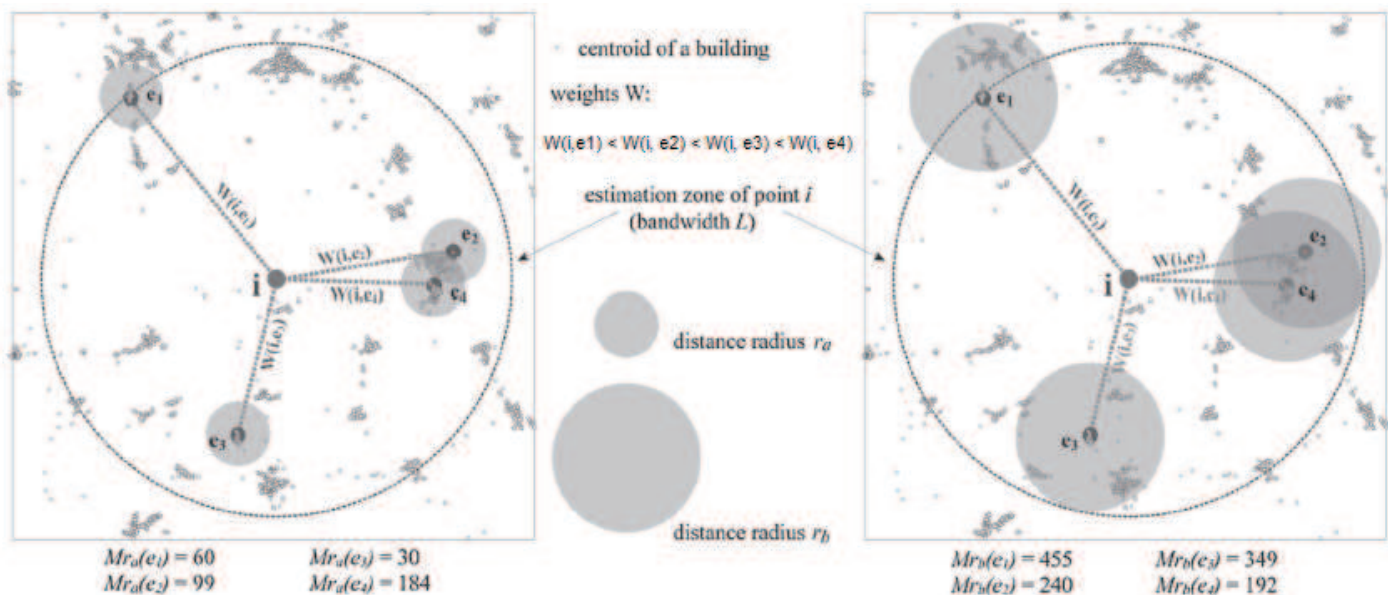


FIGURE 3.5 – Les différentes notions mises en œuvre dans la GWF. Source : SÉMÉCURBE, TANNIER et ROUX 2019

Différentes formes fonctionnelles de noyaux ont été proposées dans la littérature telles que les fonctions uniformes, gaussiennes ou quadratiques (BADDELEY, RUBAK et TURNER 2015; DIGGLE 2013). Nous avons privilégié la dernière dans nos travaux principalement pour des raisons d'optimisation des temps de calcul.

$$K(x) = (1 - x^2)^2, \text{ si } x < 1, \text{ et } 0, \text{ sinon.} \quad (3.12)$$

Le noyau quadratique ressemble à un noyau gaussien, mais contrairement à celui-ci, il a un support compact ce qui permet de restreindre les calculs aux points e situés à une distance inférieure à L du point d'observation i et ainsi d'augmenter la vitesse des traitements.

La dimension fractale locale en un point i est obtenue en estimant, comme précédemment, la pente de $\log(M_r(i))$ en fonction de $\log(r)$.

7. La fonction d'échelle présentée ici est légèrement différente de celle présentée dans l'article SÉMÉCURBE, TANNIER et ROUX 2019.

3.2.3.1 Contraintes sur les paramètres globaux

Bien qu'en théorie, nous puissions calculer la GWFA en tout point de l'espace, en pratique nous sommes forcés de discrétiser l'espace d'observation sous la forme d'une grille carroyée. Le centre de chaque carreau sert de point d'estimation. Le choix de la taille des carreaux est moins dicté par des contraintes statistiques qu'esthétiques : plus les carreaux sont petits par rapport au rayon de lissage et plus l'illusion d'observer un espace continu d'observations semble vérifiée. A titre d'information, sur la carte 3.6, les carreaux ont une taille de 2 km de côté et le rayon de lissage est de 8 km. De façon imagée, le rayon de lissage détermine la taille du support spatial des analyses locales. L'analyse est locale lorsque le rayon est petit et globale lorsque le rayon est important. Lorsque le rayon tend vers l'infini, on récupère une analyse globale sans pondération : tous les points d'estimation ont la même dimension fractale qui correspond exactement à la dimension calculée sur l'intégralité de l'espace d'intérêt étudié. La plus grande des échelles d'analyse détermine la taille minimale du rayon de lissage. Ce phénomène que l'on ne retrouve pas dans les analyses pondérées initiales (GWR, GWPCA), est propre à la GWFA. En effet, en assimilant le rayon de lissage à la taille du support de l'analyse, on comprend que lorsque celui-ci est petit par rapport aux échelles d'analyse, les effets de bords deviennent prédominants et réduisent la portée heuristique de l'analyse. En pratique, un rapport de 5 entre le rayon de lissage et la plus grande des échelles garantit la robustesse des dimensions et des fonctions d'échelles obtenues.

Les centres urbains ont des dimensions fractales supérieures à 1.5 (fig 3.6). On retrouve en cela des valeurs conformes à celles obtenues par les géographes sur le tissu bâti des espaces urbains. Les espaces ruraux très peu dense ont des dimensions inférieures à 1. Entre ces deux types d'espaces les dimensions fractales s'échelonnent entre 1 et 1,5. Le coefficient de corrélation, entre les deux cartes, supérieur à 0.9 confirme l'impression visuelle, la densité et la dimension fractale apportent une information quasi-identique.

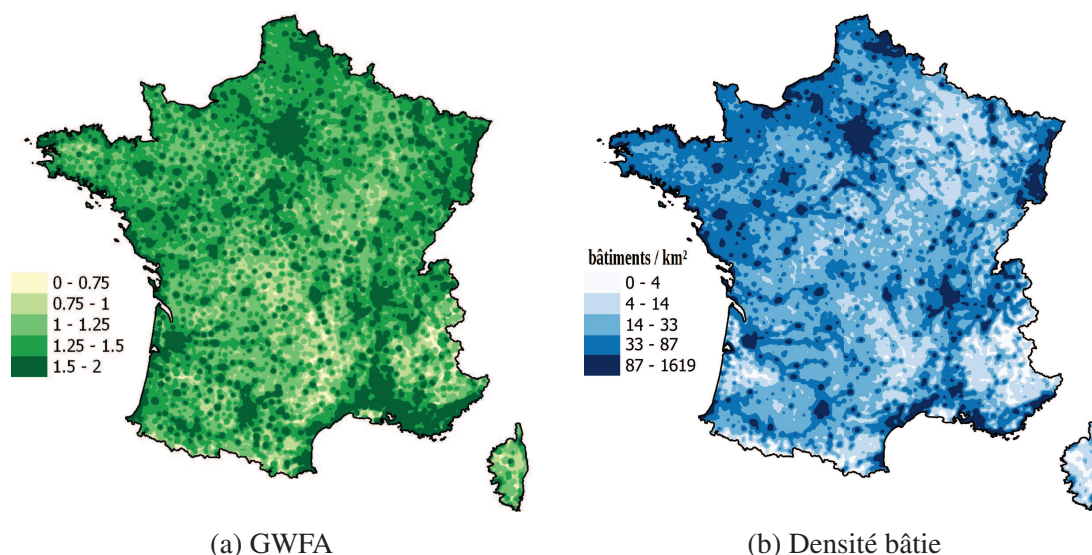


FIGURE 3.6 – Comparaison GWFA avec la dimension de corrélation et densité du bâti en France métropolitaine (les rayons considérés pour l'analyse fractale sont 50; 100; 200; 400; 800 m, le rayon de lissage pour les deux analyses est de 8 km et la taille des carreaux est de 2 km).

Remarque

L'information apportée par cette analyse fractale est redondante avec une analyse de densité classique. Ceci s'explique par le fait, que contrairement aux fractales théoriques, les points analysés représentent des objets (des bâtiments) dont l'emprise spatiale est non nulle. Pour les espaces de fortes densités, les bâtiments occupent tout l'espace, la dimension fractale est proche de 2. Au contraire dans les espaces peu denses les lacunes dominent et la dimension fractale est faible.

3.3 Analyse multifractale d'une mesure

L'analyse multifractale est une collection d'outils permettant d'analyser et de caractériser des objets invariants d'échelles plus complexes que de simples ensembles de points (APPLEBY 1996) : un seul exposant n'est pas *a priori* suffisant pour résumer les propriétés trans-échelles des objets étudiés.

3.3.1 Définition d'une mesure

La notion de mesure est à la base de l'analyse multifractale. Une mesure est un concept mathématique qui associe des poids relatifs aux différentes parties d'un ensemble (ARNEODO et al. 1996). Une répartition de population dans l'espace est une mesure de celui-ci. Les parties, dans ce cas, sont les différentes unités territoriales de l'espace considéré et le poids d'une unité territoriale est donné par la population qu'elle contient.

Une mesure μ est une fonction qui associe une grandeur numérique aux différentes parties d'un ensemble E et qui possède les propriétés suivantes :

- la mesure de la partie vide, \emptyset , est nulle, $\mu(\emptyset) = 0$.
- la mesure de la réunion de deux parties disjointes, A et B ($A \cap B = \emptyset$), est la somme des mesures de ces deux parties $\mu(A \cup B) = \mu(A) + \mu(B)$.
- la mesure du territoire englobant est bornée : $\mu(E) < +\infty$.

Une probabilité P est une mesure spécifique dont la mesure de l'ensemble global est égale à 1. On passe d'une mesure à une probabilité à l'aide d'une opération de normalisation :

$$P(x) = \frac{\mu(x)}{\mu(E)} \quad (3.13)$$

La définition de mesures auto-similaires, dans un cadre mathématique très général, est une opération complexe. Nous allons restreindre notre présentation ici aux cascades multifractales (APPLEBY 1996).

3.3.1.1 Mesures auto-similaires : cascades multifractales

Les cascades multifractales sont obtenues à l'aide d'un processus simple comparable à celui utilisé dans la construction des cascades fractales de la section précédente.

Considérons dans un premier temps que notre population d'intérêt se répartit à l'intérieur d'une grille de dimension 2×2 :

		P_1P_1	P_1P_2	P_2P_1	P_2P_2
P_1	P_2	P_1P_3	P_1P_4	P_2P_3	P_2P_4
		P_3P_1	P_3P_2	P_4P_1	P_4P_2
P_3	P_4	P_3P_3	P_3P_4	P_4P_3	P_4P_4

TABLE 3.1 – Générateur et première étape de construction d’une cascade multifractale

P_i est la probabilité associée à la cellule i . Elle caractérise la part de la population contenue dans la cellule i . Nous pouvons redécouper chacune des cellules de la grille en 4 sous-cellules dont les probabilités sont obtenues selon un processus multiplicatif (APPLEBY 1996). Pour produire une mesure auto-similaire ce procédé est reproduit une infinité de fois.

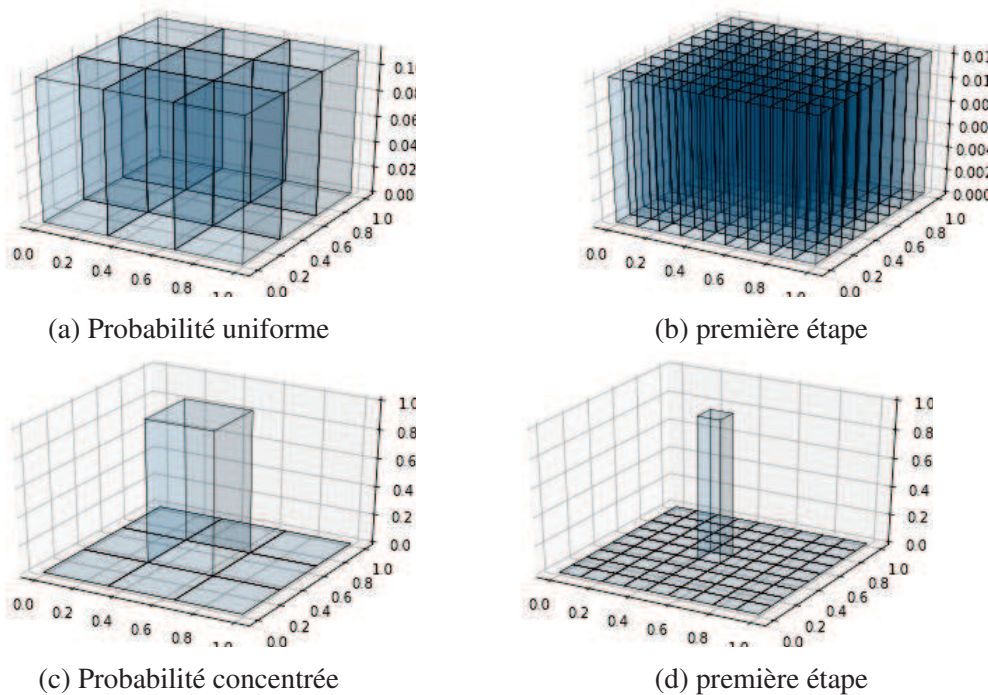


FIGURE 3.7 – Exemples de cascades multifractales remarquables

Le choix des quatre probabilités P_1 , P_2 , P_3 et P_4 conditionne la régularité spatiale de la mesure. Afin d’illustrer ce fait, nous allons présenter deux situations remarquables. Supposons, tout d’abord, que les 4 probabilités soient égales (fig. 3.7.a). La répartition de la population est uniforme et la probabilité associée à un territoire est proportionnelle à son aire. Une situation symétrique revient à considérer que la population à toutes les échelles est concentrée dans une seule cellule : elle concentre la population en un seul point (fig. 3.7.b).

Des mesures plus complexes sont obtenues en introduisant des poids inégaux dans un tapis de Sierpinski (tableau 3.2). Cette fois-ci, le générateur est de dimension (3, 3). Nous avons représenté ci-après trois mesures basées sur trois générateurs différents (de gauche à droite : cascade 1, cascade 2 et cascade 3) (fig. 3.8).

1/8	0	1/8	1/6	0	1/6	3/16	0	3/16
0	1/2	0	0	1/3	0	0	1/4	0
1/8	0	1/8	1/6	0	1/6	3/16	0	3/16

TABLE 3.2 – Probabilités associées à chaque cellule des générateurs (dans l'ordre, cas 1, cas 2 et cas 3)

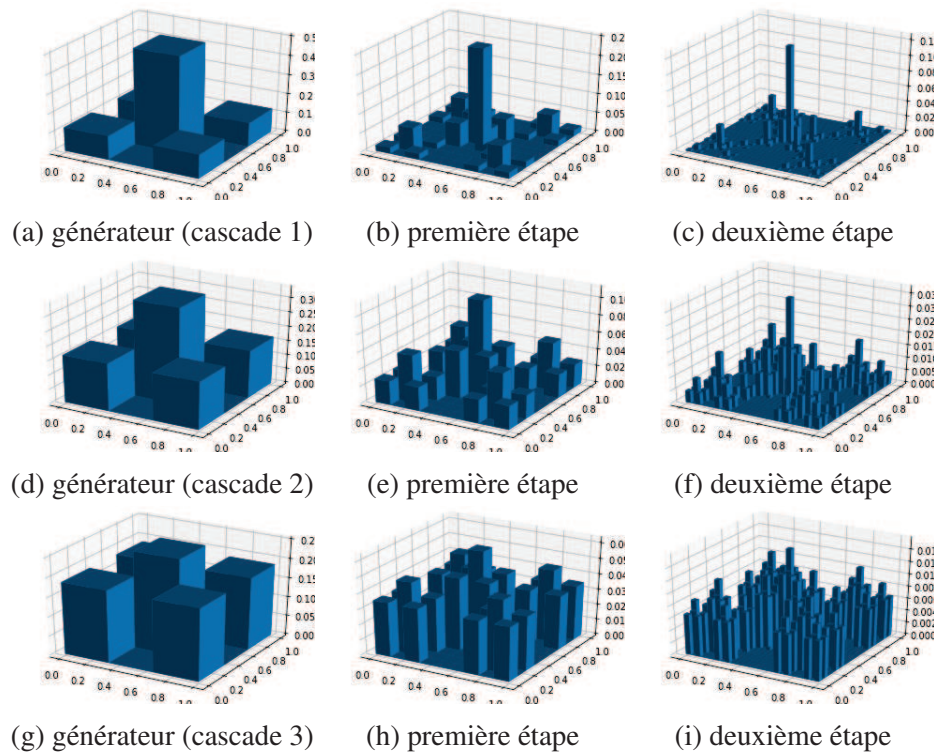


FIGURE 3.8 – Exemples de cascades multifractales

Les trois mesures ainsi construites ont le même support⁸ qui correspond exactement au tapis de Sierpinski de dimension fractale $\frac{\log(5)}{\log(3)}$. A travers les échelles, le cas 1 concentre davantage les masses que le cas 3. Le cas 2 représente quant à lui une situation intermédiaire. Si les poids de la probabilité étaient répartis uniformément ($1/5$) sur les 5 cellules du générateur, alors la répartition des probabilités serait uniforme sur le tapis de Sierpinski. Dans ce cas, la mesure serait dite monofractale.

3.3.2 Les dimensions généralisées

3.3.2.1 Présentation théorique

Les dimensions généralisées constituent une famille de dimensions qui peuvent être considérées comme une extension de la dimension de *box-counting*. Tout comme elle, leur calcul repose sur l'utilisation d'une grille régulière. On définit les fonctions d'échelle

8. Le support d'une mesure est son empreinte, à savoir les parties de l'espace qui ont une mesure non nulle. D'un point de vue mathématique, cette notion se formalise de la façon suivante : le support d'une mesure est l'ensemble des points tels que la mesure d'une boule centrée sur ceux-ci de taille quelconque est toujours strictement positive.

suivantes :

$$I_{q,\delta} = \frac{1}{1-q} \log \sum P(C_{i,\delta})^q \text{ si } q \neq 1 \quad (3.14)$$

$$I_{1,\delta} = - \sum P(C_{i,\delta}) \log(P(C_{i,\delta})) \text{ si } q = 1 \quad (3.15)$$

Où $C_{i,\delta}$ est un carreau de côté δ . Ces fonctions d'échelles sont des entropies de Rényi de la distribution de probabilités $P(C_{i,\delta})$. Pour $q = 1$, l'entropie de Rényi est appelée entropie de Shannon (BROMILEY, THACKER et BOUHOVA-THACKER 2004). Cette entropie quantifie la quantité d'information nécessaire pour localiser avec la résolution δ un événement dont l'emplacement serait aléatoire et régi par la probabilité P . Pour $q = 0$, on retrouve l'entropie de Hartley utilisée dans la définition des fonctions d'échelle de la dimension de box-counting. Les fonctions d'échelles $I_{q,\delta}$ sont décroissantes en fonction de q pour un δ donné.

Les dimensions généralisées D_q sont définies, comme précédemment, en observant le comportement de $I_{q,\delta}$ pour des échelles infiniment petites :

$$D_q = \lim_{\delta \rightarrow 0} \frac{I_{q,\delta}}{\log(\delta)} \quad (3.16)$$

La décroissance des fonctions d'échelle implique la décroissance des dimensions généralisées en fonction de q . Pour $q = 0$, on retrouve la dimension de *box-counting* du support de la mesure. Pour $q = 2$, la dimension est la dimension de corrélation (ARNEODO et al. 1996) et pour $q = 1$ la dimension est appelée dimension d'information.

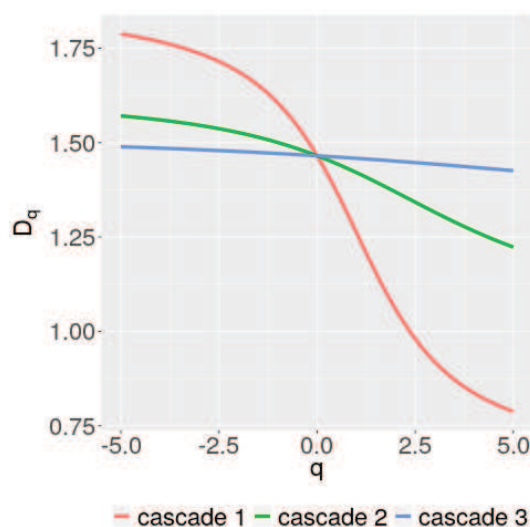


FIGURE 3.9 – Dimensions généralisées en fonction de q des 3 cascades multifractales de la figure 3.8

La vitesse de la décroissance des dimensions généralisées en fonction de q témoigne du caractère plus ou moins multifractal des mesures (fig 3.9). La cascade 1 est la cascade dont le caractère multifractal est le plus accentué. A l'opposé, la cascade 3 dont la courbe est quasi-plate se rapproche d'une mesure monofractale. Pour les mesures monofractales, toutes les dimensions généralisées sont identiques et ont pour valeur la dimension de box-counting : une seule dimension caractérise la mesure et la répartition des poids de la mesure est décrite uniquement par la forme de son support. La cascade 2 a un statut intermédiaire.

3.3.2.2 Application du calcul des dimensions généralisées à la GWFA

D'un point de vue strictement théorique, les ensembles de points ne sont pas des mesures. Leurs dimensions fractales, estimées asymptotiquement, sont uniquement déterminées par leur géométrie, à savoir le positionnement relatif des points entre eux. Cependant, dans les cas pratiques qui nous intéressent, un ensemble de points peut être considéré comme une mesure lorsque la résolution spatiale de l'analyse est suffisamment grande par rapport à l'emprise spatiale des objets que les points représentent. Par exemple, les bâtiments ayant dans leur grande majorité une surface comprise entre 100 et 200 m², le dénombrement de leur centroïde dans une grille régulière de carreaux de 50 m de côté forme une mesure non triviale de l'espace. L'analyse multifractale d'un ensemble de points peut être réalisée à l'aide de la méthode présentée ci-avant à l'aide de grilles de tailles variables tant que la résolution de la grille la plus fine est très supérieure à la granularité des points étudiés. Les dimensions généralisées décrivent à travers les échelles la façon dont les points se regroupent dans l'espace en amas de densité hétérogène.

TÉL, FÜLÖP et VICSEK 1989; VICSEK 1990; VICSEK 1992 constatent que cette manière de procéder produit des dimensions généralisées peu fiables et biaisées en particulier pour les q négatifs très sensibles aux zones très peu denses : les grilles régulières ne s'adaptent pas à l'organisation spatiale de l'objet étudié. Pour palier ce problème, les auteurs proposent une nouvelle méthode d'estimation des dimensions généralisées appelée *sandbox*. Cette méthode est une généralisation de la dimension de corrélation et consiste à remplacer les carreaux des grilles d'analyse par des disques centrés sur les points. La fonction d'échelle d'ordre q est définie de la sorte :

$$M_r^q = \left(\sum_e (M_r(e))^{q-1} \right)^{\frac{1}{q-1}} \quad (3.17)$$

où $M_r(e)$ comme précédemment est le nombre de points voisins autour du point e dans un rayon r . Les dimensions généralisées D_q sont obtenues en estimant la pente de $\log(M_r^q(i))$ en fonction de $\log(r)$. Les estimateurs *sandbox* peuvent être également « territorialisés » en appliquant une méthode identique à celle déployée dans la section 3.2.3. La fonction d'échelle en un point i devient alors :

$$M_r^q(i) = \left(\sum_e (W_{i,e} M_r(e))^{q-1} \right)^{\frac{1}{q-1}} \quad (3.18)$$

3.3.2.2.1 Analyse de sensibilité de la GWFA aux paramètres

Contrairement à la GWR où des critères de validation croisée ou d'Akaike permettent de déterminer des rayons de lissage optimaux (LOONIS et DE BELLEFON 2018), dans le cas de la GWFA le choix du rayon de lissage est principalement dicté par des considérations empiriques. Ceci implique de mener une étude de sensibilité afin de contrôler l'arbitraire des résultats obtenus.

Il apparaît que les structures spatiales révélées avec $L = 4000$ m présentent trop de détails qui nuisent à la lecture (fig. 3.10). Au contraire, un rayon $L = 16000$ m masque trop les variations locales, $L = 8000$ m représente un bon compromis pour l'analyse de la distribution spatiale locale des bâtiments.

L'étude de l'impact de la forme du noyau de lissage sur les analyses est une question récurrente en statistique non paramétrique. Dans l'objectif d'approfondir cette question, une comparaison avec une analyse fractale simple a été réalisée. Celle-ci revient à substituer

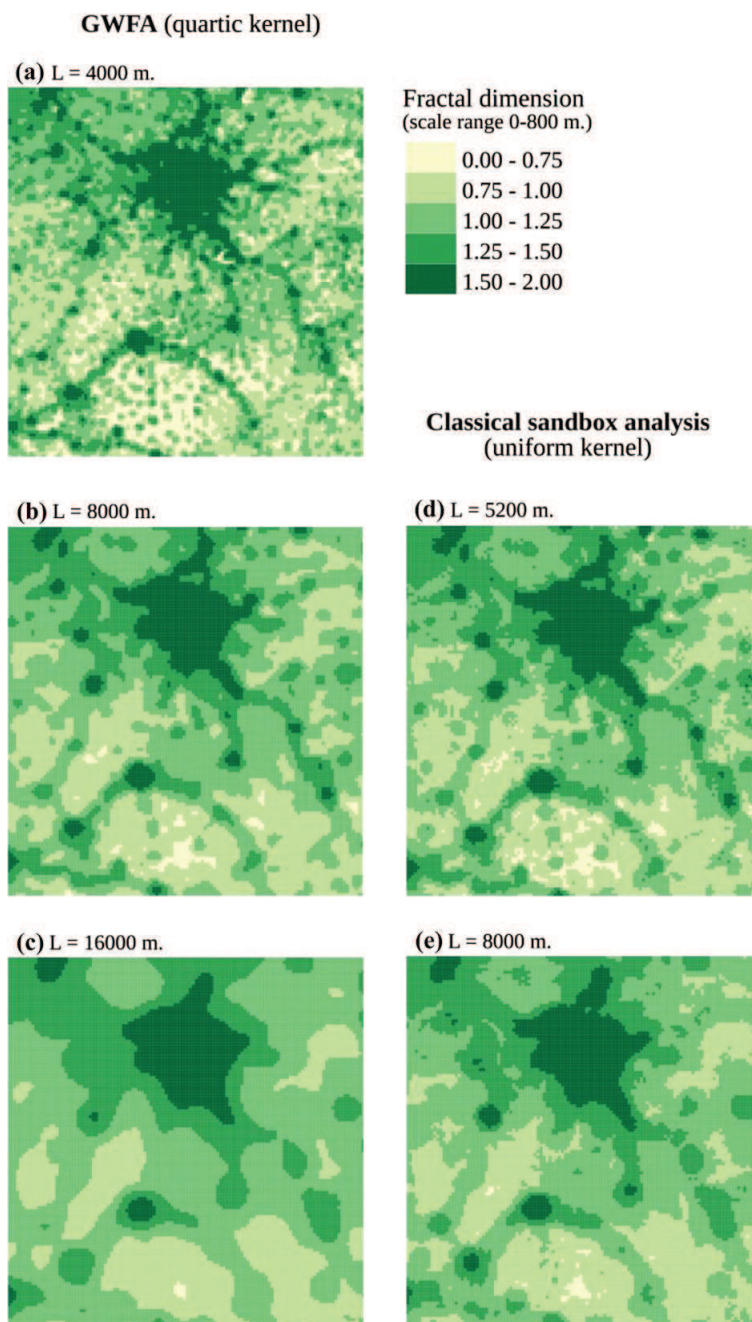


FIGURE 3.10 – Étude de sensibilité aux paramètres de la GWFA ($q = 0$). Source : SÉMÉCURBE, TANNIER et ROUX 2019

le noyau quadratique par un noyau uniforme dans les calculs de la GWFA. Avec un noyau uniforme, tous les points e situé à l'intérieur du noyau ont le même poids quelle que soit leur distance au point d'estimation i . La confrontation directe des résultats pour un même rayon n'a guère de sens car l'inégale répartition des poids entre les deux noyaux induit des biais d'interprétation. Autrement dit, en prenant le même rayon de lissage pour les deux analyses, nous n'obtenons pas exactement la même échelle de restitution des phénomènes. Afin de rendre comparable les cartes, nous avons suivi la suggestion de DIGGLE 1985 qui consiste à choisir des rayons de sorte que les variances des deux noyaux s'égalisent. En appliquant cette règle, un rayon de lissage $L = 8000$ pour un noyau quadratique correspond à un rayon $L = 5200$ pour un noyau uniforme. Selon DIGGLE 1985 et SILVERMAN 1978,

le choix du noyau est moins important que celui du rayon dans les résultats des estimations à noyaux. Cette assertion est vérifiée avec la GWFA. Les deux cartes obtenues avec des rayons comparables sont très ressemblantes (fig. 3.10b et d), bien que la carte avec le noyau uniforme présente des contours moins réguliers.

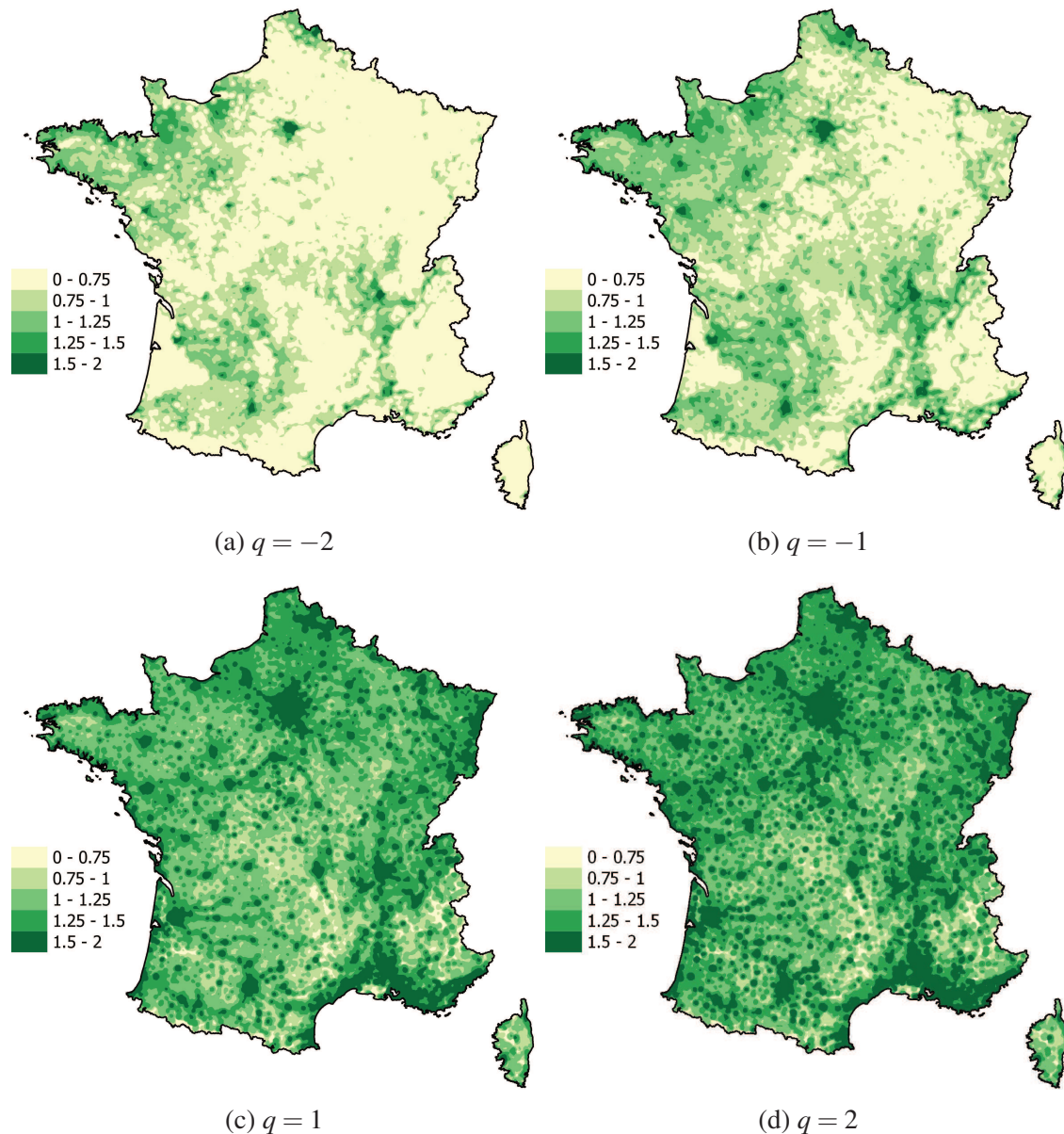


FIGURE 3.11 – Dimensions généralisées de la GWFA pour différentes valeurs de q

3.3.2.2.2 Analyse des autres dimensions généralisées de la GWFA

Dans le cas des cascades multifractales, la théorie prévoit la décroissance des dimensions généralisées au fur et à mesure de la croissance des q . On constate que, pour une majorité du territoire de France métropolitaine, cette propriété n'est pas vérifiée : les dimensions augmentent avec q (fig. 3.11). Nous pouvons émettre au moins deux hypothèses non exclusives pour expliquer ce phénomène :

- Les phénomènes observés bien qu’invariants d’échelle ne sont pas modélisables sous la forme de cascades multifractales car en certains lieux des concentrations locales de bâtiments sont trop fortes pour être engendrées par un processus multiplicatifs;
- L’invariance d’échelle n’est pas vérifiée⁹.

3.3.3 Le spectre de singularité

Le spectre de singularité est une autre manière de caractériser la distribution spatiale d’une mesure multifractale. Deux approches sont généralement utilisées pour le définir (FALCONER 2013). La première appelée *coarse multifractal spectrum* est apparentée à la dimension de *box-counting*. La seconde *fine multifractal spectrum*, beaucoup plus technique, mobilise la dimension de Hausdorff. Elle n’est adaptée qu’à l’étude des objets mathématiques théoriques. En pratique, dans de nombreux cas tels que les cascades multifractales, les deux approches fournissent des résultats identiques.

3.3.3.1 Coarse multifractal spectrum

Lorsqu’un ensemble de points est invariant d’échelle, la dimension de *box-counting* quantifie pour chaque échelle le nombre de carreaux nécessaires pour le recouvrir. Le *coarse multifractal spectrum* étend, en quelque sorte, cette idée aux mesures. Il quantifie pour une échelle donnée le nombre de carreaux C_i de taille δ dont la mesure est approximativement δ^α ¹⁰. Ce nombre est donné par la loi d’échelle :

$$\text{Card}(P(C_i) \approx \delta^\alpha) = \delta^{f(\alpha)} \quad (3.19)$$

$f(\alpha)$ appelé *coarse multifractal spectrum* est une fonction de α , exposant pouvant varier de 0 à 2. L’estimation de ce spectre s’obtient en analysant le comportement de $\text{Card}(M(c) \approx \varepsilon^\alpha)$ à travers les échelles.

$$f(\alpha) = - \lim_{\alpha \rightarrow 0} \frac{\log(\text{Card}(P(C_i) \approx \delta^\alpha))}{\log(\alpha)} \quad (3.20)$$

Le calcul de $\text{Card}(P(C_i) \approx \delta^\alpha)$ est extrêmement difficile, ce qui explique que le *coarse multifractal spectrum* ne soit jamais estimé directement de cette façon. Ce spectre bien qu’ayant vocation à décrire la variabilité et l’irrégularité de la mesure n’est pas local. En effet, un même point peut, selon la taille des grilles, appartenir à des carreaux dont les valeurs de α sont différentes.

3.3.3.2 Fine multifractal spectrum

Le *Fine multifractal spectrum* vise à quantifier la façon dont évolue localement la densité autour d’un point x_0 . On définit d’abord, l’exposant de singularité (appelé également exposant de Hölder) de la sorte (FALCONER 2013) :

$$d_H(x_0) = \lim_{r \rightarrow 0} \frac{\log P(B_{x_0}(r))}{\log(r)} \quad (3.21)$$

⁹. Cette hypothèse se trouve confirmée dans le chapitre 4 dédié à l’analyse trans-échelle.

¹⁰. Plus α est proche de 0 et plus la densité des carreaux correspondants (δ^α/δ^2) est importante.

Remarque

Le *coarse multifractal spectrum* nous renseigne sur la déformation des histogrammes de la répartition de la population dans les carreaux à travers les échelles. Son calcul théorique s'effectue en deux temps :

- Tout d'abord, on calcule pour différentes échelles, un histogramme de la distribution de la population dans les carreaux.
- Ensuite, on recherche les lois d'échelle qui relient les histogrammes entre eux.

Où $B_{x_0}(r)$ est le boule centrée sur x_0 de taille r .

Cet exposant trouve en géographie du peuplement une signification pratique à travers la notion de densité locale :

$$\text{Densité}_{loc}(x_0, r) = \frac{P(B_{x_0}(r))}{\pi r^2} \approx r^{d_H(x_0)-2} \quad (3.22)$$

$d_H(x_0, r)$ quantifie l'évolution de la densité locale :

- quand $\alpha(x_0) = 2$, la densité locale en x_0 est peu sensible au choix du rayon d'analyse et converge asymptotiquement vers une valeur finie. Le phénomène observé est régulier : l'échelle d'analyse ne conditionne pas la mesure de la densité locale.
- quand $\alpha(x_0) < 2$, la densité locale diverge vers l'infini quand le rayon d'analyse devient de plus en plus petit. La vitesse de la divergence est d'autant plus forte que l'exposant $\alpha(x_0)$ est petit. la densité locale est très tributaire de l'échelle d'analyse et ce d'autant plus que l'exposant de singularité est faible.

Remarque

L'exposant de singularité est appelé également dimension fractale locale (FALCONER 2013). FRANKHAUSER 1994 ; ARALDI et FUSCO 2018 ; KARAMI et JOHANSSON 2014 mettent en œuvre cette dimension locale pour étudier des configurations spatiales avec une résolution spatiale très fine.

Dans les cas théoriques, $\alpha(x)$ peut être supérieur à 2. En revanche, dans les cas réels, ce phénomène est impossible à observer (MUZY, BACRY et ARNEODO 1994).

Le spectre multifractal est défini en associant pour chaque valeur de α , la dimension de Hausdorff, de l'ensemble « iso-höldérien » :

$$f(\alpha) = d_H(\{x_0 \text{ tels que } d_H(x_0) = \alpha\}) \quad (3.23)$$

Contrairement, au spectre présenté précédemment, celui-ci est local et quantifie la « taille » des ensembles de points dont la densité locale se comporte de la même manière. Le recours à la dimension de Hausdorff rend en pratique très difficile l'estimation de ce spectre même pour des objets mathématiques théoriques. Pour les cascades multifractales, le *coarse multifractal spectrum* et le *fine multifractal spectrum* sont identiques.

3.3.3.3 Le formalisme multifractal

Le spectre multifractal est lié aux dimensions généralisées. Lorsque la taille des carreaux δ est suffisamment petite, l'équation 3.20 indique que le nombre de carreaux dont la valeur est approximativement égale à δ^α est $\delta^{-f(\alpha)}$ (ARNEODO et al. 1996). Ceci permet d'écrire ce que l'on appelle en statistique les moments d'ordre q sous la forme d'une somme indexée par les valeurs de l'exposant de singularité :

$$\sum P(C_i)^q \approx \sum \delta^\alpha \delta^{-f(\alpha)} = \sum \delta^{\alpha-f(\alpha)} \quad (3.24)$$

La somme $\sum \delta^{\alpha-f(\alpha)}$ est dominée par le terme $\delta^{\min(\alpha-f(\alpha))}$. On introduit les dimensions généralisées en remarquant que $\sum P(C_{i,\delta})^q$ peut s'écrire sous la forme d'une loi d'échelle $\delta^{(q-1)D_q}$. L'équation précédente devient :

$$\delta^{(q-1)D_q} \approx \delta^{\min(\alpha-f(\alpha))} \quad (3.25)$$

En passant aux logarithmes, on obtient le formalisme multifractal qui permet de déduire simplement le spectre multifractal des valeurs des dimensions généralisées à l'aide d'une transformée de Legendre (FALCONER 2013) :

$$(q-1)D_q = \min_\alpha (\alpha - f(\alpha)) \quad (3.26)$$

Remarquons que si $f(\alpha)$ est continûment dérivable, le formalisme multifractal se simplifie de la façon suivante :

$$f(\alpha_q) = q\alpha_q + (1-q)D_q \quad (3.27)$$

$$\text{avec } \alpha_q = \frac{d}{dq}(1-q)D_q \quad (3.28)$$

En théorie, le formalisme multifractal n'est pas vérifié pour toutes les mesures. En pratique, la classe des mesures qui respectent le formalisme multifractal est suffisamment large pour que nous puissions l'utiliser dans nos analyses sans trop de risque. À ce sujet, la remarque de ARNEODO et JAFFARD 2004 dans leur article est rassurante : « Il y a donc un acte de foi à appliquer le formalisme multifractal, même si les résultats mathématiques nous garantissent qu'il faudrait jouer de malchance pour qu'il ne soit pas vérifié, au moins pour la partie croissante du spectre ».

Le formalisme multifractal permet d'approfondir les propriétés du spectre. Tout d'abord, la valeur de $f(\alpha)$ est contrainte par la valeur de α :

$$0 \leq f(\alpha) \leq \alpha \quad (3.29)$$

La partie de gauche de l'inégalité est triviale et liée au fait que $f(\alpha)$ est une dimension. La partie droite est plus intéressante car elle implique que la force des singularités contraint leur répartition : une singularité forte ne peut pas se retrouver simultanément dans de nombreux lieux. Pour $q = 1$, on a exactement $f(\alpha) = \alpha$.

Le maximum du spectre est atteint en α_0 , c'est-à-dire pour $q = 0$. α_0 caractérise les densités locales les plus présentes au sein de la mesure. Plus la valeur de α_0 est élevée et plus les densités locales divergent lentement en moyenne. $f(\alpha_0)$ est égale à la dimension du support de la mesure.

Dans le cas d'une mesure monofractale, le spectre est réduit à un unique point : la répartition des poids est entièrement déterminée par l'organisation spatiale du support de la mesure. Les densités locales ont des comportements analogues en tout point du

support de la mesure. A contrario, les densités locales d'une mesure multifractale sont très hétérogènes. Plus la largeur du spectre ($\max(\alpha) - \min(\alpha)$) est importante, plus le caractère multifractal est affirmé et plus les densités locales varient au sein du support de la mesure.

Les spectres théoriques des trois cascades multifractales présentées précédemment (fig. 3.12) ont le même maximum $\log(5)/\log(3)$ correspondant à la dimension de leur support, un tapis de Sierpinski. La première cascade a le spectre le plus large. Sa multifractalité est la plus importante des trois cascades. A mesure que les échelles sont de plus en plus fines, de moins en moins de carreaux concentrent la distribution de probabilité. Avec le spectre le plus étroit, la troisième cascade a la multifractalité la moins affirmée : ses densités locales sont plus homogènes que celles des deux autres cascades. Enfin, la deuxième cascade présente une multifractalité intermédiaire.

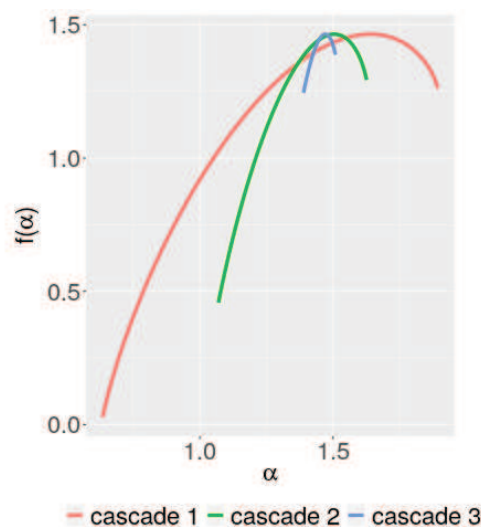


FIGURE 3.12 – Spectres multifractals des cascades multifractales de la figure 3.8

3.3.4 Quantification du *MAUP* au moyen de l'analyse multifractale : application à l'étude de la population de France métropolitaine

L'application de la géométrie multifractale aux répartitions de population permet dans un certain sens de résoudre le problème de la dépendance des résultats statistiques à l'échelle d'observation (Modifiable Areal Unit Problem). Le *MAUP* n'est dès lors plus une limite inhérente à la connaissance des phénomènes géographiques mais devient un outil pour décrire leur organisation trans-échelle.

Le *MAUP* se décompose en deux sous-problèmes interdépendants : l'effet de zonage et l'effet d'échelle. Le premier est lié à la forme des unités spatiales élémentaires tandis que le deuxième désigne la dépendance des analyses géographiques et statistiques à l'échelle d'observation. L'objectif de cette section est de quantifier les effets du *MAUP* sur des mesures telles que les répartitions de population, à l'aide de l'analyse multifractale. Cette opération nécessite de restreindre en premier lieu les découpages territoriaux admissibles. Dans le cas d'un phénomène naturel, les échelles d'observations sont limitées et les spectres de singularité et de dimension sont estimés sur des plages de tailles de grilles où la mesure respecte des propriétés d'invariance d'échelles. Sur ces plages de tailles, l'analyse

multifractale nous indique que les observations de la mesure effectuées à différentes échelles sont reliées entre elles selon des modalités déterminées par les caractéristiques de son spectre.

L'effet d'échelle impacte notre perception des distributions analysées ainsi que les indicateurs statistiques calculées sur celles-ci. L'interprétation géographique des deux spectres précédents qui se confondent pour les distributions empiriques permet d'aborder dans un cadre unifié ces deux facettes du *MAUP*. Le *fine multifractal spectrum*, indique la façon dont les densités locales varient à travers les échelles (SÉMÉCURBE, TANNIER et ROUX 2016)¹¹. Le *coarse multifractal spectrum* décrit, quant à lui, les effets du *MAUP* sur les indicateurs statistiques. La valeur d'une statistique pour une échelle donnée est entièrement déterminée par sa distribution de probabilité, autrement dit son histogramme. Comprendre les effets du *MAUP* sur les statistiques revient donc à comprendre la façon dont les histogrammes, décrivant la répartition de la population dans les carreaux, se déforment à travers les échelles.

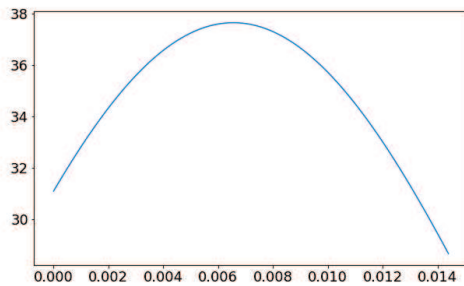
Quand le spectre de singularité est réduit au point de coordonnées (2,2), les effets du *MAUP* sont réduits. Les densités locales évoluent très lentement à la fois dans les échelles et dans l'espace, les poids de la mesure sont répartis uniformément dans l'espace. La mesure possède une densité locale finie en tout point et la modélisation du phénomène étudié sous la forme d'une fonction est plus appropriée dans ce cas (cf section 3.4). Les statistiques ne sont quasiment pas influencées par les effets d'échelle : on retrouve le même histogramme quelle que soit l'échelle considérée.

Quand la mesure est monofractale, son spectre de singularité est réduit à un seul point. Une seule loi d'échelle, ayant pour exposant la dimension de *box-counting* du support de la mesure, caractérise la répartition de la population à travers les échelles. Les différences de densités locales de population s'expliquent uniquement par l'inégale répartition des bâtiments d'habitation dans l'espace (support de la mesure). Tous les histogrammes se superposent lorsque l'on normalise les axes des abscisses et des ordonnées par une loi de puissance (d'exposant la dimension de *box-counting*)¹² (fig. 3.13 a, c, e). Ceci implique que tous les indicateurs statistiques calculés sur la répartition de la population suivent la même loi d'échelle de paramètre la dimension fractale du support. La connaissance de la valeur d'un indicateur statistique donné et de la dimension fractale du support permet d'extrapoler la valeur de l'indicateur à n'importe quelle échelle.

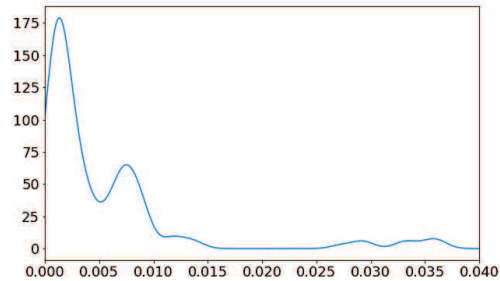
Un spectre de singularité plus large indique que la mesure étudiée est multifractale. Plusieurs comportements d'invariance d'échelle s'entremêlent. Les densités locales sont très variables dans l'espace et les échelles, et le *MAUP* résulte à la fois de l'inégale répartition des bâtiments d'habitation et du fait que chaque bâtiment soit plus ou moins peuplé. Dans ce cas, une infinité de lois d'échelle doivent être mobilisées pour décrire les transformations des histogrammes (fig. 3.13 b, d, f). Il en résulte qu'il est difficile de prévoir l'évolution des indicateurs statistiques à travers les échelles.

11. L'article est en annexe A.

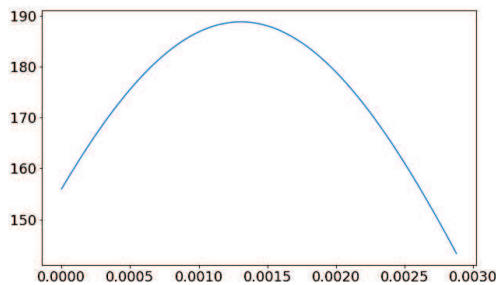
12. Nous avons volontairement choisi de prendre des tailles de carreaux qui ne correspondent à la partition naturelle du générateur pour éviter d'obtenir une distribution triviale où tous les carreaux peuplés ont exactement la même densité.



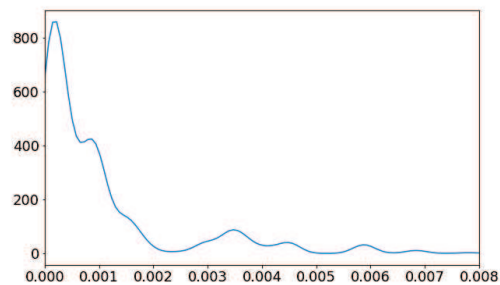
(a) à l'échelle 135



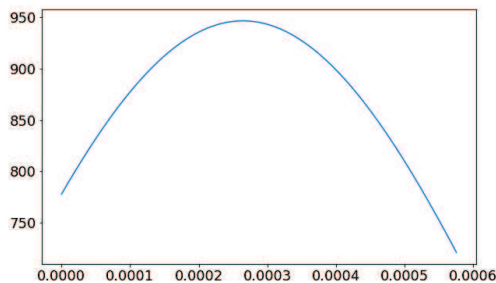
(b) à l'échelle 135



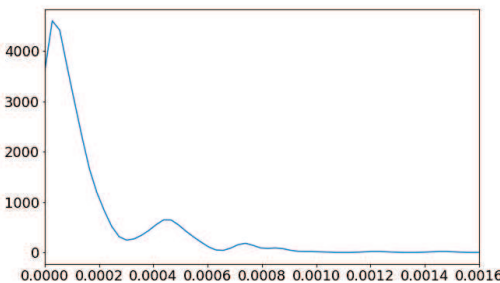
(c) à l'échelle 45



(d) à l'échelle 45



(e) à l'échelle 15



(f) à l'échelle 15

FIGURE 3.13 – Déformation des histogrammes à travers les échelles. Chaque histogramme décrit pour une échelle donnée la répartition des carreaux selon la population qu'ils contiennent. Dans la colonne de gauche, la distribution de population est issue d'une cascade monofractale ayant pour support un tapis de Sierpinski dont le générateur comporte 5 éléments non nuls. Dans la colonne de droite, elle est issue d'une cascade multifractale dont le générateur comporte également 5 éléments, mais cette fois-ci les poids des carreaux ne sont pas uniformes.

3.3.4.1 Données utilisées

Nous allons à présent appliquer l'interprétation du *MAUP* conférée par l'analyse multifractale à la population de France Métropolitaine. Pour cette analyse, nous avons mobilisé les données de population carroyée de l'Institut national de la statistique et des études économiques (Insee) de 2013. Ces données sont élaborées à partir des informations contenues dans la source « Revenus Fiscaux Localisés » 2010. Les « Revenus Fiscaux Localisés » sont établis à partir du fichier des déclarations de revenus des personnes physiques et du fichier de la taxe d'habitation fournis par la Direction générale des Finances publiques (DGFIP) et disponibles sur insee.fr. Chaque personne de cette source est géolocalisée à l'aide de l'identifiant cadastral de son lieu de déclaration. Les données carroyées sont obtenues en agrégeant dans une grille carroyée de 200 m de côté les personnes géolocalisées. Moins de 1% de la population de la source RFL est géolocalisé avec une précision supérieure à 200m. Cette part de population mal géolocalisée est suffisamment bien répartie dans l'espace métropolitain pour qu'elle n'altère pas la qualité de nos analyses.

3.3.4.2 Estimation des indicateurs multifractals

CHHABRA et JENSEN 1989 ont démontré que pour des mesures suffisamment régulières, le spectre multifractal pouvait être obtenu directement à l'aide des formules suivantes :

$$\alpha(q) = \lim_{\delta \rightarrow 0} \frac{\sum P_q(\delta, i) \log(P(C_i))}{-\log(\delta)} \quad (3.30)$$

et

$$f(\alpha(q)) = \lim_{\delta \rightarrow 0} \frac{\sum P_q(\delta, i) \log(P_q(\delta, i))}{-\log(\delta)} \quad (3.31)$$

avec

$$P_q(\delta, i) = \frac{P(C_i)^q}{\sum p_i^q} \quad (3.32)$$

La distribution de probabilité P_q est appelée distribution escorte en théorie de l'information et joue le rôle d'un microscope qui permet d'explorer les parties plus ou moins régulières de la mesure CHHABRA et JENSEN 1989. Pour les $q > 1$, la distribution P_q accentue les zones de forte densité. Pour $q < 1$, elle accentue les zones de faible densité. Et pour $q = 1$, elle réplique la mesure originale. Le numérateur de $\alpha(q)$ et $f(\alpha)$ sont deux entropies.

$\alpha(q)$ et $f(\alpha(q))$ correspondent à la pente de la courbe représentant leur numérateur respectif en fonction de $\log(\delta)$ ¹³.

3.3.4.3 Choix des paramètres de l'analyse

3.3.4.3.1 Les grilles optimales

Dans la théorie, l'origine des grilles utilisées n'affecte pas l'estimation des paramètres multifractals. En pratique, comme le souligne APPLEBY 1996, le choix de l'origine peut

13. La plage d'échelles d'estimation doit être choisie en conformité avec celle des dimensions généralisées et des fonctions d'échelles pour garantir la cohérence des analyses.

profondément modifier les estimations des indicateurs à une échelle donnée et donc les estimations des paramètres multifractals. Pour une grille de taille ($k \times 200m$), nous pouvons choisir k^2 origines correspondant à autant de façons différentes d'agréger les cellules selon leur latitude et longitude. APPLEBY 1996 propose de retenir pour une taille de grille et pour un paramètre donnés, la grille dont l'entropie est minimale. Cette sélection est effectuée pour tous les couples de tailles de grille et de paramètres. D'un point de vue méthodologique, ce choix technique est en accord avec la nature de l'entropie. En effet, comme l'entropie quantifie le nombre d'informations minimales que l'on doit transmettre pour lever toute ambiguïté sur le résultat d'une variable aléatoire, quand on choisit la grille d'entropie minimale, on conserve l'information la plus pertinente pour décrire l'organisation de la répartition de la population observée avec le paramètre q .

3.3.4.3.2 Échelles d'analyse et tailles des unités spatiales considérées

La comparaison des formes de peuplement en France nécessite de subdiviser le territoire métropolitain en unités spatiales régulières. La taille de ces unités spatiales contraint le choix de la plage d'estimation des dimensions généralisées, c'est pourquoi nous n'avons pas pu traiter séparément ces deux aspects. Ainsi, nous avons procédé à de nombreux essais croisant différentes tailles d'unités spatiales et de plages d'échelles. Néanmoins, par souci de clarté, nous présentons ici successivement le choix de la plage d'échelles d'estimation puis celui de la taille des unités spatiales considérées.

La comparaison spatiale des indicateurs multifractals requiert qu'une seule plage d'échelles ne soit utilisée pour l'ensemble des unités spatiales. Cette propriété contraint fortement notre analyse. Les formes de peuplement au sein des unités spatiales peuvent être très différentes et donc les meilleures plages d'échelle pour restituer des propriétés d'invariance d'échelle peuvent être très variables entre les unités spatiales. Le choix d'une plage d'échelles commune résulte d'un arbitrage entre les différentes plages d'échelles optimales. Nous avons observé les phénomènes suivants :

- À une résolution très fine (correspondant à des tailles de carreaux inférieures à 800 m), les données de population carroyées ne sont pas suffisamment précises pour autoriser une description satisfaisante de la population. À ce niveau d'échelle, l'analyse fractale du bâti apporte une information plus riche sur l'organisation de l'espace anthropisé.
- À une résolution grossière (correspondant à des tailles de carreaux supérieures à 6400 m), l'invariance d'échelle est dégénérée. À ce niveau d'échelle, presque chaque cellule est peuplée. La dimension du support D_0 est proche de deux.
- À une résolution intermédiaire, correspondant à des tailles de carreaux entre 800 et 6400 m, l'analyse multifractale nous renseigne bien sur l'organisation des formes de peuplement.

Les unités spatiales considérées doivent avoir une taille suffisamment grande pour que l'estimation des paramètres multifractals ne soit pas biaisée. Sinon, les indicateurs calculés à la résolution spatiale la plus grossière reflètent uniquement des effets de bord. Considérant la résolution maximale de 6400 m, nous avons choisi de prendre des unités spatiales de 25 km de côté (*fig. 3.14*). A cette échelle, chaque unité spatiale se subdivise en une grille de 4 colonnes et de 4 lignes, autrement dit 16 informations élémentaires contribuent aux calculs des fonctions d'échelles.

Le territoire métropolitain est recouvert de 992 unités spatiales de 25 km de côté. En moyenne, 9 unités spatiales intersectent les aires urbaines. La taille moyenne des 10 aires urbaines les plus peuplées est de 4900 km², correspondant approximativement à

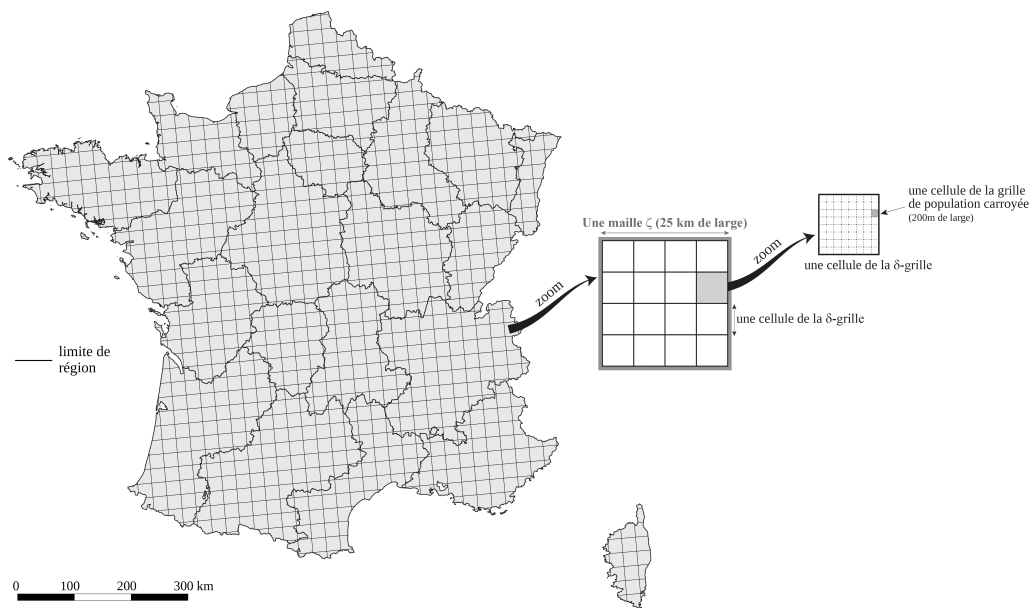


FIGURE 3.14 – Emboîtement hiérarchique des unités spatiales d’intérêt choisit en vue de l’analyse multifractale de la distribution de la population en France métropolitaine

un diamètre de 79 km. Ce partitionnement des espaces métropolitains nous permet de distinguer le pôle urbain de l’espace périurbain. De plus, il est possible avec ces unités spatiales de comparer l’organisation de la population des espaces périurbains avec celle des espaces limitrophes. Les aires urbaines peuvent être ainsi analysées dans leur contexte géographique environnant.

Concernant le paramètre q , en principe, il peut être choisi dans \mathbb{R} . Cependant, lorsqu’il est négatif, l’estimation des paramètres multifractals est très instable : l’analyse des espaces de très faible densité est très sensible au choix de l’origine des grilles utilisées. Les dimensions généralisées sont également instables lorsque q est très grand. En effet, dans ce cas, uniquement les cellules les plus peuplées sont prises en compte. Pour éviter ces problèmes d’estimation, nous avons choisi de restreindre notre analyse aux q compris dans : $\{0, 0.5, 1, 1.5, 2\}$

3.3.4.4 Résultats

Les indicateurs issus de l’analyse multifractale ont un sens uniquement si l’invariance d’échelle est vérifiée. Comme nous pouvons le voir sur le tableau présenté en figure 3.3, les coefficients de détermination R^2 sont satisfaisants ; plus de 90% sont supérieurs à 0,96. Seule la qualité des estimations de α_2 et $f(\alpha_2)$ est légèrement moins bonne.

Quantile	D_0	D_1	D_2	α_0	α_1	α_2	$f(\alpha_0)$	$f(\alpha_1)$	$f(\alpha_{1.5})$	$f(\alpha_2)$
minimum	0.94	0.90	0.81	0.89	0.90	0.53	0.94	0.90	0.63	0.00
10%	0.98	0.99	0.96	0.96	0.99	0.88	0.98	0.99	0.93	0.64
25%	0.99	1.00	0.98	0.98	1.00	0.93	0.99	1.00	0.96	0.80
50%	1.00	1.00	0.99	0.99	1.00	0.97	1.00	1.00	0.98	0.91
75%	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	0.99	0.97

TABLE 3.3 – Estimation de la qualité des analyses à l’aide des coefficients de détermination

3.3.4.4.1 Les cartes des dimensions généralisées

Les figures 3.15.b.c.d représentent les cartes des dimensions généralisées pour $q = 0, 1, 2$. Chaque carte reflète une structure spatiale spécifique qui se distingue de celle observée à l'aide des densités (fig. 3.18.a). La qualité des données étant comparable dans l'espace, les structures observées ne sont pas des artefacts liés à la source utilisée mais témoignent de l'existence de différentes formes de peuplement au sein de l'espace français métropolitain.

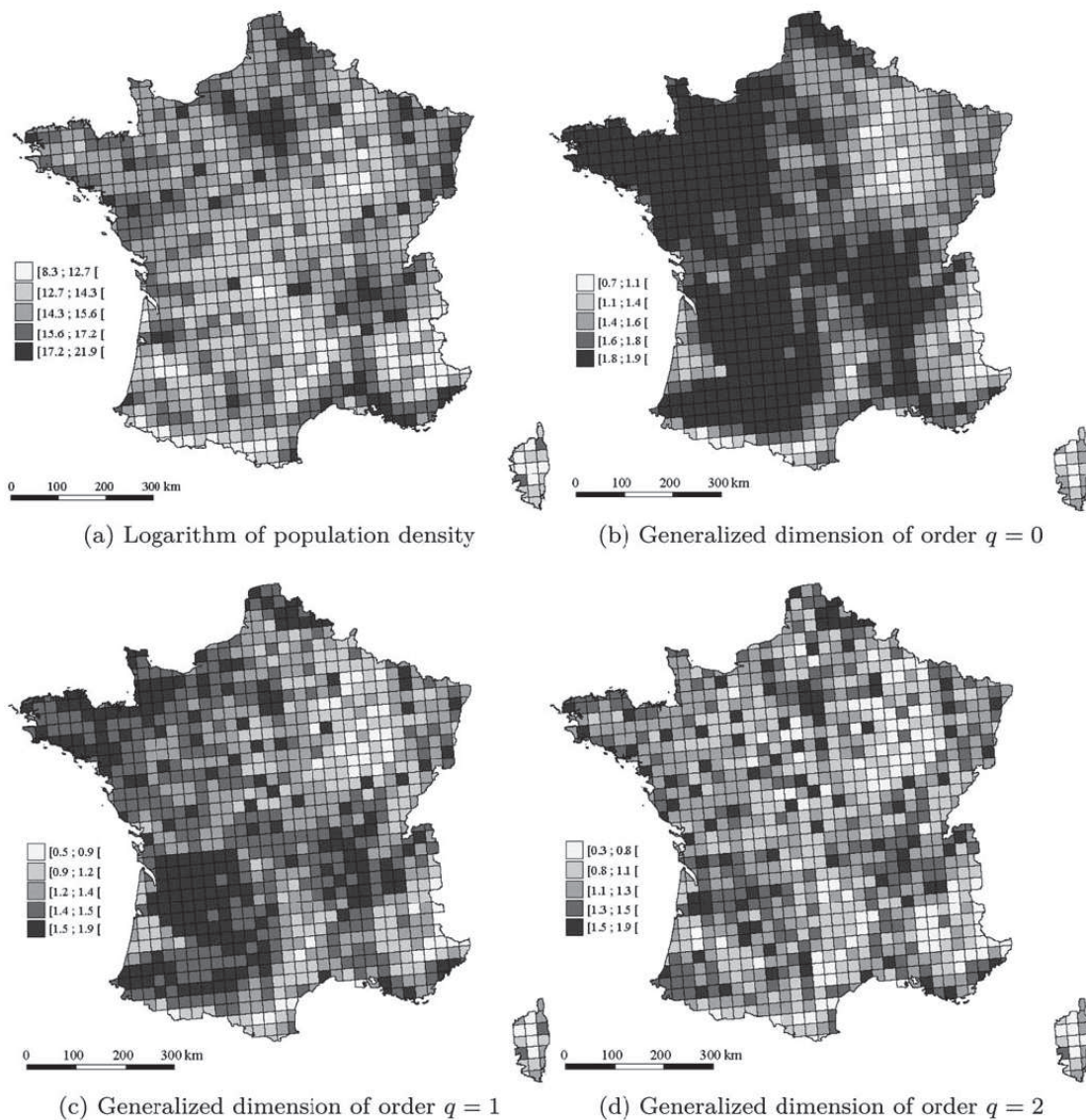


FIGURE 3.15 – Densité de population et dimensions généralisées. Source : SÉMÉCURBE, TANNIER et ROUX 2016.

La carte représentant les dimensions généralisées D_0 découpe l'espace métropolitain en de larges zones de même dimension. Les dimensions généralisées sont élevées dans l'ouest et le sud-ouest, la vallée du Rhône et dans les pôles urbains des grandes aires urbaines. Les territoires de l'Est de la France métropolitaine (Champagne, Ardennes, Lorraine et les Vosges), ainsi que les zones montagneuses et forestières ont des dimensions généralisées faibles. Le fait que les zones caractérisées par des dimensions élevées soient aussi bien

rurales qu'urbaines, montre que la dimension généralisée D_0 est indifférente aux variations de densité. Elle est sensible uniquement, comme dit précédemment, à la présence ou absence de la population dans l'espace.

Les zones urbaines denses sont clairement identifiées sur la carte représentant les dimensions généralisées d'ordre 2. Elles ressemblent à des îles de fortes dimensions entourées par des unités spatiales de faibles dimensions. L'analyse multifractale pour $q = 2$, se concentre principalement sur les cellules de forte densité. Seuls les espaces urbains ont suffisamment de zones denses proches les unes des autres pour que leurs dimensions soient importantes. Dans les zones rurales, l'éloignement entre elles des zones densément peuplées limitent les valeurs des dimensions. La carte représentant les dimensions généralisées d'ordre 1 montre des configurations spatiales intermédiaires entre les cartes D_0 et de D_2 , ce qui illustre la continuité des dimensions généralisées selon le paramètre q .

La décroissance des dimensions généralisées, en particulier dans les espaces périurbains et ruraux, témoigne du caractère multifractal des formes de peuplement que l'étude des spectres permet d'approfondir.

3.3.4.4.2 Typologie des formes de peuplement à partir des spectres multifractals

Pour réaliser notre typologie des formes de peuplement nous avons utilisé une Classification Ascendante Hiérarchique. La CAH requiert que l'on définisse une distance entre les spectres multifractals. Les spectres n'ayant pas la même amplitude en abscisse, nous ne pouvons pas calculer la distance usuelle :

$$d(\zeta_1, \zeta_2) = \left(\sum_{\alpha} (f_1(\alpha) - f_2(\alpha))^2 \right)^{1/2} \quad (3.33)$$

où ζ_1 et ζ_2 seraient deux unités spatiales.

Après plusieurs essais infructueux, nous avons choisi d'utiliser la distance de Hausdorff. Celle-ci respecte les axiomes nécessaires pour définir une distance, en particulier l'inégalité triangulaire qui est indispensable pour pouvoir mettre en œuvre une CAH. Elle est basée sur la notion de voisinage de taille λ . Pour un sous-ensemble A de points de \mathbb{R}^2 , le voisinage de A de taille λ est défini de la manière suivante :

$$A_{\lambda} = \{x \in \mathbb{R}^2 : |x - a| < \lambda \text{ pour un point } a \in A\} \quad (3.34)$$

A_{λ} est l'ensemble des points de \mathbb{R}^2 dont la distance avec au moins un des points de A est inférieure à λ (soit l'ensemble des points contenus dans la zone tampon de taille λ entourant A). La distance de Hausdorff entre les deux spectres $f_1(\alpha)$ et $f_2(\alpha)$ est la plus petite valeur de λ tel que le λ voisinage de $f_1(\alpha)$ contienne $f_2(\alpha)$ et vice versa.

Une fois l'ensemble des distances calculées entre les spectres multifractals des 992 unités spatiales, nous avons calculé la CAH en utilisant l'algorithme de Ward. Cet algorithme minimise la variance intra-classe et maximise la variance inter-classe. Le dendrogramme sert à définir le nombre de classes. La taille des branches de l'arbre quantifie la distance entre les classes. Le choix du nombre de classes résulte d'un arbitrage entre la qualité statistique et la finesse des détails que nous souhaitons conserver. Le nombre de classes est obtenu par coupure de l'arbre à un niveau donné. Le dendrogramme en figure 3.16 suggère un découpage en 5 classes :

- Classe 1 : unités spatiales dans lesquelles la population est modérément concentrée localement ;

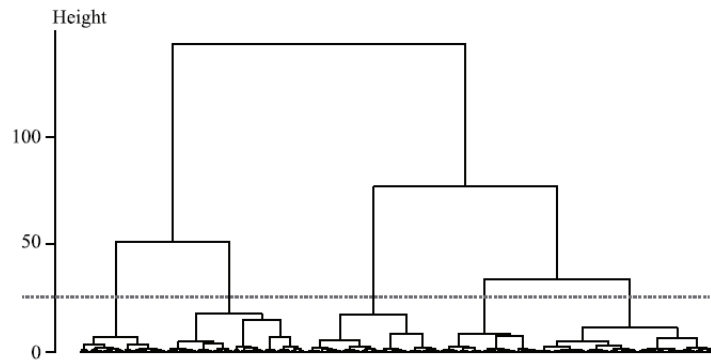


FIGURE 3.16 – Dendrogramme de la CAH des spectres multifractals de la distribution de la population en France métropolitaine. Source : SÉMÉCURBE, TANNIER et ROUX 2016

- Classe 2 : unités spatiales dans lesquelles la population est très concentrée localement ;
- Classe 3 : unités spatiales dans lesquelles la population est dispersée ;
- Classe 4 : unités spatiales dans lesquelles la population est répartie uniformément ;
- Classe 5 : unités spatiales présentant des configurations intermédiaires.

La carte de typologie des spectres multifractals synthétise les trois cartes des dimensions généralisées. Les unités spatiales caractérisées par de faibles dimensions généralisées apparaissent clairement sur la carte des spectres (classes 1 et 2). De même, les centres urbains qui apparaissent clairement sur la carte des dimensions généralisées D_2 , se retrouvent également sur la carte des spectres (classe 4). Enfin, les classes 3 et 5 de la typologie correspondent aux unités spatiales non urbaines ayant des dimensions généralisées D_0 élevées et des dimensions D_1 et D_2 élevées ou intermédiaires.

Pour chaque classe, nous avons calculé un spectre multifractal moyen :

$$\alpha_q(\text{Classe}) = \frac{\sum_{i \in \text{Classe}} \alpha_q(i)}{|\text{Classe}|} \text{ et } f(\alpha_q(\text{Classe})) = \frac{\sum_{i \in \text{Classe}} f(\alpha_q(i))}{|\text{Classe}|} \quad (3.35)$$

Les spectres multifractals moyens des unités spatiales dans lesquelles la population est concentrée (classes 1 et 2) sont caractérisés par des petites valeurs de α (fig. 3.17). Cette propriété mathématique nous indique que ces unités spatiales contiennent de nombreuses singularités spatiales. D'un point de vue géographique, ces singularités traduisent l'existence de villages ou de petites villes aux contours clairement délimités. En moyenne, le spectre des unités de la classe 2 est décalé en bas à gauche par rapport à celui de la classe 1, ce qui tend à montrer que la population de la classe 2 est plus concentrée à travers les échelles que celle de la classe 1. Sur la figure 3.18, nous constatons que la classe 2 est principalement localisée dans les massifs forestiers et montagneux (fig 3.19). La classe 1 borde les unités spatiales de la classe 2. Les caractéristiques de la classe 1 ainsi que sa répartition dans l'espace suggère que celle-ci représente une version atténuée de la classe 2.

Le spectre multifractal moyen des unités spatiales dans lesquelles la population est répartie uniformément (classes 3 et 4) est caractérisé par des valeurs de α et $f(\alpha)$ élevées. Pour $q = 0$, la valeur de α_0 est approximativement de 2 et la valeur de $f(\alpha_0)$ de 1.7. Le spectre multifractal moyen de la classe 4 est plus étroit que celui de la classe 3. Cette propriété indique que les distributions de population de la classe 4 sont moins multifractales que celles de la classe 3. La localisation spatiale des unités appartenant à

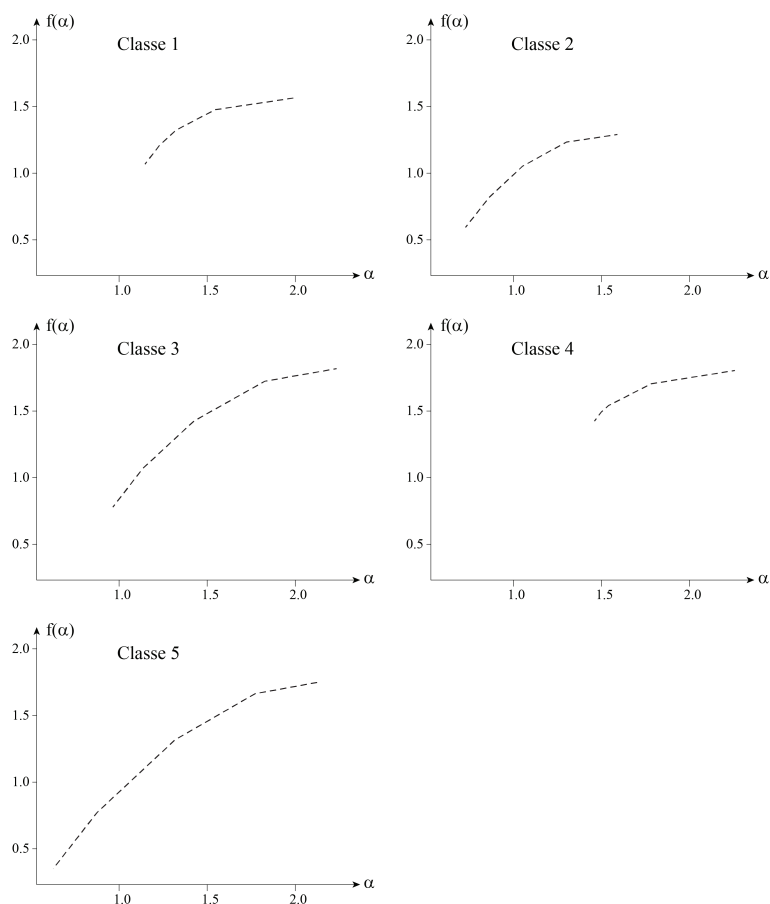


FIGURE 3.17 – Spectres multifractals moyens des différentes classes de la typologie obtenue par CAH

la classe 4 correspond à l'emplacement des principaux centres urbains (fig 3.19). Nous retrouvons le fait que les centres urbains se caractérisent par des répartitions denses et relativement uniformes de population sur de larges étendues spatiales. A contrario, les densités de population de la classe 3 varient davantage selon la résolution spatiale utilisée. Les unités spatiales appartenant à la classe 3 sont situées à la périphérie des grandes villes aussi bien que dans les espaces ruraux de l'Ouest de la France.

Le spectre multifractal moyen de la classe 5 recouvre l'ensemble des spectres précédemment discutés. La classe 5 combine les propriétés des classes 2 et 3. Cette position intermédiaire caractérise également la situation spatiale des unités de cette classe qui se positionnent à l'interface des classes 2 et 3. La classe 5 correspond à des villages entourés par des zones de faible densité de population dans lesquelles la population est dispersée. La plupart des unités spatiales caractérisées par des dimensions généralisées D_0 élevées et des dimensions généralisées D_2 faibles appartiennent à cette classe.

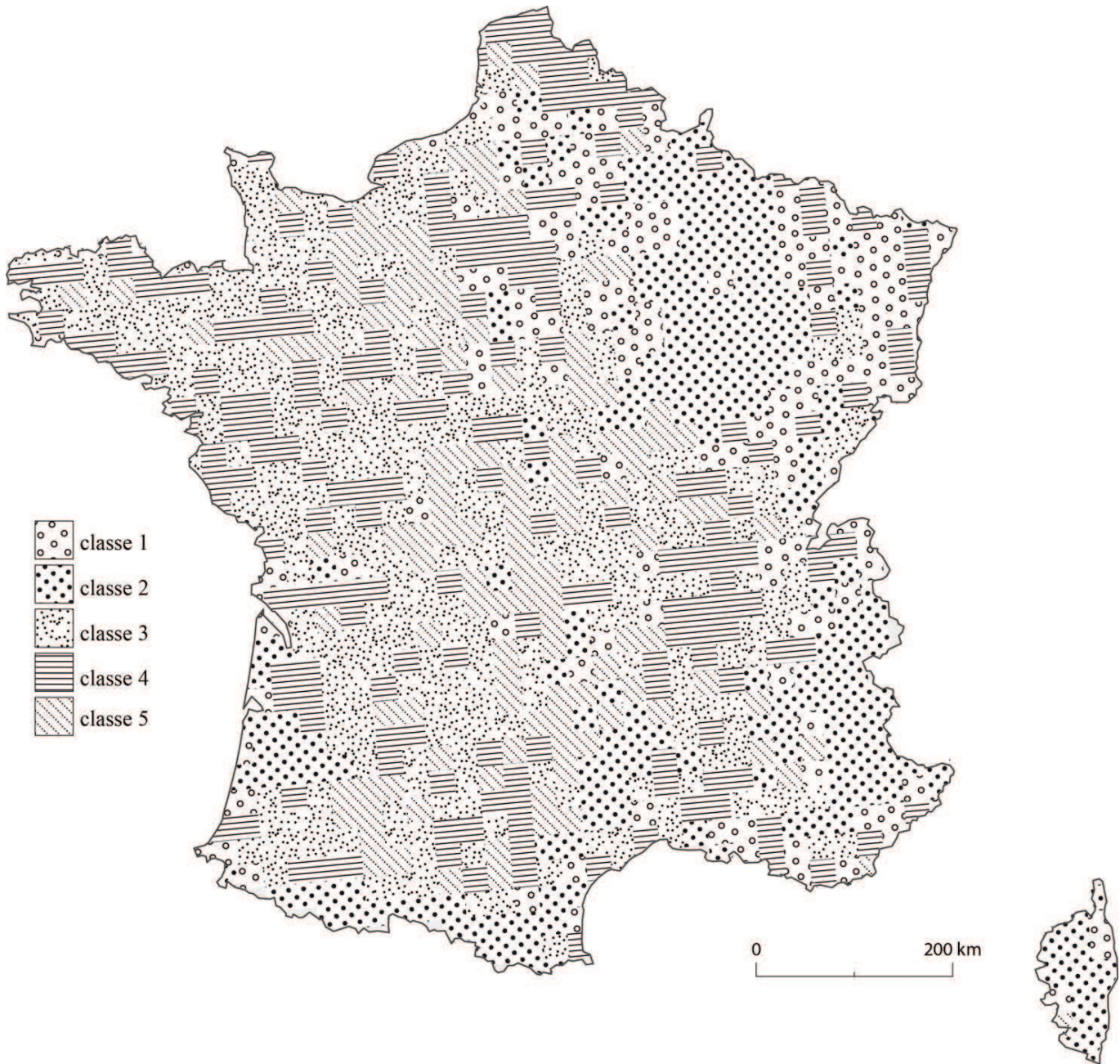


FIGURE 3.18 – Carte résultant de la classification des spectres multifractals de la distribution de la population en France. 992 unités spatiales de 25 km de large. D’après SÉMÉCURBE, TANNIER et ROUX 2016

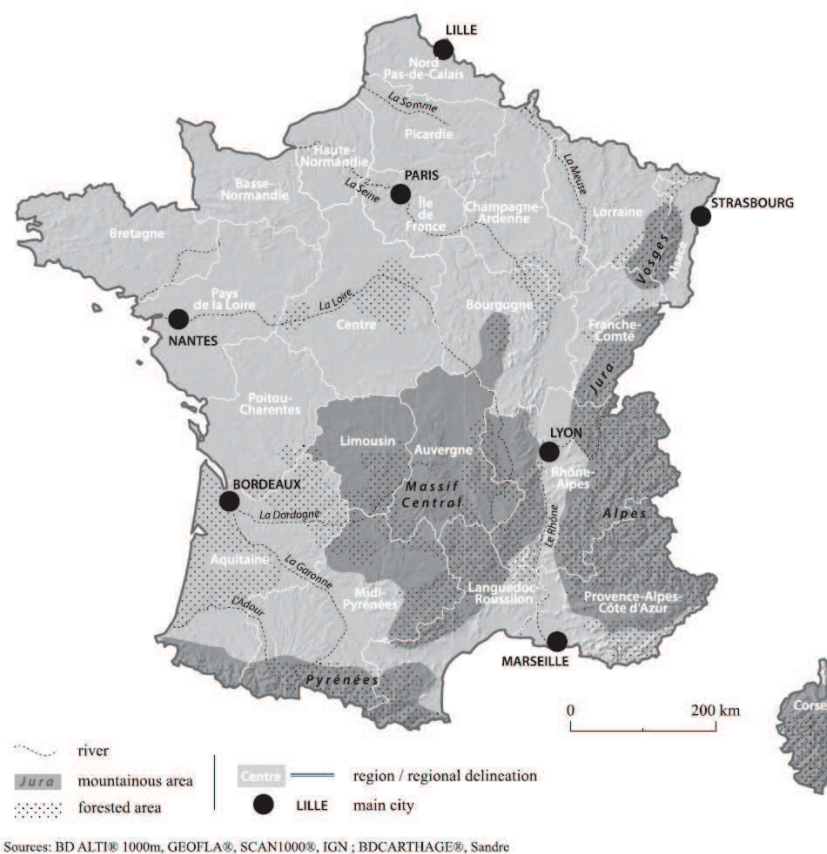


FIGURE 3.19 – Massifs montagneux et forestiers et principales villes de France métropole

Remarque

Les effets du *MAUP* sur la mesure de l'organisation du peuplement dépendent du maximum du spectre (qui décrit le support de la mesure) et de son étendue. Avec un spectre proche de 2, les grandes villes (classe 4) ressemblent davantage pour nos échelles d'observation à des structures lisses et uniformes qu'à des fractales. L'impact du *MAUP* pour cette classe est relativement limité. Les classe 3 et 5 avec un spectre très large contiennent une grande variété de formes à travers les échelles : la multifractalité est importante. Avec un maximum du spectre proche de 2, le support de leur mesure est régulier : l'emprise spatiale des carreaux contenant de la population fluctue peu selon les échelles d'observation. En revanche, les densités varient fortement dans l'espace et entre les échelles d'observation. Avec des spectres plus resserrés les classes 1 et 2 sont moins multifractales. Le *MAUP* est principalement induit par la forme du support de la mesure. Avec un maximum relativement petit, le nombre de carreaux qui contient de la population évolue fortement à travers les échelles.

3.4 Analyse multifractale d'une fonction (*i.e.* une image)

L'analyse multifractale d'une mesure ne restitue que la partie gauche du spectre ($q > 0$) qui caractérise les singularités les plus fortes. L'analyse multifractale d'une fonction à l'aide d'ondelettes a été développée pour pallier cette limite. Lorsque le phénomène analysé est suffisamment « régulier » tout le spectre peut être décrit à l'aide de celle-ci.

La première méthode d'analyse multifractale d'une fonction développée, dite méthode des Maxima du Module de la Transformée en Ondelettes (MMTO) est adaptée à l'étude des fonctions d'une variable. Elle a été mise en œuvre dans des domaines très variés tels que : l'étude du rythme cardiaque (IVANOV et al. 1999), la description des séquences d'ADN (ARNEODO, BACRY et al. 1995), la caractérisation du comportement trans-échelle des actifs financiers (MUZY, SORNETTE et al. 2000) ou encore l'analyse des variations du champs magnétique du soleil (KESTENER et al. 2010).

Une nouvelle méthode d'analyse, dite en ondelettes dominantes, beaucoup plus simple à mettre en œuvre a permis d'étendre l'analyse multifractale aux fonctions à plusieurs variables. Elle est souvent utilisée pour caractériser l'organisation trans-échelle contenue dans des images : classification des textures osseuses (ZOU et al. 2016) et analyse de la régularité des textures des peintures de Van Gogh (ABRY, WENDT et JAFFARD 2013). Nous l'avons appliquée à des photographies aériennes (orthophotographie) afin d'analyser l'agencement des objets contenus dans l'espace quelle que soit leur nature (routes, bâtiments, forêts...).

3.4.1 De l'analyse multifractale d'une mesure à celle d'une fonction

Les analyses multifractales présentées dans la section précédente sont uniquement pertinentes pour les mesures dont les valeurs de α sont inférieures à 2, dites mesures « singulières » (ARNEODO et al. 1996). Leurs densités locales divergent vers l'infini en tout point de leur support. Pour calculer les densités locales avec les méthodes de type *box-counting*, on utilise des carreaux ou boîtes. Le spectre multifractal décrit la vitesse de divergence des valeurs de ces « boîtes » à travers les échelles.

Les mesures caractérisées par des α supérieurs à 2 sont composées de points dont la densité locale est bornée, voire converge vers une valeur précise finie. Pour ces mesures, les valeurs des boîtes sont relativement constantes à travers les échelles, ce qui limite la possibilité d'en extraire des lois d'échelle. C'est pourquoi les objets présentant de telles mesures doivent être analysés sous la forme de fonctions.

L'analyse multifractale d'une fonction consiste non pas à comparer l'évolution moyenne des valeurs des boîtes à travers les échelles, mais à comparer l'évolution des différences des valeurs des boîtes à travers les échelles. En quelque sorte, l'analyse multifractale d'une fonction considère les dérivées des valeurs des boîtes tandis que l'analyse multifractale d'une mesure considère leur valeur brute. Cette extension de l'analyse multifractale à l'étude des fonctions est utilisée depuis le début des années 1990 pour caractériser des images représentant des textures.

Encadré

Les fonctions et les mesures sont deux types de relations qui associent à des objets une valeur numérique. Les objets en entrée des fonctions sont les points de l'espace et les objets des mesures sont tous les ensembles de points. Par exemple, la relation qui assigne à chaque point de l'espace une densité de population est une fonction et la relation qui associe à chaque territoire une population est une mesure. On passe d'une fonction à une mesure en intégrant et d'une mesure à une fonction en dérivant. Ceci explique que l'on peut théoriquement appliquer l'analyse multifractale d'une fonction à une mesure après l'avoir dérivée.

3.4.1.1 Exemples de fonctions invariantes d'échelle

La construction de fonctions multifractales théoriques est une opération complexe. Nous nous restreignons dans cette section à un exemple simple proposé par MANDELBROT et HUDSON 2007 et généralisé par VICSEK 1990. Les fonctions sont construites selon un processus itératif : à chaque étape du processus, les portions de courbe de même pente sont remplacées par une version adaptée du générateur (*fig. 3.20*).

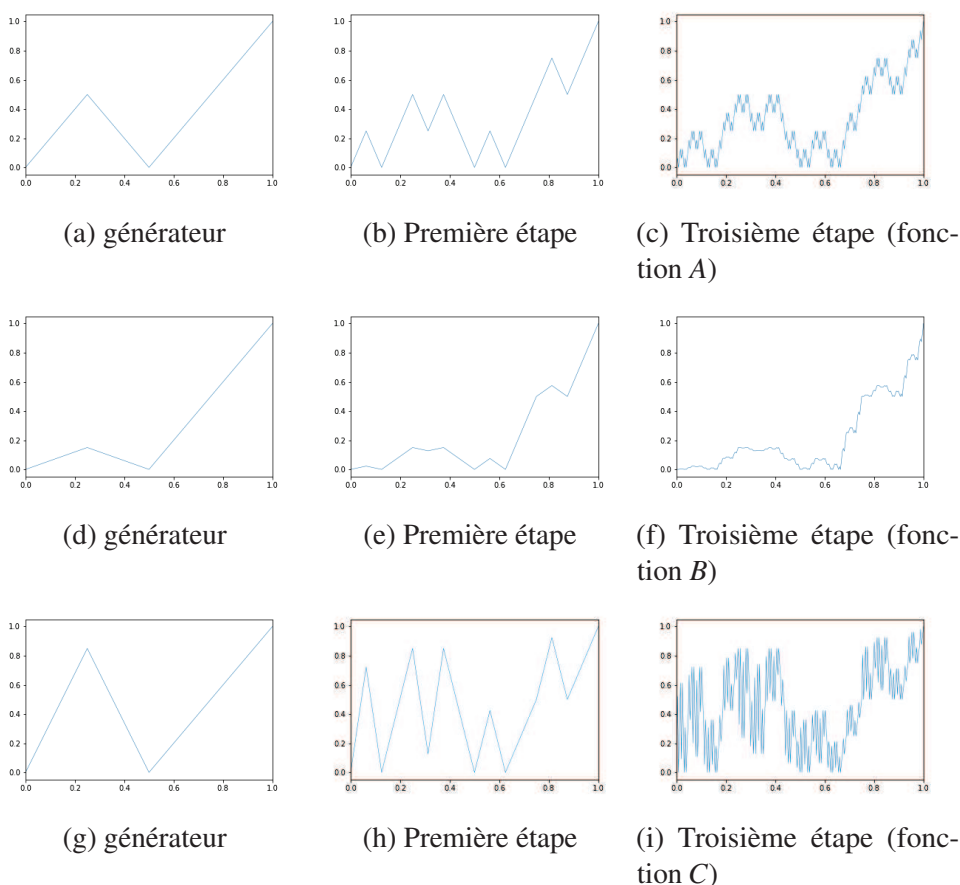


FIGURE 3.20 – Exemples de fonctions multifractales : la première colonne représente le générateur et les colonnes suivantes, les fonctions à différentes étapes du processus de construction

La fonction A est caractérisée par un seul type de fluctuations. Ses pentes locales sont en valeur absolues toutes identiques. Les fonctions B et C comportent deux types de fluctuations. Pour la fonction B , les fluctuations du premier type sont de faibles amplitudes et résultent de la forme de la première partie du générateur, tandis que celles du deuxième type sont plus importantes et sont associées à la deuxième partie du générateur. Pour la fonction C , c'est l'inverse, les fluctuations de plus fortes amplitudes sont associées à la première partie du générateur. L'analyse multifractale d'une fonction vise à quantifier les fluctuations et leur fréquence d'apparition à travers les échelles.

3.4.2 Présentation théorique de l'analyse multifractale d'une fonction

3.4.2.1 Mesure des fluctuations locales

La notion de fluctuation locale se formalise mathématiquement sous la forme de l'exposant de Hölder qui généralise les notions de continuité et de dérivabilité en un point. Une fonction f est continue au point x_0 si sa courbe peut être tracée autour de ce point sans avoir à lever son crayon. La dérivabilité est une propriété plus forte qui indique qu'au point x_0 , la fonction f a une tangente : elle se comporte à proximité de ce point comme une droite. La dérivée de la fonction f en ce point est par définition le coefficient directeur de la tangente. Une fonction dérivable au point x_0 est plus régulière qu'une fonction seulement continue : les fluctuations autour de x_0 de la première sont moins fortes que celles de la seconde.

L'exposant de Hölder quantifie de façon précise les fluctuations locales autour d'un point x_0 . Il est le plus grand exposant h tel qu'il existe un polynôme P_n de degré n et une constante c vérifiant :

$$|f(x) - P_n(x - x_0)| \leq c|x - x_0|^h \quad (3.36)$$

où n est la partie entière de h (ARNEODO et al. 1996). Plus la fonction est régulière, lisse, autour de x_0 et plus la valeur de h est importante. Pour les fonctions dérivables, h est supérieur ou égal à 1. Pour les fonctions continues et non dérivables en tout point, h est compris entre 0 et 1 et l'équation précédente devient :

$$|f(x) - f(x_0)| \leq c|x - x_0|^h \quad (3.37)$$

La valeur h de l'exposant de Hölder en x_0 correspond à l'exposant de la fonction puissance (fig. 3.21) qui approxime au mieux le comportement de f autour de x_0 .

Le spectre multifractal $d_f(h)$ d'une fonction f est défini, de manière identique au spectre multifractal $f(\alpha)$ d'une mesure, comme la dimension de Hausdorff d_H de l'ensemble des points d'exposant de Hölder h .

Les fonctions irrégulières ont des exposants de Hölder qui évoluent d'un point à l'autre. L'estimation directe du spectre pour ces fonctions, en se basant sur le calcul des exposants de Hölder, est généralement impossible et le recours à un formalisme multifractal est préconisé. La première étape de conception d'un formalisme multifractal consiste à définir des « boîtes » qui restituent correctement les valeurs des exposants de Hölder à travers les échelles.

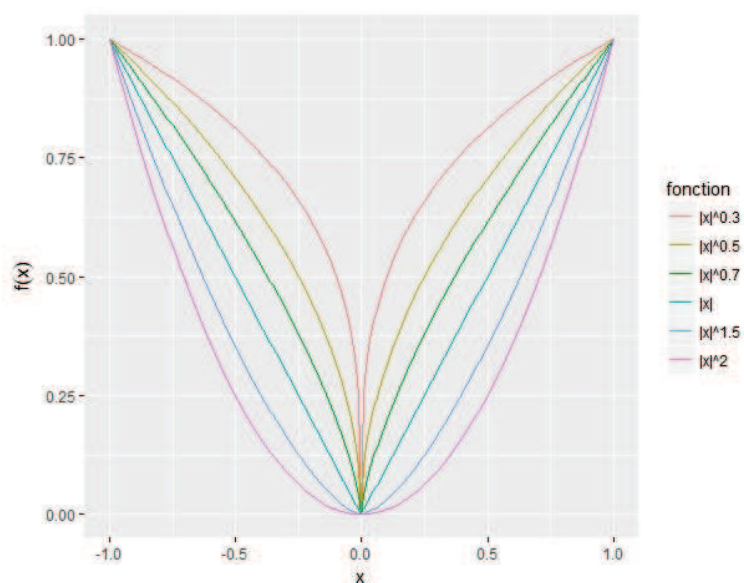


FIGURE 3.21 – Exemples de fonctions puissances non dérivables en 0

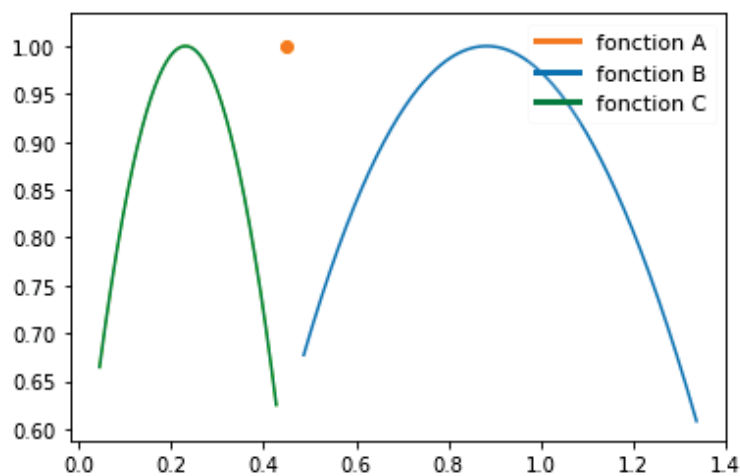


FIGURE 3.22 – Spectre multifractal des trois fonctions de la figure 3.20. Les fonctions *B* et *C* sont multifractales et la fonction *A* est monofractale car son spectre est réduit à un point. Les spectres permettent également de classer les fonctions selon leur régularité (de la moins lisse à la plus lisse) : la fonction *C* puis la fonction *A* et enfin la fonction *B*.

Remarque

h égal à 0 indique que la fonction n'est plus continue en x_0 et $h < 0$ que la fonction diverge et ressemble davantage à une mesure singulière qu'à une fonction usuelle.

3.4.2.2 Première méthode d'analyse multifractale d'une fonction : la transformée de Fourier

Au début du 19^e siècle, Joseph Fourier propose de décomposer les fonctions, tels que des signaux sonores en une somme de fonctions sinusoïdales. Cette opération, nommée transformée de Fourier, indique quelles sont les fréquences qui sont présentes dans la fonction analysée et avec quel poids elles interviennent (LESNE et LAGUËS 2015). La fonction qui associe une fréquence à son poids est appelée spectre fréquentiel.

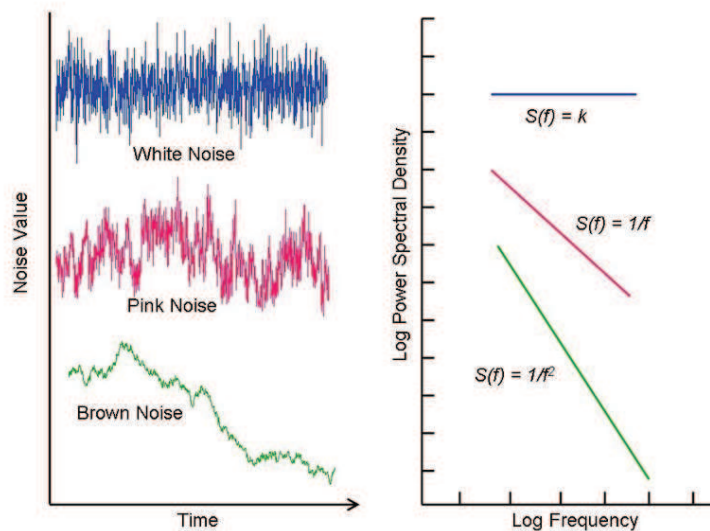


FIGURE 3.23 – Transformée de Fourier. Source : WARD et GREENWOOD 2007

En assimilant la notion d'échelle à celle de fréquence, les physiciens pensèrent pouvoir exploiter le spectre fréquentiel pour décrire les propriétés fractales des fonctions. Cette idée était si fortement partagée au sein de leur communauté que le concept d'invariance d'échelle était initialement relié aux phénomènes dit en $1/f$ (ABRY, JAFFARD et WENDT 2015), c'est-à-dire ceux dont le spectre fréquentiel suit approximativement une loi de puissance sur un intervalle de fréquences suffisamment large.

La transformée de Fourier se révéla, néanmoins, en pratique insuffisante pour approfondir le comportement multifractal d'une fonction car le spectre fréquentiel n'indique pas à quel endroit débute et se termine une fréquence mais simplement sa présence, de sorte qu'il est presque insensible aux variations locales de régularité et témoigne principalement de la régularité moyenne de la fonction analysée.

3.4.2.3 Présentation intuitive de la transformée en ondelettes

Alexander GROSSMANN et Jean MORLET mirent au point en 1984 la transformée en ondelettes. En prenant une analogie musicale, on pourrait dire que la transformée en ondelettes détermine non seulement les notes intervenant (les fréquences) mais aussi le moment où elles sont jouées (leur place dans la partition) et leur durée (blanches, noires, croches) (LESNE et LAGUËS 2015). Elle décompose une fonction en une combinaison linéaire de fonctions élémentaires, les ondelettes (cf. équation 3.38), localisées en différents points de l'espace et ayant des tailles différentes. Ces fonctions élémentaires sont toutes construites à l'aide d'une unique fonction mère par dilatation (coefficient a) et translation

(coefficient b) de celles-ci (ARNEODO et al. 1996) :

$$\Psi_{b,a}(x) = \frac{1}{a} \Psi\left(\frac{x-b}{a}\right) \quad (3.38)$$

La transformée en ondelettes $T_{\Psi}[f](b, a)$ d'une fonction f est le produit scalaire de celle-ci avec l'ondelette $\Psi_{b,a}(x)$:

$$T_{\Psi}[f](b, a)(x) = \int \Psi_{b,a}(x) f(x) dx \quad (3.39)$$

Remarque

La transformée en ondelettes en x_0 compare le comportement de f à l'ondelette $\Psi_{b,a}$. Plus f ressemble à $\Psi_{b,a}$ est plus la valeur de la transformée est importante en valeur absolue. Le paramètre b décrit le positionnement de l'ondelette et a sa taille, autrement dit l'échelle d'observation.

Le choix de la fonction mère détermine les propriétés de la transformée en ondelettes. La fonction mère (ondelette) doit être d'intégrale nulle, c'est-à-dire que son aire algébrique doit être nulle : $\int \Psi(x) dx = 0$ (fig 3.24). La courbe de la fonction mère doit dessiner au moins une oscillation, d'où le nom d'ondelette.

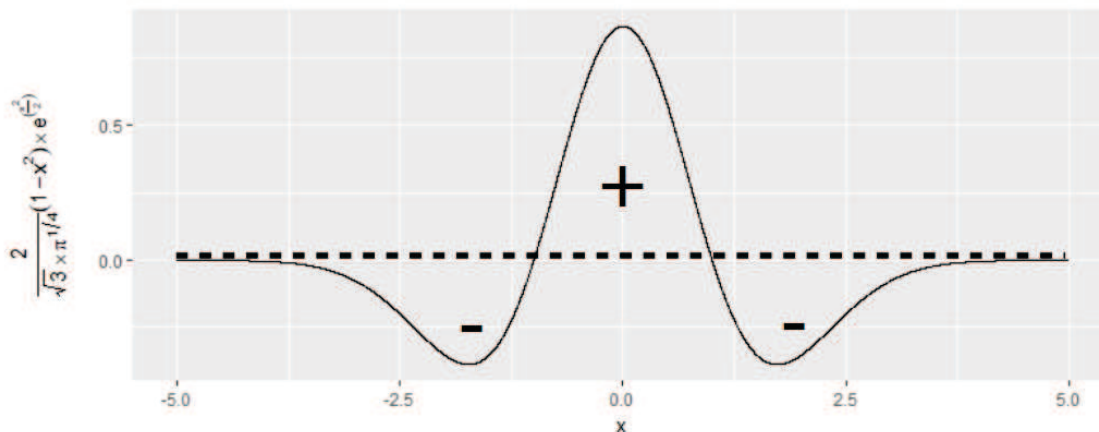


FIGURE 3.24 – Exemple d'ondelettes : le chapeau mexicain. Son aire algébrique est nulle : les deux aires (-) en dessous de l'axe des abscisses compensent l'aire centrale (+).

Les méthodes d'analyse multifractale à base d'ondelettes sont bâties sur un théorème démontré par Stéphane Jaffard en 1989 dont nous donnons ici une version très simplifiée :

$$T_{\Psi}[f](x_0, a) = a^{h(x_0)} \quad (3.40)$$

Ce théorème indique que la transformée en ondelettes calculée au point x_0 se comporte en loi de puissance en fonction de l'échelle a . $h(x_0)$ est l'exposant de Hölder en x_0 . Cependant, l'emploi direct des coefficients d'ondelettes n'est pas suffisant pour garantir une estimation robuste de l'intégralité du spectre multifractal (ABRY, JAFFARD et WENDT 2015) :

- Par construction, une grande partie des valeurs de la transformée en ondelettes sont proches de 0. Ceci rend instable, comme pour les mesures, l'estimation des paramètres multifractals pour $q < 0$.

- L'existence de singularités oscillantes¹⁴ dans la fonction analysée biaise l'estimation.

Des méthodes basées sur des transformations de ces coefficients doivent être mises en œuvre pour garantir la qualité statistique de l'estimation du spectre multifractal.

3.4.2.4 Deuxième méthode d'analyse multifractale d'une fonction : les Maxima du Module de la Transformée en Ondelettes

La méthode des Maxima du Module de la Transformée en Ondelettes (MMTO) repose sur une transformée en ondelettes dite continue : l'ondelette peut être positionnée en n'importe quel point de l'espace ($b \in \mathbb{R}$) et le paramètre de taille est positif ($a \in \mathbb{R}^+$). Étant donné qu'il n'est pas possible de construire une fonction d'échelles à partir d'une information continue, la MMTO commence par une étape de discrétisation, squelettisation, de la transformée en ondelettes.

On recherche pour chaque échelle les maxima des coefficients d'ondelettes. On relie ensuite ces derniers à travers les échelles pour former des lignes de maxima. Cette double opération revient à squelettiser la transformée et à extraire uniquement l'information pertinente et non redondante. Sur la figure 3.25, nous avons représenté deux exemples de discrétisation. Les squelettes des transformées partent du bas, soit des échelles les plus fines, pour monter vers les échelles les plus grossières. Au fur et à mesure que les échelles deviennent grossières, c'est-à-dire que a augmente, le nombre de brins diminue. Seules les plus fortes singularités sont alors suffisamment fortes pour être observées. Au contraire, aux échelles les plus fines, les maxima permettent de capter les singularités locales de petites ampleurs. Le nombre de brins à travers les échelles suit une loi d'échelle lorsque la fonction est invariante d'échelle.

Par rapport aux fonctions d'échelles de l'analyse multifractale des mesures, on introduit une innovation en ne prenant pas directement le coefficient d'ondelettes comme valeur, mais le maximum selon que l'on remonte les lignes de squelettisation. On garantit ainsi la stricte croissance des fonctions d'échelles avec l'échelle d'analyse et, in fine, la robustesse de l'analyse. Les courbes formées par les moments d'ordre q en fonction de l'échelle d'analyse suivent approximativement une droite (fig. 3.26 a et b), les phénomènes sont invariants d'échelle. L'estimation des coefficients directeurs de ces droites permet de définir des fonctions de partition (fig. 3.26 c) sur lesquelles on peut appliquer une transformée de Legendre (cf section 3.3.3.3). En sortie de cette procédure, on obtient les spectres multifractals (fig. 3.26 d).

Le spectre du brin d'ADN est monofractal : il est réduit à un seul point. L'organisation entre les échelles de ce signal est entièrement déterminé par un unique comportement d'invariance d'échelle (fig. 3.26.d). Au contraire, l'expérience de turbulence est très multifractale. Ceci confirme l'impression visuelle donnée par les figures 3.26a. et b. On constate des niveaux d'irrégularités entremêlées dans la seconde alors que la première semble plus structurée et simple à lire.

Sur le plan pratique, la mise en œuvre de la MMTO est très complexe et ce pour plusieurs raisons.

14. La notion de singularité oscillante autour de x_0 caractérise une fonction dont le comportement autour de x_0 est du type : $|f(x) - f(x_0)| = c|x - x_0|^h \times \sin(1/|x - x_0|)$. Plus on se rapproche de x_0 et plus la fonction décrit des « oscillations de plus en plus rapides ».

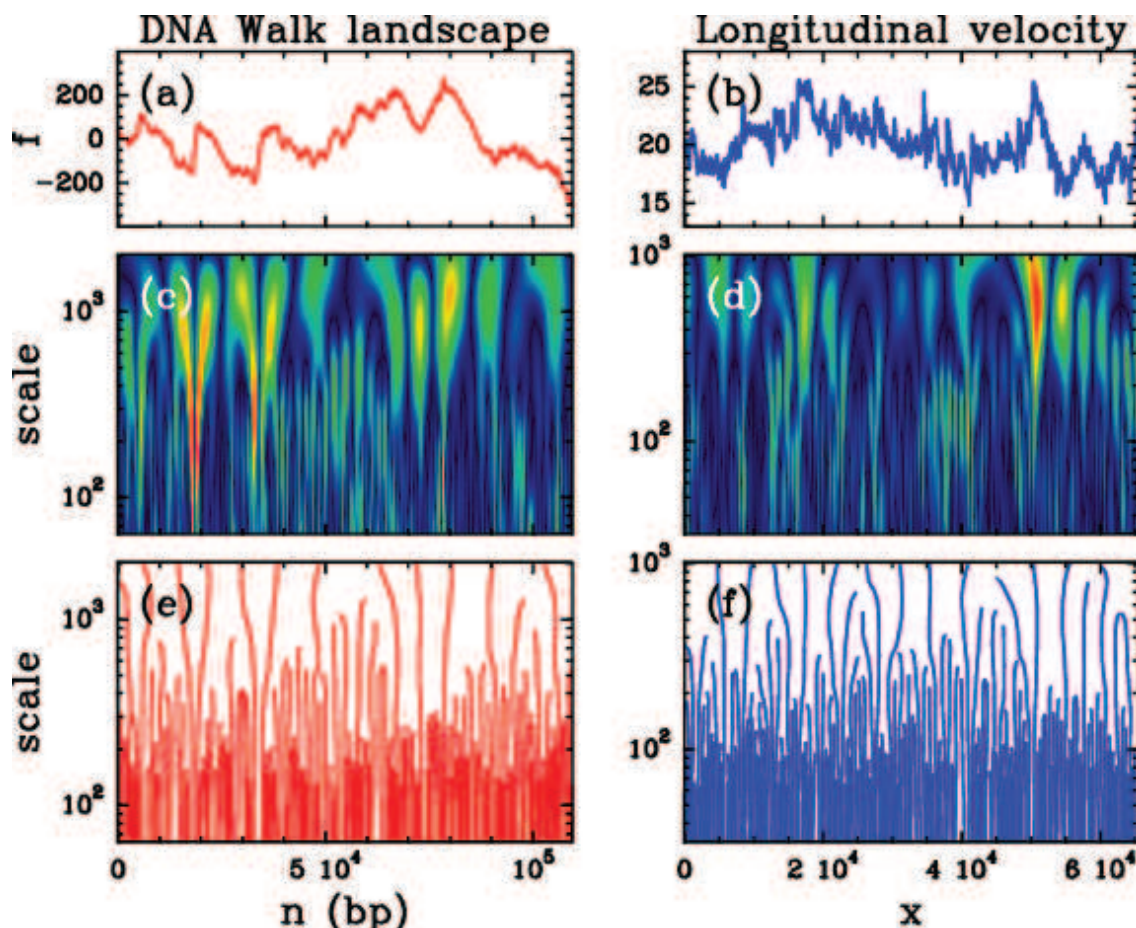


FIGURE 3.25 – Analyse MMTO de deux signaux. En haut à gauche est représentée par une courbe mathématique une portion d'ADN. L'ADN est composée de séquences de nucléotides (A, C, G et T). Aux deux premiers nucléotides (A,C) ont associé la valeur 1 et -1 aux deux derniers (G, T). En sommant cumulativement ces valeurs le long d'une séquence, on obtient une fonction caractérisant l'organisation de celle-ci. En abscisse $n(bp)$ est le numéro d'ordre des nucléotides. En haut à droite est représentée la composante longitudinale d'un champ de vitesse le long d'une direction. L'abscisse x est une position dans l'espace et l'ordonnée la valeur de la vitesse au niveau de celle-ci. Cette courbe reflète la variation des vitesses au sein d'un fluide turbulent. Au milieu, est représentée la transformée en ondelettes continues des deux signaux (la couleur rouge est associée aux coefficients de forte amplitude), en bas, la squelettisation de leur transformée d'ondelettes. Source : ARNEODO, AUDIT et al. 2008.

- Le calcul d'une transformée continue est une opération plus théorique que pratique, les ordinateurs actuels ne savent pas traiter les quantités continues. L'illusion du continu est obtenue en discrétisant l'espace et les échelles. Cette opération de discrétisation n'est pas neutre et rend approximative la recherche des maxima à une échelle donnée puis ensuite la détection des lignes de maxima.
- La recherche des maxima est difficile pour un signal unidimensionnel et devient quasi-impossible pour des fonctions à deux paramètres telles que des images.

En pratique, la MMTO a tendance à être abandonnée au profit de l'analyse en ondelettes dominantes.

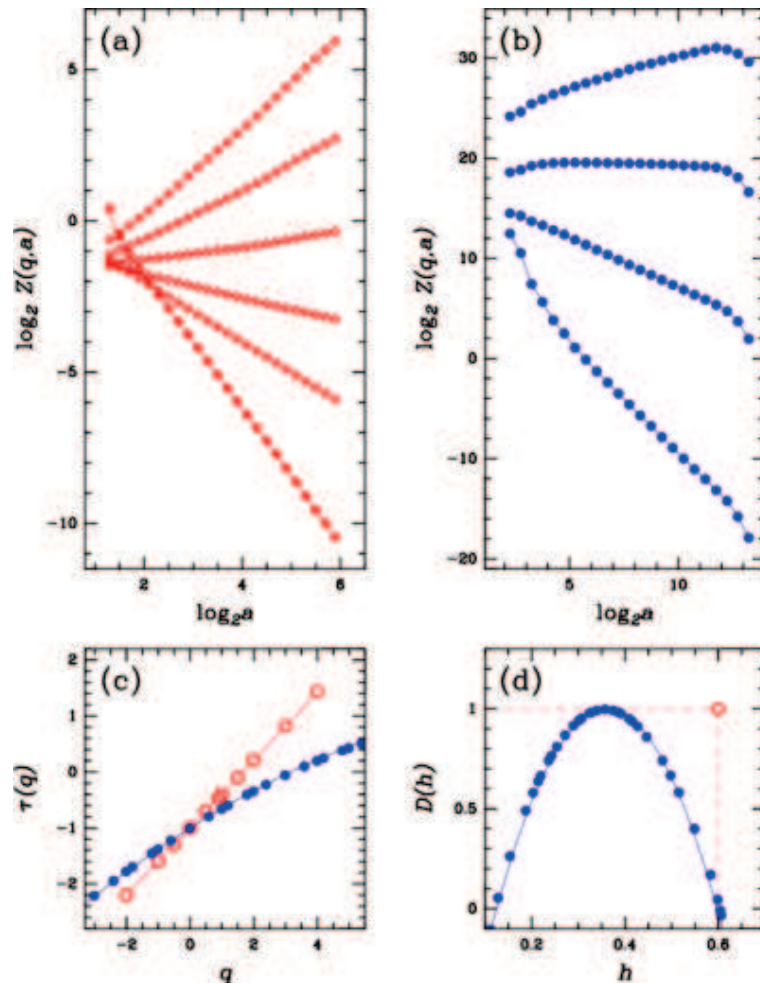


FIGURE 3.26 – Différentes étapes de l'estimation du spectre multifractal pour les deux signaux précédents. Les ronds rouges et vides sont associés à l'ADN et les ronds bleus et pleins à l'expérience de turbulence. En haut, représentation des moments d'ordre q à travers les échelles. L'alignement des points permet de calculer des coefficients directeurs : les fonctions de partition (c). On déduit les spectres (d) en appliquant la transformée de Legendre aux fonctions de partition. Source : ARNEODO, AUDIT et al. 2008.

3.4.2.5 Une méthode plus performante et plus simple à mettre en œuvre : les ondelettes dominantes

L'analyse en ondelettes dominantes (*wavelet leaders*) repose sur une transformée en ondelettes discrètes, c'est-à-dire que les tailles et les positions des ondelettes sont discrètes. Le paramètre d'échelle suit une progression géométrique : un facteur 2 sépare deux échelles successives. Les points de calcul de la transformée suivent un échantillonnage dyadique : au niveau de l'échelle la plus fine, les points de calcul sont espacés de façon régulière. Les points de calcul à l'échelle suivante sont obtenus en échantillonnant uniformément la moitié des points de l'échelle la plus fine. Cette opération est répétée à chaque changement d'échelle.

Les ondelettes de Daubechies sont utilisées dans les cas pratiques pour des raisons de performance de calcul. Elles constituent une famille d'ondelettes qui sont classées selon leur niveau de régularité (fig. 3.27). L'ondelette de Daubechies la plus simple est l'ondelette dite de Haar. Cette ondelette non continue à une seule oscillation correspond exactement

à la différence de deux « boîtes » de la *box-counting*. Par construction, elle « capte » les ruptures de continuité. Les fonctions qui nous intéressent étant continues, on privilégie dans ce cas des ondelettes plus régulières sensibles à des fluctuations plus « douces ».

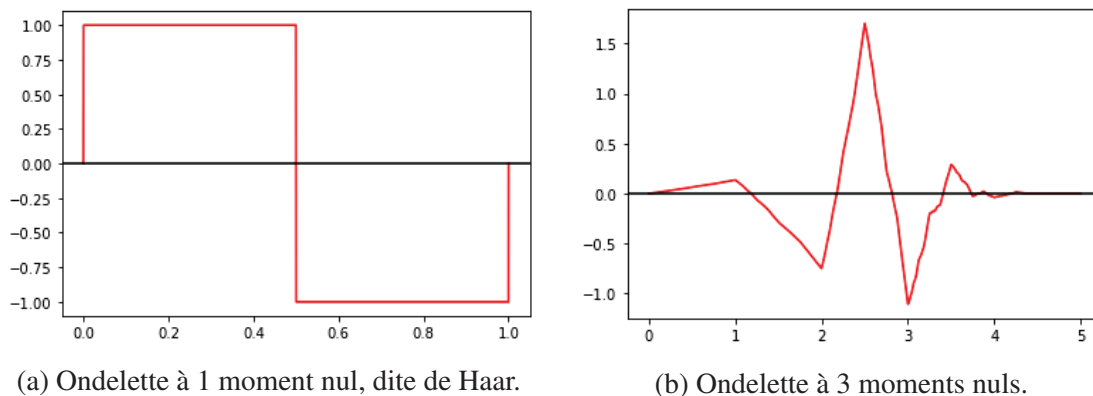


FIGURE 3.27 – Exemples d’ondelettes de Daubechies

Les ondelettes dominantes sont une adaptation de la méthode MMT0 aux ondelettes discrètes. Sur la figure 3.28, les points bleus représentent les coefficients d’ondelettes et le cercle rouge un coefficient d’ondelettes dominante. Celui-ci est déterminé en recherchant la plus grande valeur absolue parmi tous les coefficients d’ondelettes existants dans un certain voisinage spatial et à toutes les échelles inférieures (le rectangle noir). Cette opération est analogue à l’opération de squelettisation de la MMT0 mais sa simplicité algorithmique rend le calcul très rapide. Le spectre s’obtient en appliquant aux valeurs des coefficients d’ondelettes dominantes les formules proposées par CHHABRA et JENSEN 1989 (cf. section 3.3.4.2).

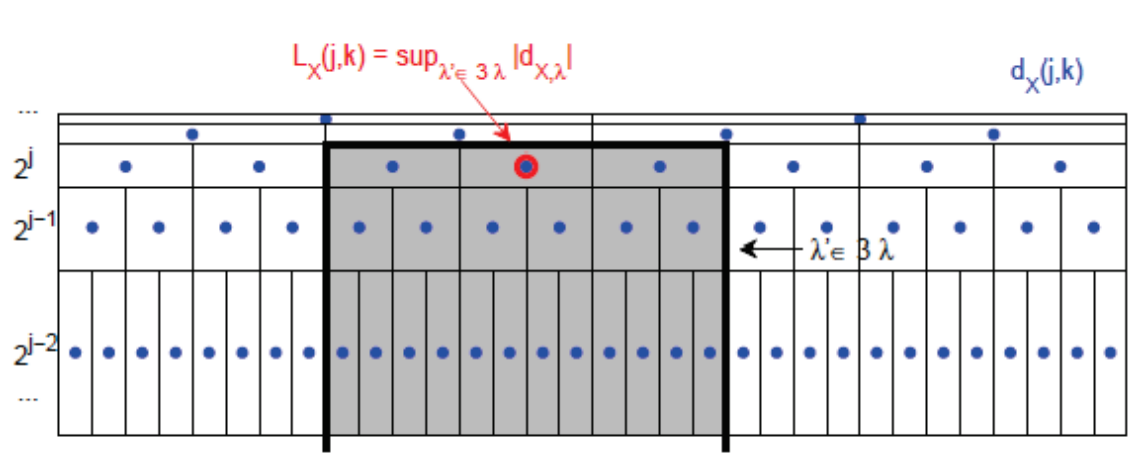


FIGURE 3.28 – Schéma de calcul des ondelettes dominantes. Source : WENDT 2008. En bas de la figure se trouve les échelles les plus fines (2^{j-2}). Chaque point bleu représente un coefficient d’ondelettes. Le cercle rouge est un coefficient d’ondelettes dominante. Il est construit en prenant le maximum des coefficients d’ondelettes (en valeur absolue) contenus dans le rectangle noir.

3.4.2.6 Les log-cumulants : résumer le spectre par deux paramètres

Nous avons utilisé précédemment la distance de Hausdorff pour comparer les spectres multifractals entre eux. Lorsque le nombre d'objets analysés devient important (plusieurs dizaines de milliers), cette approche n'est plus faisable. Une solution alternative consiste à résumer l'information contenue dans les spectres à l'aide d'un petit nombre de paramètres : les log-cumulants c_1 et c_2 . Ces paramètres peuvent être ensuite utilisés dans des analyses statistiques pour comparer et classer les textures.

Le spectre de singularité d'une fonction peut être approché par une parabole :

$$d_f(h) = d + \frac{c_2}{2} \left(\frac{h - c_1}{c_2} \right)^2 \quad (3.41)$$

où d est la dimension du support de la fonction (1 pour les fonctions à un paramètre et 2 pour celles à deux paramètres). c_1 et c_2 sont estimés respectivement à partir de l'examen des lois d'échelles caractérisant la moyenne et la variance des logarithmes des coefficients des ondelettes dominantes. Plus la valeur de c_1 est importante et plus la fonction est régulière. Avec un $c_1 = 0.88$, la fonction B est plus régulière que les fonction A et C dont les c_1 valent respectivement 0.47 et 0.23 (fig. 3.20). c_2 décrit, lui, la force de la multifractalité. Plus c_2 est important en valeur absolue (c_2 est toujours négatif), plus la fonction est multifractale et contient des points d'exposants de Hölder très différents. Avec un $c_2 = -0.25$, le caractère multifractal de la fonction B est plus accentué que celui de la fonction C dont le c_2 a pour valeur -0.05 . La fonction A a un c_2 nul, elle est monofractale.

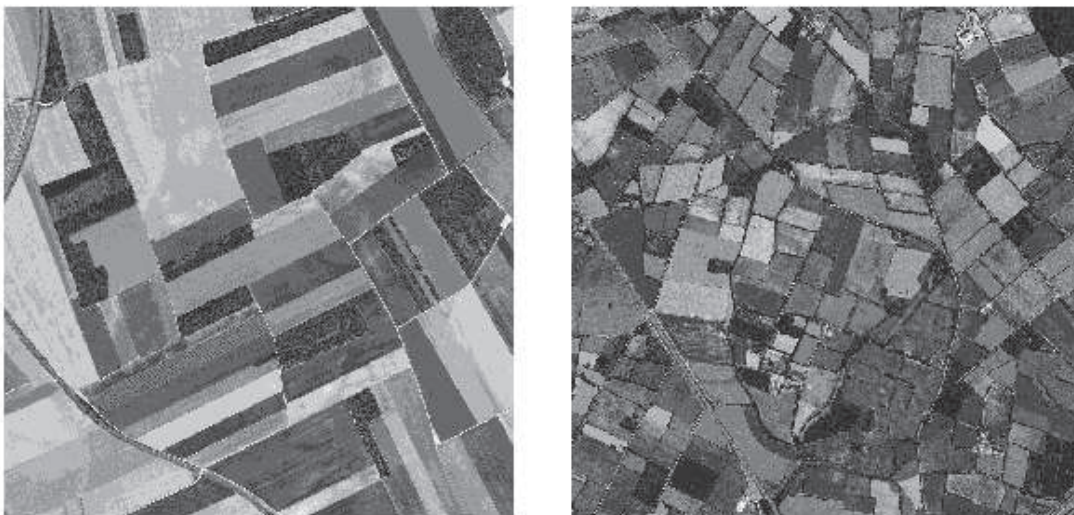


FIGURE 3.29 – Deux exemples d'orthophotographies. A gauche un paysage d'openfield et à droite de bocage. La valeur du c_1 est de 0,4 à gauche et de 0,06 à droite. La valeur du c_2 est de -0,4 à gauche contre -0.1 à droite.

Remarque

Le log-cumulant c_1 décrit l'évolution du mode (le maximum) de la distribution des coefficients d'ondelettes dominantes à travers les échelles tandis que c_2 décrit l'éparpillement des coefficients autour du maximum. Cette modélisation est générique et pourrait s'appliquer en théorie aux mesures multifractales. Les effets du *MAUP* pourraient dans ce cas être décrits par ces deux coefficients. En pratique, l'instabilité des calculs dans le cas d'analyses multifractales de mesures rend cette approche difficile à appliquer.

3.4.3 Application de l'analyse multifractale en ondelettes dominantes aux orthophotographies de France métropolitaine

3.4.3.1 Données et choix des paramètres

Nous avons utilisé, pour réaliser notre analyse, la base d'orthophotographies à 5 m de l'Institut national de l'information géographique et forestière (IGN). Afin d'analyser la variabilité spatiale des textures, nous avons découpé celui-ci en sous-images de 2560 m de côté. Chaque sous-image de taille 512 pixels * 512 pixels a été analysée individuellement. Les images étaient codées classiquement selon trois couleurs : rouge, verte et bleue. Chaque composante de couleur est définie par une matrice de taille 515 × 512 éléments. La première étape de l'analyse consiste à additionner ensemble les trois matrices pour obtenir pour chaque image une seule fonction à étudier. En tout, 65000 sous-images ont été caractérisées à l'aide de l'analyse multifractale¹⁵. Les estimations des indicateurs multifractals ont été calculées sur la plage d'échelles correspondant aux niveaux 2 et 5 de la décomposition en ondelettes, c'est-à-dire pour une plage de résolutions allant de 20 m à 160 m. Nous avons utilisé comme ondelette une Daubechies avec 3 moments nuls (fig 3.27.b).

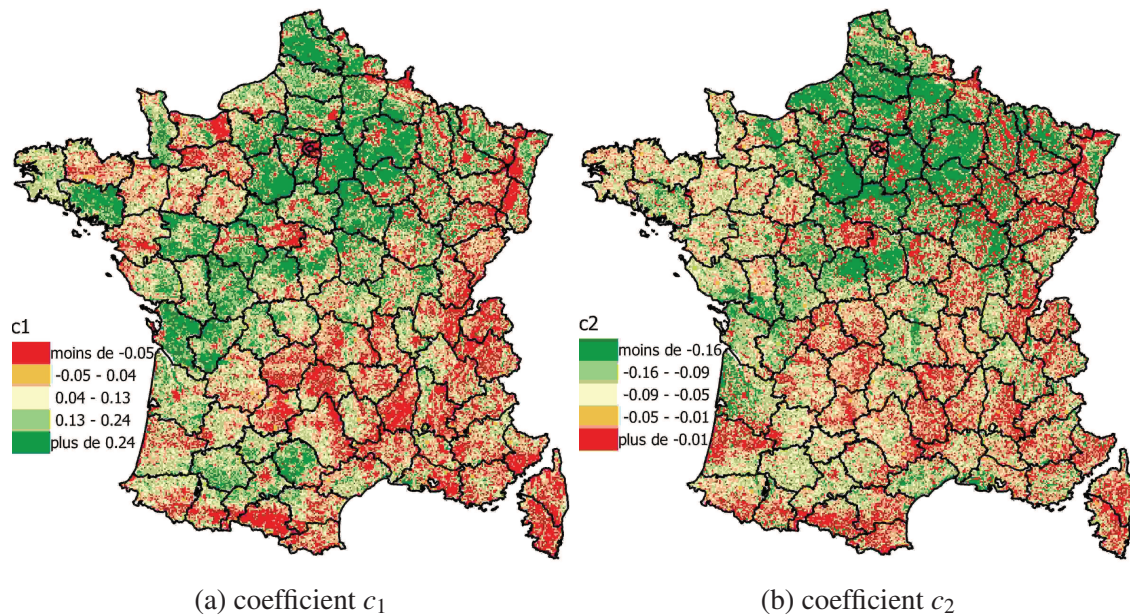
3.4.4 Résultats de l'analyse

Les résultats de l'analyse sont globalement assez décevants. Les cartes représentant le c_1 et le c_2 font bien apparaître des régularités géographiques mais celles-ci sont perturbées par des artefacts liés aux propriétés de la source de données utilisée (fig 3.30).

Nous devinons sur les deux cartes certains contours départementaux. Par exemple, les contours du département du Morbihan ressortent nettement sur les cartes par rapport aux autres départements bretons. Ce phénomène, qui est plus accentué sur la carte du c_1 que du c_2 , s'explique par les conditions de prises de vue. Les orthophotographies sont actualisées selon un cycle quinquennal : tous les ans, un cinquième des départements de la France métropolitaine est mis à jour. Les orthophotographies que nous avons téléchargées en 2015 pour réaliser notre analyse ont donc été prises entre 2009 et 2014. Au cours de cette période, plusieurs phénomènes sont susceptibles de parasiter la comparabilité des orthophotographies :

- les orthophotographies ne sont pas nécessairement prises exactement à la même période de l'année (la prise de photographies aériennes nécessite en particulier l'absence de nuages) ;

15. Remarque : il a été nécessaire d'appliquer une intégration fractionnaire d'ordre 1/2 à l'ensemble des images pour « lisser » les images et réduire les plus grandes singularités.

FIGURE 3.30 – Carte des coefficients c_1 et c_2

- le climat variant d’une année sur l’autre, la végétation peut être plus ou moins abondante ;
- il est également raisonnable de penser que l’IGN, en améliorant la qualité de ses produits, modifie les chaînes de traitement et d’acquisition des orthophotographies au cours du temps.

Les mesures que nous obtenons reflètent ainsi à la fois les propriétés du phénomène que nous étudions et les propriétés des outils d’acquisition et de traitement de l’information. Nous pouvons formuler l’hypothèse suivante pour expliquer la moindre sensibilité du c_2 aux artefacts induits par les prises de vues : contrairement au c_1 qui mesure la régularité moyenne, le c_2 mesure la variabilité autour de cette régularité moyenne et reflète ainsi des phénomènes davantage structurels tels que l’agencement des différentes entités composant le territoire (les champs, les routes, les bâtiments...) que conjoncturels tels que les effets de la météorologie sur la végétation.

Un dernier point pose également problème. Les espaces urbains et les espaces forestiers et montagneux sont indifférenciés sur les deux cartes. Les plages de valeurs prises par les deux indicateurs c_1 et c_2 se chevauchent pour ces espaces. Ainsi sur les deux cartes, l’agglomération parisienne et la Sologne apparaissent de la même couleur. De sorte que l’étude de la périurbanisation et des transitions entre les espaces ruraux et urbains à partir de ces cartes est peu pertinente.

Ces mises en garde formulées, nous pouvons néanmoins retirer certains enseignements de l’analyse des deux cartes exposées dans cette section. Tout d’abord, l’information apportée par chacune d’elles est assez proche. Le coefficient de corrélation entre le c_1 et le c_2 est de -0.52 . Nous retrouvons ainsi sur les deux cartes une opposition forte entre l’Est et le bassin parisien, d’une part, et l’Ouest et le Sud-Ouest, d’autre part (si l’on néglige les espaces forestiers et montagneux). La première zone (openfields) se caractérise par des valeurs élevées de c_1 (plus de $0,24$) et faibles de c_2 (moins de $-0,16$) contrairement à la seconde zone (bocages) qui est caractérisée par des valeurs faibles de c_1 (moins de $0,13$) et élevées de c_2 (plus de $-0,09$). La zone d’openfields est globalement plus régulière mais aussi davantage multifractale que celle de bocages. Ce schéma d’organisation de

l'espace français partage certains points communs avec ceux observés sur la carte des dimensions généralisées d'ordre 0 (fig. 3.15). Pour autant, nous ne pouvons pas, à partir de cette ressemblance, conclure à l'existence d'une relation entre les structures agraires et l'organisation du peuplement. En effet, au-delà des artefacts de mesures, les valeurs des coefficients c_1 et c_2 peuvent être induites à la fois par les tissus bâtis (contenants de la population) et par les structures agraires : les bâtiments font partie intégrantes des images et à ce titre sont susceptibles d'influencer les valeurs des paramètres multifractals de sorte qu'il est difficile de comprendre l'information contenue dans les cartes du c_1 et c_2 de façon univoque et de l'associer à un phénomène donné.

3.4.5 L'invariance d'échelle : mythe ou réalité ?

3.4.5.1 L'invariance d'échelle en tant que norme sociale

Dans son livre séminal, « Les Objets Fractals : Forme », publié en 1975, Benoît Mandelbrot regroupe dans un même ensemble des objets provenant de champs scientifiques très divers (géographie physique, économie, physique, biologie, mathématiques et astrophysique), dont les principaux points communs sont d'être à la fois très irréguliers et invariants d'échelle. La quatrième de couverture, présente la géométrie fractale comme :

« une grande extension apprenant au savant et à l'ingénieur - et à d'autres ! - À voir le monde de façon nouvelle ».

Les principaux manuels mathématiques de géométrie fractale banaliseront l'idée selon laquelle l'invariance d'échelle se retrouverait dans différents pans du monde naturel (*Fractals everywhere* de BARNSELY 2014, *Fractals* de FEDER 2013). En France, de nombreux ouvrages de vulgarisation scientifique publiés dans des grandes maisons d'édition accentuent ce propos, en évoquant explicitement l'universalité de l'invariance d'échelle. Sans être exhaustifs, nous pouvons citer les ouvrages suivants dont les titres sont sans équivoque : *Universalités et fractales : jeux d'enfant ou délits d'initié ?* de SAPOVAL 1997, *Le monde des Fractales : la géométrie cachée de la nature* de DUBOIS et CHALINE 2006. À un niveau international, on dénombre, rien qu'entre 2010 et 2019, plus de 1000 articles dans les revues associées à Nature qui contiennent dans le corps de leur texte le terme « fractal » et plus de 90 dans le titre même. Encore une fois, l'étendue des sujets traités est très vaste et s'étend de la biologie moléculaire à l'informatique. Une recherche dans les corps de textes associant les termes « fractal » et « cités » renvoie quant à elle 46 articles sur la même période.

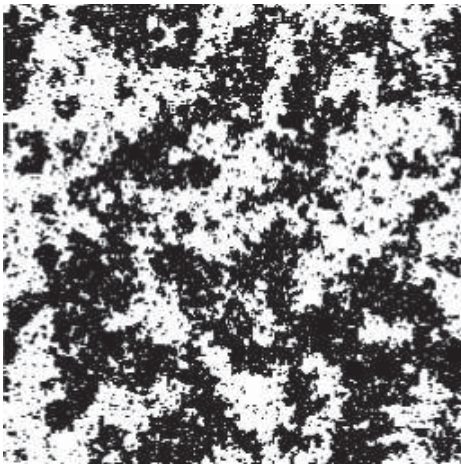
Évoquer l'absence d'invariance d'échelle au sujet des espaces urbains apparaît donc comme une position relativement marginale au sein du champ géographique et plus largement scientifique. La récente controverse à ce sujet entre TANNIER 2018 et FRANKHAUSER 2018 dans le blog de la revue Cybergéo est d'autant plus intéressante qu'elle permet d'explicitement un certain nombre de postulats et d'hypothèses qui sont rarement discutés en temps normal dans le champ géographique. Dans son billet, Cécile Tannier remet en cause l'universalité de l'invariance d'échelle des tissus urbains et ce principalement pour trois raisons : les dimensions fractales varient fortement au sein des villes, les gammes d'échelles où l'on observe un comportement scalant ne sont pas stables d'un espace à l'autre et aucune théorie expliciterait de façon satisfaisante les mécanismes physiques et sociaux qui engendreraient une régularité à travers les échelles des espaces urbains. Elle en tire comme conclusion que « rien ne prouve qu'une forme urbaine fractale soit, par nature, optimale ». Pierre Frankhauser répond à celle-ci, point par point, en insistant sur l'existence

de principes d'emboîtement d'échelles en géographie urbaine. La multifractalité, en tant que superposition de différents comportements scalants, expliquerait la variabilité spatiale des dimensions fractales au sein des espaces urbains. L'étude fractale de la ville apporterait des bénéfices théoriques et pratiques suffisamment importants pour que la question de l'invariance d'échelle parfaite des structures urbaines puisse être considérée comme secondaire (TANNIER 2018).

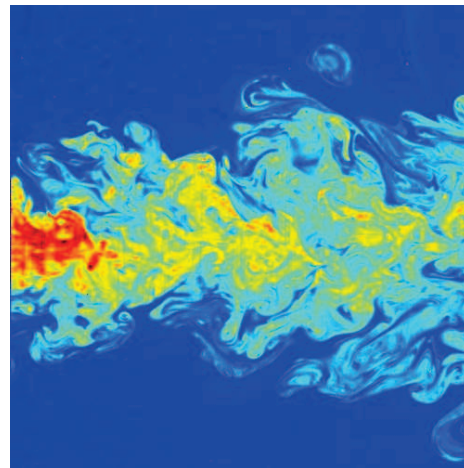
Une exploration de l'usage de l'invariance d'échelle dans les sciences permet de poursuivre cette discussion.

3.4.5.2 Le statut de l'invariance d'échelle dans les sciences

Les phénomènes de transitions de phases¹⁶ et de turbulences, étudiés par les sciences physiques, sont caractérisés par des propriétés d'invariance d'échelle. Le modèle proposé par Ising (*fig. 3.31.a*) sert à comprendre, dans les transitions de phases, les structures fractales qui apparaissent à proximité des points critiques¹⁷ (LESNE et LAGUËS 2015). En 1941, Kolmogorov décrit l'ordonnancement des tourbillons observés en dynamique des fluides turbulents à l'aide d'une modélisation invariante d'échelle encore utilisée aujourd'hui (*fig. 3.31.b*).



(a) Simulation d'un modèle d'Ising. Source : Mathworks.



(b) Visualisation par fluorescence induite par laser d'un jet turbulent. Source : Wikipédia.

FIGURE 3.31

Dans le domaine de la finance, la représentation des cours boursiers sous la forme d'une fractale, le mouvement brownien, est ancienne et remonte aux travaux de BACHELIER 1900. La crise des subprimes de 2007 entraîna la remise en cause de cette modélisation (WALTER 2008) y compris par Benoit MANDELBROT et Richard L HUDSON (2007). Des modèles fractals plus complexes autorisant des ruptures brutales dans les cours ont été

16. En physique, une transition de phase désigne la transformation physique d'un système d'une phase (ou d'un état) vers une autre, induite par la variation d'un paramètre de contrôle externe (température, champ magnétique...) (source : Wikipédia).

17. Les points critiques sont les valeurs de paramètres de contrôle externe pour lesquels on constate une transition de phases. Par exemple, 0°C est le point critique séparant pour l'eau sa phase solide de sa phase liquide.

proposés depuis pour palier le problème de la continuité des trajectoires browniennes, peu en accord avec les phénomènes observés pendant les périodes de crises (JIANG et al. 2019).

En anthropologie, DESCOLA 2006 suggère que certaines populations ont recours plus fréquemment que d'autres à des motifs fractals dans leurs arts plastiques. Cette inégale répartition serait induite par l'existence de différentes façons d'identifier les « entités du monde » par les humains qu'il appelle « ontologies ». « L'analogie »¹⁸ serait, parmi les quatre ontologies proposées par Philippe DESCOLA, la plus propice aux développements de structures invariantes d'échelle. DEHOUE 2014 note que plusieurs contre-exemples limitent l'association univoque entre l'usage des fractales et un mode de pensée spécifique.

CHRISTALLER 1933 propose un modèle pour expliquer la hiérarchisation fonctionnelle des espaces régionaux reposant sur l'accessibilité aux équipements. LE BRAS 1996 dans son ouvrage sur la répartition de la population française évoque l'organisation territoriale de l'espace pour expliquer la multifractalité du peuplement : à chaque échelon (région, département, canton et commune), le pouvoir administratif se concentre dans certaines zones ce qui tendrait à répartir inégalement les implantations humaines. Ces logiques spatiales expliquent l'emboîtement d'échelles observé en géographie (LE BRAS 1996) mais ne sont pas rapportées à des mécanismes sociaux qui contraindraient strictement les relations entre les échelles et qui justifieraient des phénomènes d'invariance d'échelle.

Ce survol rapide de l'utilisation de l'invariance d'échelle en science n'a pas prétention à être exhaustif et sert avant tout à illustrer la différence de statut de cette notion selon les disciplines. Dans les sciences de la nature, l'invariance d'échelle est intégrée dans des modèles qui permettent de prédire la valeur des paramètres trans-échelles observés alors que, dans les sciences humaines, elle est principalement un cadre de référence pour penser les formes observées.

3.4.5.3 Les lois d'échelles en géographie

Plus généralement, c'est l'existence même des lois d'échelles en tant que lois universelles qui est sujet à controverse en géographie. L'ouvrage de WEST 2017, physicien et ancien président de l'Institut de Santa Fe, « Scale : The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life », a vulgarisé auprès du grand public les relations allométriques¹⁹ concernant les villes. Nommé parmi les 100 personnes les plus influentes par le magazine américain *Time*, le poids médiatique de Geoffrey West est particulièrement important. Tout comme l'invariance d'échelle, les relations allométriques sont fort utiles pour modéliser les espaces urbains : la régularité qu'elles imposent aux objets permet de prédire leurs propriétés géométriques à toutes leurs échelles observables. Néanmoins de nombreux articles remettent en question le consensus autour de ces relations aussi bien en géographie (PUMAIN 2007, LOUF et BARTHELEMY 2014 et COTTINEAU et al. 2015) qu'en biologie (Bak (1996), Jensen (1998) et Sornette (2000)) : les valeurs des exposants sont très sensibles au choix de la définition des objets étudiés.

18. « Rappelons que l'identification analogique repose sur la reconnaissance d'une discontinuité générale des interiorités et des physicalités aboutissant à un monde peuplé de singularités, un monde qui serait donc difficile à habiter et à penser en raison du foisonnement des différences qui le composent, si l'on ne s'efforçait de trouver entre les existants, comme entre les parties dont ils sont faits, des réseaux de correspondance permettant un cheminement interprétatif » DESCOLA 2006.

19. L'allométrie est l'étude des échelles de relations entre une partie du corps et le corps dans son ensemble. Elle décrit comment les caractéristiques d'un individu changent avec sa taille (source : Wikipédia consultée le 02/08/2019).

La loi de Zipf, souvent utilisée par les géographes, met en relation la taille des villes et leur rang de classement (selon cette variable). L'écart à cette loi est souvent considéré comme une anomalie qui mérite une investigation sur son origine : par exemple « lorsque la première ville se détache clairement au-dessus de la droite (comme c'est le cas en France avec Paris), on estime être en présence d'une macrocéphalie urbaine » (Geoconfluences 2003). Encore une fois, la régularité introduite par cette loi est très adaptée aux démarches modélisatrices. Pour autant, GAN, LI et SONG 2006 en démontrant dans leur article « Is the Zipf law spurious in explaining city-size distributions ? » qu'un grand nombre de distributions aléatoires sont susceptibles de produire des résultats compatibles avec une modélisation de type loi de Zipf, ont profondément atténué sa portée théorique. Il n'est plus en effet nécessaire de postuler une théorie économique sous-jacente pour comprendre l'ajustement des systèmes urbains à une loi de Zipf. La régularité est alors un effet statistique mécanique.

3.4.5.4 Quelques éléments de réflexion sur la pertinence de comparer des dimensions fractales

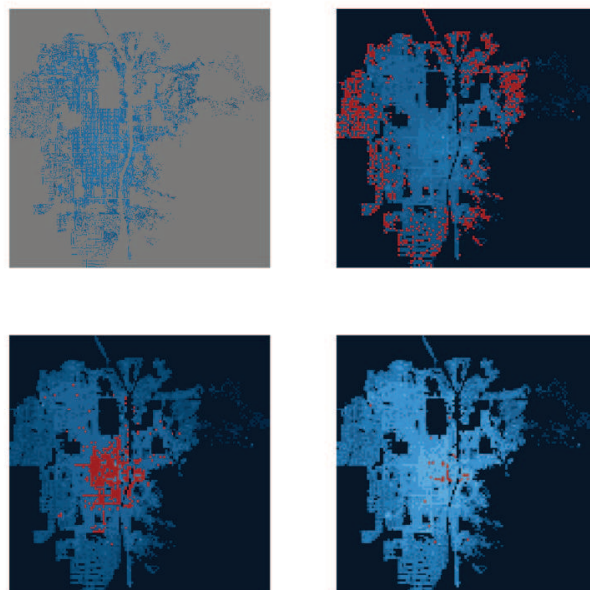
La puissance heuristique des fractales réside dans leur capacité à décrire des phénomènes complexes simplement. Mais cette capacité représente également un risque car elle est susceptible d'entraîner des mésusages si l'on oublie les conditions d'application de celles-ci.

Par définition, les dimensions fractales mesurent des propriétés orthogonales aux échelles observées. Sans unité, mais néanmoins bornées (entre 0 et 2 pour les objets du plan et entre 0 et 3 pour les objets dans l'espace usuel), elles permettent de comparer, au moins en théorie, des objets dont les tailles caractéristiques peuvent être très différentes.

En géographie humaine, la comparaison de dimensions fractales calculées sur des plages de résolutions différentes ne va pas sans poser des problèmes épistémologiques (PAVON-DOMINGUEZ et al. 2017). En effet, contrairement à la physique où l'observateur peut, au moins dans une certaine mesure, être considéré comme extérieur au phénomène analysé, le géographe est partie prenante de ce qu'il étudie. Ainsi, toutes les échelles mobilisées pour étudier les sociétés doivent être mises implicitement et explicitement en relation avec une échelle particulière qui est celle de l'être humain. Comparer des dimensions fractales sur des plages différentes revient donc à comparer des usages différents de l'espace.

Dans le cas d'analyses multifractales, la comparaison des différentes dimensions généralisées au moyen du spectre est effectuée pour des plages d'échelles identiques (cf. 3.3.3) (SALAT et al. 2018, MURCIO et al. 2015). Les analyses de SALAT et al. 2018 sur Kyoto (fig. 3.32) et FUSCO, CAGLIONI et ARALDI 2017 sur Nice montrent la décroissance des dimensions fractales locales avec la distance au centre. Or dans le cas de configurations multifractales, les dimensions locales sont discontinues et varient d'un point à l'autre. Les conditions d'applications de l'analyse multifractale sur ce type de villes ne sont donc pas vérifiées. Ceci ne veut pas dire que les spectres multifractals calculés dans SALAT et al. 2018 ne soient pas pertinents mais ils ne mesurent pas l'organisation multifractale de la ville.

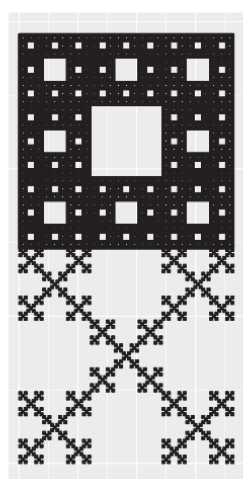
La juxtaposition de deux fractales (fig 3.33b.a) n'est généralement pas une fractale, chaque partie n'étant pas similaire au tout, ni une multifractale. L'invariance d'échelle de cet assemblage semble pourtant vérifiée si l'on se réfère au graphique bilogarithmique d'évolution du nombre de boîtes occupées selon leur taille (fig 3.33b.b). Ceci tend à montrer



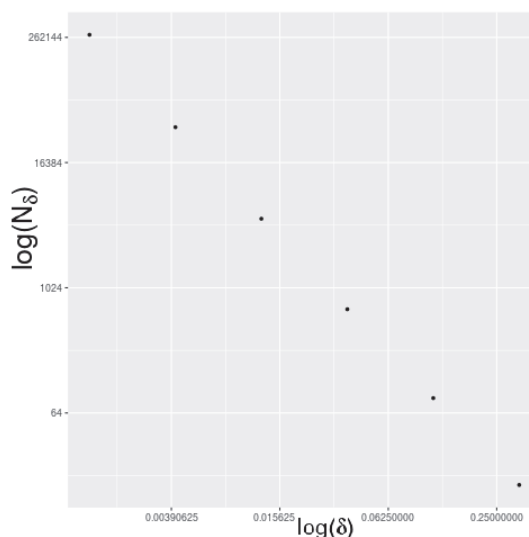
En haut à gauche : prix du foncier à Kyoto en 1912. Les autres figures représentent la distribution des coefficients α dans l'espace. En haut à droite, le rouge correspond aux valeurs de $\alpha = 3$, en bas à gauche à $\alpha = 6$ et en bas à droite à $\alpha > 7$.

FIGURE 3.32 – Analyse multifractale de Kyoto en 1912. Source : SALAT et al. 2018.

que la qualité de l'estimation d'une dimension fractale ne garantit pas l'homogénéité des formes analysées et à relativiser la capacité que l'on peut prêter aux dimensions à distinguer des formes relativement similaires.



(a) Assemblage d'un tapis de Sierpinski et d'une fractale de Vicsek



(b) Évolution du nombre d'éléments dénombrés à travers les échelles

FIGURE 3.33 – Juxtaposition de deux fractales et courbe de comptage associée

3.4.5.5 Le pouvoir symbolique des fractales

Les difficultés conceptuelles ne sont pas suffisantes à elles seules pour expliquer la suprématie des tenants de la « fractalité » des espaces urbains. Il est indispensable de rajouter à cet effet les biais de publication. STERLING 1959 constata que les études montrant l'absence de résultats avaient moins de chances d'être publiées contrairement aux études exhibant des résultats avérés. Dans notre cas, la non invariance d'échelle des tissus bâtis s'apparente à un non résultat, ce qui justifierait également la position marginale de cette proposition dans le champ scientifique.

A cela s'ajoute, des difficultés à nommer les phénomènes non invariants d'échelle. PALMER 1988 note que :

« The fractal dimensions obtained from these analyses are generally close to 2, implying weak spatial dependence. The fractal dimension is not a constant function of scale, implying that patterns of spatial variation at one scale cannot be extrapolated to other scales ».

En toute rigueur, évoquer l'évolution d'une dimension fractale à travers les échelles est oxymorique. THOMAS et al. 2010, recourent à l'expression « comportement scalant » (*scaling behaviour*) ce qui ne va pas sans poser de problèmes également. Le terme « scalant » correspond au mot anglais *scaling* et est directement relié à une propriété d'invariance d'échelle comme en témoigne l'extrait suivant de MANDELBROT 1975 :

« J'aime dire que ces courbes dont des petites et des grosses parties ont la même forme ou structure, mais à des échelles différentes, sont "scalantes". On encore (c'est plus précis, mais bien lourd !) "à homothétie interne" ».

FLORINDO et al. 2012 ; RAGHAVENDRA et DUTT 2010 évoquent des Multi-Scale Fractal Dimension. Encore une fois, l'introduction du mot « fractal » rend paradoxale cette définition. ABRY, JAFFARD et WENDT 2015 décrivent leur analyse comme étant multi-échelle ce qui permet de garantir une forme de neutralité. Néanmoins, en employant le terme multi-échelle, tout du moins en français, on tend à induire le fait que nous examinons les objets à différentes échelles, ce qui ne reflète pas la notion d'analyse à travers les échelles de l'analyse trans-échelle.

Cela dit, la question de l'usage d'un concept pour penser le monde n'est pas uniquement réductible à des arguments scientifiques et éditoriaux, ou encore sémantiques, internes aux champs considérés (LATOUR 1984 et PIERRE 2001). La controverse portée par SOKAL et BRICMONT 1997 au sujet de l'emploi qu'ils jugent abusif de termes mathématiques ou physiques dans les sciences sociales a profondément questionné celles-ci (FLEURY et LIMET 1997). Paul Virilio, urbaniste et essayiste, est cité, en particulier, par les deux physiciens pour son usage du terme fractal dont voici un exemple (VIRILIO 1984) :

« la compression temporelle a lieu, la fragmentation de la société qui en est issue finit par créer une société fractale ».

De prime abord, il semble difficile d'imaginer une société invariante d'échelle et l'on peut dès lors légitimement se questionner sur l'effet recherché de l'emploi d'un tel terme dans cette affirmation. BOURDIEU 2014 évoque le pouvoir symbolique des mots et BOUVERESSE 1999 le « littéralisme » qui consiste à croire que la science ne devient intéressante qu'une fois retranscrite dans un langage littéraire et « métaphorique » pour expliquer ces emprunts sémantiques. Plus près de nous, deux géographes non « quantitativistes » ont également eu recours au terme fractal :

« Ainsi s'explique aussi l'émiettement fractal de l'urbanisation, car dans chaque hameau, chaque écart peut exister des propriétaires qui demandent leur part » CHARMES 2019.

« La périurbanisation est un processus fondamental, transversal à toutes les situations territoriales observables : elle concerne Paris, les métropoles régionales mais aussi toutes les unités urbaines. Il s'agit d'un phénomène fractal, irrégulier et fragmenté » LUSSAULT 2018.

Les deux usages proposés ci-dessus du terme fractal contrairement à la citation de Paul Virilio, sont relativement cohérents avec ceux des géographes « quantitativistes » à la fois d'un point de vue thématique et théorique. Pour autant, sommes-nous certains que la totalité des acteurs publics et du grand public auxquels s'adressent ces deux auteurs comprendront pleinement le terme et les implications théoriques qui en découlent ? Cette remarque est d'autant plus forte que la suppression du mot « fractal » dans ces deux extraits ne changerait rien à leur pertinence. On perçoit à travers ces exemples que l'emploi d'un concept issu de la physique n'est jamais anodin et est susceptible de produire des effets contraires à ceux recherchés lorsque l'on ne distingue pas les usages qui relèvent d'une analogie stricte et ceux plus métaphoriques.

3.5 Conclusion

Dans ce chapitre, nous avons exploré l'organisation spatiale de différents phénomènes géographiques à l'aune de la géométrie fractale. Les résultats obtenus varient en terme de pertinence selon les données utilisées et les axes d'analyses privilégiés.

L'analyse multifractale des populations sur une gamme de tailles intermédiaires (entre 800 et 6400 m) à l'aide de données carroyées fine a été particulièrement satisfaisante. A ce niveau d'analyse, la modélisation du peuplement sous la forme d'une mesure multifractale est pertinente et confère au MAUP un nouveau statut. Au lieu d'être un problème qui masque notre capacité à comprendre le phénomène étudié, il devient un outil en tant que tel pour le décrire.

L'analyse multifractale des espaces paysagers à l'aide des orthophotographies s'est révélé décevante. Cette approche permet d'explorer l'ensemble des éléments (bâtiments, champs, routes...) qui constituent le paysage sans à avoir à les définir en amont de l'analyse. Cependant, la collecte hétérogène de l'information dans le temps et l'espace nous a posé des problèmes insolubles quant à l'interprétation des résultats. Ce constat peut être nuancé par le fait que l'on retrouve malgré tout une opposition entre l'Est et l'Ouest de la France métropolitaine conforme à celle observée dans le cadre de l'analyse du peuplement.

Dans le prochain chapitre, nous allons approfondir les résultats issus de la GWFA. En effet, il s'avère qu'en dehors des espaces urbains, l'invariance d'échelle n'est pas vérifiée. L'analyse trans-échelle utilise l'écart à l'invariance d'échelle pour produire de la connaissance sur le caractère multiscalaire des phénomènes analysés.

Chapitre 4

Dépasser le présupposé d'invariance d'échelle via l'analyse des signatures trans-échelles

Résumé

Nous proposons dans ce chapitre, une méthode d'analyse trans-échelle, construite à partir du formalisme de la géométrie fractale, mais qui ne présuppose aucune régularité à travers les échelles. Au lieu de résumer l'information trans-échelle à l'aide d'un unique indicateur, à savoir la dimension fractale, cette méthode en conserve l'intégralité et l'exhaustivité. L'enjeu est ensuite de savoir traiter cette information trans-échelle pour pouvoir dégager des faits qui soient à la fois révélateurs d'un ordonnancement des phénomènes à travers les échelles et qui soient aisément interprétables.

4.1 Introduction

L'analyse fractale postule que l'information pertinente pour décrire un phénomène se situe dans la façon dont ses caractéristiques observées à différentes échelles sont reliées entre elles. Sa mise en œuvre requiert l'existence d'une régularité à travers les échelles : l'invariance d'échelle. En pratique, cette propriété n'est pas toujours vérifiée, même sur des plages d'échelles restreintes. Dans SÉMÉCURBE, TANNIER et ROUX 2019, nous avons utilisé la GWFA avec la dimension d'ordre $q = 0$ pour décrire l'organisation spatiale locale des tissus bâtis de la France métropolitaine pour les échelles 50; 100; 200; 400; 800 m (fig. 4.1). Les dimensions fractales s'échelonnent entre 0 et 2 selon que l'on se trouve dans un espace rural ou urbain conformément aux résultats obtenus précédemment (cf. section 3.3.2.2.2). En revanche, la qualité de l'ajustement ayant permis le calcul des dimensions fractales décrit une autre géographie : les coefficients de détermination opposent les espaces ruraux du Nord-est à ceux de l'Ouest de la France. L'invariance d'échelle est uniquement vérifiée dans le Nord-est¹. Dans le reste des territoires ruraux, la dimension fractale ne restitue qu'imparfaitement l'organisation trans-échelle des tissus bâtis.

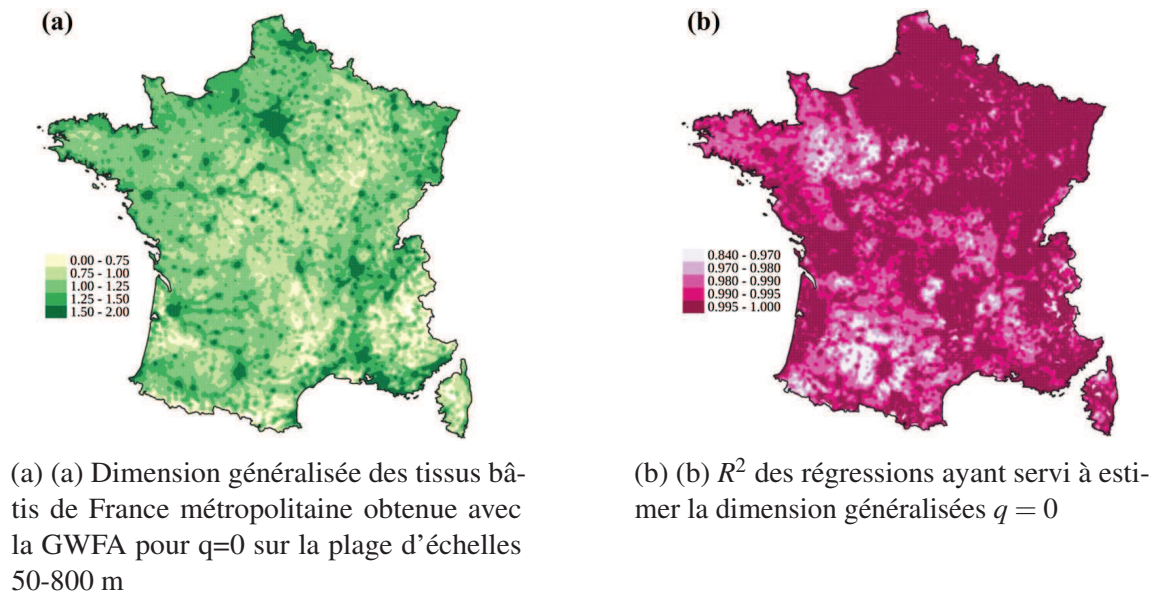


FIGURE 4.1 – Résultat de la GWFA pour $q = 0$. Source : SÉMÉCURBE, TANNIER et ROUX 2019

Le calcul des dimensions locales sur les plages 50 – 200 m et 200 – 800 m permet d'approfondir cette observation. Les deux cartes de dimensions locales obtenues sont miroirs l'une de l'autre (fig. 4.2) : les espaces de faible dimension de l'une sont les espaces de forte dimension de l'autre. Dans l'ouest de la France, les dimensions fractales locales croissent avec les échelles, tandis que dans l'est, elles décroissent ou restent stables. Ces résultats bien que contraires aux hypothèses de l'analyse fractale, ne sont pas le fruit du hasard et présentent des structures et des régularités spatiales bien établies. L'objectif de l'analyse trans-échelle est d'exploiter ces écarts à l'hypothèse d'invariance d'échelle afin de produire une connaissance scientifique. L'analyse trans-échelle reprend les concepts de l'analyse fractale mais ne suppose aucune régularité entre les échelles d'observation.

1. Ceci explique la non décroissance des dimensions généralisées observées sur la figure 3.11

Purement descriptive, elle propose un cadre dans lequel l'écart à l'invariance d'échelle n'est plus un handicap mais l'objet même de l'analyse, de sorte que tous les objets, fractals ou non, sont susceptibles d'être analysés selon cette méthode.

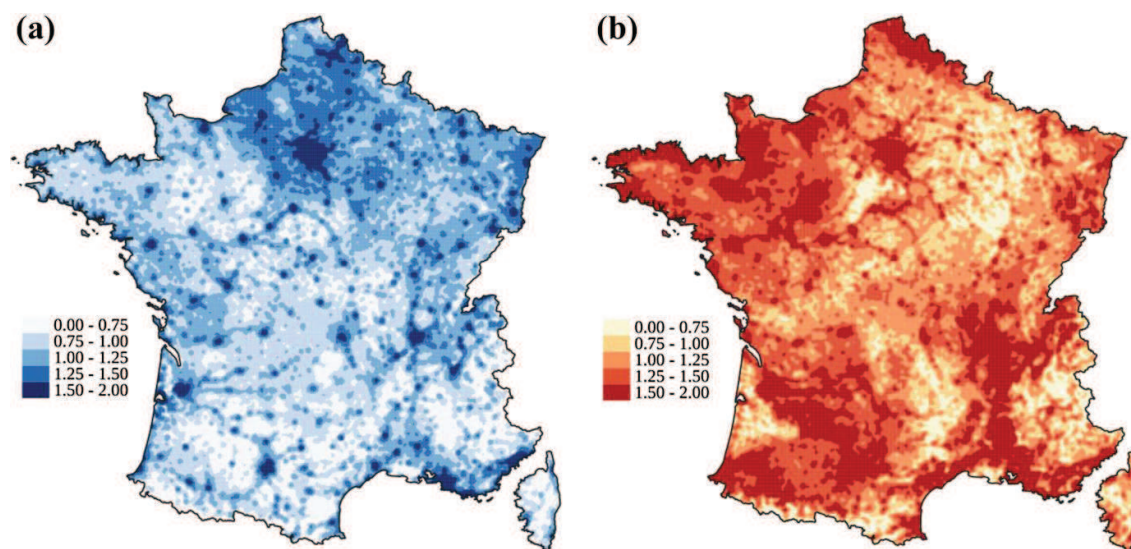


FIGURE 4.2 – Dimension fractale des tissus bâtis de France métropolitaine obtenue avec la GWFA pour $q=0$: (a) entre 50 et 200 m ; (b) entre 200 et 800 m

Le contenu de ce chapitre repose principalement sur la valorisation des signatures trans-échelles. Ces indicateurs sont calculés à partir du même matériau brut que les dimensions fractales : les courbes bilogarithmiques des indicateurs d'échelles. Mais au lieu de résumer ces courbes par un taux de variation global, on estime des variations ponctuelles, à savoir des pentes locales, ce qui d'un point de vue théorique revient à calculer une fonction dérivée. Empiriquement, on approxime la notion de dérivée en calculant des taux de variations entre deux échelles que l'on jugera suffisamment proches. Par construction, les signatures trans-échelles décrivent l'exhaustivité des caractéristiques des organisations trans-échelles. L'information issue d'une analyse trans-échelle est donc très riche et nécessite l'usage de méthodes statistiques de réduction de l'information pour rendre celle-ci claire et lisible. Deux méthodes mathématiques complémentaires sont proposées pour décrire l'écart à l'invariance d'échelle : une typologie a posteriori des signatures trans-échelles et une décomposition a priori de celles-ci. La première méthode décrit précisément les différences dans l'organisation trans-échelle des objets au sein d'un territoire. La seconde décompose les signatures selon des axes définis préalablement et dont on maîtrise parfaitement l'interprétation thématique. À la fin de ce chapitre, nous évoquons la simulation de configurations spatiales qui correspondent à des signatures trans-échelles spécifiques.

4.2 Les signatures trans-échelles

BATTY 1991 définit en géographie la notion de signature trans-échelles et l'utilise pour décrire des formes urbaines simulées à travers les échelles. FRANKHAUSER 1994 la considère comme un sous-produit de l'analyse fractale et s'en sert pour repérer les ruptures dans l'organisation des tissus bâtis. L'algorithme MorphoLim (TANNIER, THOMAS et al. 2011), recherche pour chaque tissu bâti analysé l'échelle où l'écart à l'invariance d'échelle est la plus forte. Cette échelle caractéristique propre à chaque tissu bâti analysé est obtenu par une modélisation polynomiale implicite des signatures trans-échelles.

Au-delà du champ géographique, PALMER 1988 utilise en écologie les signatures trans-échelles pour résumer la façon dont les plantes s'organisent dans l'espace. Plus récemment, FLORINDO et al. 2012 ; RAGHAVENDRA et DUTT 2010 la mettent en œuvre pour décrire des phénomènes biologiques et discriminer des formes telles que des lettres. En 2015, ABRY, JAFFARD et WENDT 2015 appliquent l'analyse trans-échelle sur des indicateurs multifractals par ondelettes pour différencier des supports papiers photographiques.

La signature trans-échelle d'un objet est constituée des taux de variation successifs $Sc_{r_a-r_b}$.

$$Sc_{r_a-r_b} = \frac{\log(M_{r_b}) - \log(M_{r_a})}{\log(r_b) - \log(r_a)} \quad (4.1)$$

Les notations employées sont celles de la GWFA (cf. 3.3.2.2) : M_r est la fonction d'échelle et r le rayon des disques.

Comme dit précédemment, pour les objets strictement invariants d'échelle, les taux sont tous égaux à la dimension fractale. Les signatures n'apportent aucune information supplémentaire sur l'organisation trans-échelle des objets étudiés. En revanche, les taux des objets non invariants d'échelle, oscillent autour d'une valeur moyenne, leur dimension fractale estimée. Les signatures restituent toutes les oscillations autour de cette valeur moyenne, échelle par échelle.

Sur la figure 4.3 nous avons représenté les 4 taux de variation de deux configurations bâties idéaltypiques : Sc_{100-50} , $Sc_{200-100}$, $Sc_{400-200}$ et $Sc_{800-400}$. Alors que la signature de la configuration d'openfield est quasi-plate, la signature de bocage est strictement croissante à partir du deuxième indicateur. Cette différence très forte nous permet de décrire l'organisation locale des espaces bâtis de France métropolitaine. Nous avons calculé les taux de variation sur les fonctions d'échelle $M_r(i)$ obtenues en sortie de la GWFA (cf section 3.3.2.2) (fig. 4.4). Dans les territoires d'openfield de l'Est de la France métropolitaine, nous observons une décroissance des signatures qui s'opposent à la croissance de celles-ci dans les paysages de bocage de l'Ouest de la France.

Par rapport à la dimension fractale, les taux de variation multiplient mécaniquement le nombre d'informations permettant de décrire un objet à travers les échelles. Au lieu d'avoir un unique indicateur, nous en disposons à présent de quatre. Si l'on souhaite profiter pleinement de cette richesse descriptive supplémentaire, il est nécessaire d'effectuer des traitements adhoc afin d'extraire les informations que l'on jugera pertinentes. Dans la section suivante, nous nous proposons d'exploiter à cet effet des méthodes statistiques qui ont pour objectif de classer les objets selon leurs différences relatives.

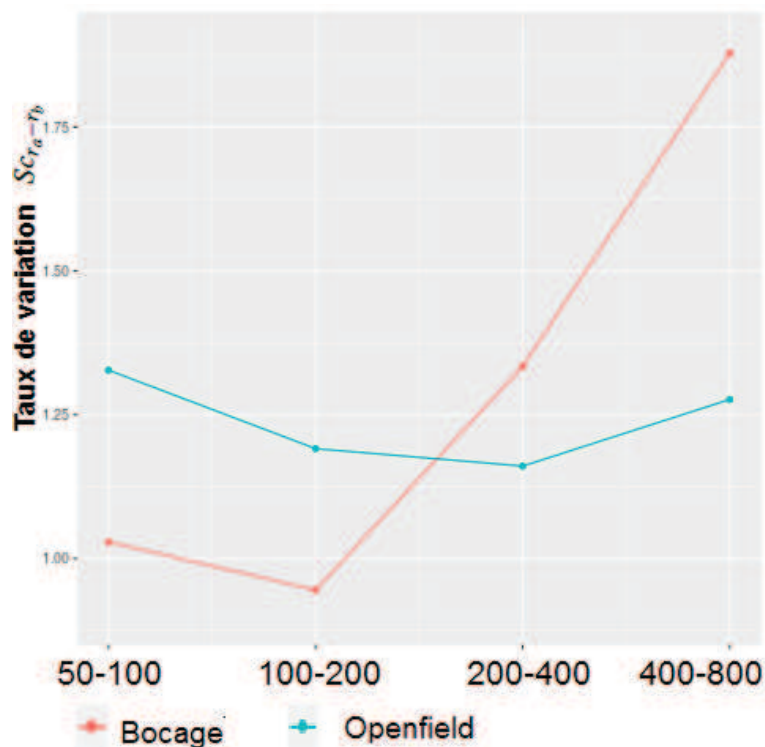


FIGURE 4.3 – Signature trans-échelle de deux configurations bâties idéaltypiques

4.3 Analyses statistiques des signatures trans-échelles

4.3.1 Analyse en composantes principales (ACP)

Nous mobilisons ici l'Analyse en Composante Principale (ACP) pour classifier les signatures trans-échelles des tissus bâtis de France métropolitaine. Nous justifions ce choix par le fait que l'ACP est à la fois très connue et très souvent utilisée, et que son approche basée sur la recherche de ressemblances et de dissemblances pour restituer les traits généraux d'un ensemble d'objets est exemplaire et révélatrice de la démarche statistique.

L'ACP consiste à déduire des corrélations au moyen de nouvelles variables décorréliées les unes des autres, appelées axes principaux. Ces axes sont déterminés de telle sorte qu'ils restituent au mieux la variabilité du positionnement relatif des objets dans l'espace². Dit plus simplement, ces axes explicitent les directions privilégiées sur lesquelles les objets se positionnent (figure 4.5).

L'inertie est le critère mathématique utilisé dans l'ACP pour définir les axes principaux. Intuitivement, elle correspond à une variance calculée selon une direction. L'idée de l'ACP est de chercher dans un premier temps, la direction qui maximise cette variance. On obtient, le premier axe de la décomposition en axes principaux. Les axes suivants sont définis en recherchant dans les espaces orthogonaux à ce premier axe, les directions qui contiennent le plus d'inertie. A la fin du processus, on récupère une liste comprenant autant d'axes que de variables ordonnées selon la décroissance de leur inertie. La somme des informations

2. L'espace considéré ici est géométrique et décrit la façon dont les objets se situent les uns par rapport aux autres selon la valeurs de leurs variables. Sauf mention contraire cet espace n'est pas l'espace géographique.

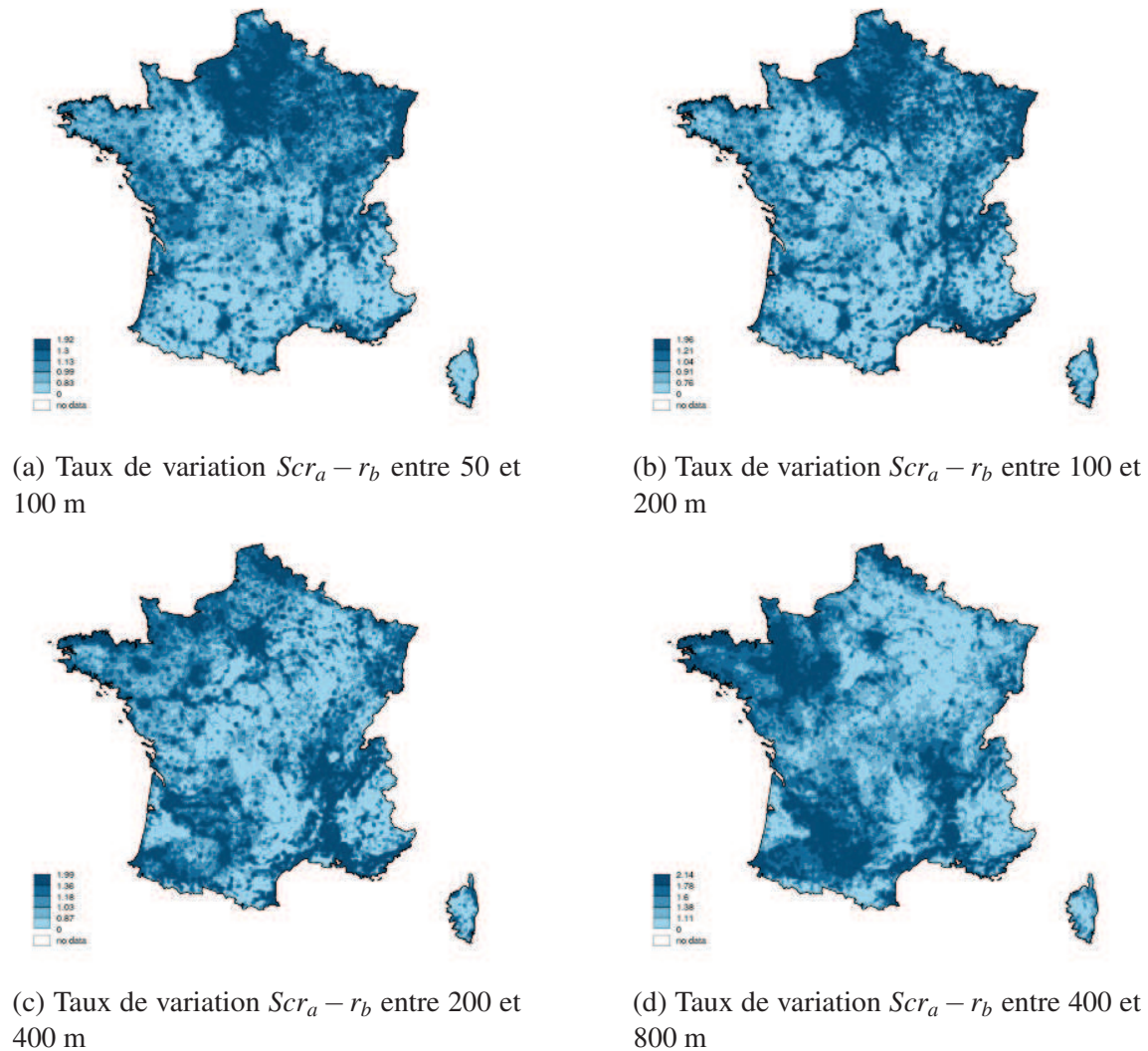


FIGURE 4.4 – Représentation des taux de variation $Scr_a - r_b$ calculés à partir de l’analyse faite avec la GWFA sur le bâti de la France métropolitaine pour des carreaux de 2 km de côté

apportées par tous les axes permet de restituer exactement l’organisation des objets. En revanche, en sélectionnant uniquement les premiers axes qui contiennent le plus d’inertie, on réduit l’information et l’on extrait les faits structurels qui semblent pertinents.

4.3.2 Application aux tissus bâtis de France métropolitaine (GWFA avec $q = 0$)

L’information contenue dans les quatre taux de variations (50 – 100, 100 – 200, 200 – 400, 400 – 800) est résumée presque intégralement par les deux premiers axes de l’ACP dont l’inertie cumulée est de 90% (fig. 4.6). Les quatre taux de variation sont très corrélés entre eux : on retrouve le fait que les cartes correspondant aux deux indicateurs extrêmes des signatures (Sc_{100-50} et $Sc_{800-400}$) sont presque miroirs l’une de l’autre. La restriction de l’ACP aux deux premiers axes garantit une représentation des signatures satisfaisantes tout en réduisant considérablement la complexité du phénomène étudié.

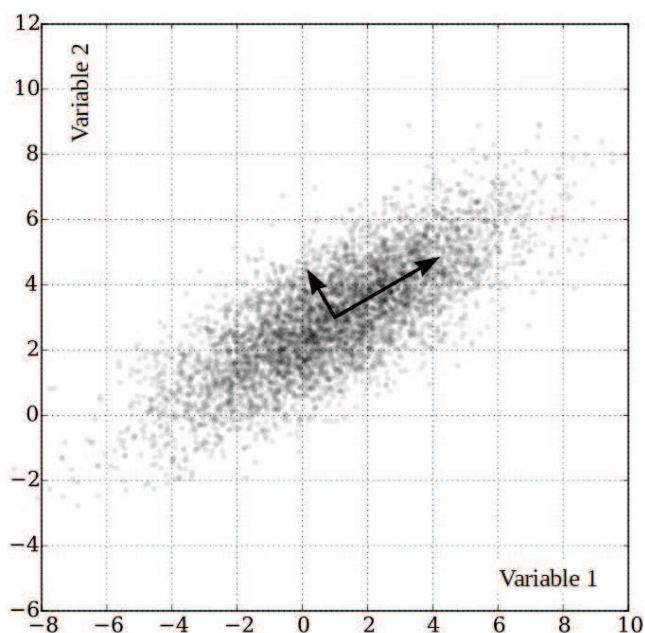


FIGURE 4.5 – ACP d'un nuage de points décrits par deux variables. Sur cette figure les deux axes de l'ACP sont représentés par les deux flèches noires.

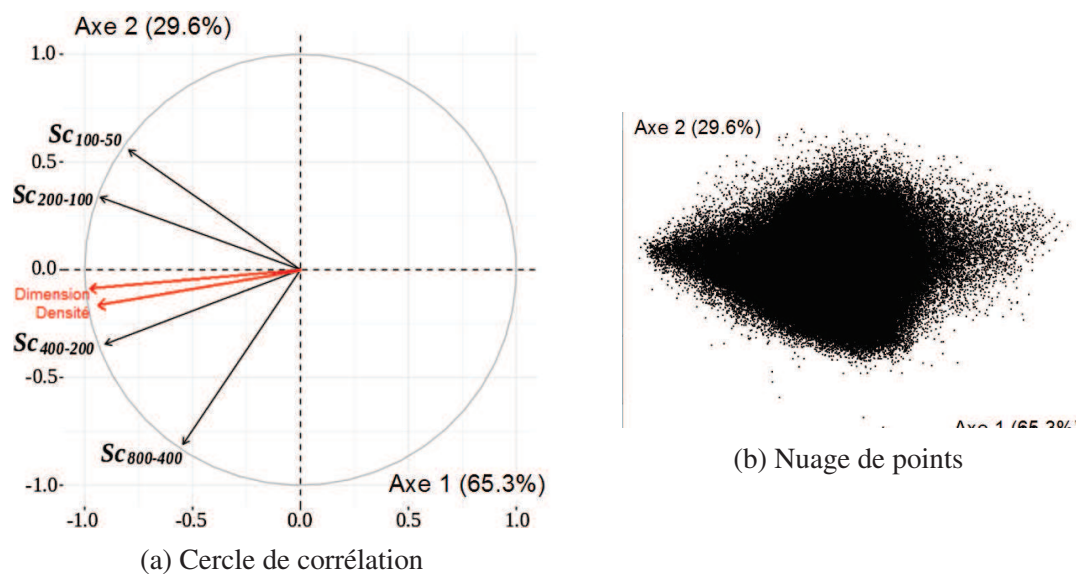


FIGURE 4.6 – Résultats de l'analyse en correspondante principale des signatures trans-échelles des tissus bâtis de France métropolitaine. Source : SÉMÉCURBE, TANNIER et ROUX 2019

Le premier axe (65% de l'inertie) est très corrélé positivement aux quatre signatures trans-échelles (fig. 4.7) et caractérise ce que les statisticiens nomment un « effet taille » : il ordonne les zones selon la valeur de leurs taux. On ne s'étonnera donc pas qu'il soit très corrélé à la dimension fractale calculée sur l'ensemble des échelles et à la densité (fig.

4.7b). Nous retrouvons visuellement cette corrélation sur la carte du premier axe (fig. 4.7a) qui reproduit le découpage classique entre les espaces urbains denses et les espaces ruraux, forestiers et montagnards peu denses.

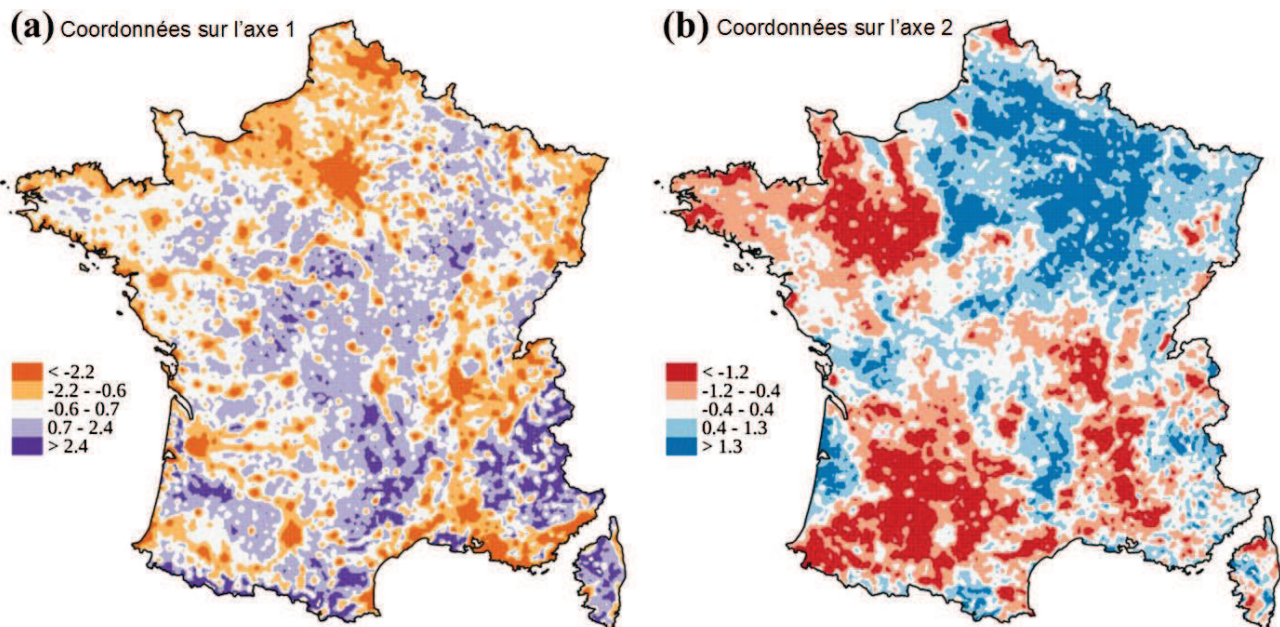


FIGURE 4.7 – Cartographie des deux premiers axes factoriels. Source : SÉMÉCURBE, TANNIER et ROUX 2019

Le deuxième axe (30% de l'inertie) est plus intéressant et décrit la façon dont les territoires dévient vis-à-vis d'une situation d'invariance d'échelle. Il oppose les indicateurs Sc_{100-50} , $Sc_{200-100}$ aux indicateurs $Sc_{400-200}$, $Sc_{800-400}$. Autrement dit, si nous contrôlons l'« effet taille » en prenant des espaces de mêmes dimensions fractales, à une valeur élevée sur le deuxième axe correspondent des Sc_{100-50} , $Sc_{200-100}$ élevés et des $Sc_{400-200}$, $Sc_{800-400}$ faibles, et réciproquement. De sorte que cet axe synthétise l'opposition entre les signatures calculées sur les plages extrêmes.

Dans les espaces urbains, les valeurs sur l'axe 2 sont faibles. Nous retrouvons le fait que les villes et les secteurs d'openfield pour les plages d'échelles de notre analyse (entre 50 et 800 m) sont approximativement invariants d'échelle. En revanche, l'amplitude des valeurs sur cet axe est grande dans les espaces de bocage de l'Ouest de la France.

Afin de simplifier encore notre observation des tissus bâtis et ainsi mieux comprendre leur organisation, nous avons réalisé une typologie de l'information contenue dans les deux axes de l'ACP. À cette fin, nous utilisons l'algorithme des nuées dynamiques (fig. 4.8). Cet algorithme de classification est itératif. À la première étape, on tire aléatoirement autant de centres que de classes que l'on souhaite obtenir. On affecte chaque point à classer au centre le plus proche. On recalcule de nouveaux centres correspondant aux barycentres des points classés. On itère le processus tant que des points changent de groupes. En sortie, les centres décrivent les caractéristiques moyennes des classes de points.

Le choix du nombre de classes est effectué à l'aide d'un arbitrage entre d'une part, l'hétérogénéité maximale que l'on autorise au sein des classes, c'est-à-dire le degré de ressemblance minimale des éléments d'une même classe, et d'autre part, la variabilité entre les classes que l'on juge pertinente, c'est-à-dire le niveau de détails que l'on souhaite conserver. Le scree plot (fig. 4.9) représente la somme des distances au carré de chaque point à son centre d'affectation en fonction du nombre de classes. L'ordonnée de ce

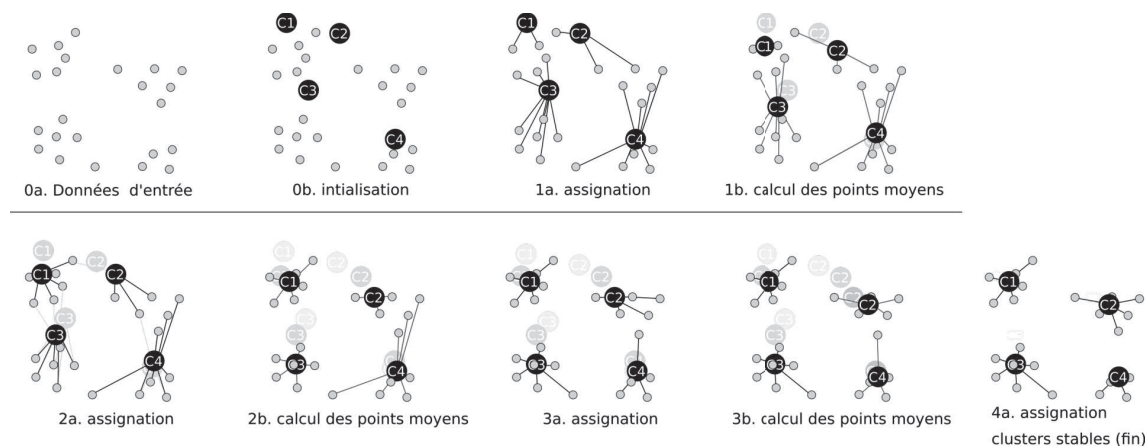


FIGURE 4.8 – Algorithme de calcul des nuées dynamiques (K-moyenne). Source : Wikipédia (fr.wikipedia.org/wiki/K-moyennes, consultée le 21/08/2019)

Remarque

THOMAS et al. 2010, dans l'objectif de classer des tissus bâtis, mesurent directement des distances deux à deux entre les signatures. La distance utilisée autorise la comparaison de taux de variation calculés à des échelles différentes. Au sein d'une même classe de leur typologie, les signatures se ressemblent modulo une translation horizontale. Par exemple, deux signatures qui décroissent puis qui croissent vont être considérées comme similaires même si leur minimum respectif n'est pas atteint au même endroit. Ceci revient à considérer comme comparables des textures dont les formes caractéristiques apparaissent à des échelles différentes.

graphique mesure l'homogénéité des éléments d'une même classe. La règle du coude est généralement employée pour choisir le nombre de classes et consiste à dire que celui-ci se situe approximativement au niveau où la courbure de la courbe est la plus forte. Autrement dit, au moment où le gain marginal en homogénéité lorsque l'on rajoute une classe supplémentaire devient moins important. Dans notre cas, le scree plot suggère un découpage en 6 classes.

La classe 1 regroupe les unités spatiales de fortes dimensions fractales qui correspondent approximativement aux villes (*fig. 4.10*). Les classes 2a et 3a correspondent à des organisations de bâtiments dispersés : de nombreux bâtiments se retrouvent en dehors des espaces villageois. Les bâtiments de la classe 2a occupent de façon plus uniforme l'espace que ceux de la classe 3a. À l'opposé, les classes 2b et 3b recouvrent des territoires où les bâtiments sont concentrés principalement dans les espaces villageois. Les villages dans la classe 2b sont plus gros et nombreux que dans la classe 3b. Les différences entre les éléments (a) et (b) de la nomenclature reflètent un positionnement spécifique sur l'axe 1 : les éléments des classes a ont des dimensions fractales supérieures à ceux appartenant à la classe b. La classe 4, caractérisée par de faibles dimensions fractales, correspond aux espaces très peu denses tels que les massifs forestiers et montagnards.

La carte de la typologie est très bien structurée alors même qu'aucune contrainte de continuité n'a été introduite dans la classification (*fig. 4.11*). Les unités spatiales appartenant à une même classe forment des ensembles connexes. Les classes intermédiaires 2a, 2b, 3a and 3b sont situées en périphérie des espaces urbains et traduisent une grande variabilité des formes périurbaines et rurales.

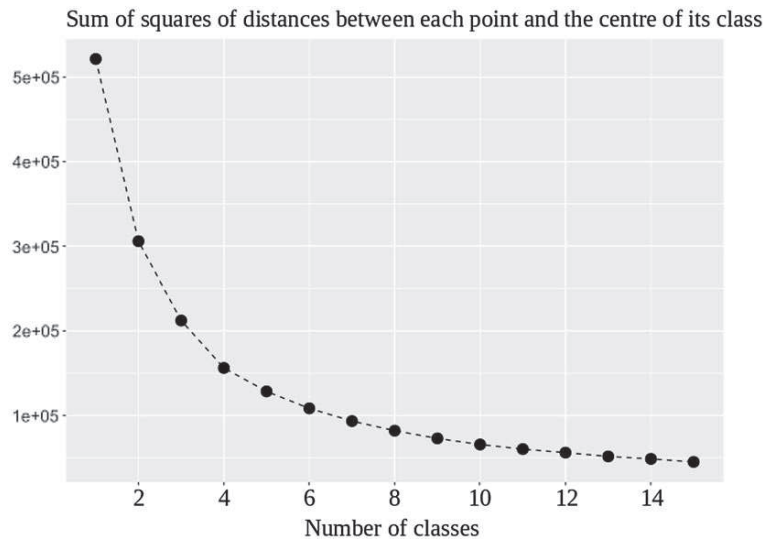


FIGURE 4.9 – Scree plot de la K-moyenne appliquée aux résultats de l'ACP présentée en figure 4.7. Source : SÉMÉCURBE, TANNIER et ROUX 2019

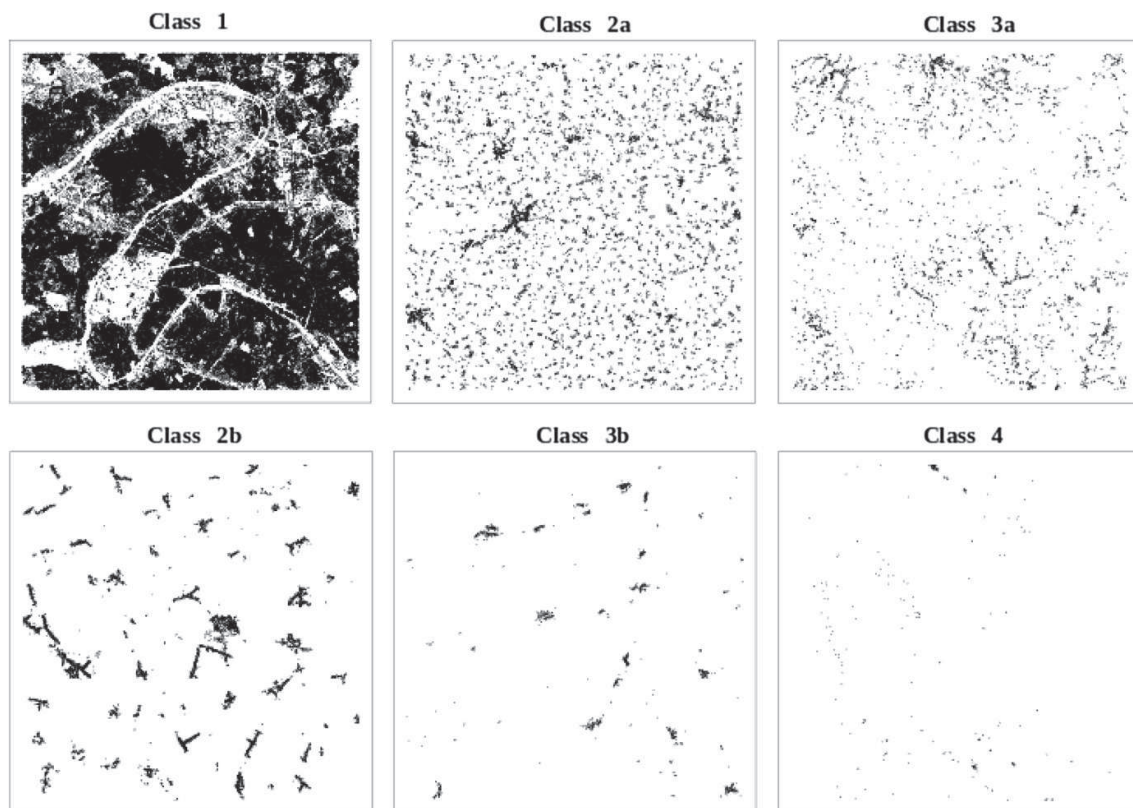


FIGURE 4.10 – Représentant idéaltypique de chaque classe issue de la typologie des signatures trans-échelles des espaces bâtis de France métropolitaine. Source : SÉMÉCURBE, TANNIER et ROUX 2019

La croissance des signatures dans les espaces de bocage s'explique simplement. Aux grandes échelles entre 50 et 100 m la probabilité de trouver autour de chaque bâtiment de nombreux bâtiments voisins est faible, les maisons sont principalement isolées les unes des autres : les taux de variation sont faibles. Au contraire, pour les petites échelles (400 et 800 m), chaque bâtiment est entouré de nombreux bâtiments : les taux sont alors élevés. Dans

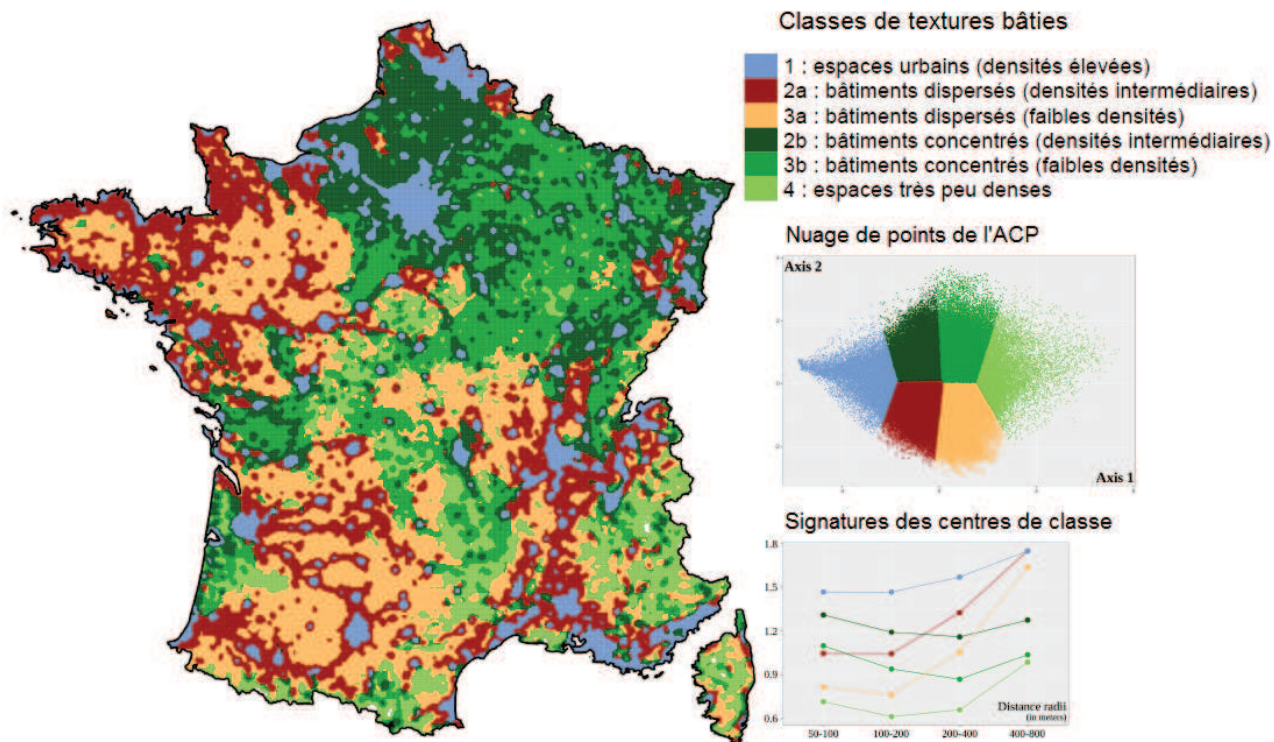


FIGURE 4.11 – Typologie de 145 178 tissus bâtis de France métropolitaine basée sur la caractérisation de la façon dont la distribution spatiale des bâtiments dévie localement d’une distribution invariante d’échelle. Source : SÉMÉCURBE, TANNIER et ROUX 2019 (séparer en deux pour faire deux figures)

les espaces d’openfield, les bâtiments étant agglomérés dans des villages de taille plus ou moins modeste, pour les grandes échelles chaque bâtiment à de nombreux voisins : les taux sont élevés. Ensuite, les signatures ont tendance à décroître progressivement jusqu’à un certain seuil de distance. En dessous de ce seuil, à chaque fois que l’on passe d’une échelle fine à une échelle plus grossière, la part relative de vide augmente autour de chaque bâtiment : les taux diminuent. Une fois dépassé ce seuil de distance, chaque bâtiment devient voisin de ceux des autres villages, ce qui explique que $Sc_{800-400}$ est plus élevé que $Sc_{400-200}$ dans les espaces de population concentrée.

4.3.3 Implications épistémologiques

Les analyses statistiques mises en œuvre ci-avant ont vocation, par construction, à produire des représentations de l’espace qui opposent les objets entre eux. Dans notre cas, ceci se manifeste par l’existence d’une opposition entre le premier axe, qui ordonne les territoires selon une « dimension fractale moyenne », et le deuxième axe qui sépare approximativement les objets selon la croissance-décroissance de leurs signatures trans-échelles. Ces oppositions franches détectées par l’ACP résultent de l’existence de formes de peuplement très différentes en proportions comparables.

L'approche statistique est très tributaire de l'espace étudié. Implicitement cette remarque met en question l'universalité des représentations statistiques obtenues. En effet, appliquer une ACP sur les signatures trans-échelles des tissus bâtis de l'ensemble de l'Europe ou bien sur un petit sous-ensemble morphologiquement homogène du territoire français, aurait produit des typologies potentiellement très différentes.

4.3.4 Comparaison des tissus bâtis locaux et de la forme des enveloppes morphologiques de 20 agglomérations urbaines françaises

Sur la base de cette typologie des tissus bâtis français, il a été possible d'évaluer l'influence des formes bâties périurbaines sur la forme globale des villes (SÉMÉCURBE, TANNIER et ROUX 2019). Pour ce faire, nous avons sélectionné des aires urbaines de taille moyenne. Ce choix était dicté par deux contraintes. La première était d'éviter de prendre les grandes aires urbaines qui contiennent parfois plusieurs unités urbaines morphologiquement séparées. La seconde était d'éviter de prendre des aires urbaines trop petites et donc de raisonner sur des espaces couverts par un très faible nombre de carreaux de 2000 m de côté, ce qui risquait de rendre particulièrement instables les résultats. Au final, 20 aires urbaines ont été retenues pour l'analyse, dont le nombre d'habitants s'échelonne entre 110 000 habitants pour Saint-Quentin et 898 000 pour Nantes.

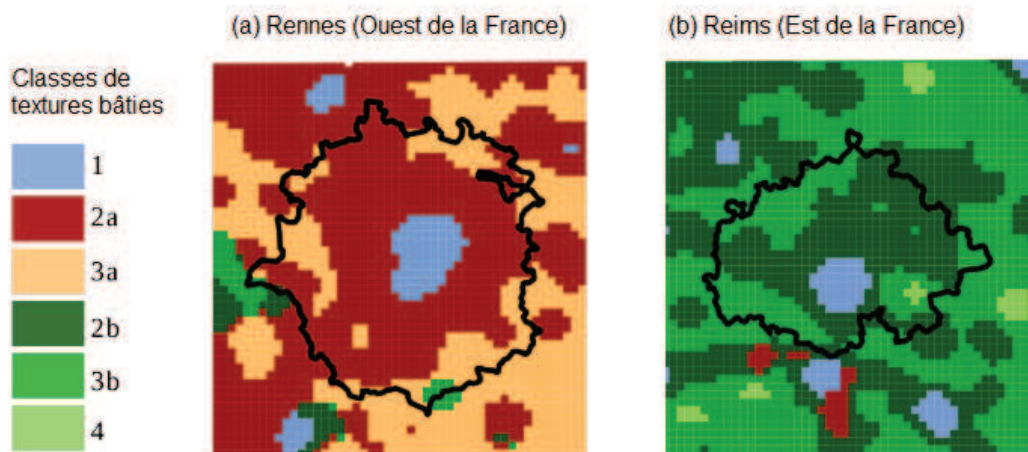


FIGURE 4.12 – Caractérisation des tissus bâtis de deux aires urbaines françaises : Rennes et de Reims. Le trait noir représente la délimitation de l'agglomération morphologique identifié au moyen de MorphoLim. Source : SÉMÉCURBE, TANNIER et ROUX 2019

Pour chaque aire urbaine, nous avons calculé un seuil de distance caractérisant une discontinuité dans l'organisation trans-échelle du bâti à l'aide du logiciel MorphoLim (TANNIER, THOMAS et al. 2011). Cette distance décrit au sein d'une ville la différence d'organisation morphologique entre les espaces urbains denses et les espaces périphériques périurbains. Elle varie fortement entre les aires urbaines et s'échelonne entre 78 m pour Avignon jusqu'à 529 m pour Troyes (tab. 4.1).

Nous observons que la proportion de types de peuplement concentrés ou dispersés dans les aires urbaines est fortement liée à la forme globale des agglomérations, soit dans notre cas au seuil de distance issu de MorphoLim (tab. 4.2). Plus les espaces périurbains sont composés de peuplement concentré et plus le seuil de distance est élevé. A l'inverse,

Nom de l'agglomération	Seuil de distance issu de Morpholim (m)	Surface de l'agglomération morphologique (km ²)	% de tissus bâtis localement dispersés (classes 2a et 3a)	% de tissus bâtis localement concentrés (classes 2b et 3b)
Avignon	78	2396	47	3
Angers	127	2744	87	2
Bayonne	143	1440	71	0
Le Mans	143	2364	89	0
Rennes	143	4236	90	2
Vannes	165	940	70	9
Nantes	171	3884	73	10
Caen	181	2168	51	26
Poitiers	207	2648	18	62
Clermond-Ferrand	242	2764	19	60
Niort	245	1652	20	69
Bourges	256	1912	9	79
Nancy	290	2832	2	82
Orléans	318	3572	33	48
Metz	389	1940	1	76
Amiens	395	2424	0	93
Dijon	415	3844	0	87
Saint-Quentin	420	952	0	88
Reims	434	2752	0	93
Troyes	529	2296	0	91

TABLE 4.1 – Caractéristiques morphologiques de 20 agglomérations urbaines. Source : SÉMÉCURBE, TANNIER et ROUX 2019

plus le peuplement périphérique est dispersé dans des bâtiments isolés, dans de petits villages ou des hameaux et plus le seuil est faible. Ceci confirme l'hypothèse que la forme des agglomérations urbaines françaises est profondément dépendante de celle de leurs espaces périphériques. En revanche, la forme globale des agglomérations urbaines n'est pas liée aux types de bâtis extrêmes que sont les classes 1 (espace urbain) et 4 (espace de très faible densité).

Variable	Coefficient de corrélation de Pearson
% d'unités spatiales appartenant aux classes 2a et 3a (tissu bâti localement dispersé)	-0.84***
% d'unités spatiales de la classe 2a	-0.84***
% d'unités spatiales de la classe 3a	-0.69***
% d'unités spatiales appartenant aux classes 2b et 3b (tissu bâti localement concentré)	0.89***
% d'unités spatiales de la classe 2b	0.81***
% d'unités spatiales de la classe 3b	0.62**
% d'unités spatiales appartenant à la classe 1 (tissu bâti du centre-ville)	-0.46*
% d'unités spatiales appartenant à la classe 4 (tissu bâti diffus)	0.10
Surface de l'agglomération morphologique	-0.03

TABLE 4.2 – Corrélations statistiques entre le seuil de distance caractérisant chaque agglomération morphologique et les types de tissus bâtis qu'elles contiennent, pour 20 aires urbaines françaises de taille moyenne. p-values *** p<0,001 ** p<0,01 * p<0,05. Source : SÉMÉCURBE, TANNIER et ROUX 2019

4.4 Décomposition des signatures trans-échelles au moyen de polynômes orthogonaux

4.4.1 Présentation théorique

Contrairement à la partie précédente, nous proposons à présent de décrire les signatures trans-échelles à l'aide d'une modélisation mathématique plus complexe. L'idée n'est plus de rechercher a posteriori des indicateurs qui résumeraient au mieux la variété des phénomènes observés, mais de décomposer *a priori* les signatures en composantes préalablement définies dont on maîtrise la signification thématique. L'intérêt de cette démarche est de produire des indicateurs qui ne soient pas tributaires des territoires étudiés.

FORRIEZ et NOTTALE 2010 modélisent la fonction d'échelle $M_r(i)$ sous la forme d'un polynôme de degré 2 :

$$M_r(i) = a_0 - a_1 \times \log(r) - a_2 \times (\log(r))^2 + \varepsilon \quad (4.2)$$

Où $M_r(i)$ est la fonction d'échelle et a_1 est la dimension fractale si et seulement si l'invariance d'échelle est vérifiée.

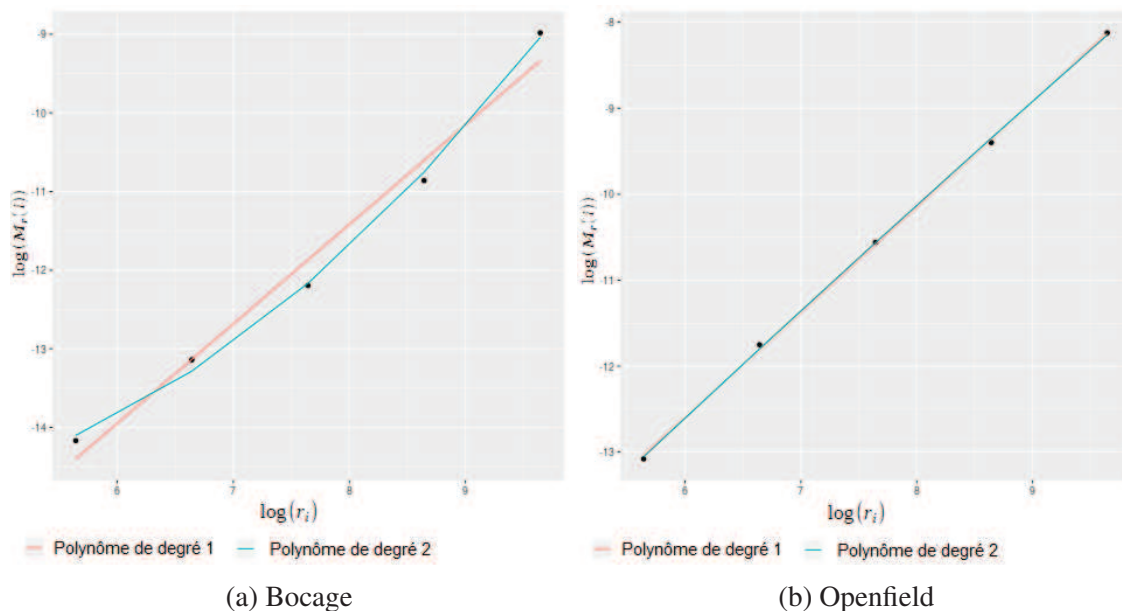


FIGURE 4.13 – Modélisation des fonctions d'échelles à l'aide d'un polynôme d'ordre 1 et d'un polynôme d'ordre 2

Cette formule revient à introduire dans les méthodes d'estimation usuelles des dimensions fractales un terme en $\log(r)^2$. Ce nouveau terme permet de capter l'écart à une situation d'invariance d'échelle stricte. Il a pour effet d'améliorer mécaniquement l'ajustement des fonctions d'échelles des configurations idéaltypiques, au moins pour la configuration de bocage (fig 4.13). Dans le cas de la configuration d'openfield la propriété d'invariance étant approximativement vérifiée, l'amélioration est marginale.

Dans le tableau 4.3, les différences de valeur de a_1 entre le modèle classique de polynôme de degré 1, et le modèle polynomial de degré 2 (0.669 contre 1.265 dans le bocage et 1.264 contre 1.226 dans l'openfield) s'expliquent par la colinéarité

Paramètre	polynôme de degré 1		polynôme de degré 2		polynômes orthogonaux	
	Bocage	Openfield	Bocage	Openfield	Bocage	Openfield
a_1	1.265	1.226	0.669	1.264	1.265	1.226
a_2	-	-	1.968	-0.124	0.558	-0.035

TABLE 4.3 – Estimation de la dimension fractale selon différents modèles statistiques

entre les vecteurs $\log(r) = (\log(50), \log(100), \log(200), \log(400), \log(800))$ et $\log(r)^2 = (\log(50)^2, \log(100)^2, \log(200)^2, \log(400)^2, \log(800)^2)$. Leur coefficient de corrélation est supérieur à 0.95, la régression ne réussit pas à séparer les effets que l'on peut attribuer à chacun d'entre eux : $\log(r)^2$ « capture » une partie du pouvoir explicatif apporté par $\log(r)$. Il en résulte que a_1 ne peut plus être considéré comme une dimension fractale.

KENNEDY et GENTLE 2018 suggèrent de recourir à une régression orthogonale pour garantir la stabilité des estimations des paramètres indépendamment du degré des polynômes introduit dans le modèle. L'idée de cette méthode consiste à remplacer $\log(r)$ et $\log(r)^2$ dans l'équation précédente par des combinaisons linéaires de ceux-ci, appelés polynômes orthogonaux, disposant de bonnes propriétés mathématiques d'orthogonalité :

$$M_r(i) = a_0 + a_1 \times P_1(\log(r)) + a_2 \times P_2(\log(r)) + \varepsilon \quad (4.3)$$

La première combinaison est un polynôme de degré 1 en $\log(r)$, $P_1(\log(r)) = b_{1,0}\vec{1} + \log(\vec{r})$, la deuxième un polynôme de degré 2 : $P_2(\log(r)) = b_{2,0}\vec{1} + b_{2,1}\log(\vec{r}) + \log(\vec{r})^2$, ou $b_{1,0}$, $b_{2,0}$ et $b_{2,1}$ sont des constantes que l'on détermine grâce aux équations d'orthogonalité suivantes :

$$P_1(\log(\vec{r})) \cdot \vec{1} = 0 \quad (4.4)$$

$$P_2(\log(\vec{r})) \cdot \vec{1} = 0 \quad (4.5)$$

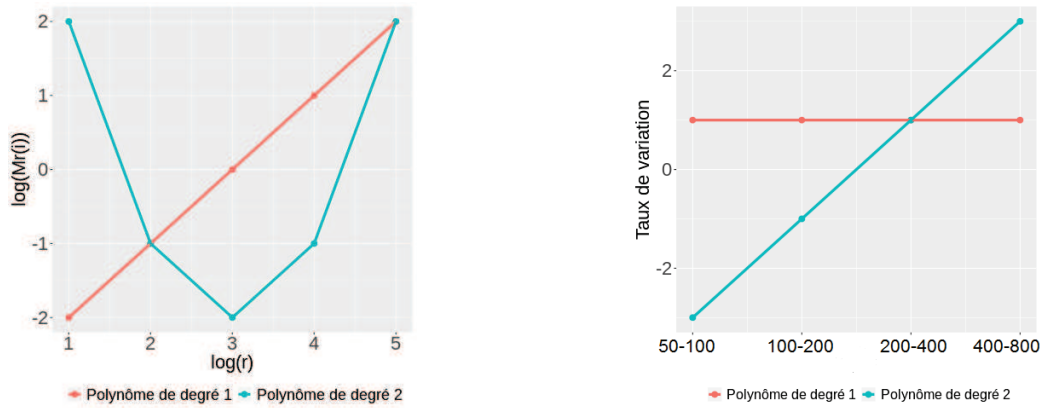
$$P_1(\log(\vec{r})) \cdot P_2(\log(\vec{r})) = 0 \quad (4.6)$$

La notation \cdot désigne le produit scalaire usuel entre deux vecteurs. Ici nous considérons 5 échelles (r), pour mémoire 50, 100, 200, 400 et 800 m. Ce faisant, le produit scalaire $P_1(\log(\vec{r})) \cdot P_2(\log(\vec{r}))$ est égal à $P_1(\log(50)) * P_2(\log(50)) + P_1(\log(100)) * P_2(\log(100)) + P_1(\log(200)) * P_2(\log(200)) + P_1(\log(400)) * P_2(\log(400)) + P_1(\log(800)) * P_2(\log(800))$.

L'algorithme de Gram-Schmidt est généralement utilisé pour résoudre ce système d'équations, ce qui permet d'obtenir les valeurs de $P_1(\log(\vec{r})) = (-2, -1, 0, 1, 2)$ et $P_2(\log(\vec{r})) = (2, -1, -2, 1, 2)^3$. Alors que la courbe $P_1(\log(r))$ est juste un décalage en ordonnée de la courbe $\log(r)$, $P_2(\log(r))$ est une parabole centrée sur l'échelle médiane 200 m (fig 4.14).

Les fonctions d'échelle estimées à l'aide de la modélisation avec des polynômes orthogonaux (équation 4.3) sont identiques à celles obtenues par l'introduction d'un polynôme de degré 2 (équation 4.2). En revanche, la signification des coefficients a_1 et a_2 diffère. Dans le cas d'une modélisation avec des polynômes orthogonaux a_1 correspond à la dimension fractale même si l'invariance d'échelle n'est pas vérifiée et a_2 est un indicateur trans-échelle dont nous allons approfondir la signification dans la partie suivante.

3. Les constantes $b_{1,0}$, $b_{2,0}$ et $b_{2,1}$ n'ont pas d'intérêt en soi.



(a) Courbes des polynômes orthogonaux de degré 1 et 2

(b) Taux de variation des polynômes orthogonaux de degré 1 et 2

FIGURE 4.14 – Polynômes orthogonaux calculés sur la plage d'échelle 50 – 800 m

4.4.2 Relation entre la décomposition au moyen de polynômes orthogonaux et les signatures trans-échelles

Dans le cas de la modélisation par polynômes orthogonaux, si l'on néglige le terme d'erreur ε dans l'équation 4.3, les taux de variation des $M_r(i)$ estimés à la base des signatures trans-échelles se décomposent ainsi :

$$Sc_{r_a-r_b} = \text{Dimension fractale estimée} + \text{Ecart à l'invariance d'échelle} \quad (4.7)$$

D'un point de vue formel, les taux de variation entre échelles correspondent à des dérivées (fig. 4.14.b.)

- Dimension fractale estimée = a_1 car $a_1 \times P_1(\log(r))$ étant un polynôme de degré 1, ses taux de variation entre deux échelles sont constants et strictement égaux à la dimension fractale moyenne calculée sur l'ensemble des échelles. Sa signature trans-échelle est une droite horizontale.
- Écart à l'invariance d'échelle = $a_2 \times Sc(P_2(\log(r)))$. $P_2(\log(r))$ étant un polynôme de degré 2, sa signature est un polynôme de degré 1, c'est à dire une droite croissante.

Le premier terme caractérise le comportement trans-échelle moyen de l'objet. Lorsque l'objet est quasi-invariant d'échelle, celui-ci correspond exactement à la dimension fractale. Le deuxième terme indique comment change ou ne change pas la dimension fractale à travers les échelles. Le signe de a_2 décrit le sens global d'évolution des signatures à travers les échelles : un signe positif est associé à une croissance des signatures et négatif à une décroissance de celle-ci⁴. La valeur absolue de a_2 traduit l'intensité de l'écart à une situation d'invariance d'échelle.

4. Le coefficient a_2 résume totalement l'information contenue dans les signatures trans-échelles dont les variations sont régulières (c'est-à-dire les signatures dont les points sont parfaitement alignés). Dans ce cas, a_2 est proportionnel à la variation entre deux taux de variation consécutifs. Pour les signatures non régulières, a_2 quantifie la variation moyenne entre les taux.

La figure 4.15 illustre graphiquement ceci. La signature est la succession des taux de variation $Sc_{r_a-r_b}$ calculés sur toutes les échelles. La modélisation de cette signature au moyen des polynômes orthogonaux permet d’extraire la dimension fractale estimée, qui correspond au coefficient a_1 , et un indicateur d’écart à l’invariance d’échelle centré graphiquement autour de a_1 ⁵.

Dans le cas de la configuration de bocage (fig 4.15.a), la valeur de $a_2 = 0.558$ est importante et reflète de façon relativement satisfaisante la croissance des signatures. En revanche, pour la configuration d’openfield (fig 4.15.a), $a_2 = -0.035$ est quasiment nul : l’ajout d’un polynôme d’ordre 2 n’apporte donc quasiment rien. On retrouve ici le fait que cette configuration d’openfield est, pour les plages de rayons d’analyse considérées, presque parfaitement fractale.

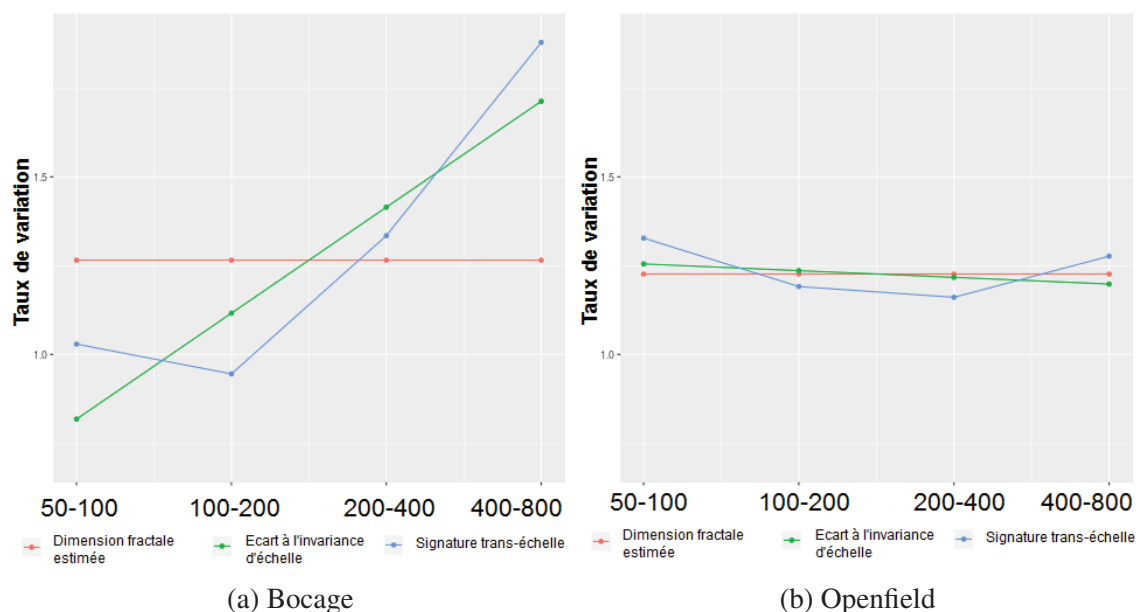


FIGURE 4.15 – Modélisation des signatures trans-échelles de deux configurations bâties idéaltypiques à l’aide de polynômes orthogonaux

Finalement, l’intérêt d’une décomposition par polynôme orthogonaux réside moins dans ses propriétés statistiques à proprement parler que dans l’interprétabilité de ces résultats.

4.4.2.1 Mise en œuvre sur le bâti de la France métropolitaine

La distribution spatiale des coefficients du deuxième polynôme a_2 (fig 4.16) est quasiment identique à celle des coordonnées sur l’axe 2 de l’ACP calculée sur les signatures trans-échelles (fig. 4.11) : le coefficient de corrélation est de 0.96. Ce phénomène s’explique par le fait que l’axe 2 oppose les taux de variation entre les grandes et les petites échelles et indique ainsi l’évolution globale de ceux-ci à travers les échelles.

La valeur absolue de a_2 est très importante dans les espaces de bocage de l’Ouest de la France, l’invariance d’échelle n’y est pas vérifiée, et quasi-nulle dans les territoires d’openfield de l’Est de la France où la modélisation fractale est satisfaisante. Ces résultats sont cohérents avec ceux de la distribution des coefficients de détermination de l’analyse fractale faite avec la GWFA pour $q = 0$ (fig. 4.1.b).

5. Si l’on ne fait pas cela, l’écart à l’invariance d’échelle serait centré autour de 0.

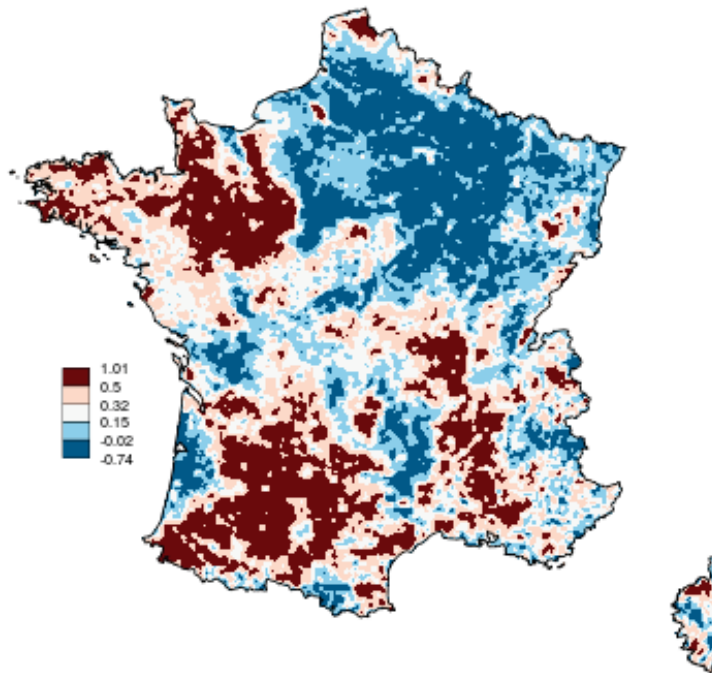


FIGURE 4.16 – Décomposition en polynômes orthogonaux des signatures trans-échelles calculées avec la GWFA sur la plage d'échelles 50-800 m pour $q = 0$: valeur du coefficient a_2

4.4.2.2 Mise en œuvre sur le bâti de la petite couronne parisienne

Nous nous interrogeons ici sur la capacité de l'analyse trans-échelle à décrire d'autres types de configurations bâties que celle de la France métropolitaine dans son ensemble. Pour répondre à cette interrogation, nous avons appliqué l'analyse trans-échelle au tissu bâti des quatre départements de la petite couronne parisienne (Paris, Hauts-de-Seine, Seine-Saint-Denis et Val-de-Marne). À ce niveau d'analyse, nous ne pouvons pas nous satisfaire d'une approximation des bâtiments par leur centroïde. La GWFA étant dédiée aux données ponctuelles, nous avons été contraints de transformer le fond de carte vectoriel en une image rasterisée à 1 mètre de résolution. L'ensemble des pixels non vides de cette image forme notre ensemble de points. Afin d'observer la variabilité interne de la morphologie du bâti parisien, nous avons opté pour la plage d'échelles suivante : 5 m, $5\sqrt{2}$ m, 10 m, $10\sqrt{2}$ m, 20 m, $20\sqrt{2}$ m et 40 m. Contrairement à l'analyse menée sur la France métropolitaine, cette plage d'échelles suit une progression géométrique de raison $\sqrt{2}$. Nous avons fait ce choix afin de décrire finement les spécificités locales en sachant que nous sommes contraints par l'échelle supérieure qui conditionne le rayon de lissage de l'analyse (400 m). Nous avons concentré notre analyse sur les espaces suffisamment denses à l'intérieur de l'unité urbaine de Paris. En effet, dans les zones de faible densité à la limite de la zone agglomérée le nombre de bâtiments dans des rayons de 400 m est trop faible.

Les coefficients de détermination de l'estimation classique de la dimension fractale sont très élevés (supérieurs à 0.99) dans les espaces denses (fig. 4.17). À l'intérieur de Paris intramuros, la grande majorité des R^2 sont supérieurs à 0.9999. On retrouve ici le fait que dans les villes, les textures bâties sont invariantes d'échelle. La dimension fractale est proche de 2 au sein de la commune de Paris puis décroît avec l'éloignement au centre.

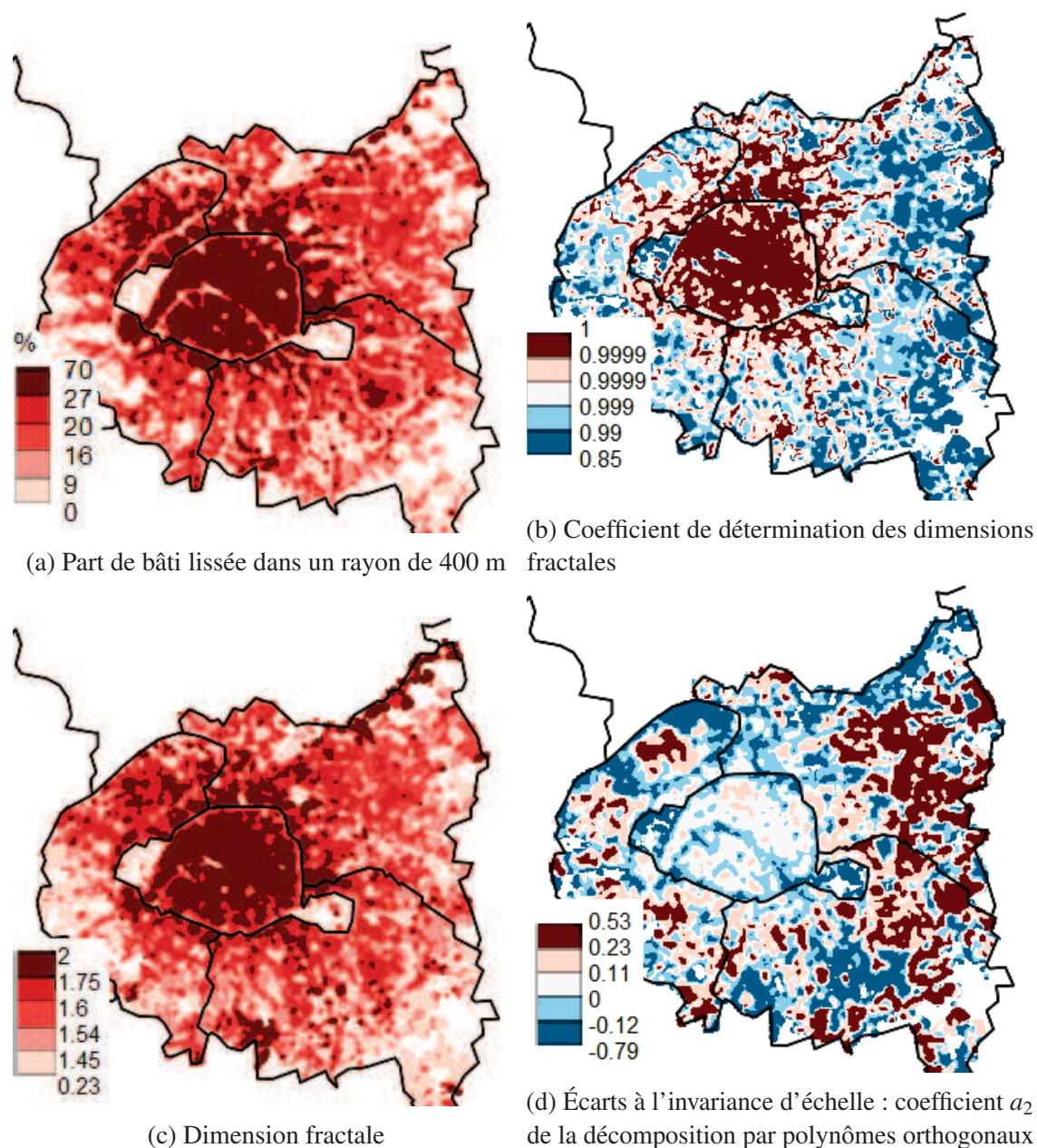
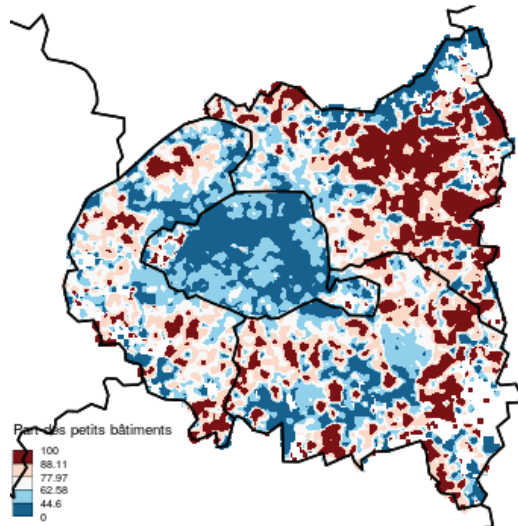


FIGURE 4.17 – Analyse trans-échelle (GWFA) du bâti de la petite couronne parisienne. Une discrétisation par quantiles a été appliquée à chaque carte.

Globalement, l'information apportée par la dimension fractale est très similaire à celle de la densité lissée. La corrélation entre le logarithme de la densité et la dimension fractale est de 0.85.

L'écart à l'invariance d'échelle (coefficient a_2) est quasi-nul dans Paris intramuros ce qui est cohérent avec les remarques faites précédemment. Dans la banlieue, il varie fortement et alterne entre des valeurs positives et négatives. Globalement il est corrélé, quoique moins fortement que la dimension fractale, avec le logarithme de la densité lissée (coefficient de corrélation égal à 0.4). Une étude visuelle des textures bâties permet de proposer une explication à l'organisation spatiale du coefficient a_2 : les valeurs élevées caractérisent des quartiers de petites maisons individuelles et les valeurs faibles caractérisent des quartiers comprenant principalement de grands bâtiments. Afin d'approfondir l'étude

FIGURE 4.18 – Part des bâtiments de surface au sol inférieure à 140 m^2

de ces relations, nous avons utilisé un indicateur morphologique simple : la proportion locale de bâtiments de petite taille. Nous avons choisi comme seuil pour définir les bâtiments de petite taille, le troisième quartile des surfaces de bâtiments calculées sur l'ensemble de la zone, soit 140 m^2 (fig 4.18). Le coefficient de corrélation entre les deux indicateurs (proportion de bâtiments de petite taille et coefficient a_2) est de 0.62. Nous avons ensuite estimé la régression suivante : $a_2 \approx \log(\text{densit}) + \text{part des bâtiments de petite taille}$. Avec un R^2 de 0.6, l'écart à l'invariance d'échelle se révèle bien expliqué par cet indicateur morphologique simple, mono-échelle.

Contrairement à THOMAS et al. 2010 qui réussissaient à décrire et caractériser finement une grande variété de quartiers, l'apport conceptuel de la GWFA et de l'analyse trans-échelle à l'étude de la morphologie parisienne semble moindre : nos résultats sont comparables à ceux obtenus avec des analyses simples basées sur des indicateurs de la géométrie euclidienne. Néanmoins, les deux études ne sont pas directement comparables. Celle de THOMAS et al. 2010 est basée sur le choix préalable d'une collection de quartiers provenant de différentes villes et examinés ensuite au moyen d'une analyse trans-échelle. Ce faisant, THOMAS et al. 2010 n'explorent pas l'intégralité d'une ville mais extraient des quartiers morphologiquement homogènes d'un environnement bâti urbain potentiellement non homogène. Sans cette action de sélection manuelle, il n'est pas certains que leurs résultats soient aussi pertinents.

Dans notre étude, la corrélation observée entre le coefficient a_2 et l'indicateur morphologique simple de proportion de bâtiments de petite taille ne témoigne pas nécessairement du fait que ce type d'analyse est sans intérêt et apporte peu d'informations supplémentaires sur l'organisation de l'espace. Notre choix du seuil de taille des bâtiments a été défini après plusieurs essais : l'idée étant de conserver celui qui maximise la corrélation avec l'indicateur trans-échelle. En procédant de la sorte, nous ne mettons pas en opposition les analyses trans-échelle et mono-échelle, mais utilisons les premières pour nous guider dans le choix des paramètres des secondes. L'avantage des indicateurs mono-échelle étant leur simplicité qui facilite leur appropriation auprès des acteurs publics et de la société civile.

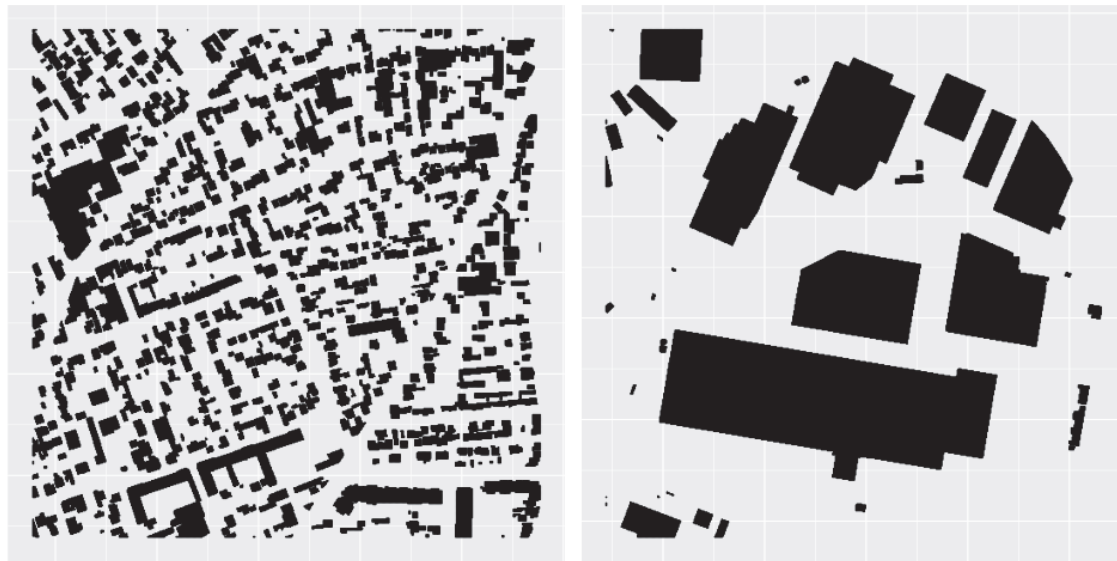
L'algorithme de la GWFA se décompose en deux phases : dans la première est calculée pour chaque point son nombre de voisins pour les différents rayons de voisinage, puis cette information est agrégée localement sous la forme d'une dimension fractale. Autrement

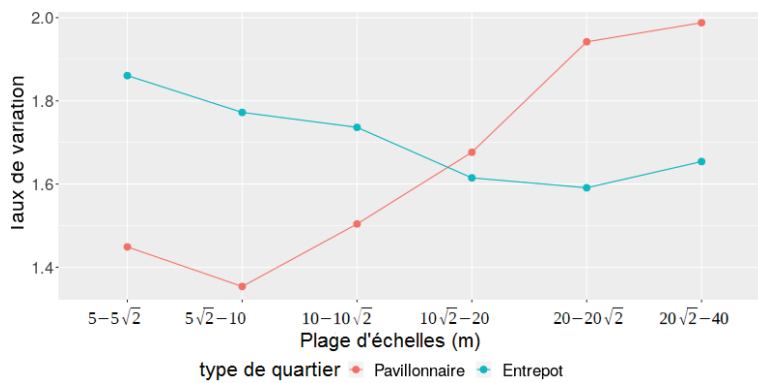
dit, les résultats en sortie de la GWFA sont en quelque sorte une moyenne des dimensions locales. KANEVSKI et PEREIRA 2017 utilisent eux directement des dimensions locales pour explorer l'organisation spatiale des feux de forêts. ARALDI et FUSCO 2018 utilisent également des dimensions locales pour caractériser la distribution spatiale de commerce de détail. Leur calcul de dimensions locales est effectué sur la base de distances mesurées sur le réseau viaire et non de distances à vol d'oiseau.

Avec plus de 141 millions de points (pixels) que comporte la carte du bâti de la petite couronne à partir de laquelle nous travaillons, nous ne sommes pas en mesure de représenter une carte des dimensions locales. Nous avons donc sélectionné deux quartiers de la commune de Colombes située au Nord-Ouest de Paris pour explorer à une échelle fine leur tissu bâti (fig 4.19). Le premier quartier est un quartier pavillonnaire de dimension fractale globale égale à 1,63, et dont l'écart à l'invariance d'échelle (coefficient a_2) est 0,32. Sa signature trans-échelle est croissante à partir de la plage de rayons $5\sqrt{2} \sim 10$ m. Le second quartier est une zone d'entrepôts de dimension fractale globale égale à 1,7 et d'écart à l'invariance d'échelle (coefficient a_2) égal à $-0,12$. Sa signature trans-échelle est légèrement décroissante.

Nous avons calculé pour chaque pixel de nos images rastérisées à 1 m, la dimension locale et l'écart à l'invariance d'échelle (coefficient a_2) (fig. 4.19). Globalement, au sein du quartier pavillonnaire, les dimensions locales sont plus faibles (< 1.5) que celles du quartier d'entrepôts (> 1.5). Les dimensions fractales globales reflètent ces différences locales : la dimension fractale du quartier pavillonnaire est plus faible que celle du quartier d'entrepôts. En revanche, les écarts à l'invariance d'échelle locale sont nettement plus importants pour le quartier pavillonnaire que pour le quartier d'entrepôts, ce que confirment également les analyses globales. Sur la figure 4.20, a été représentée la signature d'un pixel situé au centre d'une maison individuelle dans le quartier pavillonnaire de Colombes. A très grande échelle la signature décroît : les disques de voisinage sont plus grands que l'étendue spatiale de la maison individuelle, les taux de variation reflètent le vide entre les bâtiments. Lorsque les rayons sont suffisamment importants, les voisinages intersectent les autres bâtiments : les taux de variations croissent. La tendance générale est à l'accroissement des taux de variation comme l'indique la valeur du coefficient a_2 .

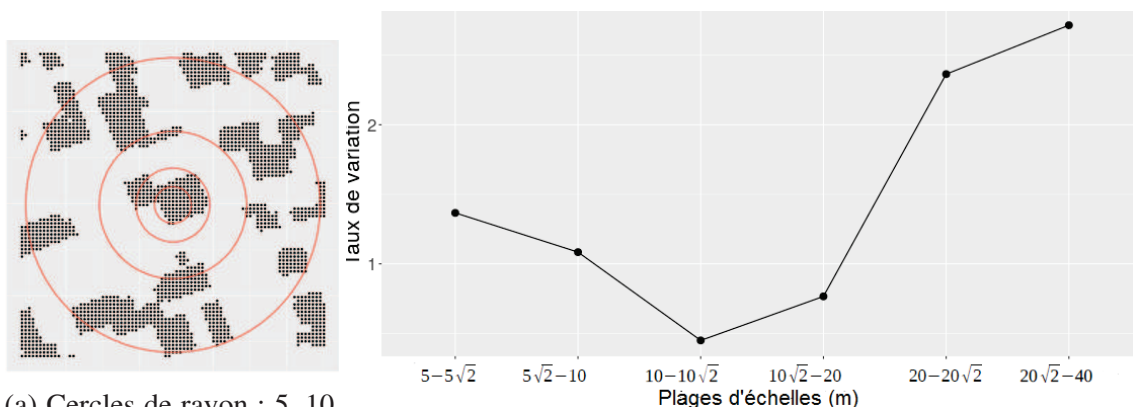
L'étude de ARALDI et FUSCO 2018 sur le commerce de détail porte sur des données (l'emplacement des magasins) que l'on peut considérer ponctuelles au regard des échelles d'analyse utilisées. Dans notre application, les données sont surfaciques : un même bâtiment est décrit par plusieurs points. Nos analyses locales sont donc très redondantes car deux points d'un même bâtiment ont de grandes chances de suivre des comportements similaires (fig. 4.21). Pour autant, le recours à des analyses locales permet de mieux comprendre les indicateurs agrégés. La signature locale d'un bâtiment croît puis décroît selon une forme en V qui dépend de sa taille et de sa distance par rapport aux bâtiments voisins (fig 4.20.b). L'écart à l'invariance d'échelle locale caractérise cette évolution et la moyenne de celui-ci sur une zone, l'écart à l'invariance globale, reflète la taille et la position relative moyenne des bâtiments. Contrairement à l'étude ARALDI et FUSCO 2018, nous ne filtrons pas les phénomènes observés selon leur présumée invariance d'échelle, mais considérons que l'écart à cette situation est signifiant et pertinent.


 (a) Pavillonnaire : $a_1 = 1.63$ et $a_2 = -0.32$

 (b) Entrepôts : $a_1 = 1.7$ et $a_2 = -0.12$


(c) Signatures trans-échelles des deux quartiers

FIGURE 4.19 – Deux quartiers de banlieue parisienne (commune de Colombes) et leur signature trans-échelle. Chaque quartier est inclus dans un carré de 500 m de côté. Source : BD Topo 2017



(a) Cercles de rayon : 5, 10, 20 et 40 m.

(b) Signature du pixel sélectionné.

 FIGURE 4.20 – Calcul de la signature trans-échelle d'un pixel appartenant à une maison individuelle du quartier pavillonnaire de Colombes. La dimension locale est de 1.31 et $a_2 = 0.75$. Source : BD Topo 2017

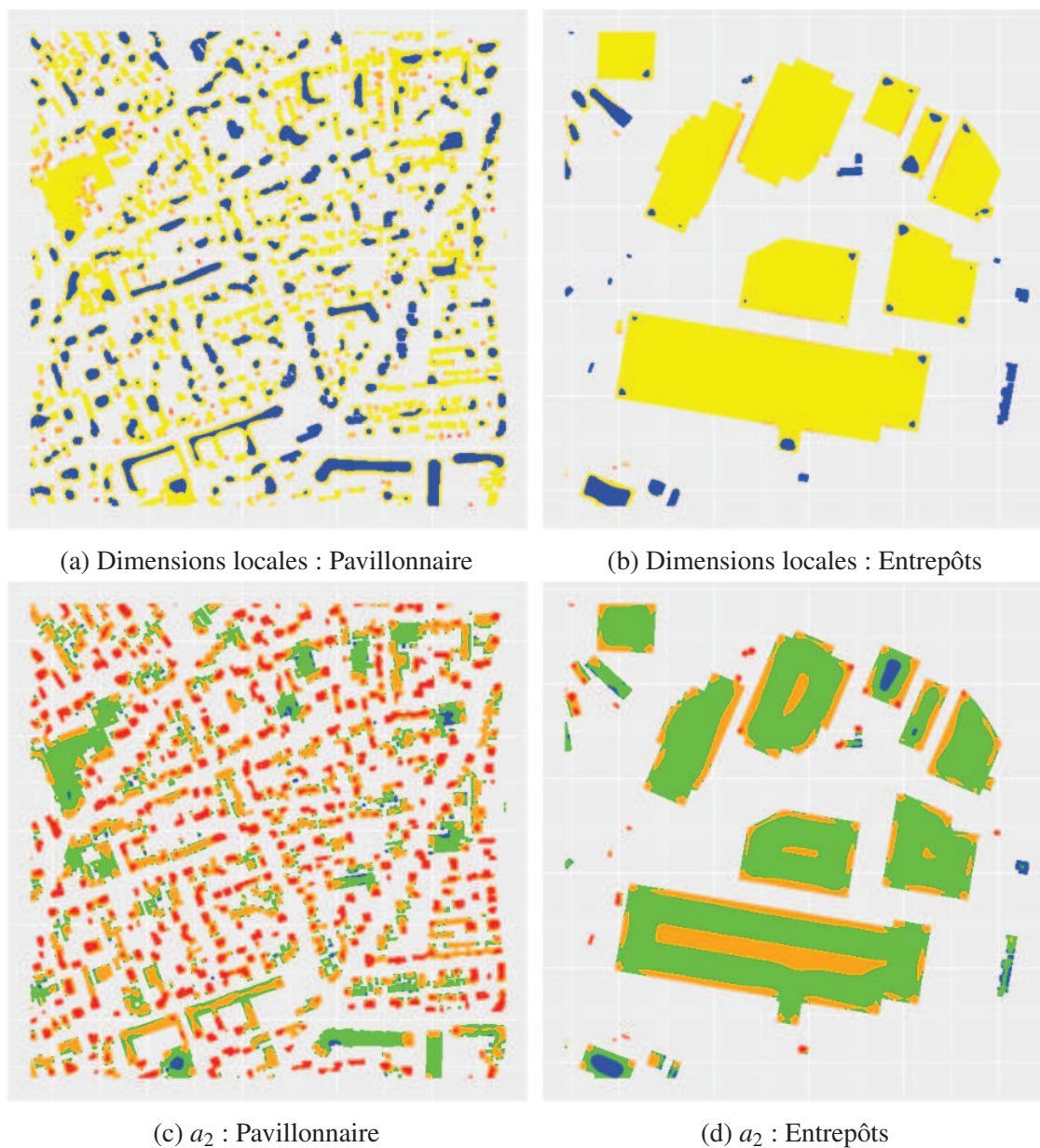


FIGURE 4.21 – Analyse trans-échelle des deux quartiers de Colombes. En haut les dimensions locales et en bas les écarts à l’invariance d’échelle. Légende pour les dimensions locales : en bleu, les dimensions locales inférieures à 1.5, en jaune, comprises entre 1.5 et 2, en orange et rouge, les dimensions supérieures à 2. Légende pour le coefficient a_2 : en bleu les valeurs inférieures à -0.5 , en vert comprises entre -0.5 et 0 , en orange entre 0 et 0.5 et en rouge supérieures à 0.5 . Source : BD Topo 2017

4.5 Simulation de formes bâties rurales à l'aide de l'analyse trans-échelle

L'application de la géométrie fractale à l'étude des espaces urbains a montré son efficacité pour penser l'aménagement des villes (Frankhauser 2018). L'absence de mécanismes sociaux ou physiques garantissant l'invariance d'échelle, n'est nullement une limite dans ce cas. La fractalité sert avant tout à filtrer et ordonner les formes structurantes des tissus bâtis, autrement dit à les modéliser. Des outils tels que MUP-City (TANNIER, VUIDEL et al. 2010) permettent d'identifier des zones prioritaires de développement résidentiel selon différents critères en se basant sur des logiques d'emboîtement d'échelles fractales. De façon analogue, on peut se demander si l'analyse trans-échelle ne pourrait pas servir également à proposer des schémas de développement dans les espaces ruraux. Cette question, qui dépasse très largement le cadre de cette thèse, peut servir néanmoins de point de départ à une réflexion plus spécifique sur l'utilisation de l'analyse trans-échelle pour simuler des configurations qui ressembleraient à celles que nous avons observées dans le présent chapitre.

À cette fin, nous nous sommes inspirés des cascades trans-échelles présentées en section (3.3.1.1). Nous considérons dans un premier temps un territoire carré de 16 000 m de côté (comme nos configurations idéaltypiques). Nous subdivisons ce territoire en une grille de 10×10 carreaux. Ces carreaux de 1600 m sont analogues aux disques de rayon de 800m utilisés dans nos analyses. Le taux de variation entre 800 et 400 nous indique qu'en moyenne chaque carreau se subdivise en 4 carreaux dont seulement $2^{S_{c_{400-800}}}$ sont occupés. Cette valeur n'étant pas entière, nous calculons le nombre global de sous carreaux occupés $100 \times 2^{S_{c_{400-800}}}$ que nous répartissons aléatoirement sur l'ensemble de la zone. On réitère récursivement cette procédure sur les nouveaux carreaux afin de simuler l'ensemble des signatures trans-échelles. Contrairement aux cascades trans-échelles classiques dans notre cas, la part des espaces considérés comme bâtis entre deux échelles successives n'est pas constante. Autrement dit, nous cherchons à produire explicitement des situations d'écart à l'invariance d'échelle.

Les configurations simulées d'openfield et de bocage (fig 4.22) ont des signatures identiques à leurs homologues observées. L'aspect dispersé et concentré des points simulés est indiscutable, pour autant leur organisation s'accorde mal avec les configurations observées. Dans les deux simulations, on observe un manque de structure de grande ampleur correspondant à des noyaux villageois ou à des petites villes. Ce déficit s'explique par l'aspect simpliste des mécanismes de construction inter-échelles mis en œuvre. On part d'une échelle grossière pour définir, en itérant à travers les échelles, des détails de plus en plus en plus fins. En se référant à ce schéma, on postule implicitement un sens des interactions : seules les échelles grossières agissent sur les échelles plus fines.

La notion de dépendance d'échelle, utilisée en aménagement du territoire, permet de penser des systèmes dynamiques dont les échelles interagissent entre elles :

« Dans ce cas, on considère que le mécanisme génératif, à l'origine de la dépendance d'échelles est la superposition et l'imbrication de multiples voisinages spatiaux (appartenance à une même maille territoriale, proximité géométrique ou temporelle) ou autres (voisinage social, culturel...), dont chacun définit l'enveloppe d'une catégorie d'interactions potentielles » TANNIER 2020.

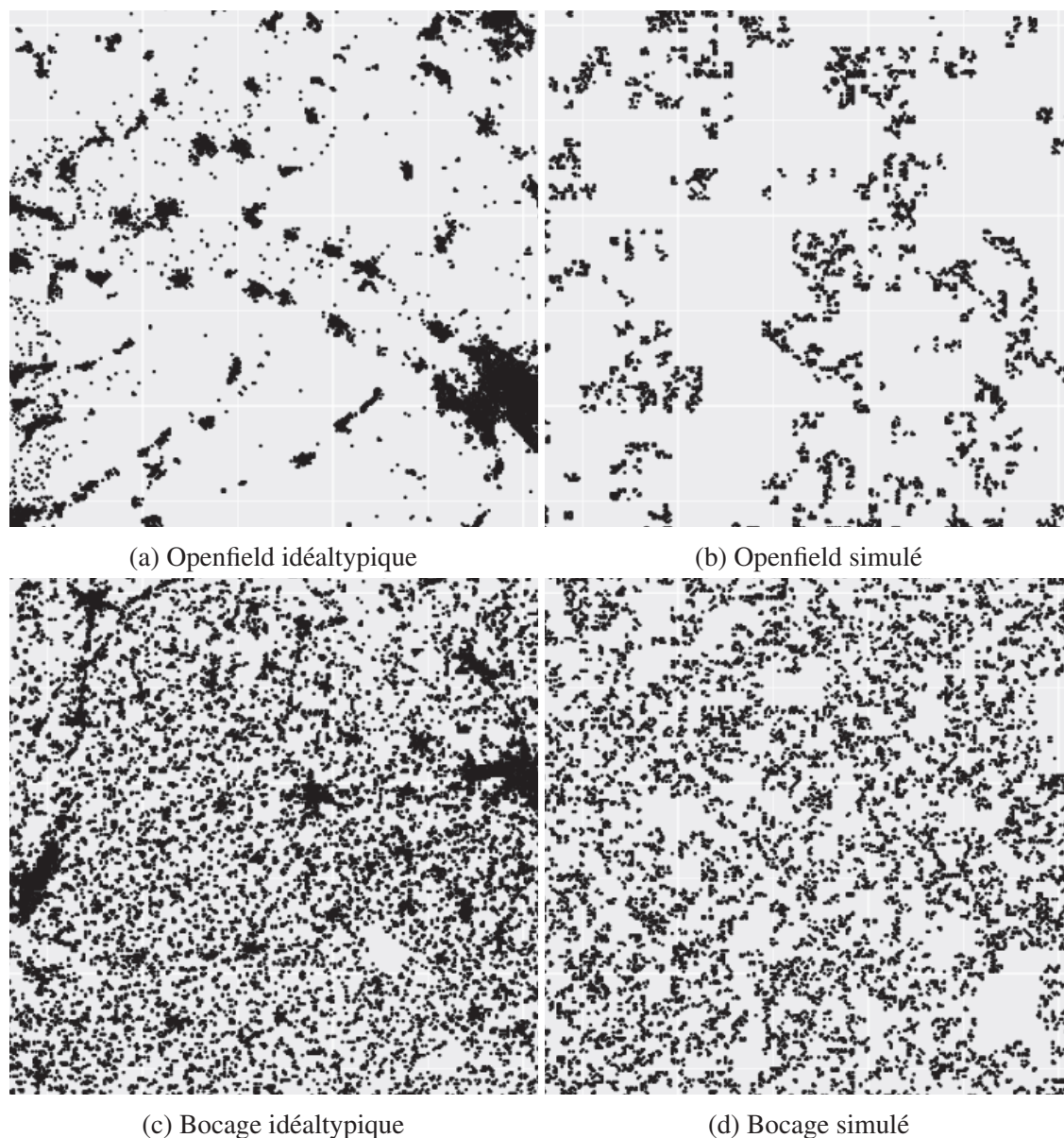


FIGURE 4.22 – Comparaison entre les configurations simulées et observées

Le modèle MobiSim-DR (TANNIER, BOURGEOIS et FRÉMOND 2017), dédié à la simulation prospective de nouveaux bâtiments et logements dans une région urbaine formalise explicitement ce type de relations entre les échelles (fig 4.23). Les règles de création des logements et des bâtiments relèvent de trois niveaux d'organisation : macro, méso et micro-géographiques. La demande de logements à l'échelle macro-géographique détermine le nombre de logements à créer dans les agrégats méso-géographique qui détermine à son tour à l'échelle micro-géographique, les cellules spatiales (l'unité spatiale la plus fine) devant être urbanisées. En retour, un processus non présenté sur la figure 4.23 relie l'échelle micro à l'échelle méso : « Chaque année de simulation, les nouvelles constructions réduisent le nombre de cellules potentiellement urbanisables au niveau micro-géographique et modifie la configuration spatiale de niveau méso-géographique. Ceci contraint en retour la localisation des nouveaux bâtiments créés » TANNIER 2020.

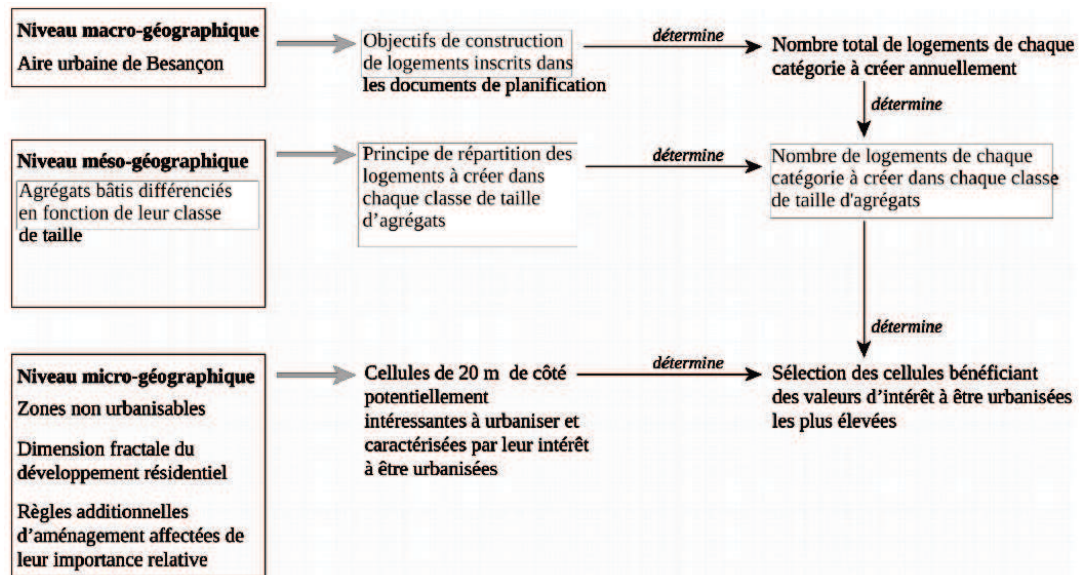


FIGURE 4.23 – Modélisation trans-échelle dans MobiSim-DR. Source : TANNIER 2020

Les rétroactions introduites par le modèle MobiSim-DR sur la localisation des logements et des bâtiments entre les échelles permettent de simuler une émergence : les structures micro-géographiques rétroagissent sur les structures méso-géographiques. Rien n'interdit de penser que, pour produire des agglomérations villageoises, des phénomènes ascendants de concentration et de regroupement sont également mobilisés. De sorte que la simulation trans-échelle impose de réfléchir à des relations entre les échelles beaucoup plus complexes si l'on souhaite produire des résultats plus réalistes.

4.6 Conclusion

Les analyses trans-échelles poursuivent les objectifs scientifiques qui ont concouru à l'introduction de la géométrie fractale en géographie humaine :

« la dimension fractale étant déterminée à partir des comptages effectués selon différents niveaux de résolution d'une distribution spatiale, ce calcul met en évidence des variations systématiques de la répartition d'un fait géographique donné à travers les échelles. Ces variations sont non proportionnelles, non linéaires, elles n'apparaîtraient pas si on se contentait de mesures plus simples des répartitions comme celles des densités, si souvent employées dans la géographie classique (...) » (PUMAIN 2010).

En ne supposant pas l'existence d'une propriété d'invariance d'échelle, l'analyse trans-échelle, nous a permis d'explorer la grande diversité d'agencement des bâtiments à travers les échelles de la France métropolitaine. Cette diversité est très structurée. On retrouve l'opposition entre le Nord-Est et l'Ouest de la France décrite dans la partie précédente à l'aide des analyses multifractales de la localisation de la population dans l'espace, mais l'utilisation d'une description très locale des propriétés trans-échelles nous a permis d'aller plus loin dans les analyses et d'observer très précisément tous les espaces y compris ceux périurbains.

L'algorithme MorphoLim⁶ développé par le laboratoire ThéMA fournit quelques pistes de réponse. MorphoLim est dédié à la recherche d'une échelle caractéristique : le seuil de distance entre deux logiques d'organisation de l'espace. Nous avons utilisé cette méthode pour caractériser l'enveloppe morphologique de 20 agglomérations urbaines (SÉMÉCURBE, TANNIER et ROUX 2019). Les seuils de distance ainsi obtenus oscillent autour de 200 m. On retrouve cette échelle implicitement dans le deuxième axe de l'ACP des signatures trans-échelles des tissus bâtis français. En opposant les petites échelles aux grandes échelles, celui-ci met en exergue le rôle pivot de l'échelle médiane. Nous avons utilisé explicitement cette échelle médiane dans notre deuxième modélisation par polynômes orthogonaux pour décrire l'écart à l'invariance d'échelle : le polynôme orthogonal d'ordre 2 est centré autour de 200m (*fig. 4.14*). Il capte les différences entre les grandes et les petites échelles et est peu sensible aux variations médianes.

Ce seuil de distance est associé à l'urbain en Europe (BRETAGNOLLE 2015). Proposé en 1959 lors de la conférence des statisticiens européens réunie à Prague, il sert de référence à l'Insee pour la délimitation des unités urbaines. Le monde académique se l'est approprié pour définir l'urbain et identifier des agrégats bâtis morphologiquement cohérents (NORDBECK 1971 ; DONNAY, BARNSLEY et LONGLEY 2000 ; CARUSO, HILAL et THOMAS 2017). Finalement cette échelle caractéristique des espaces villageois structure notre façon de concevoir l'écart à l'invariance d'échelle.

6. sourcesup.renater.fr/www/morpholim/

Chapitre 5

Conclusion

La conversation entre TANNIER 2018 et FRANKHAUSER 2018 déjà citée dans la section 3.4.5.1 montre à quel point deux chercheurs ayant longtemps travaillé ensemble peuvent entretenir des divergences nettes sur le rôle des normes d'aménagement. D'un côté Pierre Frankhauser soutient l'intérêt du recours à des modèles de villes fractals à la fois pour minimiser l'impact des déplacements et maximiser le bien-être des habitants qui souhaitent accéder aux aménités vertes et de loisir. De l'autre, Cécile TANNIER remet en cause l'existence même d'une norme objective et absolue, qu'elle soit fractale ou non : « une forme n'est pas, en elle-même, optimale : elle ne l'est que par rapport aux processus (i. e. aux modes de fonctionnement) qu'elle permet d'optimiser. De plus, l'absolu n'existe pas en termes de distribution spatiale des implantations humaines en raison de la diversité des contextes sociaux, politiques, économiques, naturels, etc. desquels résultent des fonctionnements des hommes dans l'espace également très divers ».

La remise en cause de l'existence des normes d'aménagement objectives et absolues fait partie intégrante du fonctionnement des sciences sociales. Mais en même temps, elle prive l'acteur public de l'allié potentiel que serait le géographe prescripteur d'une politique publique. Ces questionnements théoriques et épistémologiques se retrouvent sur des sujets apparemment plus simples, tel que le lien entre la taille d'une agglomération et sa production de CO_2 (LOUF et BARTHELEMY 2014). En fonction de la méthode choisie pour définir et délimiter la ville, les villes les plus vertueuses en terme de production de CO_2 peuvent devenir les moins vertueuses. Ceci suggère aux auteurs la mise en garde suivante à destination des acteurs publics :

« From a policy point of view, now, what should one do to curb CO₂ emissions? Favour the growth of large urban areas or the repartition of population in less populated cities? Both can be argued by considering data analysis alone. It should therefore be obvious that, until they have a satisfactory understanding of the mechanisms responsible for the observed behaviours, scientists should refrain from giving policy advice that might have unforeseen, disastrous consequences. If they choose to do so anyway, policy makers should be wary about what is, at best, a shot in the dark ».

Cette dépendance des résultats d'analyses géographiques aux découpages utilisés met également en cause la territorialisation de l'action publique. Le concept de *piège territorial* développé par AGNEW 1994 désigne la confusion entre le territoire du statisticien et celui de l'acteur public. Les statistiques produites par l'État requièrent de définir une base spatiale pour quantifier les phénomènes sociaux considérés. Mais trop souvent, du fait de la force du nombre à objectiver les faits sociaux, le découpage spatial choisi est

également pensé comme le réceptacle naturel de la société. La territorialisation de l'action publique tend de fait à occulter le rôle des échelles macro-géographiques sur la production des inégalités socio spatiales observées aux échelles micro-géographiques : « Il n'est pas question de nier l'importance des facteurs urbains dans le développement des problèmes sociaux. Mais en centrant les analyses de cette façon on a sans doute hypertrophié la responsabilité des urbanistes et des architectes dans les difficultés actuelles des " quartiers ". Cette manière de s'exprimer, devenue usuelle, est d'ailleurs révélatrice du déplacement opéré. Ce sont les quartiers qui posent problème, et plus le chômage, la formation, la famille, l'intégration. Tout se passe comme si la dégradation des quartiers était posée comme cause avant d'être examinée comme conséquence » PINÇON et PINÇON-CHARLOT 1994. Dès lors, on peut se demander si les notions de villes compactes et étalées ne participent pas à occulter la complexité des phénomènes géographiques en jeu dans le traitement de la question du développement durable. Les fonctionnements urbains diffèrent à travers les échelles. Il n'apparaît donc pas forcément pertinent de postuler *a priori* qu'une forme, qu'elle soit compacte ou non, fractale ou non, serait plus vertueuse qu'une autre pour toutes les échelles d'analyse.

Ces réflexions nous amènent à approfondir trois points dans cette conclusion. Tout d'abord, nous souhaitons revenir sur la prise en compte de l'espace dans les analyses multi-échelles et trans-échelles. Ce questionnement théorique est utile pour mieux faire le lien entre les échelles des acteurs publics et celles des géographes. Dans un deuxième temps, nous souhaitons aborder plus précisément la mesure du phénomène de mitage. L'objectif est d'illustrer la difficulté à produire des messages à destination des acteurs publics qui soient à la fois simples et pertinents. Enfin, nous terminons en présentant rapidement le *deep learning* (une sous-branche de l'intelligence artificielle) et en discutant de ses applications en géographie. Ce dernier point nous semble indispensable pour démythifier des méthodes qui, contrairement au sens commun, ne sont pas intelligentes et qui consistent principalement à reproduire les classements opérés par les êtres humains (JULIA 2019; DESSALLES 2019).

5.1 Retour sur la prise en compte de l'espace dans les analyses multi-échelles et trans-échelles

5.1.1 Mesurer la concentration et la dispersion d'une population

Dans cette thèse, nous avons étudié et caractérisé la distribution spatiale des implantations humaines à l'aide de trois approches multi-échelles différentes : l'analyse spatiale de semis de points, l'analyse fractale et l'analyse trans-échelle. Ces trois approches se distinguent moins par leur manière de traiter les échelles que par leur modèle spatial sous-jacent, qui détermine la façon dont on associe aux objets une portion de l'espace à travers les échelles.

Le modèle de référence des méthodes d'analyse spatiale classiques est la répartition uniforme. Chaque objet se voit attribuer la même quantité d'espace pour une échelle donnée et ce, quelle que soit sa position relative par rapport aux autres objets. Cette hypothèse est particulièrement forte et explique probablement l'inadaptation des méthodes d'analyse spatiale pour caractériser des distributions spatiales très irrégulières. Des méthodes d'analyse spatiale plus complexes ont été proposées pour dépasser le modèle de répartition uniforme. Elles reposent principalement sur l'estimation *a priori* d'une répartition de référence basée

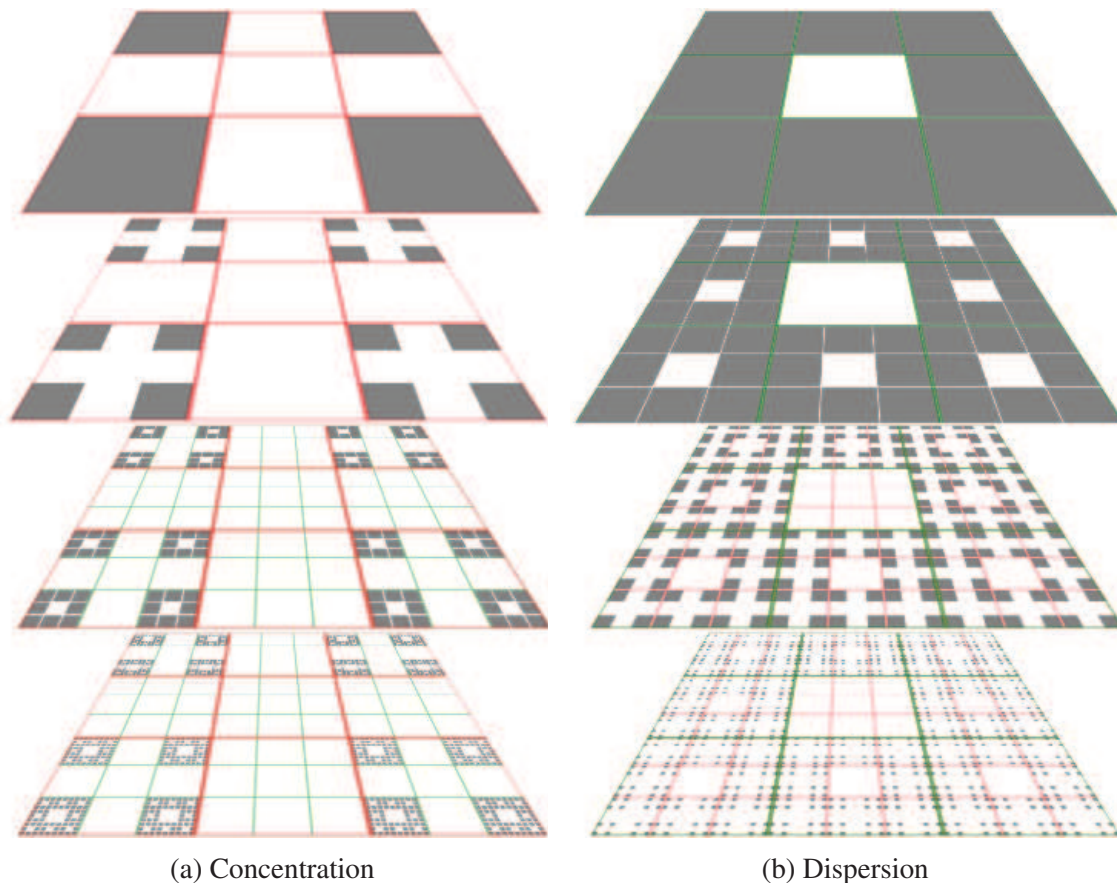


FIGURE 5.1 – Concentration et dispersion. Les deux configurations sont construites par désagrégrations spatiales successives (cf section 4.5) à l'aide de deux générateurs. Le premier est associé à la dimension fractale $\frac{\log(4)}{\log 3}$ et le deuxième, à la dimension $\frac{\log(8)}{\log(3)}$. L'échelle la plus grossière est en haut. Le changement de régime a lieu entre la deuxième et la troisième étape : on change de générateur.

sur la distribution des objets analysés. Mais même avec ce raffinement, les résultats ne sont pas satisfaisants. La répartition de référence est définie pour une échelle donnée et est encore trop régulière pour décrire convenablement les relations qu'entretiennent les objets avec l'espace.

Alors qu'avec l'analyse spatiale, on compare une configuration spatiale analysée à des distributions de référence exogènes à celle-ci, avec les analyses fractales et trans-échelles on considère que la configuration est sa propre distribution de référence. Plus précisément, avec ces approches, l'information pertinente pour caractériser un objet ne se situe pas à une échelle donnée mais émerge de la comparaison des propriétés de la configuration entre deux échelles distinctes. Les espaces *contenant*¹ les objets à une échelle fine deviennent des espaces *contenus* à une échelle grossière. L'arbitraire induit par l'affectation d'une portion de l'espace à chaque objet est en partie levé car la configuration analysée est son étalon de mesure. Dans le cas d'une configuration spatiale parfaitement fractale, le *MAUP* n'existe pas : la configuration est strictement identique à elle-même quelles que soient les échelles considérées. La dimension fractale nous renseigne sur la manière dont les espaces *contenant* les objets deviennent des espaces *contenus* observés à des échelles plus

1. Attention, dans cette sous-section, les notions de contenant et de contenu ont des sens imagés différents de ceux adoptés dans l'introduction. Cette formalisation a été proposée initialement par Cécile Tannier.

grossières. Plus celle-ci est élevée et proche de 2 et plus *contenant* et *contenu* sont de même nature. Au contraire, pour des dimensions fractales de faible valeur, le *contenant* s'éclate en *contenus* isolés les uns des autres. Lorsque l'invariance d'échelle n'est pas vérifiée, la nature des relations entre *contenant* et *contenu* évolue continuellement entre les échelles.

L'invariance d'échelle est une propriété rare que l'on ne rencontre en pratique que pour des plages d'échelles très restreintes. Dans le cas d'invariance d'échelles, les notions de concentration et de dispersion n'ont pas de sens, car pour qu'il y ait dispersion il faut qu'il y ait concentration à une échelle plus fine. Les distributions non invariantes d'échelles sont d'une certaine manière plus intéressantes. Elles sont les seules, d'un point de vue théorique, à pouvoir être considérées comme concentrées ou dispersées de par la présence d'un changement de régime dans la façon dont les *contenants* se transforment en *contenus* à travers les échelles (fig 5.1). Une distribution concentrée idéal-typique est définie par une échelle caractéristique qui marque la rupture de régime. Pour les échelles plus grossières à celle-ci, les *contenants* se disloquent en *contenus* très isolés, puis pour les échelles plus fines, *contenants* et *contenus* sont quasi-identiques. Les distributions dispersées idéal-typiques sont définies par une inversion de ce schéma d'organisation à travers les échelles. On constate donc que les états de concentration et de dispersion ne sont pas définis pour une échelle donnée mais pour une plage d'échelles. L'échelle caractéristique n'existe pas dans l'absolu mais est fonction de la plage d'échelle considérée.

Le fait de remplacer une échelle par une plage d'échelles pour définir un état de concentration ou de dispersion ne doit pas nous faire oublier que la notion d'échelle est en elle-même également un artefact géographique qui n'existe pas en dehors de l'action d'observation effectuée par le géographe. L'échelle, en géométrie fractale, est toujours associée à un outil, à une règle. Observer une distribution à une échelle donnée consiste à supprimer de celle-ci tous les détails dont la taille est inférieure à celle de l'échelle mesurée à l'aide de l'outil choisi. Si l'on change d'outil, en passant d'une « boîte » simple à une ondelette par exemple, la notion d'échelle change également. Il est donc important, lorsque l'on vulgarise les résultats d'une analyse trans-échelle, de ne pas trop réifier cette notion et d'admettre que l'état de concentration ou de dispersion se réalise autour d'une plage d'échelles aux contours flous.

Remarque

Les cascades fractales sont utilisées par les mathématiciens pour créer des distributions spatiales théoriques dont les dimensions sont entièrement déterminées par les propriétés du générateur. On peut généraliser ce type de procédé en appliquant alternativement deux générateurs une infinité de fois : après n_1 itérations avec le premier générateur, on réalise n_2 itérations avec le second (ceci revient à « boucler » à travers les échelles les cascades de la figure 5.1).

Au bout de k répétitions, le nombre de carreaux occupés est $N_k = \delta^{-k(n_1 d_1 + n_2 d_2)}$ où d_1 et d_2 sont les dimensions fractales associées aux deux générateurs. La dimension de l'ensemble de points généré selon ce processus est une moyenne des deux dimensions fractales pondérées par la fréquence d'apparition des deux générateurs : $\frac{n_1 d_1 + n_2 d_2}{n_1 + n_2}$.

Malgré l'usage de deux générateurs, l'ensemble de points généré converge encore vers une fractale théorique parfaitement invariante d'échelle. Pour obtenir une rupture asymptotique de l'invariance d'échelle, on est obligé de modifier fortement la fréquence d'apparition des générateurs à travers les échelles en recourant, par exemple, à la fonction factorielle : on commence par itérer une fois avec le premier générateur, puis $2! = 2$ avec le second générateur, puis on recommence $3! = 6$ avec le premier générateur et $4! = 24$ avec le second et ainsi de suite. Dans ce cas, la limite utilisée pour définir la dimension de *box-counting* n'est pas définie. En revanche, elle oscille entre les valeurs des deux dimensions d_1 et d_2 .

On constate à travers ce petit développement à quel point les fractales définies sur une plage restreinte d'échelles se différencient des fractales théoriques : la notion d'invariance d'échelle utilisée pour quantifier les formes naturelles n'a pas les mêmes propriétés que sa version asymptotique.

5.1.2 Concentration et dispersion dans le cas de la ségrégation socio-spatiale

L'étude des causes qui engendrent la ségrégation socio-spatiale² dans les grandes métropoles et de ses conséquences sur les populations qui la subissent nécessite de quantifier précisément son organisation spatiale. Dès les premiers travaux de DUNCAN 1957 sur la mesure de la ségrégation, la variabilité des résultats selon la taille des mailles d'analyse est mise en exergue. Cette dépendance de la ségrégation aux échelles d'analyse ne concerne pas uniquement pas sa quantification mais également ses causes et conséquences :

« *Segregation can exist at several levels simultaneously, ranging from specific households to neighborhoods to nation-states to the world. Methodologically, scale affects how we can measure and/or represent segregation. Moreover, the very nature of segregation—the forces that create and maintain it, as well as the material and cultural consequences—differ[s] across scale. An adequate understanding of segregation must be grounded in a framework that explicitly recognizes its inherently scalar nature* ». (KAPLAN et HOLLOWAY 2001)

Il a fallu néanmoins attendre l'article de REARDON, MATTHEWS et al. 2008 pour qu'il y ait une prise en compte explicite de l'échelle dans la mesure de la ségrégation. S. FOWLER 2016 explique ce décalage par le fait que la dépendance à l'échelle d'analyse était plutôt considérée comme un problème à résoudre (en choisissant, par exemple, des unités spatiales adaptées) qu'une source d'information sur l'organisation spatiale des populations analysées.

5.1.2.1 Mesure multi-échelle de la ségrégation

Dans les études récentes, la ségrégation est mesurée à l'aide de l'indice H basé sur l'entropie définie dans REARDON et O'SULLIVAN 2004 :

$$E = \sum_{m=1}^M \pi_m \ln\left(\frac{1}{\pi_m}\right) \quad (5.1)$$

$$H = \frac{1}{E} \sum_{m=1}^M \pi_m \sum_{j=1}^J \frac{t_j}{T} r_{jm} \ln(r_{jm}) \quad (5.2)$$

où m est l'index des M sous-populations, j l'index des J unités spatiales, t_j la population totale dans l'unité j , $T = \sum t_j$ la population totale dans la zone d'étude, π_m est la part de la population du sous-groupe m dans la population totale et $r_{jm} = \frac{\pi_{jm}}{\pi_m}$ la surreprésentation de la sous-population m dans l'unité j . L'indice H est calculé à l'aide d'un découpage spatial et est donc, par construction, mono-échelle. La ségrégation est minimale lorsque les sous-populations sont réparties uniformément dans l'espace.

REARDON, MATTHEWS et al. 2008 au lieu de résumer la ségrégation à l'aide d'un seul indice, construisent des profils de ségrégation qui sont fonction de l'échelle d'analyse. On n'utilise alors plus directement les données brutes collectées sur les unités spatiales, mais des versions lissées de celles-ci. Le rayon de lissage détermine l'échelle d'analyse³. Les profils sont définis en associant à chaque rayon la valeur de H calculée sur les données lissées correspondantes.

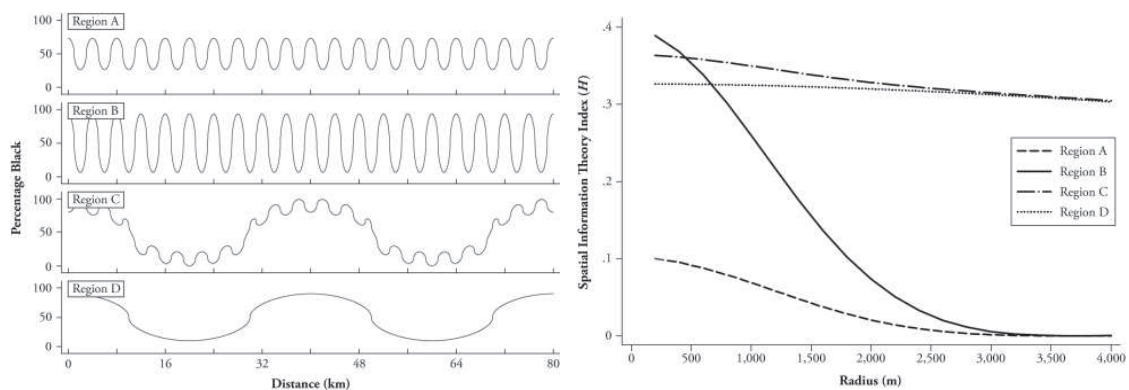
2. Selon Geoconfluences, la ségrégation désigne la « séparation physique des lieux de résidence des groupes sociaux dans l'espace urbain » (geoconfluences.ens-lyon.fr/glossaire/segregation).

3. Le noyau utilisé est le noyau *biweight* (cf. 3.11).

En procédant de la sorte, on introduit dans l'analyse de la ségrégation un positionnement centré sur les individus davantage en accord avec leur vécu⁴ :

« *This approach is operationalized by assuming each individual inhabits a local environment whose population is made up of the spatially weighted average of the populations at each point in the region of interest. Note that these local environments are conceived as over-lapping egocentric environments, rather than discrete spatially bounded neighborhoods (as defined by census tracts, for example). Typically, the population at nearby locations will contribute more to the local environment of an individual than will more distant locations (a “distance-decay” effect)* » (REARDON, MATTHEWS et al. 2008).

Sur la figure 5.2.a sont représentées 4 distributions stylisées de la population noire dans l'espace provenant de REARDON, MATTHEWS et al. 2008. On suppose que l'espace est réduit à une dimension, que la densité de population totale est constante tout le long de cette dimension et que la population noire représente 50% de la population totale. Dans la région A, la part de la population noire varie entre 30% et 70%. La région B a une organisation spatiale similaire à la région A mais la variation de la concentration de la population noire est beaucoup plus forte. La région D est une dilatation horizontale de la région B : les espaces où les populations noires sont majoritaires sont plus larges que ceux de la région B. La région C contient deux échelles de ségrégation : des petites « poches » de ségrégation se superposent à la distribution spatiale de la région D.



(a) Part de population noires dans l'espace pour quatre distributions stylisées (b) Profils de ségrégation pour les quatre distributions stylisées

FIGURE 5.2 – Analyse multi-échelle de la ségrégation. Source : REARDON, MATTHEWS et al. 2008

Les quatre profils de ségrégation sont décroissants, en effet plus le rayon augmente et plus l'on mélange les sous-populations et plus la ségrégation diminue mécaniquement. La ségrégation est la plus faible pour la région A. Dès que l'on dépasse le rayon de 2 km, il n'y a plus de ségrégation, les populations sont réparties uniformément dans l'espace. La région B a un profil plus exacerbé que la région A. En dessous de 2 km la ségrégation est très forte. La ségrégation décroît lentement pour la région D. Ceci s'explique par le fait que la valeur du rayon maximal considérée est bien inférieure à l'emprise des zones de concentration de la population noire : les sous-populations restent séparées. Le profil de la

4. Cette démarche a inspiré l'application hubblo.fr développée par François Sémécurbe pour la description multi-échelle des environnements sociaux.

région C , décroît rapidement et converge vers le profil de la région D car, une fois que l'on a dépassé la taille des « poches », les répartitions des sous-populations de la région C sont comparables à celles de la région D .

L'analyse de la ségrégation proposée par REARDON, MATTHEWS et al. 2008 bien que multi-échelle ne permet pas de définir explicitement des échelles de ségrégation qui caractériseraient la concentration et la dispersion des sous-populations analysées. Ceci s'explique par le caractère cumulatif des profils : l'information portée par un rayon donné contient également celles des rayons plus petits. On retrouve en cela les limites propres aux analyses multi-échelles discutées dans le chapitre 2.

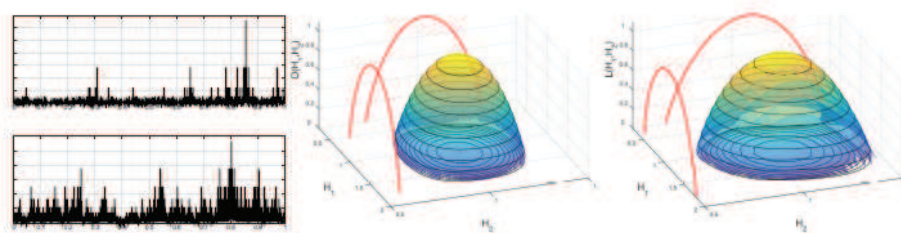


Figure 1: **Independent stochastic binomial measures.** Time series (left), theoretical multivariate spectrum (center) and estimated multivariate spectrum (right; red solid lines show the 1-variable marginal spectra).

(a) Les deux distributions de probabilité sont totalement indépendantes. Le spectre est une fonction de deux variables. La connaissance du comportement d'une distribution à travers les échelles ne permet pas d'en déduire celui de l'autre distribution.

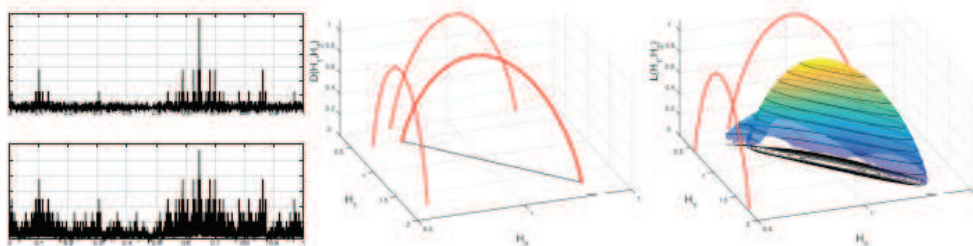


Figure 2: **Fully dependent stochastic binomial measures.** Time series (left), theoretical multivariate spectrum (center) and estimated multivariate spectrum (right; red solid lines show the 1-variable marginal spectra).

(b) Les deux distributions de probabilité sont totalement dépendantes. La connaissance des valeurs d'une distribution de probabilité permet de connaître exactement celle de la seconde. Le spectre se réduit à une courbe.

FIGURE 5.3 – Exemple de spectre bivarié théorique et estimé. Source : JAFFARD et al. 2019

Des travaux pourront être entrepris pour transformer l'analyse multi-échelle de la ségrégation en une analyse trans-échelle. Ceci impliquerait d'utiliser les entropies de la formule de REARDON, MATTHEWS et al. 2008 pour calculer des signatures trans-échelles. La répartition d'une sous-population dans l'espace ne serait plus directement décrite par son profil mais par sa fonction dérivée. Par ailleurs, au lieu d'utiliser un lissage pour caractériser l'organisation d'une sous population à une échelle donnée, on pourrait recourir à une approche de type *sandbox*, c'est-à-dire à des disques centrés sur les individus. En procédant de la sorte, on introduit explicitement dans la mesure de la ségrégation un point de vue egocentré. L'ensemble de ces travaux pourraient bénéficier des apports théoriques proposés par JAFFARD et al. 2019 ; ABRY, WENDT, JAFFARD et DIDIER 2019 qui ont

étendu l'analyse multifractale aux données bivariées afin de quantifier les relations qui relient deux quantités à travers les échelles. Le spectre multifractal n'est plus, dans ce cas, une courbe mais une surface (fig. 5.3).

5.2 Le mitage : résultat de l'étalement urbain contemporain ou héritage du passé ?

Dans l'introduction de cette thèse, nous avons montré, à l'aide d'une juxtaposition de deux cartes, une étonnante relation entre d'un côté, le mitage et de l'autre, les structures agraires observées au 18^e siècle par l'agronome Arthur Young. Les territoires qui sont les plus enclins au mitage des terres agricoles recouvrent approximativement les terres de « bocage » et de population dispersée de l'ouest de la France. Arrivé à la fin de notre thèse, il nous semble utile de revenir dessus à l'aune de l'expérience acquise.

ALBIZZATI, POULHES et SULTAN PARRAUD 2017 définissent le mitage comme l'ensemble des permis de construire débutés entre 2005 et 2013 et dits « isolés », « c'est-à-dire ne se situant ni en zone bâtie, ni en périphérie d'une zone bâtie ». Les zones bâties sont déterminées à l'aide de deux couches d'occupation du sol : la BD Carto de l'IGN et CORINE Land Cover dont les objets ont une emprise spatiale minimale de 8 ha, ce qui correspond approximativement à un cercle de 90 m de rayon. Un permis de construire est isolé lorsqu'il est éloigné de plus de 300 m de toutes zones bâties (fig. 5.4).

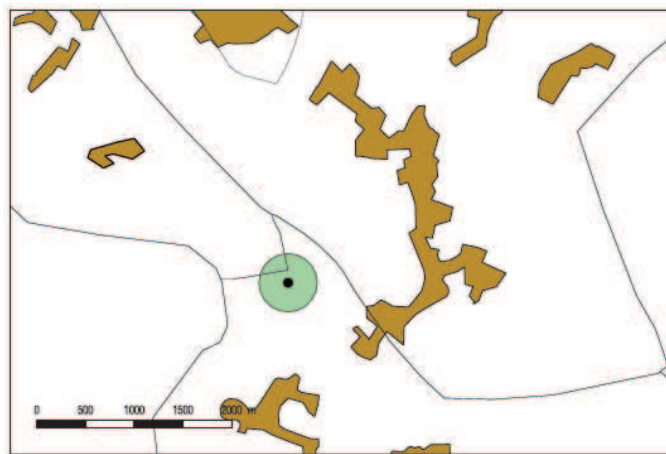


FIGURE 5.4 – Mitage. Construction isolée à plus de 300 mètres d'une zone bâtie en 2005. Source : ALBIZZATI, POULHES et SULTAN PARRAUD 2017

Les zones bâties servant à définir le mitage sont des agrégats délimités au moyen de zones tampons calculées à l'aide d'un système d'information géographique. Seules les zones tampons dont l'emprise spatiale est suffisamment grande sont conservées en tant que zones bâties. Les choix de la taille des tampons et de la taille minimale des zones conditionnent la quantification du mitage :

- Plus la taille des tampons est petite et plus l'emprise spatiale des zones agrégées est faible et donc la part des bâtiments en dehors de celles-ci est importante ;
- Si l'on augmente la taille minimale que les zones tampons doivent respecter pour être considérées comme zones bâties, il y a moins de zones bâties et donc plus de bâtiments en dehors de leur emprise.

Avec une emprise spatiale de 8 ha, une grande part de l'habitat dispersé des espaces de l'ouest de la France est considérée comme isolée des zones bâties. Ceci relativise le pouvoir heuristique de la définition du mitage utilisée : bien avant que le mitage devienne un enjeu des politiques publiques, les territoires de l'ouest de la France étaient fortement « mités » du fait même de leur histoire.

L'indicateur utilisé par ALBIZZATI, POULHES et SULTAN PARRAUD 2017 mesure en fait moins le mitage que l'éloignement des nouvelles constructions aux espaces villageois. Les auteurs remarquent d'ailleurs à ce propos que « pour les constructions correspondant à du mitage (type 4), les locaux agricoles représentent 23% de la surface consommée par le bâti alors qu'ils sont absents dans l'artificialisation de masse et ne concernent que 5% des espaces consommés en continuité de bâti. La construction agricole tend à se faire le plus loin possible des autres constructions, en partie car les prix des terrains sont plus élevés à proximité du bâti ». Cette information est intéressante d'un point de vue géographique et suggère deux questions imbriquées :

- Pourquoi, dans les territoires d'habitat dispersé, construit-on davantage en dehors des espaces villageois que dans les espaces d'habitat groupé ?
- Ou sont situées les nouvelles constructions ?

Ne disposant pas de la base des permis de construire, nous ne pouvons pas répondre précisément à ces questions. En revanche, nous pouvons proposer quelques pistes de réflexion en nous appuyant sur d'autres éléments.



(a) Openfield

(b) Bocage

FIGURE 5.5 – Carte de Cassini des deux configurations archétypales. Les carrés bleus représentent l'emprise spatiale des deux territoires. Source : geoportail.gouv.fr

Notamment, la comparaison des anciennes orthophotographies de l'IGN (entre 1950 et 1965) avec les plus récentes suggère que les nouveaux bâtiments d'habitation, y compris ceux éloignés des espaces agrégés, se construisent principalement à proximité des bâtiments préexistants. On retrouve ainsi le fait que les structures de peuplement « subsistent aux causes qui les ont façonnées et n'évoluent que lentement » (LE BRAS 1996). Les cartes de Cassini publiées à la fin du 18^e siècle illustrent cette inertie du peuplement (fig 5.5). Dans la configuration de bocage les lieux-dits et les villages sont disséminés dans l'espace, alors

qu'ils sont quasiment inexistant dans la configuration d'openfield. L'inertie n'est pas en soi une cause mais une conséquence des mécanismes sous-jacents qui favorisent l'installation des nouveaux bâtiments principalement à proximité des bâtiments préexistants.

L'imbrication des réseaux et des tissus bâtis apparaît comme un autre élément fondamental pour comprendre les dynamiques territoriales (*fig. 5.6*). Tous les bâtiments étant reliés entre eux par des réseaux (viaires, énergétiques, d'eau), la construction d'un nouveau bâtiment est plus ou moins aisée selon sa localisation et les coûts de raccordement qu'elle implique. Le raccordement d'un bâtiment à un nœud du réseau déjà existant est quasi nulle alors qu'étendre les réseaux nécessite de les reconfigurer.

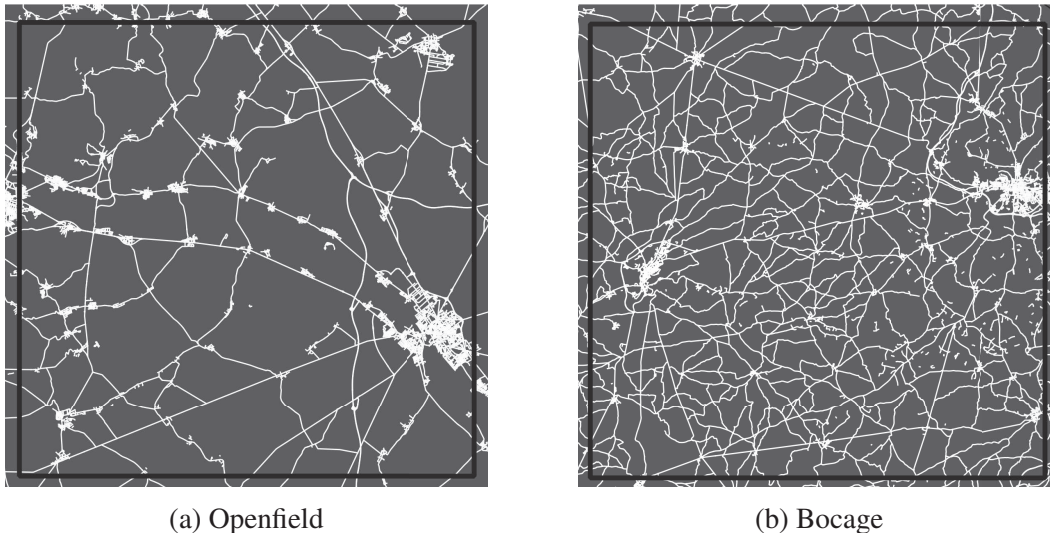


FIGURE 5.6 – Réseau viaire des deux configurations archétypales. Source : OpenStreetMap

Comme on s'y attend, le réseau viaire du territoire de bocage est nettement plus dense que celui du territoire d'openfield. Mais dépasser cette comparaison visuelle est délicate du fait de l'hétérogénéité de la notion de route sur le territoire français. Le classement des routes selon leur statut (chemin, route communale, route départementale, route nationale) est autant fonction de leurs caractéristiques géographiques que de choix politiques entre les différents échelons territoriaux.

Le réseau électrique semble davantage normalisé et plus facile d'utilisation. Il se décompose en deux sous-réseaux. Le premier, le réseau de Haute Tension sert à transporter l'électricité depuis les lieux de production, à savoir les centrales électriques, vers les réseaux de distribution. Le réseau de distribution a, lui, vocation à alimenter l'ensemble des consommateurs. Entre les deux, les postes HTA/BT permettent de faire la jonction. Composés d'un transformateur, ces postes ont pour rôle d'abaisser la tension à 230 V (par exemple). Cette architecture articulant un réseau de transport et de distribution résulte d'un compromis coût-efficacité (APPLEBY 1996). La construction d'un nouveau bâtiment induira, selon sa localisation, une transformation différente du réseau électrique. Si la nouvelle construction est proche d'un bâtiment existant, seul le réseau de distribution va être affecté, sinon ce sont les deux réseaux qui doivent être adaptés, avec en particulier la construction d'un poste HTA/BT spécifique.

Au niveau de la France métropolitaine, on observe une densité de postes HTA/BT nettement plus importante dans l'ouest de la France (*fig. 5.7.a*). Cependant, la qualité des données ne permet pas d'affiner ce constat. En effet, les postes ne sont pas tous mentionnés car Enedis n'est pas le seul distributeur d'électricité et certaines collectivités

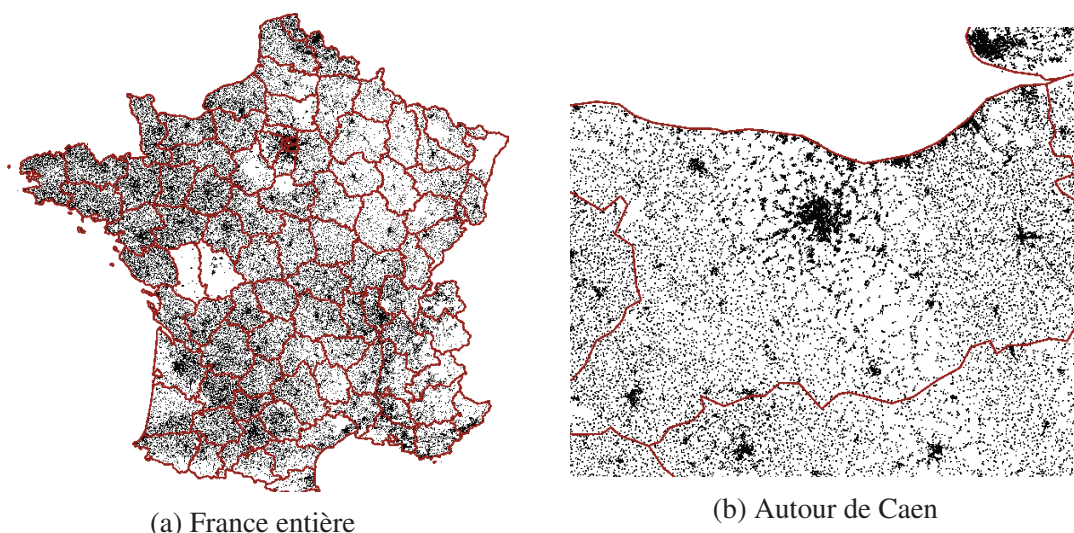


FIGURE 5.7 – Convertisseur Haute basse Tension. Les contours rouges représentent les départements. Source : Open Data Enedis

ont historiquement ou suite à l'introduction d'un partenariat privé recours à des entreprises locales de distribution d'électricité. Les postes appartenant à ces entreprises ne sont pas renseignés dans la base Enedis, ce qui explique les vides observés sur la carte. Toutefois, à l'échelle de l'agglomération de Caen où Enedis est le seul distributeur officiel (fig. 5.7.b), on observe une différence dans la distribution spatiale des postes entre la plaine de Caen et les espaces de bocages ; dans ces derniers, le semis de postes est nettement plus dense. Ceci semble prouver que le réseau électrique tend à suivre l'organisation du bâti existant.

S'il est confirmé que le mitage se produit principalement le long des axes routiers et à proximité de bâtiments isolés, cela devrait questionner le discours normatif selon lequel la construction de lotissements en continuité des villages serait préférable au mitage. En effet, rien ne permettrait de conclure *a priori* que le premier soit plus durable que le second. Une analyse plus précise de la localisation des nouvelles constructions serait nécessaire afin d'espérer obtenir des résultats pertinents qui puissent être utiles à l'action publique.

5.3 Perspectives : *deep learning* et réseaux convolutifs

Le *deep learning*, s'il n'a pas encore totalement changé notre présent, a en revanche profondément transformé notre vision du futur. Les robots capables de suppléer les hommes dans certaines tâches ne sont plus de simples spéculations futuristes mais apparaissent progressivement dans notre quotidien. Au début de nos travaux de recherche, aucun outil simple ne permettait de mettre en œuvre les méthodes de *deep learning* sur des cas pratiques. Depuis lors, les bibliothèques Python et R, Keras et Pytorch, ont rendu accessibles ces méthodes à l'ensemble des chercheurs et ce, quelle que soit la nature de leurs données : textes, indicateurs statistiques, images, sons... Dans la conclusion de cette thèse, il nous semble intéressant de présenter et discuter ce type de méthodes car elles partagent des points communs avec l'analyse multi-échelle mais aussi parce qu'elles pourraient bientôt être utilisées couramment, en complément ou en remplacement de certaines méthodes traditionnelles en géographie.

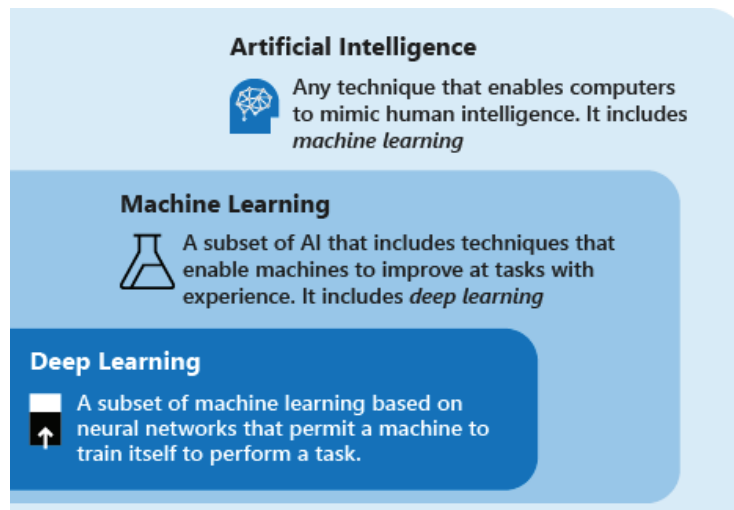


FIGURE 5.8 – L’intelligence artificielle, le *machine learning* et le *deep learning*. Source : Microsoft.

L’intelligence artificielle (IA) est « l’ensemble des théories et des techniques mises en œuvre en vue de réaliser des machines capables de simuler l’intelligence »⁵. SAMUEL 1959 définit le *machine learning* comme « le domaine d’étude qui donne aux ordinateurs la capacité d’apprendre sans être explicitement programmés ». Le *machine learning* fait partie de l’intelligence artificielle (fig 5.8) et comprend un ensemble de méthodes fréquemment utilisées par les géographes et, plus largement les chercheurs en sciences sociales, telles que les régressions linéaires et les méthodes de classification automatique.

Le *deep learning* est une sous-partie du *machine learning*. Il traite l’information à l’aide de réseaux de neurones. Ce nouveau champ méthodologique a émergé progressivement au cours de la seconde moitié du 20^e siècle. Initialement, les chercheurs se sont inspirés de l’organisation des cerveaux biologiques constitués de neurones et de synapses, pour rendre intelligents les systèmes informatiques. Progressivement, les chercheurs s’émancipèrent de cette référence pour produire des réseaux de plus en plus originaux et adaptés à des situations spécifiques.

L’évolution des réseaux de neurones et plus généralement de l’intelligence artificielle est en dents de scie, passant de phases d’essor vigoureux à des périodes de recul appelées par les spécialistes du domaine, les hivers de l’intelligence artificielle. A partir du début des années 2010, la conjonction de puissances de calcul inégalées et de la mise à disposition de grandes bases de données labellisées explique le dynamisme du *deep learning*. Des tâches jugées quasi impossibles à traiter avec les algorithmes classiques, tels que la reconnaissance des objets contenus dans une image, la traduction automatique et instantanée entre deux langues ou encore l’automatisation de la conduite automobile, sont devenues réalités. On note toutefois dans ce dernier cas, que les prédictions les plus ambitieuses sur l’arrivée massive des voitures autonomes dans nos vies quotidiennes à l’horizon des années 2020 se sont révélées fausses. Cet échec annonce pour certains scientifiques l’arrivée d’un nouvel hiver de l’intelligence artificielle (LEPRINCE-RINGUET 2020).

5. Définition dictionnaire Larousse.

L'intelligence artificielle intéresse les géographes depuis le milieu des années 80 (COUCLELIS 1986). En 1997, Stan OPENSHAW et Christine OPENSHAW publient : *Artificial Intelligence in Geography*. Ce livre traite des principaux sujets de l'intelligence artificielle y compris des réseaux de neurones. En France, Didier JOSSELIN en 1995 dans sa thèse sur « La Déprise agricole en zone de montagne : vers un outil d'aide à la modélisation couplant systèmes d'induction et d'information géographique », utilise les réseaux de neurones pour modéliser un espace géographique. L'usage du *deep learning* se généralise en géographie avec le développement des systèmes d'information géographique et la mise à disposition massive de données spatiales au début des années 2010. Les données d'occupation du sol, les orthophotographies et les images satellites proposées par les instituts cartographiques, OpenstreetMap et les acteurs du web tel que Google permettent aux géographes et aux ingénieurs d'analyser l'ensemble de la surface terrestre et ce, avec une résolution de l'ordre du mètre.

L'usage le plus commun du *deep learning* est l'interprétation des images satellites, aériennes ou prises par des drones. L'objectif est de classer automatiquement les objets contenus dans celles-ci. Par exemple, on peut utiliser un réseau de neurones afin de détecter et délimiter dans une image, les piscines, les voitures, les champs ou les panneaux solaires (ABDISHAKUR 2019 ; DOUILLARD 2018). Ceci nécessite de l'avoir entraîné au préalable à l'aide d'une base d'images labellisées par des humains. La *GeoAI* rassemble l'ensemble des méthodes de *machine learning* et de *deep learning* qui ont vocation à extraire et produire des fonds de cartes vectoriels pertinents directement à partir de l'interprétation de données rasters. JANOWICZ et al. 2020 postulent la transformation de l'ensemble de la production cartographique à l'horizon 2030 :

« Can we develop an artificial GIS analyst that passes a domain-specific Turing Test by 2030? Put differently, can we design a software agent that takes a user's GIS-related domain question, understands how to gather the required data, how to analyze them, and how to present the results in a suitable form? Imagine a user asking for available undeveloped spaces that are most suitable for community-based solar panel installations. The artificial GIS analyst would find the required data layers using SDI, perform operations such as insolation analysis, and return a suitability map, thereby (ideally) becoming indistinguishable from a human analyst. The key here is to open up GIS to Siri-like interaction for the masses, not to replace highly-trained GIS analysts performing complex analysis. Several large-scale projects such as the NSF funded EarthCube and, more recently, the Open Knowledge Networks track of NSF's Convergence Accelerator address the challenge of designing human and machine-readable and reasonable data repositories. Methods-wise, the above-mentioned research by Scheider et al. (2019) and other research teams can already be seen as contributions to this moonshot. From the industry's side, companies such as Esri have long experimented with automatically suggesting analysis and visualization options for common datasets and types. Hence, while ambitious, the envisioned artificial GIS analyst is a realistic goal if GeoAI research continues at today's speed ».

Dans des domaines plus proches de ceux évoqués dans cette thèse, l'état de l'art mené par GREKOUSIS 2019 identifie les champs d'application susceptibles de bénéficier du *deep learning* en géographie urbaine :

« Therefore, complex and interdisciplinary urban problems such as simulating urban growth and dynamics, analyzing urban pollution- climate change, studying urban poverty, urban unemployment, urban health, analyzing urban resilience and urban sustainability can be addressed through comprehensive artificial intelligence modeling ».

Des analyses morphologiques réalisées à l'aide du *deep learning* émergent progressivement dans le champ géographique. MOOSAVI 2017 analyse le tissu bâti d'un million de villes réparties sur toute la planète à l'aide d'un réseau convolutif non supervisé, un *autoencodeur*. Les *autoencodeurs* sont une hybridation entre les méthodes de classification et les réseaux de neurones convolutifs et servent à définir des classes d'images comparables. Afin de couvrir un large territoire, l'étude exploite la base de données du réseau viaire issue d'OpenStreetMap. DUBEY et al. 2016 modélise la perception du paysage en recourant aux données provenant de *StreetView* de Google. L'objectif est de décrire sur de larges territoires, les paysages et leur perception par la population afin de promouvoir des politiques d'aménagement de l'espace globales. Plus récemment, LAW et NEIRA 2019 analysent à la fois des images issues de *StreetView* et de données d'occupation du sol à l'aide de réseaux convolutifs pour étudier la relation entre les paysages et l'organisation de l'espace. Ces éléments de bibliographie ne sauraient être exhaustifs et tendent juste à montrer que le *deep learning* a intégré la boîte à outils standard du géographe au même titre que l'analyse fractale.

Une présentation des principes de base des réseaux de neurones est donnée en annexe C.

5.3.1 Exemple d'application : typologie des styles parcellaires

Nous avons appliqué les méthodes de *deep learning* à l'étude de la forme des parcelles cadastrales. Ce choix a été motivé par le fait que les archéologues des espaces ruraux constatent localement une coévolution entre les *styles parcellaires*⁶ et l'habitat : « Il [le bocage] meurt ; mais le dessin cadastral ne change pas ; les champs gardent leur forme trapue, l'habitat reste dispersé » (Meynier, 1943, p. 163) cité par WATTEAUX 2012.

Afin de vérifier la véracité de cette coévolution à une large échelle, nous avons réalisé une carte des styles parcellaires à l'aide d'un réseau de neurones (HIMPENS, SÉMÉCURBE et POULHES 2019). La France métropolitaine a été découpée en carreaux de 1024 m de côté dans lesquels on cherche à décrire l'agencement spatial des parcelles. Les fonds de cartes des parcelles cadastrales ont été rasterisés à 4 m afin de pouvoir mettre en œuvre un réseau convolutif. Les images en entrée du réseau font 256 pixels par 256 pixels. Les parcelles des départements de la Manche et de la Marne ont servi de base d'apprentissage (fig 5.9)). Dans le département de la Manche, les parcelles sont irrégulières et compactes alors que dans celui de la Marne, elles forment des bandes régulières. Enfin, une troisième classe recouvrant les territoires urbains de ces deux départements a été rajoutée à l'analyse.

Sans avoir introduit de contrainte de contiguïté spatiale, la carte des styles parcellaires laisse apparaître des espaces homogènes recouvrant approximativement l'opposition entre les territoires d'habitat dispersé et concentré (fig 5.10). Ce résultat ne doit pas être surinterprété, il s'agit plus d'une corrélation statistique que d'une causalité. Les deux départements que nous avons choisis pour définir les poids du réseau sont idéal-typiques à la fois en termes de répartition de la population et de styles parcellaires, ce qui tend à surimposer

6. CHOUQUER 2000 les définit comme l'« agencement du dessin parcellaire dont on pense qu'il indique des héritages morphologiques »

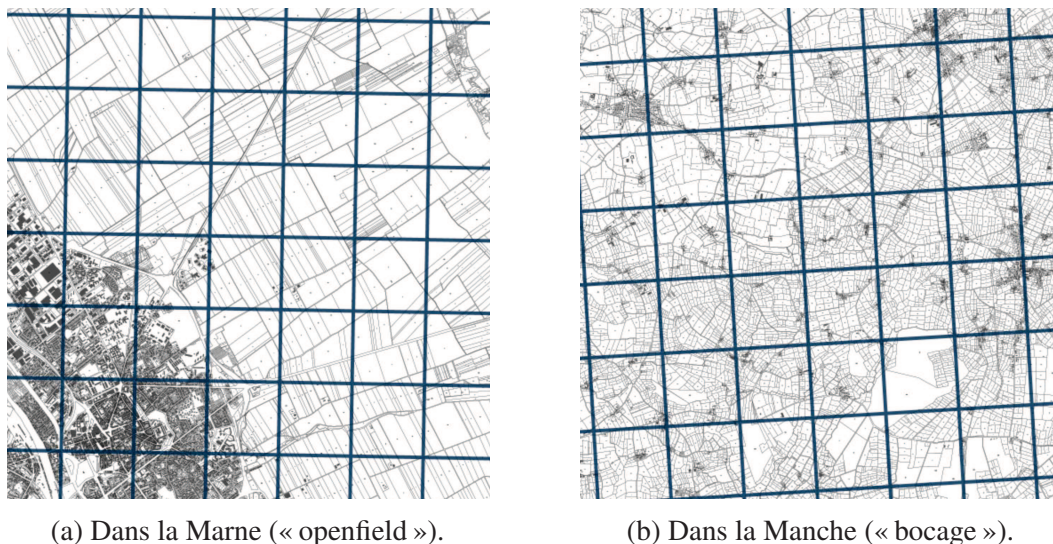


FIGURE 5.9 – Exemples de styles parcellaires. Source : Géoportail

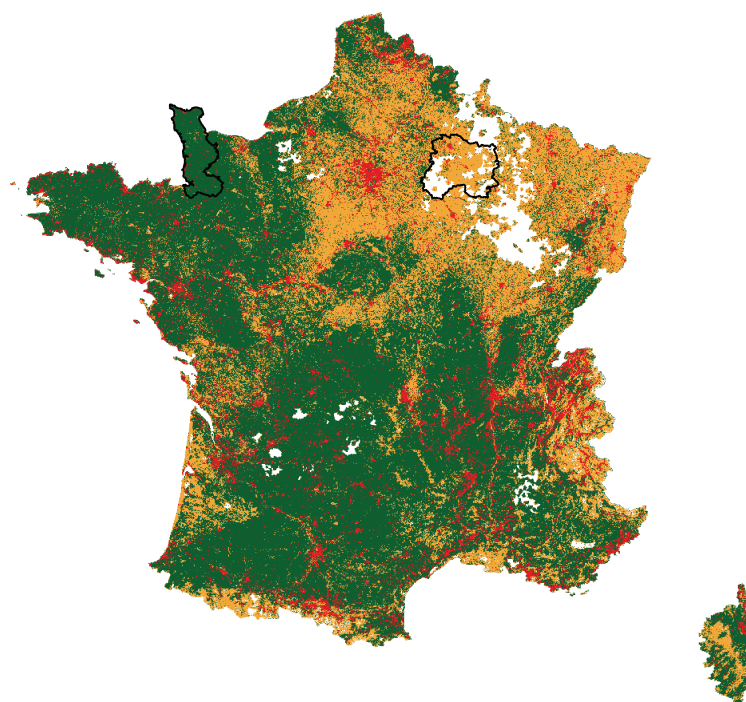


FIGURE 5.10 – Carte des styles parcellaires. En vert : carreaux dont l’agencement des parcelles est similaires à celles du département de la Manche, en orange à celles du département de la Marne. Le rouge correspond aux parcelles urbaines (HIMPENS, SÉMÉCURBE et POULHES 2019). Source : Etalab 2019

une vision partielle et partielle de la forme des parcelles. D’autres classes d’apprentissages auraient pu produire des cartes différentes et la corrélation aurait pu être nettement moins prononcée.

5.3.2 Questions autour de l'usage du *deep learning*

5.3.2.1 Les limites théoriques du *deep learning*

Les réseaux de neurones que nous avons évoqués précédemment sont dits supervisés : il est nécessaire de calibrer leur poids à l'aide d'une phase d'apprentissage avant de les utiliser. Cette opération requiert de disposer d'une base de données labellisées comprenant par exemple des textures bâties et une classe d'appartenance. Dans de nombreuses applications commerciales, ce travail de labellisation est opéré par les utilisateurs d'un service, avec ou sans leur consentement éclairé, ou bien par des « travailleurs du clic » comme les appelle CASILLI 2019⁷. Avec la première solution, la base est constituée par les traces numériques produites par les usagers d'Internet, telles que les notes qu'ils donnent aux produits achetés sur les sites de vente en ligne. L'autre solution consiste à acheter auprès de plateformes dédiées de la « labellisation ». Dans une démarche comparable à celle proposée par Uber pour les déplacements, ces plateformes ont pour objectif de mettre en relation les entreprises et des travailleurs autonomes comme en témoigne la page du *Mechanical Turk* d'Amazon, l'un des principaux leaders du marché⁸ :

« Amazon Mechanical Turk (MTurk) is a crowdsourcing marketplace that makes it easier for individuals and businesses to outsource their processes and jobs to a distributed workforce who can perform these tasks virtually. This could include anything from conducting simple data validation and research to more subjective tasks like survey participation, content moderation, and more. MTurk enables companies to harness the collective intelligence, skills, and insights from a global workforce to streamline business processes, augment data collection and analysis, and accelerate machine learning development. »

La rémunération à la tâche restreint principalement l'intérêt de cette activité aux travailleurs des pays en voie de développement regroupés dans des « fermes à clics » ou à des personnes précaires dans les pays développés qui cherchent à compléter leurs revenus depuis chez eux (300 000 micro-travailleurs en 2019 en France selon CASILLI 2019).

Ce problème éthique n'est pas le seul que pose l'usage des réseaux et plus généralement du *machine learning*. Les bases d'apprentissage sont également source de nombreux biais. Par exemple, lorsque l'on utilise une base de labellisation provenant de sources administratives ou commerciales, l'algorithme aura tendance à reproduire les préjugés genrés ou racistes en cours dans la société (LEDFOORD 2019). Ceci est accentué pour les réseaux de neurones, de par leur côté boîte noire inhérent à leur définition : ils ne reposent sur aucune modélisation formelle pour prédire un résultat. Puisque l'on ne connaît pas explicitement le rôle des variables ni des structures spatiales qui participent à la réalisation d'un phénomène, il est dès lors difficile de proposer des préconisations et encore moins des leviers d'action à partir des résultats d'un réseau de neurones GREKOUSIS 2019 :

« When the focus lies on the explanatory power of the model and on the identification of casual relationships, then ANNs might not be as appropriate as other techniques, as their explanatory capabilities are weak. For example, when the objective is to explain the causes of urban growth (or in general a cause and effect relationship), ANNs are not particularly descriptive as they do not provide equations that link input to output variables, so causal

7. Faute de moyen, nous avons été contraint à recourir à une labellisation très simple reposant sur deux départements dans notre application du *deep learning*.

8. Consultée le 22/03/2020

direction is not clear. In this case, some conventional methods, as for example logistic regression or spatial econometrics, offer a better explanation in the identification of the driving forces behind urban land use dynamics ».

GREKOUSIS 2019 suggère d'utiliser des méthodes hybrides combinant un réseau de neurones et une approche statistique classique pour inférer des relations explicites entre les entrées et les sorties d'un réseau de neurones.

5.3.2.2 Comment contourner l'étape de labellisation ?

Au delà des questions éthiques, l'étape de labellisation demande des ressources financières ou en temps humains importants et nécessite de définir *a priori* les concepts que l'on souhaite chercher. Cette démarche n'est pas la seule utilisée en science. Les méthodes non supervisées consistent à décrire un ensemble d'objets en les regroupant dans des classes pertinentes définies *a posteriori*. Celles-ci sont construites de sorte qu'elles regroupent ensemble des objets similaires et que les objets de deux classes distinctes soient les plus différents possible. Les auto-encodeurs (fig. 5.11) sont des réseaux non supervisés qui visent à projeter un ensemble d'images dans un espace vectoriel de faible dimension. Autrement dit, on recherche à décrire une image sous la forme d'un vecteur dont chaque composante décrirait une propriété caractéristique de l'ensemble étudié. Les auto-encodeurs présentent une structure en « double entonnoir ». Le premier « entonnoir » recherche à réduire l'information contenue dans l'image, le deuxième à recréer à partir de cette information restreinte l'image complète. A l'interface des deux « entonnoirs » se trouve la couche d'auto-encodeur sur laquelle on peut appliquer des méthodes de partitionnement de données (nuées dynamiques ou classification ascendante hiérarchique). Les poids d'un auto-encodeur s'obtiennent à partir d'une étape d'apprentissage où l'on essaie de reconstruire les images en entrée (la sortie est l'image en entrée). La couche du milieu aura ainsi tendance à capturer les propriétés structurelles des images analysées et s'apparente, d'un point de vue théorique, aux résultats d'une Analyse en Composantes Principales (BOURLARD et KAMP 1988). Elles possèdent d'ailleurs les mêmes limites pratiques que l'ACP : les structures observées sont fonction de l'ensemble d'images étudié. Ainsi, une caractéristique très discriminante associée à une sous-classe d'images très peu représentée au sein de l'ensemble analysé peut être totalement ignorée par ce type de méthodes⁹.

Une autre solution pour contourner l'étape de labellisation des données et le côté boîte noire du *deep learning* consiste à entraîner les réseaux sur des images synthétiques. MAUDUIT et al. 2020 apprennent ainsi à des réseaux à caractériser les propriétés multifractales d'une image. La base d'apprentissage est constituée par des images multifractales dont on connaît exactement les régularités moyennes et l'intensité de la multifractalité. Les résultats satisfaisants observés par les auteurs tendent à montrer que les réseaux de neurones sont bien adaptés à l'étude des objets fractals. Nous allons dans la section suivante discuter plus amplement cette relation entre *deep learning* et analyse fractale.

9. Nous avons été confronté à ce type de biais lors de notre étude des styles parcellaires. Les classes obtenues à l'aide des auto-encodeurs étaient inintelligibles

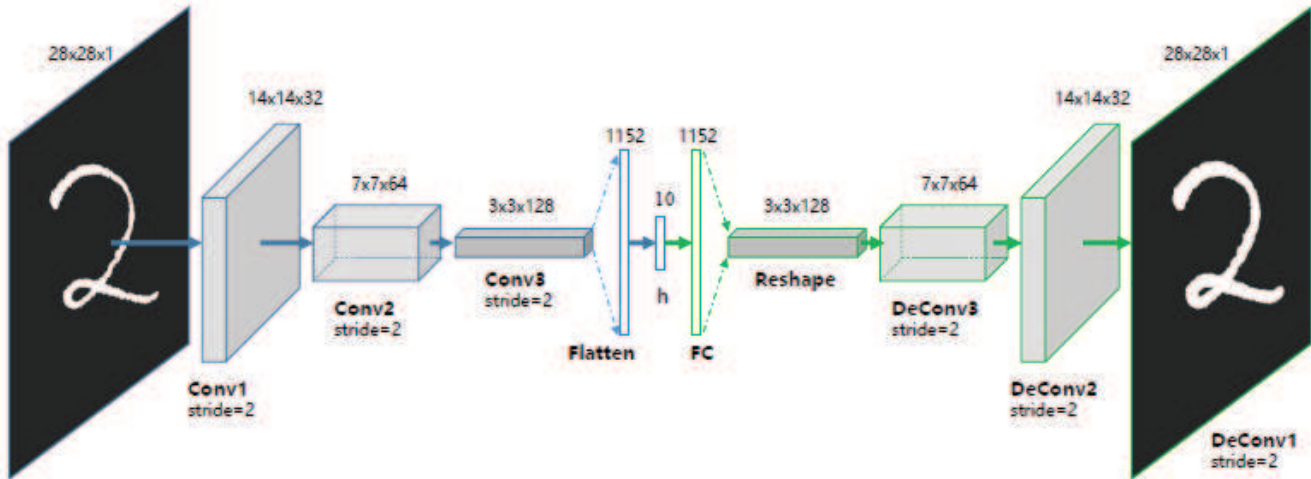


FIGURE 5.11 – Exemple d’auto-encodeur. Source : GUO et al. 2017

5.3.3 Relation entre l’analyse fractale et le *deep learning*

La recherche des raisons sous-jacentes aux succès des réseaux profonds pour le traitement de tâches complexes a nourri une littérature abondante. HOANG et GUERRAQUI 2018 émettent l’hypothèse que leur réussite dans des domaines jugés autrefois difficiles résulterait de leur capacité à traiter des structures fractales. Les notions mathématiques dans cet article sont très complexes et nous nous sommes inspirés d’une vidéo de vulgarisation proposée par l’un des auteurs pour restituer leur raisonnement (HOANG) 2018).

Les fonctions d’activation en sortie des neurones rompent la linéarité et découpent l’espace contenant les données analysées en sous-zones très fines qui peuvent être traitées indépendamment (RAGHU et al. 2017). Supposons que l’on dispose d’une base d’images de paysages que l’on a labellisée en trois classes. En sortie d’un neurone, l’espace d’observation est découpé en deux. Cette structure simple est insuffisante pour séparer de façon satisfaisante les trois classes. En revanche, si au lieu d’un neurone, on prend n neurones positionnés sur une couche, on découpe alors l’espace en approximativement $n^2/2$ zones, ce qui rend d’autant plus facile la classification. Mais le gain le plus important intervient lorsque l’on concatène plusieurs couches les unes derrière les autres. Dans ce cas, le nombre de sous-zones est proportionnel à taille des couches^{profondeur du réseau}. Les sous-zones ainsi définies ont des « frontières » qui s’apparentent à des fractales (POOLE et al. 2016) et sont donc à même de séparer des classes d’objets fractals. Selon l’hypothèse formulée par HOANG et GUERRAQUI 2018, la supériorité du *deep learning* par rapport aux méthodes traditionnelles de classement de textures et de formes proviendrait du fait que la grande majorité des classes d’appartenance des objets serait fractale. De façon imagée, ce ne serait pas les espaces urbains (ou ruraux) eux-mêmes qui seraient fractals, mais l’ensemble des images de paysages urbains (ou ruraux) qui le serait.

Il est difficile d’évaluer la véracité de cette hypothèse à l’heure actuelle. En revanche, elle amène à se questionner sur l’une des principales activités réalisées en science : classer des objets en catégories homogènes. L’analyse fractale a permis aux géographes de caractériser des distributions spatiales complexes pour lesquelles les méthodes classiques étaient inopérantes. Mais l’étape de classification, qui se base sur les résultats de l’analyse fractale, est similaire à celles réalisées avec des indicateurs mono-échelles : elle ne s’effectue pas directement sur les tissus bâtis mais sur des indicateurs qui les caractérisent. Des classes trop petites ou aux contours très irréguliers sont donc susceptibles d’être ignorées

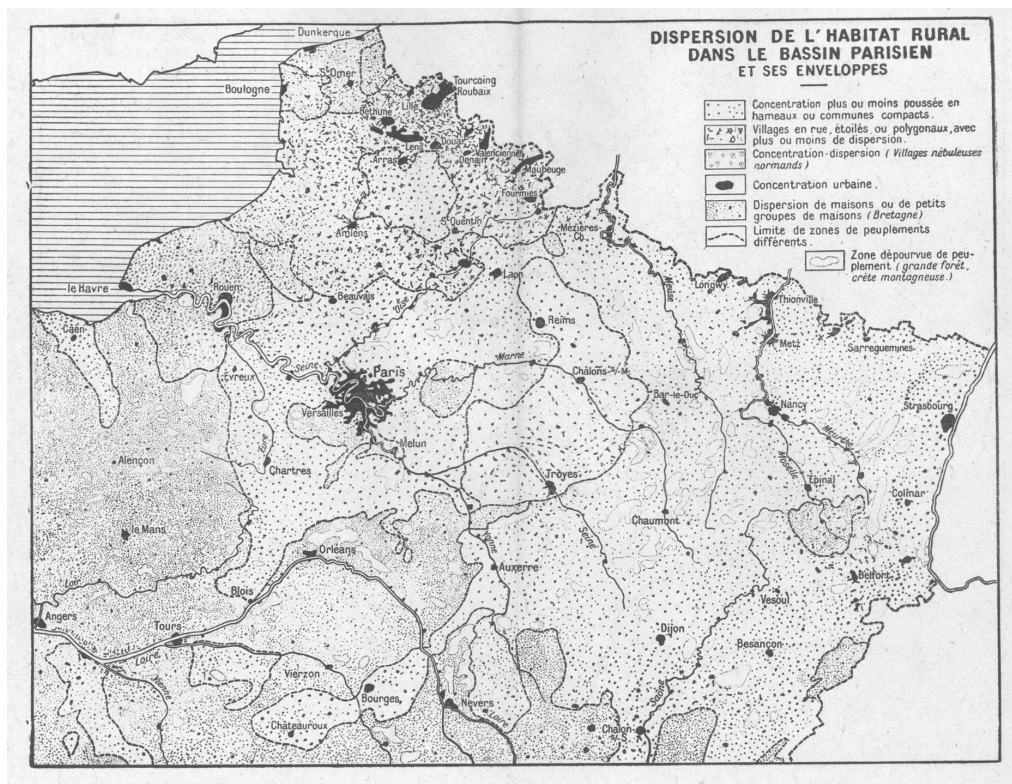


FIGURE 5.12 – Dispersion de l’habitat rural dans le bassin parisien. Source : COLAS 1945

par les méthodes de classification statistiques. Les classes qui émergent implicitement à l’issue d’une phase de labellisation avec les réseaux de neurones sont plus subjectives, mais aussi plus à même de restituer la richesse des phénomènes géographiques qui sont toujours protéiformes. Cette démarche n’est toutefois pas nouvelle en géographie. Avant l’avènement de la géographie quantitative, les géographes élaboraient « à la main » des cartes décrivant la variabilité de distributions spatiales. COLAS 1945 réalise de cette manière une carte de la répartition de l’habitat rural (fig. 5.12) très ressemblante à celles construites à partir de l’écart à l’invariance d’échelle (SÉMÉCURBE, TANNIER et ROUX 2019). Il justifie ce travail de la façon suivante :

« On conçoit qu’il soit difficile de faire entrer dans une formule simple tous les types de l’habitat avec leurs divers caractères. La constitution d’une formule est impossible lorsqu’on essaie de tenir compte de la morphologie des groupes. Même la représentation la plus satisfaisante mathématiquement, celle qui utiliserait les courbes de fréquence, ne peut tenir compte de ce dernier facteur. En conséquence, nous avons été amenés à concevoir une carte à échelle suffisamment grande pour pouvoir représenter en même temps l’intensité du semis et la forme physique des groupes, généralisant ainsi la méthode adoptée par Mlle Lefèvre au Congrès de Florence pour la représentation de l’habitat en Belgique. (...) La plus grande difficulté a été de représenter par un piqueté identique des régions qu’il n’a pas été possible de traiter en même temps. La carte ayant été faite en peu de jours, en soixante heures de travail effectif, nous pensons avoir conservé à peu près "la même main", ce qui est très important pour un travail de ce genre, et avoir obtenu un résultat assez homogène. »

Pour conclure, *Deep learning* et analyse fractale n'ont pas vocation à être opposés. L'analyse fractale peut être mise en œuvre pour objectiver à travers les échelles les structures définies dans les phases de labellisation et le *deep learning* peut servir à détecter des formes rares ignorées par les analyses fractales. Dans tous les cas, l'exploration visuelle des distributions des implantations humaines à l'aide de cartes restera un impératif du métier de géographe, ce que nous ne pouvons qu'approuver tant nous avons pris de plaisir à le faire durant ce travail de thèse.

Chapitre 6

Annexes

Annexe A

Spatial Distribution of Human Population in France : Exploring the Modifiable Areal Unit Problem Using Multifractal Analysis

Spatial Distribution of Human Population in France: Exploring the Modifiable Areal Unit Problem Using Multifractal Analysis

François Sémécurbe¹, Cécile Tannier², Stéphane G. Roux³

¹ThéMA UMR 6049, CNRS-Univ. Bourgogne Franche-Comté, Besançon Cedex, France, ²Chrono-Environnement UMR 6249, CNRS-Univ. Bourgogne Franche-Comté, La Bouloie - UFR Sciences et Techniques, Besançon Cedex, France, ³Laboratoire de Physique UMR 5672, CNRS-Ecole Nationale Supérieure de Lyon, Lyon Cedex 07, France

Case studies in geography are strongly dependent on the size of the spatial units used for the analysis. This has been expressed as the Modifiable Areal Unit Problem (MAUP): whatever the phenomenon under consideration, it is impossible to identify a single spatial partition that would be most appropriate to analyze it. In this respect, multifractal analysis may be an interesting tool for geographers. It integrates not just a series of nested spatial resolutions, as fractal analysis does, but also a series of points of view about the quantity of information contained in each spatial unit. In this article, we first expose the mathematical bases of multifractal analysis and we describe how it applies to geographical analyses. We insist on the mathematical notion of dimension, which allows us to describe how multifractal parameters can be used to quantify the MAUP. Then, we use the method to characterize the spatial distribution of population density in France. The main result is a typology map of population density that uses the MAUP as a descriptive tool. This map allows the joint identification of several phenomena: the main cities, the rural settlement patterns, and several types of periurban settlement patterns.

Introduction

In settlement geography, scholars observe the spatial distribution of population by dividing space into more or less regular partitions and then calculating the population density in each spatial unit of the partition. This allows them to analyze the spatial variation of settlement intensity.

However, when analyzing the spatial distribution of population, we observe that population density varies greatly with the spatial resolution of the space partition under consideration. For instance, the average population density in mainland France in 2010 is 120 inhabitants per km². In the Île de France region (around Paris), the population density is 1,000 inhabitants per

Correspondence: Cécile Tannier, Chrono-Environnement UMR 6249, CNRS-Univ. Bourgogne Franche-Comté, La Bouloie - UFR Sciences et Techniques, 16, route de Gray, F-25 030 Besançon Cedex, France

e-mail: cecile.tannier@univ-fcomte.fr

Submitted: January 12, 2015. Revised version accepted: December 18, 2015.

km² whereas population density in the Limousin region is 44 inhabitants per km². Considering a finer spatial resolution, differences in population density are greater: the subregion (i.e., *département*) of Paris has a very high population density (20,000 inhabitants per km²) whereas the subregion Creuse in the Limousin region is almost deserted (22 inhabitants per km²).

Thus, case studies in geography are highly dependent on the size of the spatial units used for the analysis. This has been expressed by Openshaw (1983) as the Modifiable Areal Unit Problem (MAUP): whatever the phenomenon under consideration, it is impossible to identify a single spatial partition that would be most appropriate for analyzing it.

In this respect, fractal analyses are interesting because they allow us to analyze a phenomenon across spatial units of varying sizes. In geography, fractal analyses are usually performed on binary images representing the presence or absence of a phenomenon. They are more often used for analyzing the shape of built patterns within urban agglomerations or urban regions (see e.g., Batty and Xie 1996, 1999; Benguigui et al. 2000; Shen 2002; Benguigui, Blumenfeld-Lieberthal, and Czamanski 2006; Thomas, Frankhauser, and Biernacki 2008; Feng and Chen 2010). Fractal dimensions enable us to distinguish easily between built-up patterns characterized by high diversity in the size of the built clusters and in the distances separating these clusters (fractal dimension close to 1.5) and uniform built-up patterns in which buildings may be either scattered or concentrated (fractal dimension close to 2) (Thomas, Frankhauser, and De Keersmaecker 2007). As set by Chen (2013), fractal indexes characterizing urban patterns are consistent with one another only if the radial fractal dimension falls between 1.5 and 2.

Considering a 2D built pattern, the built density ρ is reduced when passing from a spatial partition at a given spatial resolution i to a finer one at a resolution $i+1=i/r$ according to the equation:

$$\frac{\rho_i}{\rho_{i+1}} = r^{-2-D} \quad (1)$$

where D is the fractal dimension and r is the reduction factor from a given spatial resolution to a finer one (Tannier et al. 2012). The built density measured for any spatial resolution remains the same only if the fractal dimension of the built pattern is equal to 2.

Such fractal analyses consider the self-similarity of built patterns under the assumption that each pattern is characterized by a single self-similarity dimension. Yet, several scholars have shown that the self-similarity dimension is not always constant in reality (Goodchild 1980); it is usually constant over a limited range of scales but varies somewhat over successive ranges of scales (Lam 1990; Lam and Quattrochi 1992; White and Engelen 1994; Tannier and Pumain 2005). Other scholars have suggested that urban patterns are characterized by various fractal properties (Batty and Longley 1994; Frankhauser 2008), which explains why a single fractal dimension may not always fully describe them. As an alternative, they propose using multifractal analysis.

Whereas monofractal analyses are performed on binary images representing the presence or absence of a phenomenon (e.g., built-up surfaces), multifractal analyses also take into account the relative amount of the phenomenon in each place. The intensity of the phenomenon under consideration varies with the spatial aggregation level. Multifractal analysis can be thought of as an extension of monofractal analysis by the addition of information about the relative intensity in each place of the phenomenon under consideration. It integrates not just a series of nested spatial resolutions (i.e., analyzing the phenomenon using spatial units of nested sizes), as monofractal analysis does (Frankhauser 1998; Stanley et al. 1999), but also a series

of points of view about the quantity of information contained in each spatial unit. Therefore, the effects of the MAUP are fully taken into account.

In urban geography, Chen and Wang (2013) have applied multifractal analysis to characterize the built pattern of Beijing at several points in time; Ariza-Villaverde, Jiménez-Hornero, and Ravé (2013) have used it to characterize urban street patterns; other scholars have used it to characterize systems of cities and rank-size distributions of settlements (Haag 1994; Chen and Zhou 2004). To our knowledge, only two applications of multifractal analysis for analyzing the spatial distribution of population have been published. Appleby (1996) used multifractal analysis to measure the spatial distribution of population in the United States and Great Britain. He considered the multifractal organization of each settlement pattern to be homogeneous within each country. That strong assumption was contested by Orozco, Golay, and Kanevski (2015), who compared the settlement patterns of three Swiss regions and found differences in the multifractal organization of the population distribution between the regions.

In this article, we use multifractal indexes for quantifying the MAUP. Our aim is to use the MAUP as a descriptive tool for identifying local differences in the spatial distribution of population density in France. By varying the size of the spatial units under consideration, we analyze the population density through a series of nested spatial partitions. By changing the focus on the local population densities, we consider either places where population is highly concentrated, or places where population is not highly concentrated. As a limiting case, we also consider the mere presence or absence of population regardless of local density (i.e., monofractal analysis).

The article is organized as follows. “Multifractal analysis: mathematical bases” section sets out the mathematical bases of multifractal analysis. “Multifractal analysis: practical application” section describes how it applies to geographical analyses. “Results” section reports the analysis results of the spatial distribution of population density in France. “Discussion” section discusses the results in the light of the centuries-old history of French rural settlements and recent changes in periurban areas. “Conclusion” section concludes the article.

Multifractal analysis: mathematical bases

Multifractal analysis is about studying local variations of a measure (Jaffard, Lashermes, and Abry 2007). A measure is similar to a probability: it attributes a size to a set of points. Considering any set of points F of the Euclidean plane, the measure $\mu(F)$ of the spatial distribution of population is simply the number of individuals contained in F :

$$\mu(F) = \text{Population contained in } F \quad (2)$$

The population density of the set of points F corresponds to the ratio between the number of inhabitants in F and its area. When the supporting space considered (i.e., the spatial footprint of the set F) is small enough, the population density is said to be local. Exploring the MAUP consists in studying the way the local density changes when this supporting space becomes smaller.

In terms of mathematics, the local density is said to be regular if it varies slowly when the area of the supporting space decreases. Conversely, the local density is irregular when it increases suddenly -it diverges to infinity- when the area of the supporting space decreases. The regularity of local density around the point x of the Euclidean plane can be measured with the Hölder exponent $\alpha(x)$. The closer to 0 the exponent is, the more irregular the local density is. Very strong local irregularities are called singularities. Conversely, high Hölder exponents

characterize regular places where the variations in the local density are weak when the area of the supporting space decreases.

The measure $\mu(F)$ can exhibit either only local regularities, or only irregularities, or a mix of both. Multifractal analysis allows us to quantify this. The first step in the analysis consists in dividing the space supporting the measure into subsets of the same regularity α . Those subsets are called iso-Hölder sets. The area of iso-Hölder sets cannot usually be measured because their shape is too complex. Consequently, each iso-Hölder set is characterized by its Hausdorff dimension (see “Fractal box-counting dimension and Hausdorff dimension” section). The relation between each Hölder exponent value and its corresponding Hausdorff dimension gives the multifractal spectrum $f(\alpha)$ of the spatial distribution of the measure $\mu(F)$ (see “Hölder exponent $\alpha(x)$ and singularity spectrum $f(\alpha)$ ” section). The Hausdorff dimension is almost impossible to use in practice. Consequently, the multifractal spectrum is obtained through a Legendre transform of two other multifractal parameters: the scaling exponent τ_q and the generalized dimension D_q , which allow the quantification of the MAUP for a measure $\mu(F)$ (see “Multifractal quantification of the modifiable areal unit problem” section).

NB: *although the mathematical definitions and examples are presented in \mathbb{R}^2 , they remain valid and can be generalized for any vector space \mathbb{R}^n .*

Fractal box-counting dimension and Hausdorff dimension

In mathematics, a dimension measures the size of a set of points. In \mathbb{R}^0 , \mathbb{R}^1 , and \mathbb{R}^2 , the size of a set of points corresponds respectively to its length, its area and its volume.

One possible way to measure the area of a set of points F in \mathbb{R}^2 is to cover F by objects whose area is perfectly known. The sum of the areas of the objects covering F gives approximately the total area of F . The smaller the covering objects are, the more precise the approximation of the area of F is. When the covering objects are infinitely small, we may hope to obtain the exact area of F .

Yet, mathematicians have long been confronted with the difficulty of measuring the length, area, and volume of complex shapes. For instance, it is impossible to measure the length of a line by covering it with line segments (i.e., one-dimensional shapes) if the line is contained in \mathbb{R}^2 but not in \mathbb{R} . Similarly, it is impossible to measure the area of a surface by covering it with circles (i.e., two-dimensional shapes) if this surface is contained in \mathbb{R}^3 but not in \mathbb{R}^2 . Fractal dimensions are used for measuring the size of such complex shapes.

Box-counting dimension

One way to define the box-counting dimension calls on the use of a δ -grid (Falconer 2013). A δ -grid is obtained by partitioning space into ordered square grid cells of side length δ . Considering the Euclidean plane \mathbb{R}^2 , the cells of a δ -grid are defined as follows:

$$[o_x + m_x\delta, o_x + (m_x + 1)\delta] \times [o_y + m_y\delta, o_y + (m_y + 1)\delta] \quad (3)$$

where o_x and o_y are real numbers representing respectively the origin of the grid’s x -axis and y -axis. For a given δ -grid, o_x and o_y are the same for all cells. m_x and m_y are positive integers corresponding to the row and column numbers of the grid cells. Cells are numbered from 1 to n .

The box-counting dimension of a set of points F of the plane is given by:

$$\dim_B(F) = \lim_{\delta \rightarrow 0} \frac{\log(N_\delta(F))}{-\log(\delta)} \quad (4)$$

where $N_\delta(F)$ is the number of δ -grid cells that intersect F .

Hausdorff dimension

The problem with the box-counting dimension is that $\lim_{\delta \rightarrow 0}$ does not always exist for any set of points F . The Hausdorff dimension can overcome this problem using a δ -cover instead of a δ -grid (Falconer 2013). A δ -cover is a partition of F into parts U_i , which may have any shape and whose diameter is at most equal to δ . U_i are called hyperballs. The diameter of a hyperball is defined as the distance between the two elements of it that are the farthest apart. For any F , δ -covers always exist. The Hausdorff measure of the dimension of the set F varies according to the chosen δ -cover. The Hausdorff measure of dimension d (i.e., for a given δ -cover) is defined as:

$$\mathcal{H}^d(S) = \lim_{\delta \rightarrow 0} \inf \left\{ \sum_{i=1}^{+\infty} |U_i|^d : \{U_i\} \text{ is a } \delta\text{-cover of } F \right\} \quad (5)$$

If $\mathcal{H}^d(S)$ is zero then the dimension of F is lower than d . Accordingly, the Hausdorff dimension $\dim_{\mathcal{H}}(F)$ of the set F is the smallest value of d for which the Hausdorff measure $\mathcal{H}^d(F)$ is nonzero:

$$\dim_{\mathcal{H}}(F) = \inf \{d : \mathcal{H}^d(F) = 0\} \quad (6)$$

δ -covers are extremely difficult to use in practice, which explains why the Hausdorff dimension is only used for theoretical purposes. Contrarily, the box-counting dimension is usable in practice for measuring the size of real objects although it does not exist for any set F . The box-counting dimension is estimated by calculating the slope of the curve $\log(N_\delta)$ as a function of $-\log(\delta)$

(see equation 4) for the range of δ -values for which the curve is linear.

Hölder exponent $\alpha(x)$ and singularity spectrum $f(\alpha)$

The Hölder exponent $\alpha(x)$ measures the irregularity of the local density of a measure μ (for instance, a local population density) around a point x .

$$\alpha(x) = \lim_{\delta \rightarrow 0} \frac{\mu(B(x, r))}{-\log(\delta)} \quad (7)$$

where $B(x, r)$ is the disc of radius r centered on the point x . This definition is valid only if the $B(x, r)$ is always positive for any radius. The set of points having this property is the support of the measure μ .

The Hölder exponent is directly related to the local density density (x, r) :

$$\text{density}(x, r) = \frac{\mu(B(x, r))}{\pi r^2} \simeq \frac{r^{\alpha(x)}}{\pi r^2} \simeq r^{\alpha(x)-2} \quad (8)$$

$\alpha(x) < 2$ In this case, $\text{density}(x, r)$ tends to infinity when r tends to 0. As $\alpha(x)$ comes close to 0, $\text{density}(x, r)$ diverges faster toward infinity;

$\alpha(x) > 2$ In this case, density(x, r) tends to 0 when r tends to 0. As $\alpha(x)$ increases, density ($B(x, r)$) converges faster toward 0;

$\alpha(x)=2$ In this case, the behavior of density(x, r) cannot be determined. As r tends to 0, density (x, r) tends either toward a nonzero value or to infinity, or even converges to 0. The divergence of density (x, r) toward infinity or its convergence toward 0 is much slower than in the case of $\alpha(x) \neq 2$.

By measuring the dimension of the sets of points x characterized by the same Hölder exponent $\alpha(x)$, we obtain the singularity spectrum (also called multifractal spectrum) $f(\alpha)$, namely the Hausdorff dimension of iso-Hölder sets of points $F_\alpha = \{x : \alpha(x) = \alpha\}$:

$$f(\alpha) = \dim_{\mathcal{H}}(F_\alpha) \quad (9)$$

$f(\alpha(0))$ gives the maximum value of the singularity spectrum. It indicates the strength of the irregularities of the measure. Complementarily, the range of the spectrum informs us about the multifractality of the spatial distribution of the measure (more or less strong, even nonexistent in the case of a monofractal distribution) (Wendt et al. 2007).

The singularity spectrum $f(\alpha)$ complies with the following property:

$$0 \leq f(\alpha) \leq \alpha \quad (10)$$

The left-hand side of the inequality is trivial and results from the properties of fractal dimensions. The right-hand side of the inequality is much more interesting because it implies that the intensity of local singularities constrains their spatial distribution: a very singular behavior cannot occur in too many places.

As the Hausdorff dimension is not usable in practice, the direct calculation of the singularity spectrum using equation (9) is not possible. Instead, the multifractal formalism can be used (Jaffard, Lashermes, and Abry 2007). It enables the determination of the singularity spectrum through a Legendre transform of the scaling exponent τ_q :

$$f(\alpha) = \inf_{q \in \mathbb{R}} (q\alpha - \tau_q) \quad (11)$$

Multifractal quantification of the Modifiable Areal Unit Problem

Scaling exponent τ_q

A map of population density displays graphically the population contained in spatial units. In the case of a raster map built on a δ -grid $G_{\delta,i}$, the population contained in a given cell i corresponds to the measure $\mu(G_{\delta,i})$. In such a case, the MAUP only depends on the variation in spatial resolutions (i.e., variations of the values of δ) and not on the shape of the spatial units under consideration. The analysis of the behavior of the measure $\mu(G_{\delta,i})$ through varying δ -values allows the effect of the MAUP on the mapped population density to be determined. For this, one possible way is to abstract the spatial distribution of the measure $\mu(G_{\delta,i})$ for each δ -value using an appropriate function, and then to observe the result of the function when δ changes. Obviously, a single function is not enough to fully characterize the spatial distribution of the measure $\mu(G_{\delta,i})$. Alternatively, a group of functions, namely the structure functions $S_{q,\delta}(\mu)$, are suitable for this:

$$S_{q,\delta}(\mu) = \sum_{\mu(G_{\delta,i}) > 0} \mu(G_{\delta,i})^q \quad (12)$$

Multifractal analysis requires us to set the assumption that the structure functions $S_{q,\delta}(\mu)$ follow a scaling law when the values of δ are small:

$$S_{q,\delta}(\mu) \approx \delta^{\tau_q} \quad (13)$$

Here τ_q is a scaling exponent. It can be calculated using the following formula:

$$\tau_q = \lim_{\delta \rightarrow 0} \frac{\log(S_{q,\delta}(\mu))}{-\log(\delta)} \quad (14)$$

As for estimating the box-counting dimension, it is impossible to reduce δ toward 0 indefinitely. Consequently, the estimation of τ_q is given by the slope of the curve of $\log(S_{q,\delta}(\mu))$ as a function of $\log(\delta)$.

Generalized dimensions D_q

On the basis of τ_q , it is possible to define the generalized dimensions D_q :

$$\begin{cases} D_q = \frac{1}{1-q} \tau_q \text{ pour } q \neq 1 \\ D_1 = \lim_{\delta \rightarrow 0} \sum_{\mu(G_{\delta,i}) > 0} \mu(G_{\delta,i}) \log(\mu(G_{\delta,i})) \end{cases} \quad (15)$$

D_q with respect to q is called the dimension spectrum. The coefficient q acts as a filter when calculating the generalized dimensions. For $q \rightarrow +\infty$, only the cells in which high measures are concentrated are taken into account. Conversely, for $q \rightarrow -\infty$ generalized dimensions describe the spatial distribution of the cells in which low measures are concentrated. For $q = 1$, the filter does not give more or less importance to the cells with respect to the intensity of the measure. For $q = 0$, the generalized dimension does not take into account the spatial variation of the measure; only the fact that the cells are populated or not is taken into account. Because of this sort of filtering operated by q , generalized dimensions D_q allow a characterization of the spatial distribution of population that is both multiscale (via the introduction of the δ -grid) and multiviewpoint (via the coefficient q).

Singularity spectrum, dimension spectrum, and the MAUP

When the multifractal parameters τ_q , D_q and $f(\alpha)$ are differentiable, the Legendre transform of equation (11) becomes:

$$\alpha(q) = \frac{d}{dq} (q-1)D_q \quad (16)$$

$$f(\alpha(q)) = \alpha q - (q-1)D_q \quad (17)$$

These two equations show that both the singularity spectrum $f(\alpha(q))$ and the dimension spectrum (D_q with respect to q) contain the same information. Their relationships and possible uses are discussed in (Chen and Wang 2013). Our interest here is that both spectrums enlighten the MAUP.

- The absence of MAUP means that the spatial distribution of population is almost uniform. In this case, the dimension spectrum is a horizontal line with the y -intercept equal to 2 and the singularity spectrum is the point of coordinates (2, 2).
- When the measure is monofractal then the singularity spectrum $f(\alpha(q))$ is only a point of coordinates $\alpha(q)=f(\alpha(q))=D_0$. In this case, the MAUP only results from the spatial distribution of the populated places. Basically, the study area contains identical buildings all having the same number of inhabitants and being spatially distributed in a monofractal manner. In this case, local differences in population density come only from the spatial distribution of the buildings.
- A singularity spectrum having a range of $\alpha(q)$ values indicates a multifractal behavior. In this case, the spatial footprint of the buildings varies (e.g., small single-family houses, multifamily houses, high and large collective buildings) as well as their number of inhabitants: the MAUP results from both.

The singularity spectrum $f(\alpha)$ is easier to use than both τ_q and the generalized dimensions D_q because $f(\alpha)$ is like a probability density function (Chevallard et al. 2012). For this reason, the singularity spectrum $f(\alpha)$ is used in image processing for analyzing the texture of images (Wendt et al. 2009), and more particularly for image classification. In this article, we use the singularity spectrum in the same way to obtain a typology map of the spatial distribution of human population in France (see “Classification of spatial units according to their multifractal spectrum $f(\alpha)$ ” section).

Multifractal analysis: practical application

Data

For this study we used the 200 m grid population 2010 data set provided by the French National Institute of Statistics and Economic Studies (INSEE) for mainland France. These data result from counting georeferenced households within georeferenced quadrilateral grid cells having a side length of 200 m. One advantage of this data set is that counted households are precisely located: the data of the INSEE’s “Revenus Fiscaux Localisés” database, which contains the taxable income of all households in France characterized by the address of their tax domicile (i.e., their principal place of residence) and the number of individuals in the household, are georeferenced by matching them against geographic information layers of the land registry. Fewer than 0.4% of the households contained in the “Revenus Fiscaux Localisés” database cannot be located with 200 m precision. In general, the reason is that their address is not related to a land parcel in the land registry. This share of 0.4% of households is sufficiently dispersed spatially for the quality of the 200 m grid population data set to be entirely satisfactory for the present study.

NB: data used as well as the R code written to make the analyses are available online as “Supporting Information”.

Estimation of multifractal indexes

Estimation of the generalized dimensions D_q

In practice, the generalized dimensions D_q are defined as:

$$D_q = \lim_{\delta \rightarrow 0} \frac{\frac{1}{1-q} \log\left(\sum_{i=1}^n \mu(G_{\delta,i})^q\right)}{-\log(\delta)} \quad (18)$$

By convention, $0^q = 0$ for any $q \in \mathbb{R}$.

The numerator of a generalized q -dimension is the generalized entropy defined by Renyi (Appleby 1996). When q tends to 1, the generalized dimension converges on the information dimension. When $q = 0$, the generalized dimension D_0 is the fractal box-counting dimension of the support of the measure $\mu(G_{\delta,i})$. When $q = 2$, the generalized dimension D_2 is called the correlation dimension. It is associated with the probability that two individuals selected randomly will be a distance δ apart (Seuront 2009).

In theory, the generalized dimensions are obtained by studying the asymptotic behavior of the generalized entropy when the resolution δ of the δ -grid tends to 0. However, it is impossible to reduce δ toward 0 indefinitely: when applying the box-counting method on real data, the minimum possible size of a δ -grid depends on the resolution of the data themselves. Consequently in practice, generalized dimensions are estimated through a linear regression relating the generalized entropy to $-\log(\delta)$ by the ordinary least squares method. Generalized dimensions correspond to the slope of the regression line. Of course, the goodness-of-fit of the linear regression has to be satisfactory. Visually, the plotted points relating entropies (on the y -axis) with $\log(\delta)$ (on the x -axis) have to form a straight line. This requires selecting a range of δ values for which the goodness-of-fit of the linear regression is satisfactory.

Estimation of the multifractal spectrum $f(\alpha(q))$

To estimate the multifractal spectrum, we do not use the Legendre transform of τ_q , which is unstable in practice: we estimate the multifractal spectrum directly from the formula proposed by Chhabra and Jensen (1989):

$$\alpha(q) = \lim_{\delta \rightarrow 0} \frac{\sum_{i=1}^n P_q(\delta, i) \log(\mu(G_{\delta,i}))}{-\log(\delta)} \quad (19)$$

and

$$f(\alpha(q)) = \lim_{\delta \rightarrow 0} \frac{\sum_{i=1}^n P_q(\delta, i) \log(P_q(\delta, i))}{-\log(\delta)} \quad (20)$$

with

$$P_q(\delta, i) = \frac{\mu(G_{\delta,i})^q}{\sum_{i=1}^n \mu(G_{\delta,i})^q} \quad (21)$$

where $\{G_{\delta,i}\}$ is a δ -grid.

As when estimating the generalized dimensions, the multifractal spectrum is obtained by a linear regression relating the numerator of equation (19) (respectively 20) to $-\log(\delta)$ considering the same range of δ values as when estimating the corresponding generalized dimensions.

Choice of the parameters involved in the multifractal analysis of the spatial distribution of population in France

To compare local settlement patterns with multifractal analysis, we need to divide mainland France into regular spatial units. We choose a regular square partitioning into spatial units ζ

(Figs. 1 and 2). The size of the square spatial units ζ depends on and determines the scale range (i.e., series of spatial resolutions δ) for estimating the multifractal indexes. Consequently, both the size of the spatial units and the scale range of analysis have been chosen jointly. Yet, for the sake of clarity these two aspects are set out successively in what follows.

Choice of the scale range of δ -values for estimating multifractal indexes

Comparing multifractal indexes across a large series of spatial units ζ involves setting a single scale range of δ -values for all spatial units. This requirement strongly constrained our study because the settlement pattern in each spatial unit may be very different and, consequently, the scale range that allows the best estimation of multifractal indexes for each spatial unit may be also different.

Considering this, the choice of the scale range of δ values for estimating multifractal indexes results in a tradeoff between statistical goodness-of-fit and geographical purposes. We conducted a series of tests consisting in calculating multifractal indexes for several portions of mainland France with very different spatial distributions of population. In general, we observed the following phenomena:

- at very fine spatial resolutions (side length of the δ -grid cells less than 800 m), grid population data are not precise enough to allow a good multifractal description of population distributions. At those scales, simple fractal analyses of built patterns performed on detailed data (e.g., land registry data - see for instance (Tannier and Thomas 2013)) provide more information;
- at coarse spatial resolutions (side length of the δ -grid cells larger than 6,400 m), the scaling behavior corresponds to a fractal degenerated behavior. At this scale range, every grid cell of a spatial unit ζ is populated. The dimension of the support D_0 is very close to 2. Besides this, the use of coarse spatial resolutions would imply defining very large square spatial units ζ to ensure a good statistical estimation of multifractal indexes, which is not relevant for analyzing local spatial distributions of population;
- by contrast, considering medium spatial resolutions (side length the δ -grid cells between 800 and 6,400 m) the multifractal analysis of grid population data interestingly informs us about the effect of urban sprawl on regional settlement patterns.

Choice of the size of the square spatial units ζ

According to the chosen scale range of δ -values, the square spatial units ζ must have a minimum size to allow a good estimation of the slope of regression lines used for calculating multifractal indexes. Knowing that the maximum resolution δ is equal to 6,400 m, the size of the square spatial units ζ has to be large enough. Otherwise, calculated entropies would be null or quasi-null, which would bias estimations of multifractal indexes. Considering this, the size chosen for the square spatial units ζ is 25 km side length (Fig. 1). The spatial units ζ are therefore almost four times larger than the larger δ -grid cells when the resolution δ is 6,400 m.

With a side length of 25 km, the square spatial units ζ divide the largest French metropolitan areas quite finely: on average, the spatial units ζ partition large metropolitan areas into nine portions. The average size of the 10 most populated metropolitan areas is 4,900 km², which corresponds approximately to a diameter of 79 km. This partitioning of metropolitan areas enables us to distinguish the urban centers from their periphery. Moreover, it is possible to



Figure 1. Nine hundred ninety two square spatial units regularly dividing the area of mainland France.

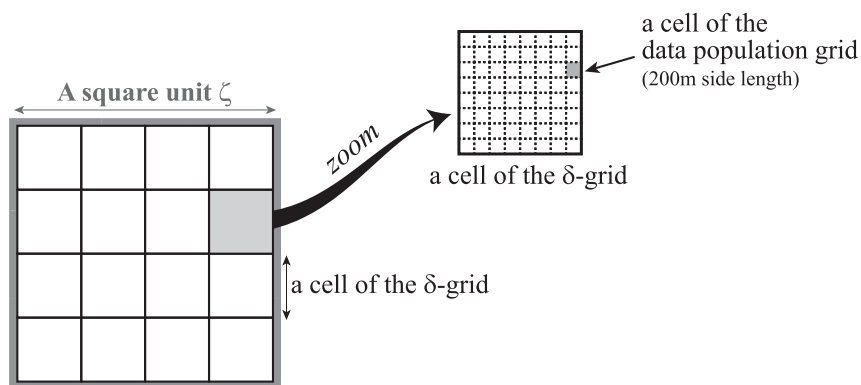


Figure 2. A square spatial unit ζ , δ -grid cells, and 200 m grid population cells.

compare the spatial distributions of population in periurban areas to those of areas bordering on periurban areas. Thus metropolitan areas can be analyzed in their surrounding context.

No scale range can be unequivocally associated with a given geographical phenomenon. Yet from the arguments exposed above, it can be surmised that geographical phenomena change differently with different scales of analysis. Built patterns change quickly at fine scales as a consequence of urban sprawl. Conversely, the macro-structure of the French settlement system remains unchanged at large scales. Finally, an intermediate scale range (intermediate size of both δ -grids and square spatial units ζ) seems to be well adapted for studying changes in the spatial distribution of population. According to Orozco, Golay, and Kanevski (2015), this intermediate scale range is suitable for studying the influence of societies on space.

Choice of the appropriate δ -grid position for calculating entropies for each dimension

Each grid cell of a δ -grid results from the aggregation by longitude and latitude of k cells of the grid population data set. Accordingly, considering a given k corresponding to a given

resolution δ , a set of k^2 δ -grids can be built by varying the origins o_x and o_y of respectively the grid's x -axis and y -axis 200 m by 200 m.

In theory, the position of the δ -grid does not influence the value of multifractal indexes. Appleby (1996) notices, however, the variation in entropies depending on the chosen δ -grids. As he suggests, we selected, among the k^2 possible δ -grids, the δ -grid for which we obtained the lowest entropy value. This selection was made at each resolution δ for each q value and each entropy. Such a methodological choice is based on the consideration of the very nature of entropy. In information theory, a source transmitting information can be characterized by a probability distribution, namely optimal coding, which is the shortest information coding by which the result of a given action can be transmitted (for instance, transmitting the information that the result of a throw of dice is the number 2). The entropy measures the length of this coding. In geographical multifractal analysis, entropy represents the minimum amount of information required for locating an individual in a δ -grid. When choosing the δ -grids associated with the lowest entropy value we consider the optimal information coding that characterizes the population distribution.

Selection of a set of q values

In principle, q can be chosen in \mathbb{R} . However, when q is negative, the multifractal parameter estimates are very unstable: when focusing on the cells with very little population, the generalized entropies are very sensitive to the positioning of the δ -grid. The generalized dimensions are also unstable when q values are very high because only few cells of maximum measures are taken into account in the calculation. To avoid these estimation problems, we chose here to restrict the range of multifractal analyses to q in $\{0, 0.5, 1, 1.5, 2\}$.

Results

When estimating multifractal indexes, we need to verify the quality of the estimation. As shown in Table 1, R^2 values calculated are satisfactory: more than 90% are higher than 0.96. Only the quality of the estimation of $\alpha_2, f(\alpha_2)$ is slightly lower.

Maps of generalized dimensions D_q

Fig. 3 displays three maps of generalized dimensions for $q = 0, 1, 2$. Each map exhibits a clear spatial structure, which differs from one map to the other. As explained in “Multifractal analysis: practical application” section, the quality of data used does not vary spatially. Hence we assume that the revealed spatial structures do not result from calculation artefacts but correspond to real structures of the French settlement pattern.

On the map of generalized dimensions D_0 , the spatial organization of mainland France consists of very large areas of the same dimension. Generalized dimensions are high (in general higher than 1.5) in the West and the South-West, the Rhone valley, and the largest metropolitan areas (see Fig. 4 for the location of each element). By contrast, generalized dimensions are low in the East of France (Champagne, Ardennes, Lorraine, and Vosges) as well as in mountain and forest zones. The fact that areas characterized by high dimensions are either rural zones or metropolitan areas illustrates the indifference of the generalized dimensions D_0 to the spatial variations of the density measure. As stated above, when $q = 0$ the generalized dimensions take into account only the fact that the cells are populated or not, but ignore their population density.

Table 1. Distribution of R^2 Values Obtained by Estimating the Multifractal Indices

Quantile	D_0	D_1	D_2	α_0	α_1	α_2	$f(\alpha_0)$	$f(\alpha_1)$	$f(\alpha_{1.5})$	$f(\alpha_2)$
minimum	0.94	0.90	0.81	0.89	0.90	0.53	0.94	0.90	0.63	0.00
10%	0.98	0.99	0.96	0.96	0.99	0.88	0.98	0.99	0.93	0.64
25%	0.99	1.00	0.98	0.98	1.00	0.93	0.99	1.00	0.96	0.80
50%	1.00	1.00	0.99	0.99	1.00	0.97	1.00	1.00	0.98	0.91
75%	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	0.99	0.97

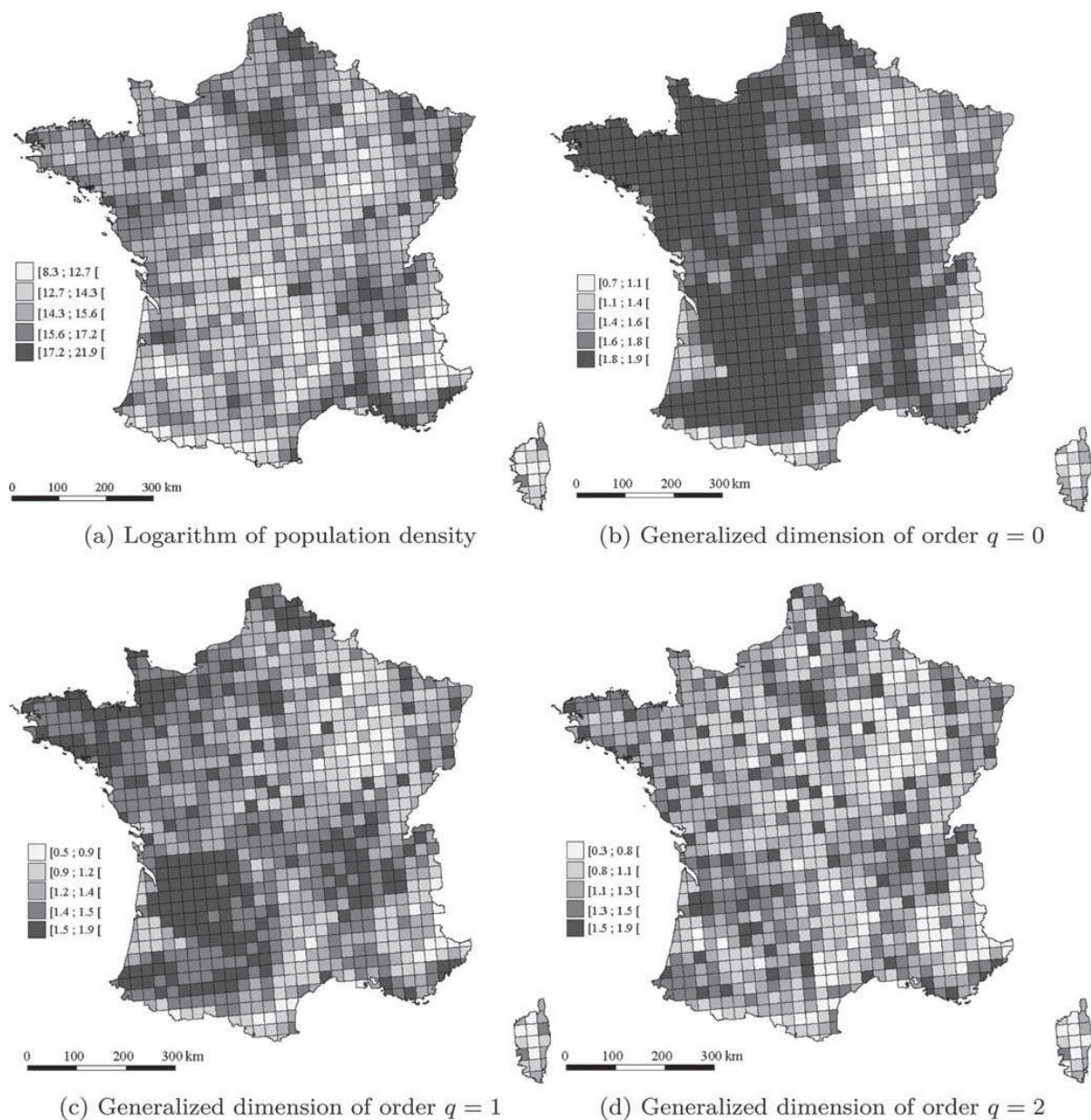
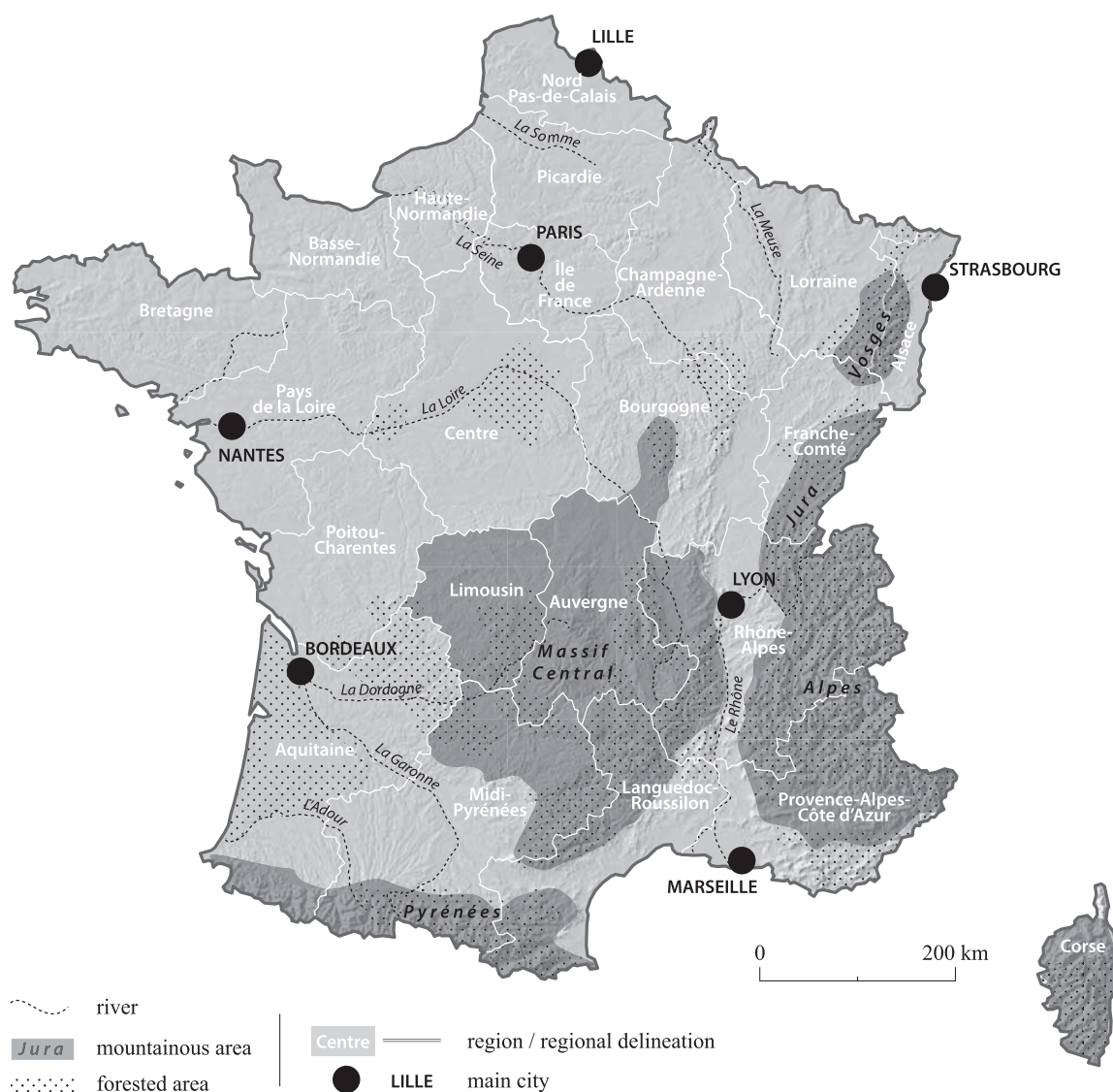


Figure 3. Population densities and corresponding multifractal generalized dimensions in France. *Source: 200 m grid population 2010 data set provided by the French National Institute of Statistics and Economic Studies (INSEE).*

In contrast, dense urban metropolitan areas (see Fig. 5) are very clearly identified on the map of generalized dimensions D_2 . They look like islands of high dimension D_2 surrounded by spatial units of low dimension. Multifractal analysis for $q = 2$ concentrates on very populated



Sources: BD ALTI® 1000m, GEOFLA®, SCAN1000®, IGN ; BDCARTHAGE®, Sandre

Figure 4. Mountains, forested areas, rivers, regions, and main cities in mainland France.

cells. Only urban spatial units have enough highly populated cells close enough together to reach high D_2 values. In rural zones, densely populated urban centers are too far apart to increase the value of D_2 . Last but not least, nonurban areas with high values of D_0 (in the West, the South-West, and in the Rhone valley) have lower values of D_2 . They stand out much less clearly on the map of generalized dimensions D_2 than on the map of generalized dimensions D_0 .

The map displaying the generalized dimensions D_1 exhibits an intermediate spatial configuration between the maps of D_0 and D_2 . This illustrates the continuous change of D_q with respect to q .

The fourth map represents the logarithm of population density, which has been here preferred to the raw population density because when mapping the latter, the spatial unit containing the city of Paris with its population density overburdens the other spatial units. Besides this, the logarithm of population density is in accordance with the rank-size (in population) distribution of urban areas: when transforming the population density of each spatial unit into its logarithm, spatial units are approximately classified according to their rank order considering their population.

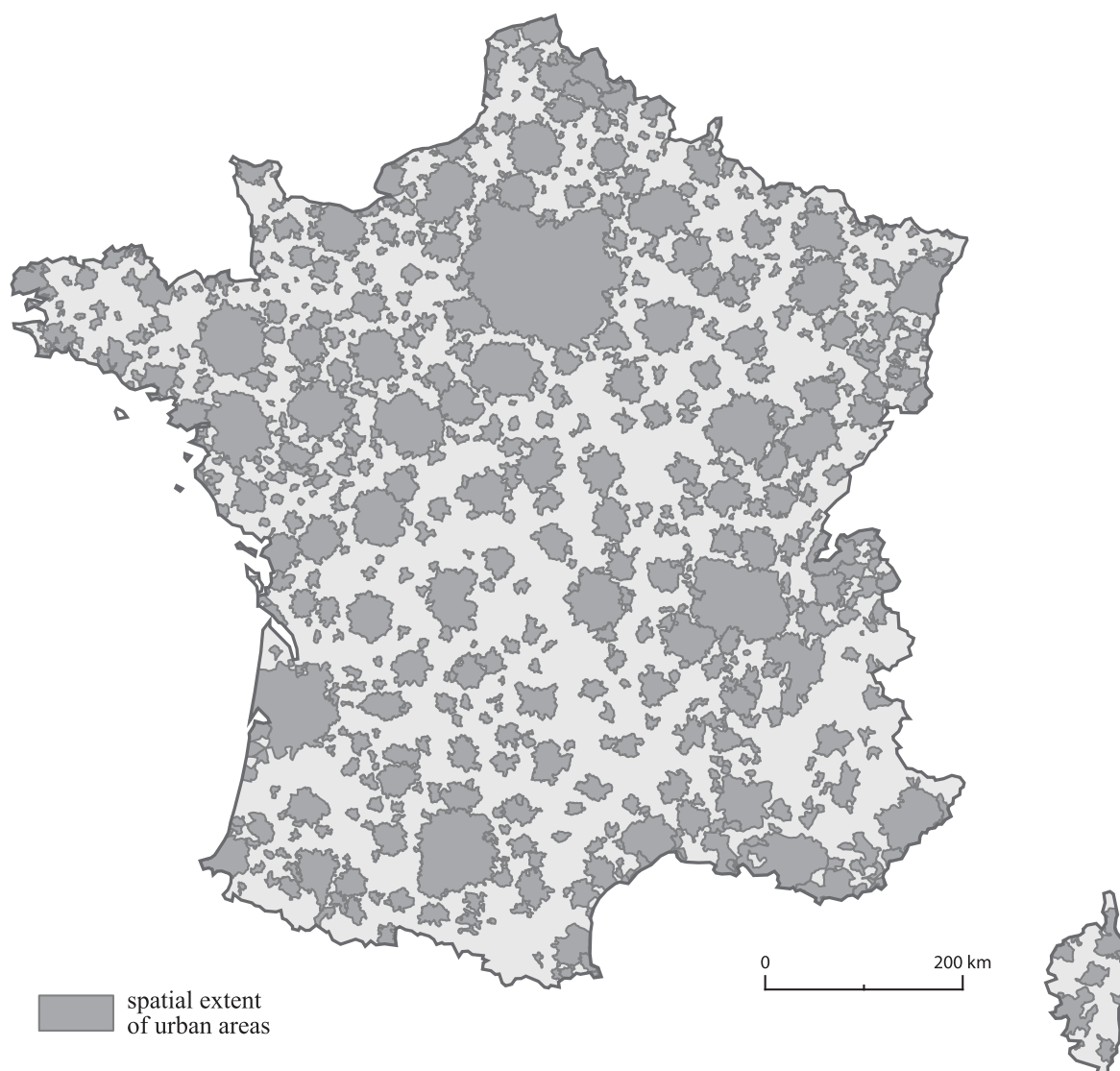


Figure 5. Delineation of French *aires urbaines* (metropolitan areas). In France, an *aire urbaine* encompasses a densely built *unité urbaine*—which is akin to a U.S. urban area—and its commuter belt. *Source: French National Institute of Statistics and Economic Studies (INSEE) 2010.*

Maps of q -dimensions yield different and nonredundant information than the map of population density. The main metropolitan areas stand out clearly on the population density map but not so clearly as on the map of generalized dimensions D_2 . Moreover, maps of generalized dimensions can be used for detecting a variety of settlement patterns in nonurban areas but not so the population density map. The fact that density is unable to capture such morphological variations is logical: density is an averaging index, which hides the variations of population concentrations across scales (Thomas, Frankhauser, and De Keersmaecker 2007). By contrast, generalized dimensions measure the variations of population density across scales.

Classification of spatial units according to their multifractal spectrum $f(\alpha)$

For defining a typology of spatial units according to their multifractal spectrum, we used a Hierarchical Cluster Analysis. HAC first involves measuring the discrepancy between the multifractal spectra of each couple of spatial units. Because multifractal spectra do not all have the

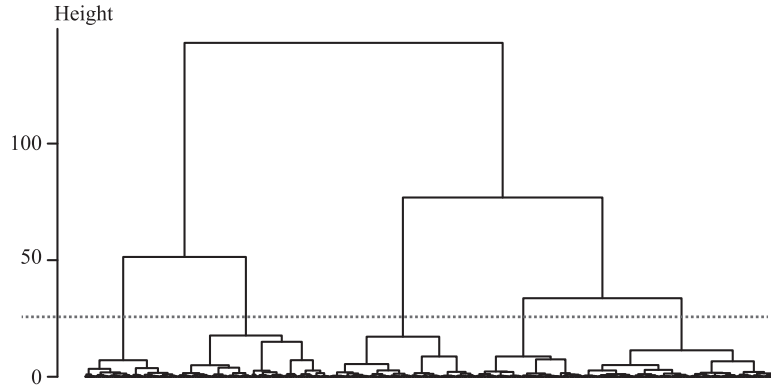


Figure 6. Dendrogram resulting from the classification of spatial units according to their multifractal spectrum.

same extent on the y -axis, we cannot use the usual distance: $d(\zeta_1, \zeta_2) = \sum_{\alpha} (f_1(\alpha) - f_2(\alpha))^2$ where ζ_1 and ζ_2 are two spatial units with identical α values.

After several unsuccessful trials, we chose to use the geometrical Hausdorff distance. The Hausdorff distance respects the axioms laid down for defining a distance mathematically, in particular the triangle inequality, which is convenient for a HAC. The Hausdorff distance relies on the notion of δ -neighborhood. A δ -neighborhood of a set A is the set of points within a distance δ of A (i.e., the points belonging to a buffer zone of width δ around the set A):

$$A_{\delta} = \{x \in \mathbb{R}^2 : |x - a| \leq \delta \text{ for some } a \in A\} \quad (22)$$

On this basis, the Hausdorff distance between two given multifractal spectra $f_1(\alpha)$ and $f_2(\alpha)$ is the least δ for which the δ -neighborhood (i.e., the buffer zone of width δ) of $f_1(\alpha)$ contains $f_2(\alpha)$ and vice versa (Falconer 2013).

Once all Hausdorff distances between the multifractal spectra of the 992 spatial units had been computed, we calculated the HAC using the Ward algorithm. The Ward algorithm minimizes intra-group variation and maximizes inter-group variation. Fig. 6 displays the dendrogram. Classes are defined by cutting off branches the dendrogram. The dendrogram suggests grouping the spatial units into five classes.

- Class 1: spatial units in which population is moderately concentrated locally
- Class 2: spatial units in which population is highly concentrated locally
- Class 3: spatial units in which population is dispersed (nonurban spatial units)
- Class 4: spatial units in which population density is uniformly high (urban spatial units)
- Class 5: spatial units exhibiting an intermediate spatial configuration

For each class, we calculated a representative multifractal spectrum by averaging for every q the Hölder coefficient and the Hausdorff dimension of all spatial units of the class (Fig. 7).

$$\alpha_q(\text{Class}) = \frac{\sum_{i \in \text{Class}} (\alpha_q(i))}{|\text{Class}|} \quad (23)$$

$$f(\alpha_q(\text{Class})) = \frac{\sum_{i \in \text{Class}} (f(\alpha_q(i)))}{|\text{Class}|} \quad (24)$$

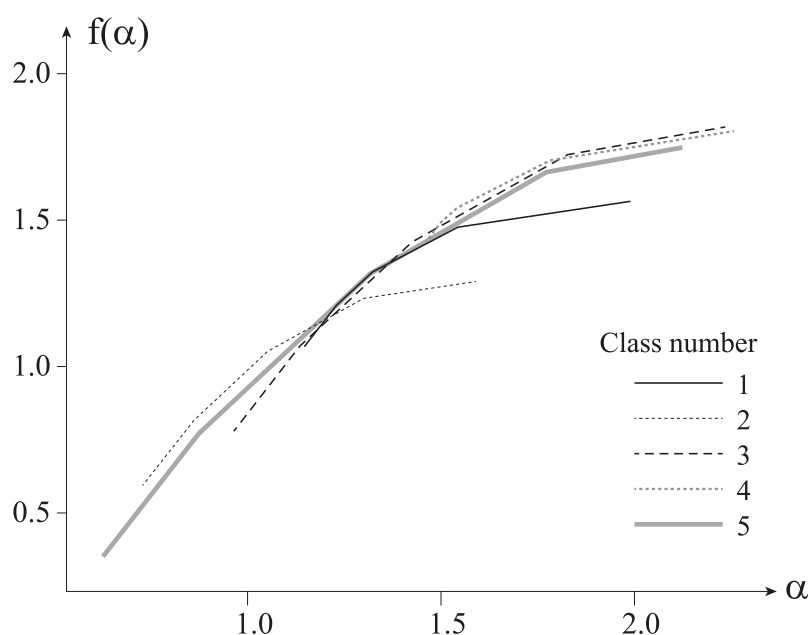


Figure 7. Average multifractal spectrum characterizing each class of the HAC.

Logically, the map of the multifractal spectra typology (Fig. 8) synthesizes the three maps of generalized dimensions displayed in Fig. 3. Spatial units characterized by low generalized dimensions D_0 , which appear clearly on the corresponding map, appear also clearly on the typology map (classes 1 and 2). Urban centers, which appear clearly on the map of generalized dimensions D_2 , also appear clearly on the typology map (class 4). Classes 3 and 5 of the typology map correspond to nonurban spatial units having high generalized dimensions D_0 and high or moderately high generalized dimensions D_1 and D_2 .

The average multifractal spectrum of spatial units in which population is locally concentrated pertains to low α values. This reflects the existence of numerous singularities in the spatial units concerned. In geographical terms, a multifractal spectrum limited to low α values characterizes clearly delineated villages or small towns surrounded by areas containing few (or even no) population. The average multifractal spectrum of class 2 is located at the links of the spectrum of class 1. This shows that population in spatial units belonging to class 2 is more concentrated across scales than population in spatial units of class 1. On the map displayed in Fig. 8, we see that class 2 spatial units are mainly located in the East of France as well as in mountain and forested areas. Class 1 spatial units surround class 2 spatial units. Both the geographical location of class 1 spatial units and their average multifractal spectrum suggest that class 1 is a mitigated version of class 2.

The average multifractal spectrum of spatial units in which population is uniformly distributed associates high α and $f(\alpha)$ values. For $q = 0$, the value of α_0 is approximately 2 and the value of $f(\alpha_0)$ is approximately 1.7. Such a multifractal spectrum characterizes a homogeneous and uniform spatial distribution of population in the spatial units considered. The average multifractal spectrum of class 4 is narrower than the spectrum of class 3. This indicates that the spatial distribution of population in class 4 is less multifractal than in class 3. Fig. 8 shows that the geographical location of spatial units belonging to class 4 corresponds to the location of the main urban centers. These exhibit high densities of population spreading uniformly across large portions of space. By contrast, the population density varies more according to the spatial resolution in class 3 spatial units. Spatial units belonging to class 3 are located on the periphery of large cities as well as in rural areas in the West of France.

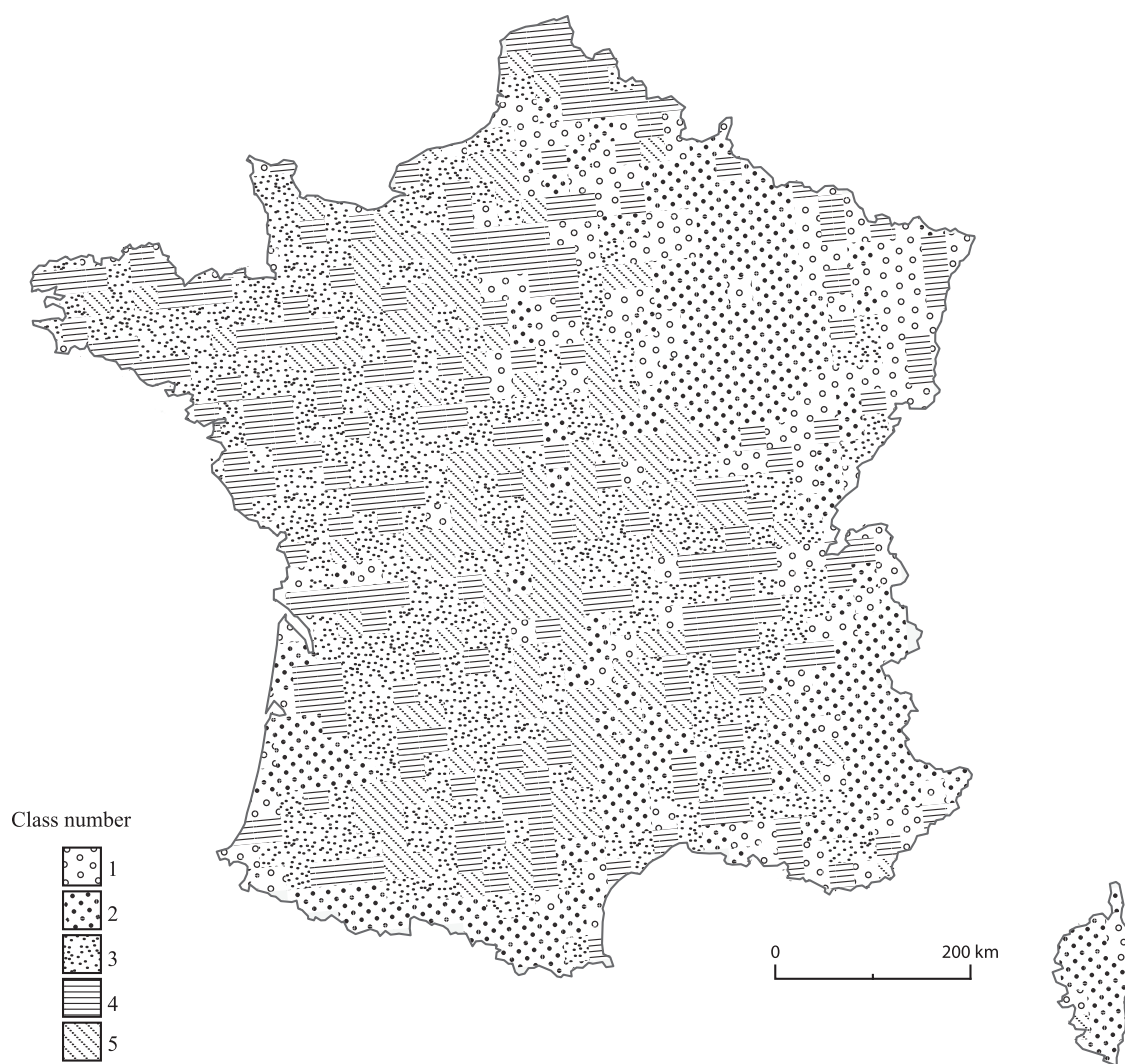


Figure 8. Map of the spatial units classified according to their multifractal spectrum.

The average multifractal spectrum of class 5 spatial units encompasses all the spectra described above. Class 5 combines the characteristics of classes 2 and 3. This intermediate position also characterizes the geographical location of the spatial units belonging to it: they are in general located between spatial units of class 2 and class 3 (Fig. 8). Class 5 corresponds to villages surrounded by areas of low population density in which population is uniformly dispersed. Most spatial units characterized by high generalized dimensions D_0 and low generalized dimensions D_2 belong to class 5.

Discussion

Traces of century-old agrarian systems

In the early twentieth century, renowned French geographers and historians analyzed the geography of rural areas (Demangeon 1927; Bloch 1931; Dion 1934). Their research led to a classical division of mainland France into two agrarian systems:

- enclosed field systems associated with scattered settlements in the center and the West of France;
- openfield systems associated with grouped settlements in the East of France.

Although actual research in history currently questions the relevance of the distinction between openfield systems and enclosed field systems (Lavigne 2003), this distinction interestingly enlightens the multifractal typology map of Fig. 8.

Classes 1 and 2 of the multifractal typology (locally concentrated population) are located in the East of France as well as in mountain and forested areas. In those areas the traditional agrarian systems were openfield systems.

In nonurban areas (i.e., outside the limit of urban areas), spatial units belonging to class 3 of the multifractal typology (dispersed population) are located in areas where the traditional agrarian systems were enclosed field systems (mainly in the West of France). Class 5, which corresponds to an intermediate configuration, is also mainly located in the West of France as well as in Burgundy.

Urban sprawl in France: a single process resulting in multiple settlement patterns

Ideally, the analysis of the effect of urban sprawl on settlement patterns should be based on the comparison of grid population data at several points in time. The first point in time should be just before the beginning of the urban sprawl process. Regrettably, there are no such time series of data for France and even if there were, it would be very difficult to deduce from them which changes in settlement patterns really stem from urban sprawl and which from other processes, in particular rural exodus. Indeed, urban sprawl and rural exodus were concomitant in France over a more or less long time period. Moreover, urban sprawl did not begin at the same point in time everywhere in France.

Urban sprawl is basically defined as the expansion of the cities beyond their initial limits. Yet, urban delineations used for creating statistical data change over the course of time to adjust to the outline of changing urban forms. This makes it difficult to use such statistical data to analyze urban sprawl. To overcome this difficulty, we did not introduce any a priori urban delineation. We also adopted an ergodic point of view by saying that observed differences in actual settlement patterns are due (at least partly) to local differences in the way the process of urban sprawl occurred. In adopting this point of view, the risk is that cause and effect might be inferred where mere contingencies exist. To avoid this, we did not posit a priori morphological criteria by which to define periurban settlement patterns.

With the multifractal analysis applied to the 992 French spatial units, the largest cities stand out clearly (except in Burgundy and Champagne): they all have high generalized dimensions. However, the analysis does not allow us to identify a type of settlement patterns specific to periurban areas: the multifractal index values of periurban areas around the largest cities differ widely.

Three different periurban configurations are observed for France's three largest metropolitan areas (Paris, Lyon, and Marseille). Paris's urban core is made up of several spatial units categorized as class 4 of the multifractal spectra typology. The periphery of Paris corresponds to spatial units belonging to class 3 (dispersed population) and class 1 (locally concentrated population). The metropolitan area of Lyon similarly contains several spatial units categorized as class 4 of the multifractal typology, which are mainly surrounded by class 3 spatial units (dispersed population). The urban area of Marseille exhibits a third configuration: in the neighborhood of class 4 spatial units, we find spatial units of classes 1 and 3 but also class 2 spatial units (concentrated population). The local concentration of population in those periurban areas can be explained for Marseille by strong relief constraints (hilly landscapes and Calanques) that preclude urban development.

In the Champagne-Ardennes region, multifractal analysis does not allow the detection of any urban area (no class 4 spatial unit). In this region, the metropolitan areas (Reims, Troyes, Charleville-Mézières, etc.) belong to classes 1 or 2 of the multifractal typology. In the West of

France, too, the smallest urban areas belong not to class 4 of the multifractal typology but to class 3 (dispersed population).

The existence of all these different types of urban and periurban settlement patterns suggests that the urban sprawl process does not affect all places in the same way, which has previously been well documented (Frankhauser 1994). Settlement patterns resulting from urban sprawl depend not only on the location of areas considered with respect to center cities but also on the preexisting types of agrarian systems (uniformly dispersed periurban settlement patterns located in the traditional areas of enclosed field systems; periurban settlement patterns locally concentrated in the traditional areas of openfield systems), and the natural shapes (mountains, forests) in those areas.

Conclusion

Whereas classical fractal dimensions (box-counting dimension, dilation dimension, correlation dimension, etc.) only characterize the spatial organization of the support of a measure (i.e., a container), generalized multifractal dimensions D_q also characterize the spatial variations of the measure (i.e., the content). Generalized dimensions D_q are constant for any q when the spatial distribution of population is monofractal. In this case, the spatial variation of the measure is the same as the spatial variation of its support: container and content are combined. This is the case for large French metropolitan areas whose generalized dimensions are high whatever the value of q . By contrast, some spatial units in the West of France exhibit a clear multifractal behavior: they have high generalized dimensions for $q = 0$ and low dimensions for $q = 2$.

Besides the generalized dimensions, the singularity spectrum $f(\alpha)$ enables geographers to quantify the MAUP. When the singularity spectrum is reduced to the point of coordinates (2, 2), the population is spatially uniformly distributed: population densities are almost constant locally; in this case, the effect of the MAUP is weak or even nonexistent. In the other cases, the distribution of population is subject to the MAUP. Two types of phenomena can occur:

1. Container shape: the smaller the maximum of $f(\alpha)$ (reached for $\alpha = \alpha_0$), the more irregular the support of the measure, the more numerous the empty cells in the δ -grid for small δ -values, and consequently the stronger the MAUP resulting from the shape of the container.
2. Content heterogeneity: two measures with the same maximum value of $f(\alpha)$ may have different intensities of multifractality. The multifractality is weak when the singularity spectrum concentrates around α_0 : in this case, population densities vary little among the cells of the δ -grid. Only the support of the measure (i.e., the container) creates the MAUP: the spatial distribution of population is almost monofractal. Contrarily, a large singularity spectrum indicates strong multifractality. In this case, the MAUP results both from the container shape and from the content heterogeneity within the container.

Acknowledgements

The authors would like to thank Armelle Couillet (research laboratory ThéMA, Besançon, France) for her support in drawing some of the figures. This research has been partly funded by the IXXI (*Institut rhônalpin des systèmes complexes*).

References

- Appleby, S. (1996). "Multifractal Characterization of the Distribution Pattern of the Human Population." *Geographical Analysis* 28(2), 147–60.
- Ariza-Villaverde, A. B., F. J. Jiménez-Hornero, and E. G. D. Ravé. (2013). "Multifractal Analysis of Axial Maps Applied to the Study of Urban Morphology." *Computers, Environment and Urban Systems* 38, 1–10.
- Batty, M., and P. A. Longley. (1994). *Fractal Cities: A Geometry of Form and Function*. London, UK: Academic Press.
- Batty, M., and Y. Xie. (1996). "Preliminary Evidence for a Theory of the Fractal City." *Environment and Planning A* 28, 1745–62.
- Batty, M., and Y. Xie. (1999). "Self-Organized Criticality and Urban Development." *Discrete Dynamics in Nature and Society* 3(2-3), 109–24.
- Benguigui, L., E. Blumenfeld-Lieberthal, and D. Czamanski. (2006). "The Dynamics of the Tel Aviv Morphology." *Environment and Planning B: Planning and Design* 33(2), 269.
- Benguigui, L., D. Czamanski, M. Marinov, and Y. Portugali. (2000). "When and Where is a City Fractal?" *Environment and Planning B: Planning and Design* 27(4), 507–19.
- Bloch, M. (1931). *Les caractères originaux de l'histoire rurale française*. Oslo: Instituttet for sammenlignende kulturforskning.
- Chen, Y. (2013). "Fractal Analytical Approach of Urban Form Based on Spatial Correlation Function." *Chaos, Solitons and Fractals* 49, 47–60.
- Chen, Y., and J. Wang. (2013). "Multifractal Characterization of Urban Form and Growth: The Case of Beijing." *Environment and Planning B: Planning and Design* 40(5), 884–904.
- Chen, Y., and Y. Zhou. (2004). "Multi-Fractal Measures of City-Size Distributions Based on the Three-Parameter Zipf Model." *Chaos, Solitons and Fractals* 22(4), 793–805.
- Chevillard, L., B. Castaing, A. Arneodo, E. Lévêque, J.-F. Pinton, and S. G. Roux. (2012). "A Phenomenological Theory of Eulerian and Lagrangian Velocity Fluctuations in Turbulent Flows." *Comptes Rendus Physique* 13(9), 899–928.
- Chhabra, A., and R. Jensen. (1989). "Direct Determination of the $f(\alpha)$ Singularity Spectrum." *Physical Review Letters* 62, 1327–30.
- Demangeon, A. (1927). "La géographie de l'habitat rural." *Annales de Géographie* 36, 1–23.
- Dion, R. (1934). *Essai sur la formation du paysage rural français*. Tours, France: Arrault.
- Falconer, K. (2013). *Fractal Geometry: Mathematical Foundations and Applications*. New York: Wiley.
- Feng, J., and Y. Chen. (2010). "Spatiotemporal Evolution of Urban Form and Land-Use Structure in Hangzhou, China: Evidence from Fractals." *Environment and Planning B: Planning and Design* 37(5), 838–56.
- Frankhauser, P. (1994). *La fractalité des structures urbaines*. Paris: Anthropos.
- Frankhauser, P. (1998). "The Fractal Approach. A New Tool for the Spatial Analysis of Urban Agglomerations." *Population* 10(1), 205–40.
- Frankhauser, P. (2008). "Fractal Geometry for Measuring and Modelling Urban Patterns." In *The Dynamics of Complex Urban Systems*, 213–43, edited by S. Albeverio, D. Andrey, P. Giordano, and A. Vancheri. Heidelberg: Springer Physica Verlag.
- Goodchild, M. (1980). "Fractals and the Accuracy of Geographical Measures." *Journal of the International Association for Mathematical Geology* 12(2), 85–98.
- Haag, G. (1994). "The Rank-Size Distribution of Settlements as a Dynamic Multifractal Phenomenon." *Chaos, Solitons and Fractals* 4(4), 519–34.
- Jaffard, S., B. Lashermes, and P. Abry. (2007). "Wavelet Leaders in Multifractal Analysis." In *Wavelet Analysis and Applications, Applied and Numerical Harmonic Analysis*, 201–46, edited by T. Qian, M. Vai and Y. Xu. Basel: Birkhäuser.
- Lam, N. (1990). "Description and Measurement of Landsat TM Images Using Fractals." *Photogrammetric Engineering and Remote Sensing* 56(2), 187–95.
- Lam, N. S.-N. and D. A. Quattrochi. (1992). "On the Issues of Scale, Resolution, and Fractal Analysis in the Mapping Sciences." *The Professional Geographer* 44(1), 88–98.

- Lavigne, C. (2003). "De nouveaux objets d'histoire agraire pour en finir avec le bocage et l'openfield." *Etudes rurales* 167–168, 133–86.
- Openshaw, S. (1983). "The Modifiable Areal Unit Problem." *Geo Books* 38.
- Orozco, C. D. V., J. Golay, and M. Kanevski. (2015). "Multifractal Portrayal of the Swiss Population." *Cybergeo: European Journal of Geography* (714). Available from url: <http://cybergeo.revues.org/26829>.
- Seuront, L. (2009). *Fractals and Multifractals in Ecology and Aquatic Science*. New York: CRC Press.
- Shen, G. (2002). "Fractal Dimension and Fractal Growth of Urbanized Areas." *International Journal of Geographical Information Science* 16(5), 419–37.
- Stanley, H., L. N. Amaral, A. Goldberger, S. Havlin, P. C. Ivanov, and C.-K. Peng. (1999). "Statistical Physics and Physiology: Monofractal and Multifractal Approaches." *Physica A: Statistical Mechanics and its Applications* 270(1), 309–24.
- Tannier, C., and D. Pumain. (2005). "Fractals in Urban Geography: A Theoretical Outline and An Empirical Example." *Cybergeo* (307). Available from url: <http://cybergeo.revues.org/3275>.
- Tannier, C., and I. Thomas. (2013). "Defining and Characterizing Urban Boundaries: A Fractal Analysis of Theoretical Cities and Belgian Cities." *Computers, Environment and Urban Systems* 41(0), 234–48.
- Tannier, C., G. Vuidel, H. Houot, and P. Frankhauser. (2012). "Spatial Accessibility to Amenities in Fractal and Non Fractal Urban Patterns." *Environment and Planning B: Planning and Design* 39(5), 801–19.
- Thomas, I., P. Frankhauser, and C. Biernacki. (2008). "The Morphology of Built-Up Landscapes in Wallonia (Belgium): A Classification Using Fractal Indices." *Landscape Urban Planning* 84(2), 99–115.
- Thomas, I., P. Frankhauser, and M.-L. De Keersmaecker. (2007). "Fractal Dimension Versus Density of Built-Up Surfaces in the Periphery Of Brussels." *Papers in Regional Science* 86(2), 287–308.
- Wendt, H., P. Abry, S. Jaffard, H. Ji, and Z. Shen. (2009). "Wavelet Leader Multifractal Analysis for Texture Classification." In *Image Processing (ICIP), 2009 16th IEEE International Conference on IEEE*, 3829–32.
- Wendt, H., P. Abry, Roux, G. Stéphane, and S. Jaffard. (2007). "Analyse multifractale d'image: l'apport des coefficients dominants." In 'GRETSI 2007', Université de technologie de Troyes (UTT); Université de Reims Champagne-Ardennes (URCA), GRETSI, Troyes, France.
- White, R. and G. Engelen. (1994). "Urban Systems Dynamics and Cellular Automata: Fractal Structures Between Order and Chaos." *Chaos, Solitons Fractals* 4(4), 563–83.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Annexe B

Applying two fractal methods to characterise the local and global deviations from scale invariance of built patterns throughout mainland France



Applying two fractal methods to characterise the local and global deviations from scale invariance of built patterns throughout mainland France

François Sémécurbe¹ · Cécile Tannier¹  · Stéphane G. Roux²

Received: 18 August 2017 / Accepted: 30 November 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

In the early twentieth century, a handful of French geographers and historians famously suggested that mainland France comprised two agrarian systems: enclosed field systems with scattered settlements in the central and western France and open-field systems with grouped settlements in eastern France. This division between grouped and scattered settlements can still be found on the outskirts of urban areas. The objective of this paper is to determine whether the shape of urban areas varies with the type of built patterns in their periphery. To this end, we identify and characterise the local and global deviations from scale invariance of built patterns in mainland France. For this, we propose a new method—Geographically Weighted Fractal Analysis—that can characterise built patterns at a fine spatial resolution without making any a priori distinction between urban patterns and suburban or rural patterns. By applying GWFA to the spatial distribution of buildings throughout mainland France we identify six geographically consistent types of built patterns that are distinctive in the way buildings are either concentrated or dispersed across scales. The relationship between the local built textures and the global shape of twenty metropolitan areas is then analysed statistically. It is found that the proportion of dispersed (or concentrated) outer suburban built patterns in metropolitan areas is closely related to the distance threshold that marks the morphological limit of their urban areas.

Keywords Built textures · Fractal analysis · Suburban fringes · Scale invariance · Mainland France

JEL Classification C63 · C80 · Y80

✉ Cécile Tannier
cecile.tannier@univ-fcomte.fr

Extended author information available on the last page of the article

1 Introduction

If we are to identify the advantages and disadvantages of different city shapes for various planning goals (e. g. preserving ecological connectivity, improving access to urban and rural amenities, ensuring good ventilation of the city centre, etc.) then—among other things—the associated urban built patterns need to be more accurately described and characterised. This need has engendered a wealth of publications describing and characterising city shapes, including numerous methods of identifying different types of urban patterns along with many spatial indexes for measuring urban sprawl. In this paper, we seek to contribute to this field of research by exploring the multiscale morphological properties of urban built patterns in more depth.

In the early twentieth century, a number of renowned geographers and historians analysed the geography of rural France (Demangeon 1927; Bloch 1931; Dion 1934). Their research led to a classical division of mainland France into two agrarian systems: enclosed field systems associated with scattered settlements in the centre and the West of France and openfield systems associated with grouped settlements in the East of France. This distinction between grouped and scattered settlements can still be found on the periphery of urban areas. Obviously, the spatial expansion of cities does not take place within a blank, isotropic space, especially when cities belong to centuries-old settlement systems. As outer suburban areas are the space into which cities expand, it is to be expected that pre-existing suburban settlements will influence the shape of the urban development they come to accommodate. Regarding this, the objective of this paper is to determine whether the shape of French urban areas is determined by the shape of built-up areas on their outskirts.

Analysis of the relationship between built-up shapes surrounding urban areas and the shape of the urban areas themselves first involves characterising the built patterns at a fine spatial resolution, without making any a priori distinction between urban patterns and suburban or rural patterns. Yet, built patterns in suburban fringes are somewhat hybrid: they blend inherited features (traditional rural buildings, old village cores) with new urban developments (most often, detached housing estates). This makes them difficult to describe and characterise (Chaudhry and Mackaness 2008). A number of publications have shown the value of fractal dimensions for characterising irregular shapes of this kind. Following pioneering research supporting the assumption that cities are fractal (Batty and Longley 1986, 1987; Frankhauser 1988; Mandelbrot 1982; White and Engelen 1993), a string of publications has explored the fractal aspects of built forms (e.g. among the earliest ones, Longley and Batty 1989; Batty and Xie 1996; Frankhauser 1998; Benguigui et al. 2000; Shen 2002; Benguigui and Czamanski 2004). Fractal analysis can be used to compare the built patterns of entire urban regions (e.g. Feng and Chen 2010; Chen and Wang 2013; Tannier and Thomas 2013) or smaller spatial units (neighbourhoods, communities) (e.g. De Keersmaecker et al. 2003; Thomas et al. 2008). Fractal dimensions describe the way buildings are spatially distributed at several nested spatial resolutions. The closer the fractal dimension is to 2, the more uniform is the spatial distribution of buildings. In contrast, a fractal dimension close to 0 describes a highly concentrated spatial distribution of buildings at just a few locations across

all scales. More often than not, buildings exhibit concentrated spatial configurations at some spatial resolutions and more dispersed patterns at other resolutions (fractal dimension in-between 1 and 2). In any event, the fractal dimension is not expected to be constant in reality (Goodchild 1980); most often, it is constant over a limited range of scales but varies somewhat over successive ranges of scales (Lam 1990; Lam and Quattrochi 1992; White and Engelen 1994; Tannier and Pumain 2005).

In this paper, we propose a new method, Geographically Weighted Fractal Analysis (GWFA). As with Geographically Weighted Regression (Brunsdon et al. 1996; Fotheringham et al. 2003), GWFA uses a mathematical kernel that describes the way the neighbourhood of a point is taken into account to estimate the fractal dimension of that point. By applying GWFA to French built patterns, we analyse the local deviations from scale invariance of the spatial distribution of buildings throughout mainland France. Then the classification of analysis results enables us to map the built textures of France at a spatial resolution of 2000 m. Next, we compare the built textures thus mapped with the global shape of twenty metropolitan areas. For this, we choose to apply the method proposed by Tannier et al. (2011) to detect discontinuities in space across scales. This method can determine a distance threshold specific to each area under study. The statistical comparison of this distance threshold with the local built textures resulting from the classification indicates whether or not the global shape of French metropolitan areas varies with the built shapes in their periphery.

2 Data and method

Source data are from the vector data to the nearest 1 m provided by the French cartographic service, BD Topo[®] IGN 2011. This represents about 24 million buildings mapped in 2D. It would have been possible to use another data source providing similar kind of data, especially Open Street Map. Yet, building data from BD Topo[®] IGN are easier to gather and to use than OSM data when analysing the whole mainland France. In general, OSM data are more accurate than data from BD Topo[®] IGN in dense urban areas. Conversely, data from BD Topo[®] IGN are more accurate than OSM data in rural areas visited by a small number of individuals. For the analysis, each building is represented by its centroid in order to reduce both the size of the data base and the computation time.

In order to analyse this very detailed data set, a four-step methodology has been defined. It takes into account three guiding principles: (1) introducing no a priori distinction between urban built patterns and suburban or rural patterns; (2) focusing the analyses on the deviations from scale invariance; (3) examining and confronting both global and local deviations from scale invariance.

Step 1. Determining if the spatial distribution of buildings in mainland France is locally scale invariant or not

Analysis is based on a 2000-m spaced grid of 145,178 estimation points i that covers the whole country. On this basis, GWFA is used to estimate the fractal dimension at each point i of the study area (see Sect. 3). Fractal dimensions

describe the way buildings are spatially distributed at several nested spatial resolutions. The underlying assumption is that the spatial distribution of buildings is scale invariant, which means that the variation in the built density from one given spatial resolution to another is constant across scales. The scale invariance is proven when the goodness-of-fit of the estimation of the fractal dimension is satisfactory. This goodness-of-fit is usually very high when estimating fractal dimensions. Therefore, R^2 values between 0.95 and 0.995 do not guarantee scale invariance for the whole scale range, and R^2 values less than 0.95 are sure signs of a disruption in scale invariance.

Step 2. Identifying different types of built patterns according to the way the spatial distribution of buildings deviates locally from a scale invariant distribution.

Here the spatial distribution of buildings at each point i is not supposed to be scale invariant and each point i is characterised by a series of scaling indexes Sc calculated for successive scale intervals. These Sc indexes correspond to the $\alpha(\varepsilon)$ indexes used by Thomas et al. (2010) to characterise the scaling behaviour of built patterns of European urban wards.

In order to analyse the profile of Sc indexes for each point i , we apply a Principal Component Analysis. Then we classify the 145,178 points i according to their two coordinates on the PCA axes using the k-means method.

Step 3 For a selection of twenty French metropolitan areas, characterising the global deviation of the spatial distribution of buildings from a scale-invariant distribution. For this, calculation of a distance threshold that indicates a crucial discontinuity in space across scales.

For this, we use the free software application named MorphoLim (<https://sourcesup.renater.fr/morpholim/>) that enables us to apply the method proposed by Tannier et al. (2011). By this method a dilation of each individual building is first applied, then a distance threshold is detected on the dilation curve. This distance threshold corresponds to the dilation step at which the curve deviates the most from a straight-line. At this point on the dilation curve, distances separating buildings no longer exhibit the same fractal behaviour. Thus, this point indicates the maximum morphological difference between the inlying built patterns (i.e. the morphological agglomeration) and the outlying built patterns (i.e. the surrounding suburban areas). The distance threshold is specific to each area under study: the less an urban area differs morphologically from its surrounding built landscape (in terms of variations in distances separating buildings), the greater the distance threshold; the fewer the buildings in a study area that are connected across scales, the greater the distance threshold.

Step 4 For the twenty selected French metropolitan areas, making a statistical comparison of the distance threshold that characterises globally the spatial distribution of buildings across scales with the dominant types of local built textures identified at step 2.

3 Geographically Weighted Fractal Analysis (GWFA): detailed description

3.1 Estimation of the fractal dimension at each point i of the study area

The fractal dimension D of an object is defined through a scaling function $F\varepsilon$ that describes an aspect of the spatial distribution of the points e of the object.

$$F\varepsilon \sim \varepsilon^{-D} \text{ where } \varepsilon \text{ is the size of the counting window} \quad (1)$$

For theoretical mathematical objects, D corresponds to the asymptotic behaviour of $F\varepsilon$ when ε tends to 0. For real world objects, D is estimated for the range of scales ε for which the points of the curve of $\log(F\varepsilon)$ with respect to $-\log(\varepsilon)$ form a straight line. D is the slope of this straight line.

Each scaling function $F\varepsilon$ enables the calculation of a given fractal dimension. The chosen scaling function for Geographically Weighted Fractal Analysis (GWFA) is M_r^q , where r corresponds to ε and M corresponds to F .

$$M_r^q \sim r^{Dq} \quad (2)$$

For a given point e , $M_r(e)$ is the number of points within a distance r around e . M_r^q is obtained by aggregating the $M_r(e)$ of all points e . When a simple average is used for the aggregation, the fractal dimension obtained is the correlation dimension (Grassberger and Procaccia 1983). In contrast, when a mathematical norm is used for the aggregation (Eq. 3), the fractal dimension obtained is a multifractal generalized dimension D_q . The corresponding estimation method is the sandbox method (Vicsek 1990).

$$M_r^q = \left[\frac{1}{n} \sum_e M_r(e)^{q-1} \right]^{\frac{1}{q-1}} \sim r^{D_q} \quad (3)$$

The parameter q controls the nature of the aggregation: more importance is given either to high values of $M_r(e)$, when q is high, or to low values of $M_r(e)$, when q is low. In GWFA, q is set to 0 meaning that high or low values of $M_r(e)$ all have the same importance. The corresponding scaling function is noted $M_r(i)$ (Eq. 4) as it is calculated for each estimation point i . The resulting generalized dimension D_0 is closely related to, but more reliable than, the well-known box-counting dimension (Tél et al. 1989; Vicsek 1990). D_0 is obtained by estimating the slope of the curve of $\log(M_r(i))$ with respect to $\log(r)$ for a given set of distance radii r by OLS (ordinary least squares) regression.

$$M_r(i) = \left[\sum_e (W_{i,e} M_r(e))^{-1} \right]^{-1} \sim r^{D_0} \quad (4)$$

For counting, a circular neighbourhood of diameter L is defined around each estimation point i . Only points e located within this neighbourhood L are taken into account but points counted around each e may be located outside the neighbourhood L (Fig. 1). In this way, boundary effects related to points e located close to the boundary of the neighbourhood L are avoided.

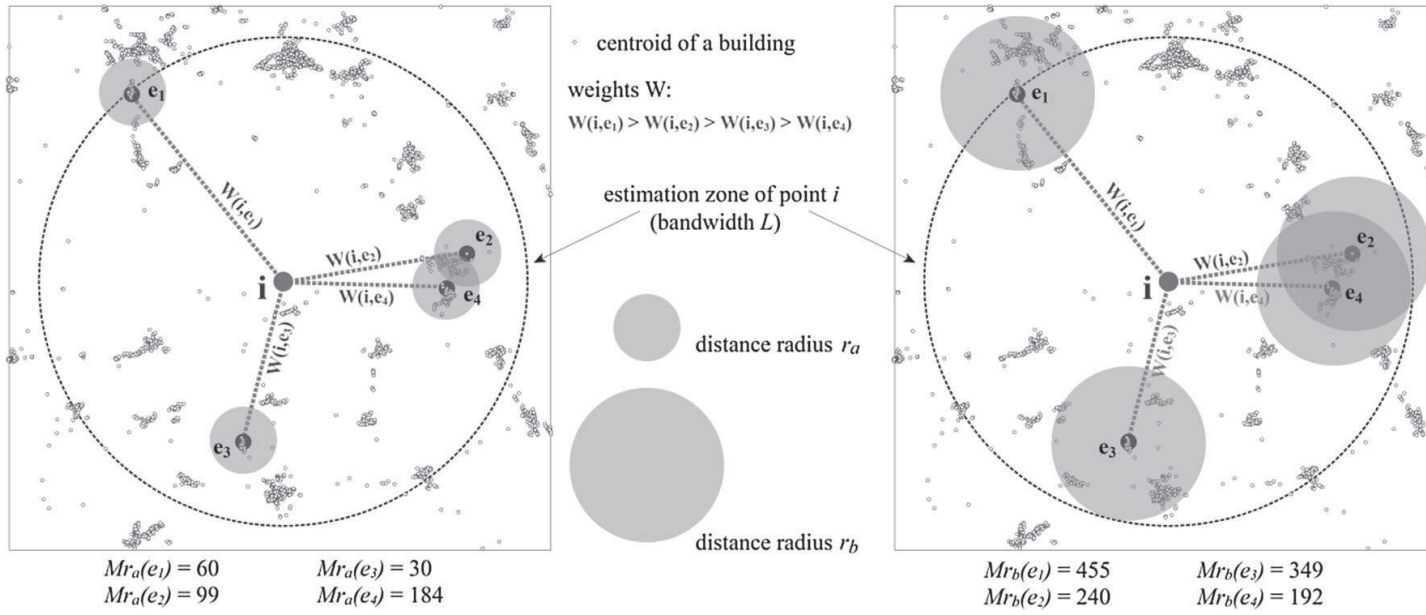


Fig. 1 Counting variables

Weights $W_{i,e}$ are attributed to the $M_r(e)$ values. Each weight corresponds to the distance of the point e under consideration to the estimation point i (Fig. 1). Thus GWFA introduces distance-based weights assigned to each pair of points (i,e) in the calculation of the scaling function $M_r(i) \cdot M_r(e)$ values for points e close to i have a higher weight in the estimation of $M_r(i)$ than $M_r(e)$ values for points e that lie far from i . Weights $W_{i,e}$ are calculated by applying a mathematical positive kernel K :

$$W_{i,e} = K(d_{ie}/L)$$

where d_{ie} is the distance between the estimation point i and the observation e ; L is the bandwidth of the analysis.

For GWFA, the quartic (bi-weight) kernel is used.

$$K(x) = (1 - x^2)^2, \text{ if } x < 1, \text{ and } 0, \text{ otherwise}$$

The quartic kernel resembles the Gaussian kernel but has a compact support and enables faster computation. In any event, a positive, continuous and decreasing kernel that tends asymptotically to 0 ensures regular and continuous estimations in space, which enables results to be mapped satisfactorily (Brunsdon et al. 1996).

GWFA can be computed using Python and/or R. The computing code written to calculate D_0 for each estimation point i is provided in Appendices A and B. The computing time required for all buildings of mainland France (i.e. about 24 million building centroids) is about five minutes with a calculating server having a low computing power. Alternatively, when the data set is quite small (less than about one million building centroids), an R package entitled GWFA can be used (<https://CRAN.R-project.org/package=gwfa>).

3.2 Setting values of parameters

The distance between any two estimation points i is set at 2000 m. This distance is small enough to analyse local variations in fractal behaviour. At the same time, it is large enough to obtain a reliable estimation of fractal dimensions for scales (i.e. distance radii r) ranging from 0 to 800 m (see below).

The size of the neighbourhood L determines the fact that the analysis is local (when L is small) or global (when L is large). In any cases, the neighbourhood L around each estimation point i has to be larger than the distance between two neighbouring points i in order to enable the analysis of built textures in an almost continuous way. Accordingly, L is set at 8000 m. Thus it includes two neighbouring estimation points i in each direction around a given point i . On average, each zone of radius L around each estimation point i contains 8600 centroids of buildings.

Five distances r are considered: 50, 100, 200, 400, and 800 m. As the target estimated curve is a straight line, five points are enough to perform a statistical estimation of the fractal dimension that is reliable. Thus introducing additional distances r between 50 and 800 m. is useless. Of course, standard errors of the estimates may be rather large but if they are, this shows that the spatial distribution of buildings is locally not scale invariant. The 1:10 ratio between $r=800$ and $L=8000$ ensures that the statistical estimations of fractal dimensions are robust. In addition, 100 circles of radius $r=800$ m can entirely

cover each estimation zone of radius L around i . Introducing additional distances r larger than 800 m. would not be relevant because this would create boundary effects. Introducing additional distances r lower than 50 m. would also not be relevant because each point e corresponds to the centroid of a building but not to its real spatial footprint.

3.3 Sensitivity of estimated fractal dimensions to the chosen values of parameters

First, we have studied the results of GWFA when changing the size L of the neighbourhood (see Fig. 2a–c). It appears that the spatial structures revealed when

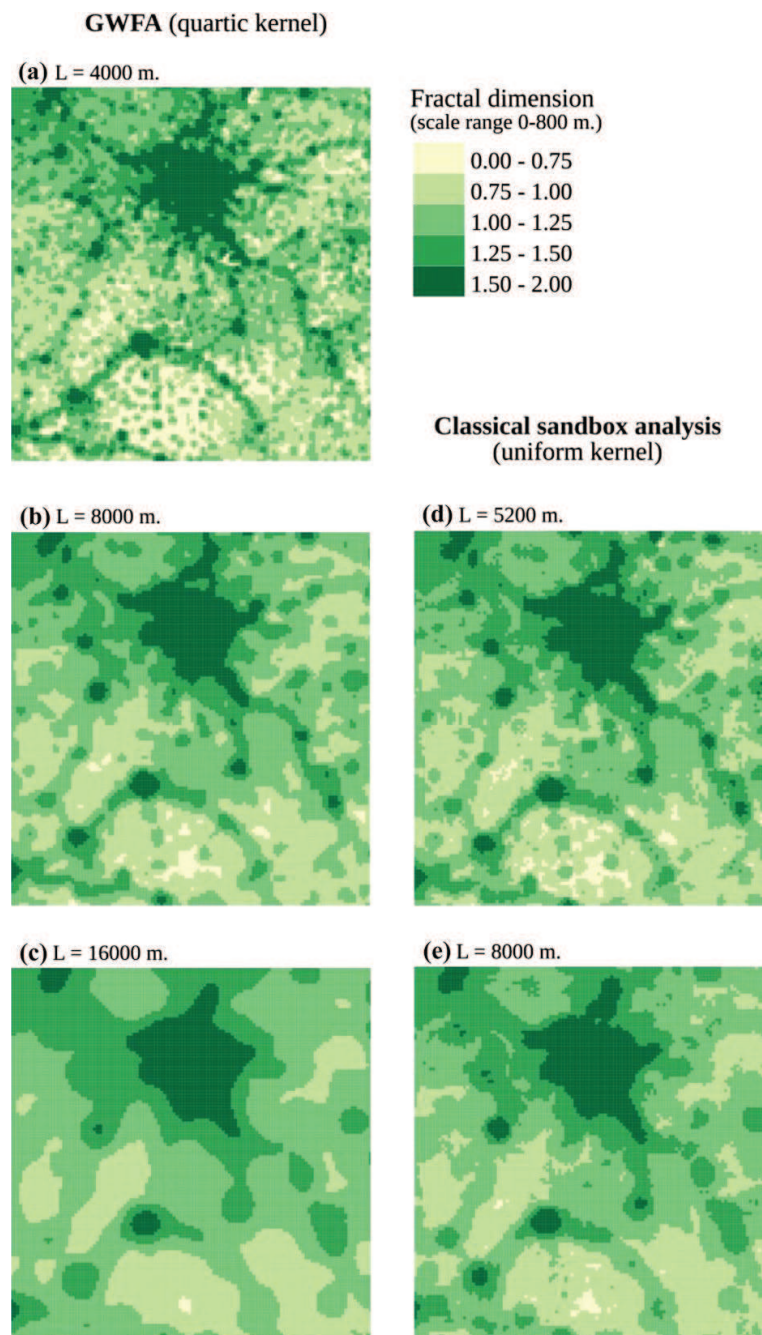


Fig. 2 Comparison of local fractal dimensions obtained with different kernels and different bandwidths L . Paris's metropolitan area and its southern periphery

$L=4000$ m. exhibit too many unnecessary details. Contrarily, the spatial structures revealed when $L=16000$ m. are too smoothed and not enough informative. In-between, the neighbourhood size $L=8000$ m. represents an interesting compromise.

Second, we have compared the results obtained with GWFA and the results obtained with the classical sandbox fractal analysis method, which does not introduce geographically weightings. The classical sandbox method is equivalent to assigning to each point e a weight $W_{i,e}$ equal to 1 when e lies within the bandwidth L around i and a weight $W_{i,e}$ equal to 0 when e lies outside L . In other words, a uniform (boxcar) kernel is used for the classical sandbox method whereas a smoothing (quartic) kernel is used for GWFA. In order to ensure the comparability of analysis results obtained with the two methods, we have followed the suggestion of Diggle (1985) that is to calibrate the two kernels by equating their variance. By applying this equivalence rule, a bandwidth $L=8000$ m. in case of a quartic kernel corresponds to a bandwidth $L=5200$ m. in case of a uniform kernel. Diggle (1985), quoting (Silverman 1978), asserted that the successful implementation of smoothing data with a kernel function rests on choosing an appropriate value of the bandwidth L , whereas the choice of the kernel function is of very secondary importance. The fact is that the two maps in Fig. 2b (GWFA with $L=8000$ m.) and Fig. 2d (classical sandbox analysis with $L=5200$ m.) look a lot alike. Yet, the spatial structures are more clearly delineated in Fig. 2b, which shows the interest of applying a smoothing quartic kernel. Additionally, spatial structures displayed in Fig. 2e seem to us too smooth (not enough detailed) regarding both Fig. 2b, d.

4 Results

4.1 Estimated fractal dimensions show local deviations from scale invariance, especially in the West and South-West of France

By applying GWFA to the building centroids for the whole of mainland France, local fractal dimensions were computed for the five distance radii r (50, 100, 200, 400, and 800 m) (Fig. 3a). Figure 3b shows that the goodness-of-fit of the estimation is higher than 0.995 only in eastern France and in the largest urban areas (see Fig. 4 for their location). In contrast, most places in central and western France exhibit R^2 values of less than 0.995. Thus, the spatial distribution of buildings there is probably not scale invariant.

Considering this, we estimated two other fractal dimensions, one for distance radii r 50, 100, and 200 m and another for distance radii r 200, 400, and 800 m. As expected, these two fractal dimensions are clearly different for places located outside the main urban areas in the West and South-West of France: in those places, estimated fractal dimensions are low for r 50, 100, and 200 m, but high for r 200, 400, and 800 m (Fig. 5a, b).

These results suggest that estimating one fractal dimension for the entire scale range (Fig. 3) or two fractal dimensions for two smaller scale ranges (Fig. 5) does not enable us to characterise the spatial distribution of buildings in each zone around i precisely enough.

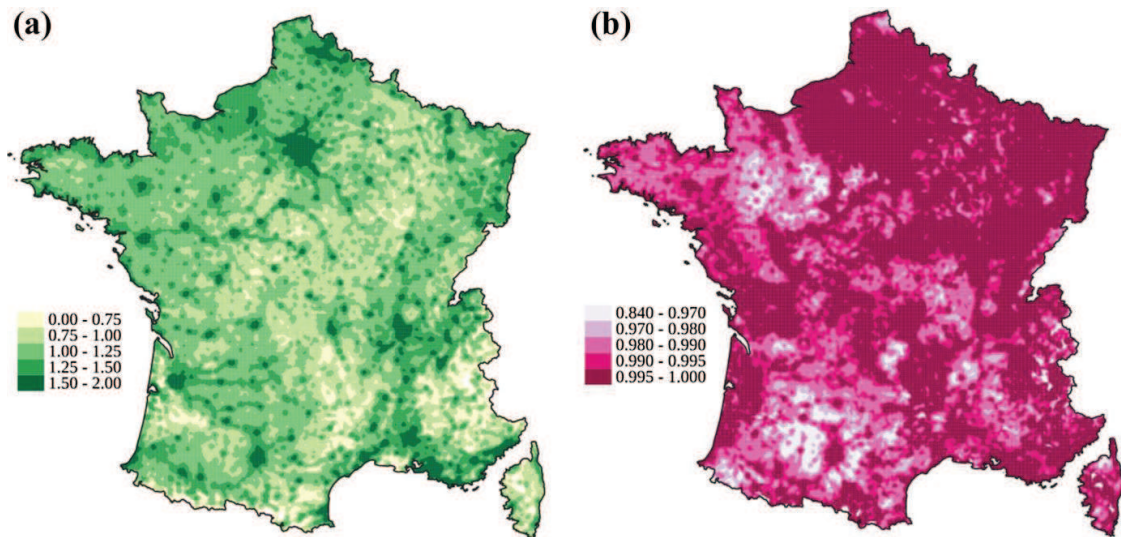


Fig. 3 **a** Fractal dimensions estimated with GWFA for the scale range 0 to 800 m. **b** Goodness of fit of the estimation (R^2). 145,178 estimation points i . On the maps, each estimation point corresponds to a 2000 m-wide square spatial unit

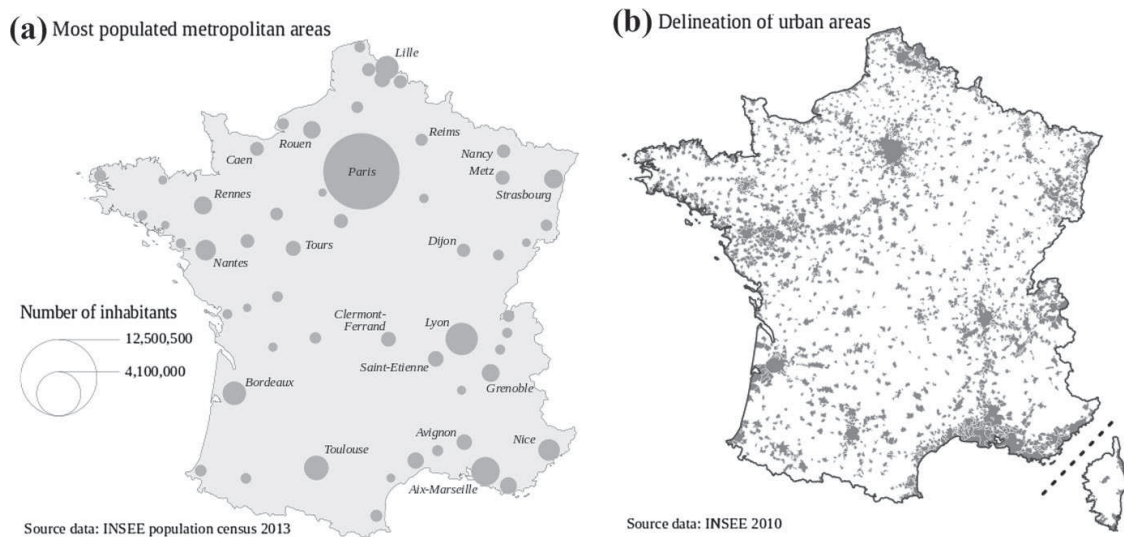


Fig. 4 Metropolitan areas and urban areas in mainland France. Note: in France, an urban area (*unité urbaine*) is an isolated *commune* (i.e. the basic administrative unit used for dividing up France) or a group of neighbouring *communes* having more than 2000 inhabitants and containing a contiguous built-up zone that gathers more than half of the population of each *commune*. A built-up zone is made up of adjacent buildings forming a whole such that no building is more than 200 metres apart from the next nearest building (Le Gléau et al. 1997)

4.2 Typology of local built textures built on scaling indexes S_c

S_c indexes are defined as the differences in the logarithms of two scaling functions calculated for two successive scale ranges, e.g.

$$S_{c_{100-50}} = \log(M_{100}(i)) - \log(M_{50}(i)) \quad (5)$$

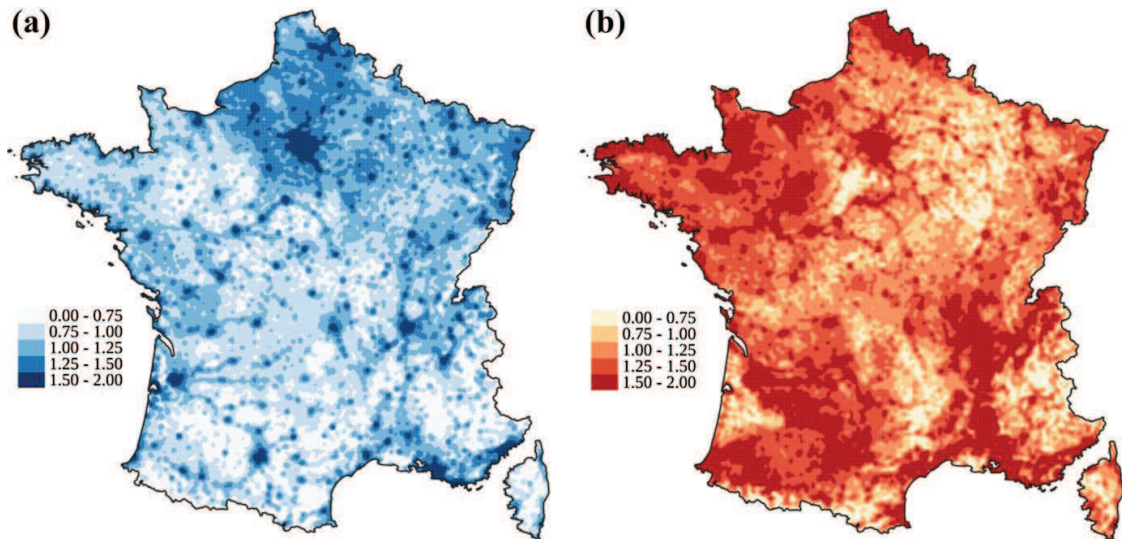


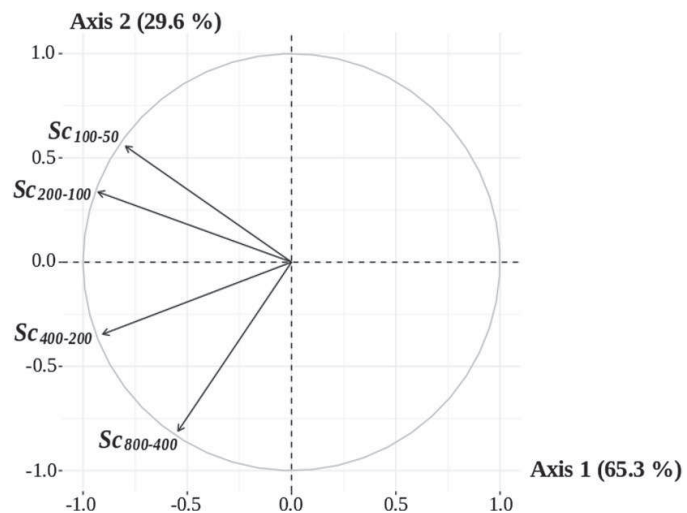
Fig. 5 Fractal dimensions estimated for the distance radii **a** 50, 100, and 200 m and **b** 200, 400, and 800 m

With five distance radii r , it is possible to calculate four scaling indexes Sc . In order to analyse them simultaneously, we applied a Principal Component Analysis. The two first components of the PCA concentrate 90% of the total inertia (Fig. 6).

On the circle of correlations, we observe that this first principal component is strongly correlated with the four Sc indexes, all oriented in the same direction with respect to the first axis. Thus the first component contrasts areas characterised by a low fractal dimension and a low built density (*i. e.* rural areas) with areas characterised by a high fractal dimension and a high built density (*i.e.* urban areas) (Fig. 7a).

The second component of the PCA reflects the fact that a built pattern may or may not exhibit a scale-invariant distribution across all scale ranges, from 0 to 800 m. It orders the scaling indexes in accordance with their distance radii: on the circle of correlations, Sc_{100-50} and $Sc_{200-100}$ lie very close to each other and far from $Sc_{800-400}$; $Sc_{400-200}$ lies in an intermediate position. In Fig. 7b, we see that the coordinate of urban areas is close to 0 for this second component. Indeed, their spatial organization is scale invariant across the whole scale range (from 50 to 800 m) and, consequently, the four indexes Sc have almost the same values. Contrarily, the values of the scaling

Fig. 6 Principal Component Analysis: circle of correlations. Four variables: Sc_{100-50} , $Sc_{200-100}$, $Sc_{400-200}$, and $Sc_{800-400}$; 145,178 individuals (2000 m-wide square spatial units)



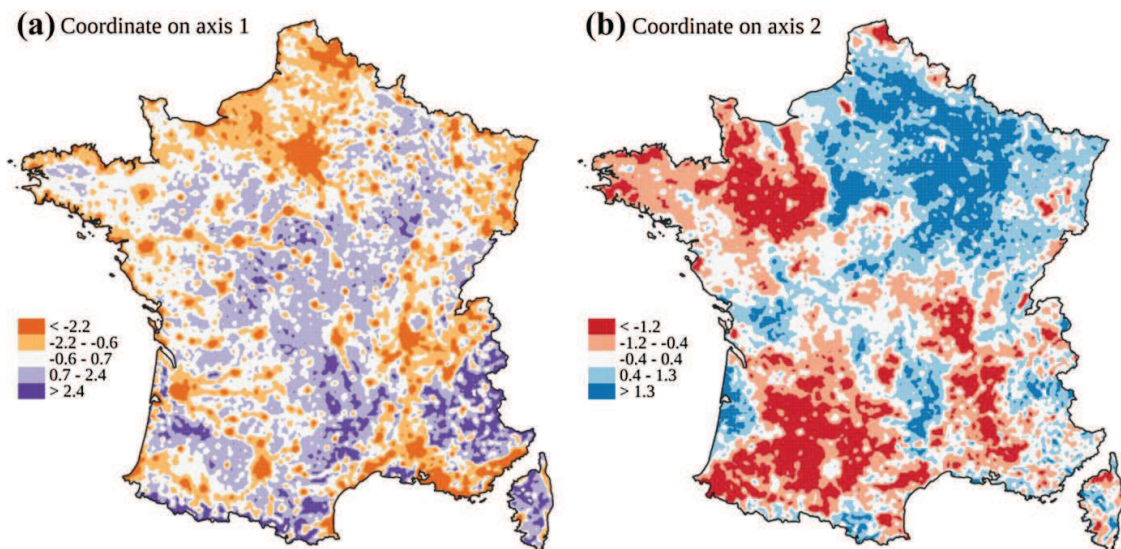


Fig. 7 Coordinates of each 2000-m-wide zone i on the two first axes of the PCA

indexes vary with the distance radii for peripheral areas in the West of France, which appear in red in Fig. 7b: obviously, their fractal dimension is low in Fig. 5a and high in Fig. 5b. This stark contrast is explained by the dispersion of clusters of buildings, namely small villages and hamlets: for large distance radii (0–400 and 0–800 m), all points e have numerous neighbours and the index $Sc_{800-400}$ is high; conversely, for small distance radii, many points e have almost no neighbours and the indexes Sc_{100-50} and $Sc_{200-100}$ are low. Built patterns in blue in Fig. 7b exhibit a more complex scaling behaviour because they are locally more concentrated. For distance radii 0–50 m and 0–100 m, each point e has numerous neighbours located within the same built cluster and the index Sc_{100-50} is high. Then, for distance radii 0–200 m and 0–400 m, there are relatively fewer neighbours of each point e ; they belong either to the built cluster of point e or to other built clusters (neighbouring villages) but the probability of finding other villages in the vicinity of the village of point e is low for such distances. The value of indexes $Sc_{200-100}$ and $Sc_{400-200}$ are accordingly lower. Finally, for large distance radii (0–400 and 0–800), the probability of finding neighbouring villages is much higher and the value of index $Sc_{800-400}$ is also higher.

Then, we classified the 145,178 spatial units according to their two coordinates on the PCA axes using the k -means method. The scree plot in Fig. 8 shows the increase in homogeneity within each class with respect to the increase in the number of classes. The point at which the slope of the curve levels off indicates the best compromise between the number of classes (that must be the smallest as possible) and the homogeneity within each class (that must be the highest as possible, i.e. the sum of the squares of the distances from each point to the centre of its class has to be the smallest as possible). Here, it seems to be relevant to take on six classes. Classes 1 and 4, which occupy contrasting positions on the first axis of the PCA, are separated in the classification from the other classes (see the PCA scatter plot in Fig. 10). Classes 2 (a and b) and classes 3 (a and b) are also separated since they occupy contrasting positions on the first axis of the PCA. Simultaneously, classes a (2 and 3) and classes b (2 and 3) are separated since they occupy contrasting positions on the second axis of the PCA.

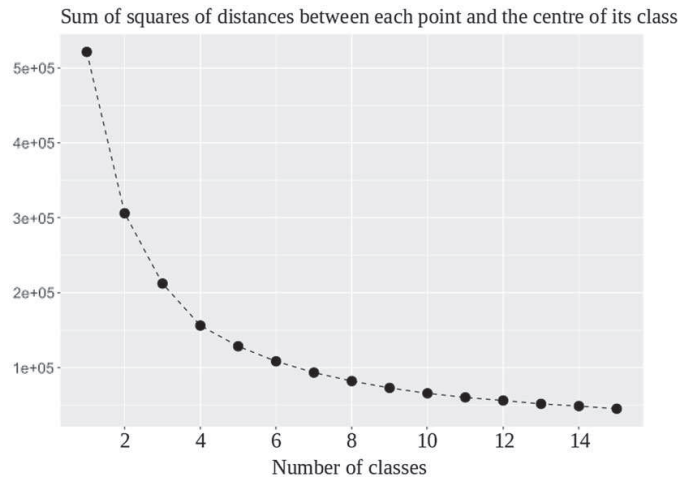
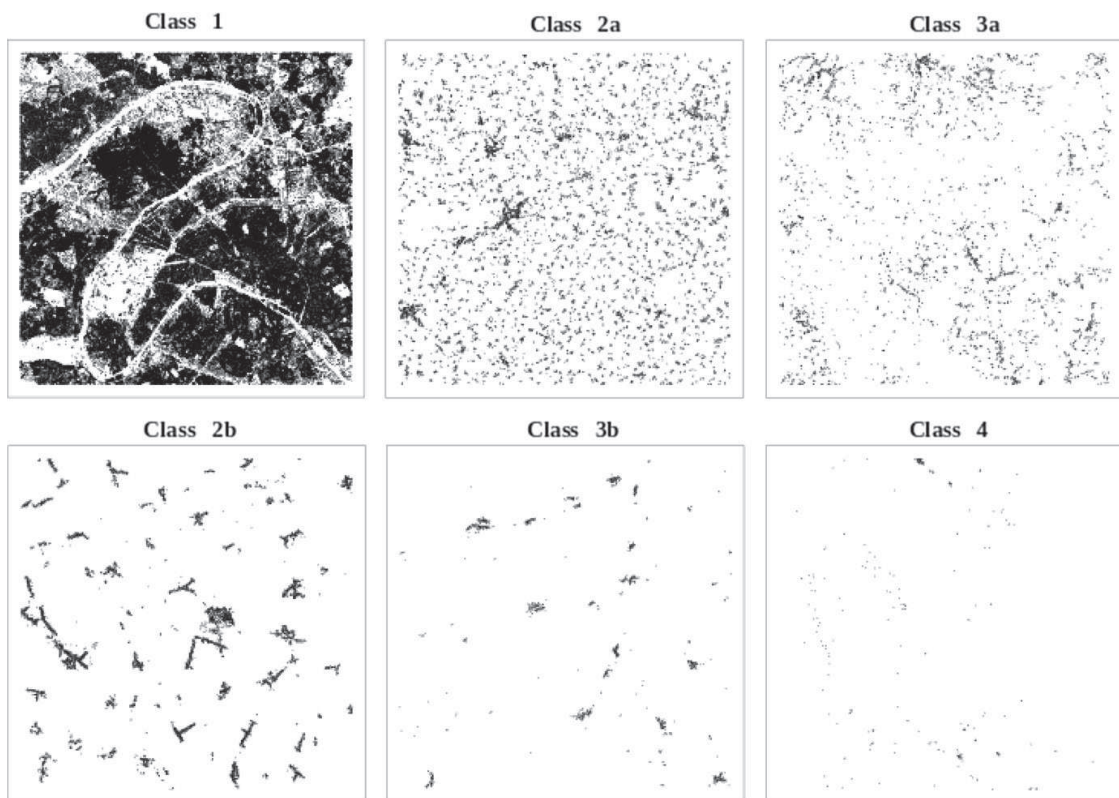
Fig. 8 Scree plot of the k -means clustering

Figure 9 shows that classes 2a and 3a correspond to locally dispersed built patterns. They are more scattered and diffuse in class 3a than in class 2a. Classes 2b and 3b correspond to locally concentrated built patterns; villages in class 2b are larger and more numerous than in class 3b. Class 1 corresponds to urban built patterns. Class 4 corresponds to sparse built patterns.

The resulting typology map is well structured although no contiguity criterion has been introduced into the classification (Fig. 10). Clusters of spatial units belonging to a same class are very consistent: class 1 spatial units are urban areas; class 4 spatial units are sparsely populated areas; class 2a, 2b, 3a, and 3b spatial units are located in the periphery of urban areas.

**Fig. 9** Built pattern typical of each class

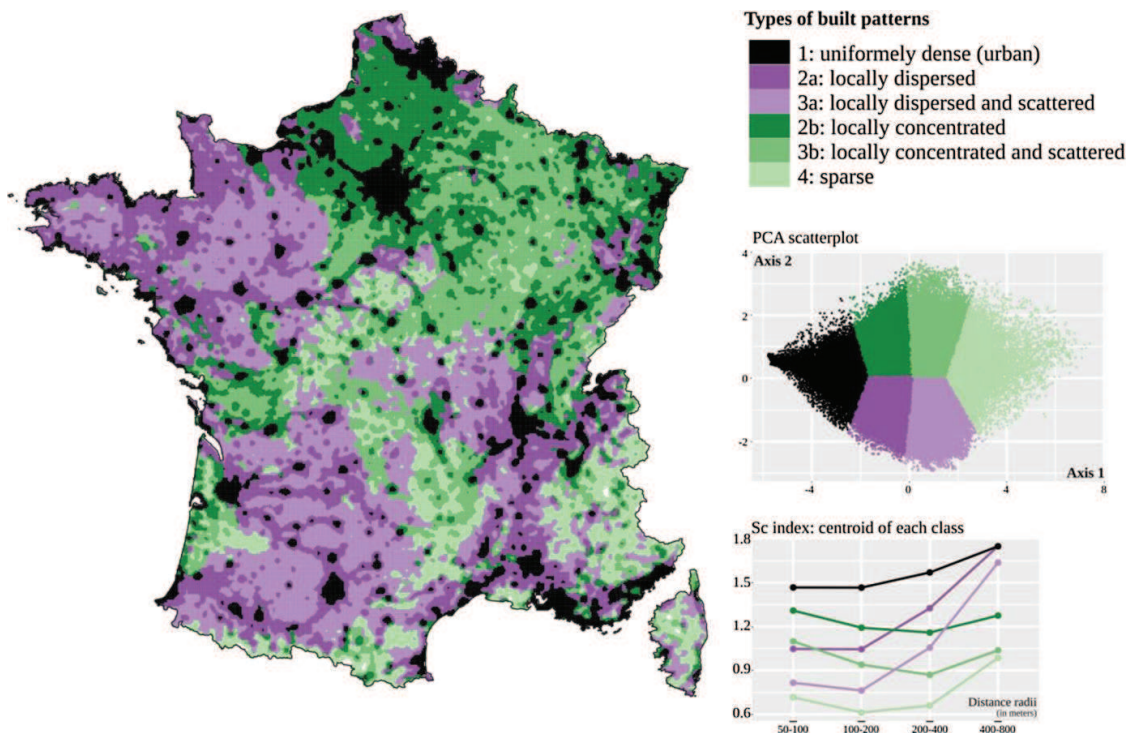


Fig. 10 Typology of built patterns in mainland France. Colours have been chosen according to the recommendations of (Jenny and Kelso 2007) using the software application ColorBrewer 2.0 (<http://colorbrewer2.org/>)

4.3 For twenty middle-size metropolitan areas, identification of the distance threshold that characterises globally how the spatial distribution of buildings deviates from scale invariance

For the purpose of the analysis, we used the delineation of French *aires urbaines* (metropolitan areas) (Fig. 11), which encompass a densely built urban area and its commuter belt. We chose to consider neither large metropolitan areas because they encompass several *aires urbaines* that are sometimes spatially separated, nor

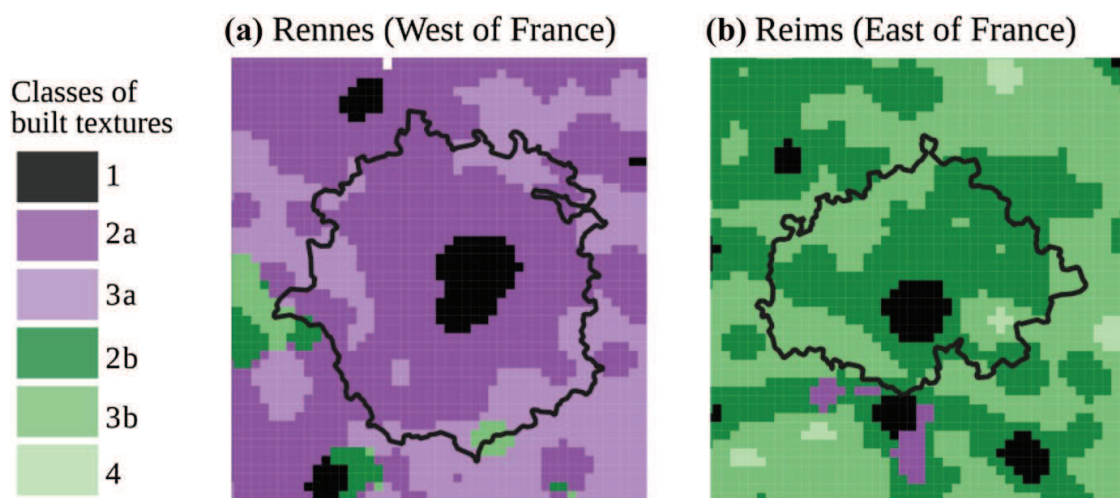


Fig. 11 Two metropolitan areas and their built textures. Black line: limit of the metropolitan area (*aire urbaine*). Colours have been chosen according to the recommendations of (Jenny and Kelso 2007) using the software application ColorBrewer 2.0 (<http://colorbrewer2.org/>)

small metropolitan areas as they do not contain enough 2000-m-wide spatial units. Thus, twenty metropolitan areas have been analysed. Their number of inhabitants is comprised between 110 946 (Saint-Quentin) and 897 713 (Nantes) (source: French National Institute of Statistics and Economic Studies 2012).

Within the perimeter of each metropolitan area, a distance threshold that indicates a crucial discontinuity in space across scales was identified using the MorphoLim software application (Tannier et al. 2011). This distance threshold indicates the maximum morphological difference between the built patterns that belong to the urban area (i. e. the morphological agglomeration) and the outlying built patterns (i. e. the surrounding suburban areas). Table 1 shows that this distance threshold varies markedly among cities, from 78 m. for Avignon (South of France) to 529 m. for Troyes (East of France).

4.4 Comparison of local built patterns with the global shape of metropolitan areas

Tables 1 and 2 show that the proportion of dispersed (or alternatively concentrated) suburban built patterns in metropolitan areas is closely related to their global shape, i.e. their distance threshold. The more the suburban built patterns are dispersed into numerous, regularly spaced, small villages and hamlets, the smaller the distance threshold; the more the suburban built patterns are locally concentrated within large villages, the greater the distance threshold. This clearly confirms the hypothesis that the shape of French urban areas strongly depends on the built-up shapes in their periphery. Conversely, neither the proportion of spatial units in classes 1 (urban areas) and 4 (sparsely populated areas) nor the area of the metropolitan areas is correlated with the distance threshold.

5 Discussion

On the typology map (Fig. 10), we can recognize several geographical features of mainland France that can explain the local shape of built patterns. First, the delineations of the main urban areas (see Fig. 4b) appear clearly on the typology map as well as the mountains and the valleys (see Fig. 12). Second, the typology map in Fig. 10 exhibits numerous similarities with the map in Fig. 13. Indeed the typology map shows the classical division of mainland France into enclosed field systems associated with scattered settlements in the centre and the West of France and openfield systems associated with grouped settlements in the East of France. Early twentieth-century historians (Bloch 1931; Dion 1934; Demangeon 1927) tried to explain local differences in rural settlement patterns in France by adding other explanatory variables to the agrarian system. They showed that mountain and forest landscapes are associated with specific settlement patterns, whereas other natural factors such as access to water or the shape of hydrographic networks are not. Analysis results presented in Fig. 10 confirm this. Valleys are either in class 2a (*locally dispersed built patterns*) or in class 2b (*locally concentrated built patterns*). Contrarily, high mountains are characterised by a mix of classes 3b (*moderately concentrated built patterns*) and 4 (*sparse built patterns*) everywhere in France. Finally, the typology map in Fig. 10 enlightens us at a fine spatial resolution about differences in the spatial distribution of buildings that result from shaping factors acting at local or regional scales and being both social and natural.

Table 1 Morphological characteristics of 20 French metropolitan areas in 2012

Name of the metropolitan area	Distance threshold (m)	Area (km ²)	% of spatial units exhibiting locally dispersed built patterns (classes 2a and 3a)	% of spatial units exhibiting locally concentrated built patterns (classes 2b and 3b)
Avignon	78	2396	47	3
Angers	127	2744	87	2
Bayonne	143	1440	71	0
Le Mans	143	2364	89	0
Rennes	143	4236	90	2
Vannes	165	940	70	9
Nantes	171	3884	73	10
Caen	181	2168	51	26
Poitiers	207	2648	18	62
Clermont-Ferrand	242	2764	19	60
Niort	245	1652	20	69
Bourges	256	1912	9	79
Nancy	290	2832	2	82
Orléans	318	3572	33	48
Metz	389	1940	1	76
Amiens	395	2424	0	93
Dijon	415	3844	0	87
Saint-Quentin	420	952	0	88
Reims	434	2752	0	93
Troyes	529	2296	0	91

Table 2 Statistical correlation between the distance threshold characterising each urban area and other morphological characteristics

Variable	Pearson correlation coefficient
% of spatial units in classes 2a and 3a (dispersed built pattern)	-0.82***
% of spatial units in class 2a	-0.82***
% of spatial units in class 3a	-0.69***
% of spatial units in classes 2b and 3b (concentrated built pattern)	0.89***
% of spatial units in class 2b	0.81***
% of spatial units in class 3b	0.62**
% of spatial units in class 1	-0.52*
% of spatial units in class 4	0.07
Area of the metropolitan areas	-0.04

20 metropolitan areas. *p*-values ****p* < 0.001 ***p* < 0.01 **p* < 0.05

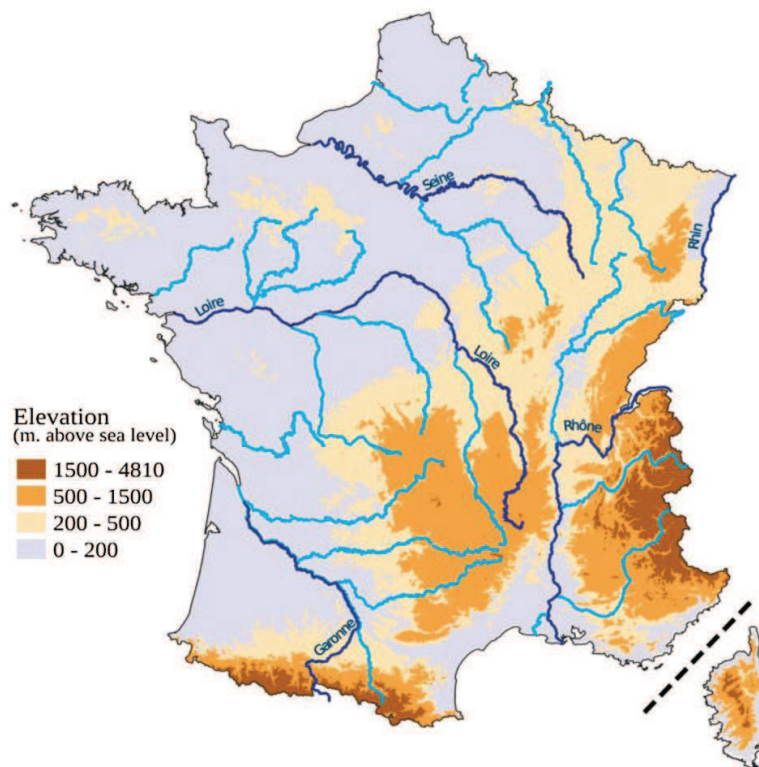


Fig. 12 Main mountains and rivers in mainland France. Sources: BD ALTI[®] IGN and BD TOPO[®] IGN. Colours have been chosen according to the recommendations of (Jenny and Kelso 2007) using the software application ColorBrewer 2.0 (<http://colorbrewer2.org/>)

The classification of scaling indexes obtained with GWFA has enabled the identification of six types of built patterns whose multiscale spatial organization is well differentiated. Sémécurbe et al. (2016) have previously identified similar types of settlement patterns by applying a multifractal analysis on fine-grained population data. That previous study was coarser grained as mainland France was divided into 992 square spatial units 25 km wide. (The present study considered 145,178 estimation points *i* 2 km apart.) The method was also different: the typology (a Hierarchical Ascending

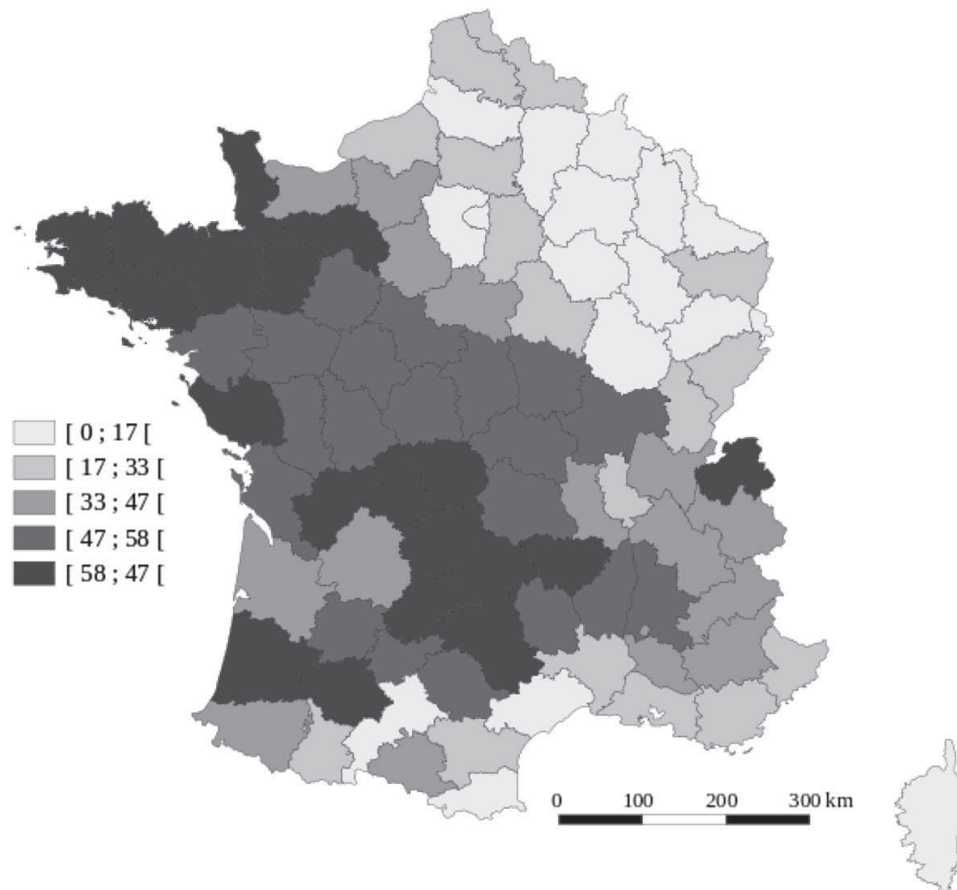


Fig. 13 Share of sparse population (in %) in the total population of each French sub-regions (*départements*) in 1876. Source: French National Institute of Statistics and Economic Studies (SGF data). Note: the sparse population of each local community is the population located outside the main population cluster of the community, which includes in general the town hall, church, and marketplace

Classification) was based on the multifractal spectra of each spatial unit. Interestingly, despite the differences in methods and spatial resolutions, those two studies detected the same phenomena. In particular, the map displaying the value of each point i on the first axis of the PCA in the present study, which contrasts areas characterised by a low fractal dimension and a low built density (*i.e.* rural areas) and areas characterised by a high fractal dimension and a high built density (*i.e.* urban areas) (Fig. 7a), resembles the map of generalized fractal dimension of order $q=2$ that highlights the strongest spatial singularities [Fig. 3d in (Sémécurbe et al. 2016)]. It should be recalled here that GWFA calculates a generalized fractal dimension of order $q=0$. Conversely, the map displaying the value of each estimation point i on the second axis of the PCA in the present study (Fig. 7b that highlights the deviation with respect to scale invariance) resembles the map of generalized fractal dimension of order $q=0$ for which only the presence or absence of population in each cell is taken into account [Fig. 3b in (Sémécurbe et al. 2016)]. Thus, the two studies look at the same phenomena, namely the concentration and dispersion of human settlements at several scales, from two different points of view, namely the departure from scale invariance in the present paper and the more or less multifractal aspect of the spatial distribution in (Sémécurbe et al. 2016).

6 Conclusion

In this paper, we have proposed a new method for spatial analysis, namely Geographically Weighted Fractal Analysis. We have used it to explore the local deviations from scale invariance of spatial distributions of buildings in mainland France. On this basis, without making any a priori distinction between urban patterns and suburban or rural patterns and without imposing any contiguity criterion, we have identified six geographically consistent built patterns that differ in the way buildings are concentrated and dispersed across scales. Then, for twenty middle-size metropolitan areas, we have compared the local built textures with the distance threshold that characterises globally the spatial distribution of buildings. This comparison has shown that the global shape of cities is closely related to their surrounding built patterns: the urban sprawl process has affected places differently according to their pre-existing rural built patterns (typically, numerous small dispersed hamlets or fewer large villages), which correspond to centuries-old agrarian systems (typically, enclosed field systems and openfield systems). Finally, the results obtained from this original four-step methodological process are more informative than just calculating a local fractal dimension. (The comparison of Fig. 3 with Fig. 10 illustrates this very clearly.)

One point of interest in the present study is that the relatively small size of elementary spatial units (2000 m-wide) has allowed us to compare and contrast the local built textures with the global shape of cities (i.e. their distance threshold). Because of the use of both a k-means classification and the MorphoLim method, the results obtained may be specific to mainland France: on the one hand, the result of the statistical classification of built textures depends on the diversity of built textures within the study area (in mainland France, the built textures are highly contrasted); on the other hand, the distance threshold characterising each urban area depends on the definition of metropolitan areas, which may vary among countries. Thus, it would be interesting to apply the same methodology as proposed in this paper to analyse built patterns in other countries, with a view to making international comparisons.

The fact that outer suburban patterns are inherited (in part at least) from past settlement patterns is well recognized: like many complex systems, settlement systems are generally characterised by strong path dependency (Andersson 2008). Nevertheless, in some cases, the urban sprawl process has not preserved features of old agrarian systems, but has made a more or less clean sweep of pre-existing rural built patterns. For instance, in the Champagne-Ardennes region of France, when land was controlled by cereal farmers and winegrowers, residential developments occurred mainly as extensions to the larger villages. Conversely, housing estates with stereotyped architecture were constructed where successive land reparcelling made sites available or where farmers were behind the projects so as to cash in on their landholdings or prepare for their retirement (Mancebo and Salles 2014). Such spatial variation in the intensity of path dependency could be explored quantitatively by applying the method and tools proposed in this paper.

Acknowledgement The authors would like to thank Armelle Couillet, cartographer at the research laboratory IDEES (Rouen, France), for her help in making maps readable by people with red–green colour blindness.

Appendix A: Computing GWFA using R

```

library(sf) # to use shapefiles
library(dbscan) # to identify the neighbours of each point e
library(btb) # to smooth the data with a quartic kernel (with the function msf using a bandwidth of size 2000 m. to 8000
m.)
library(dplyr) # to manage data.frame
library(cartography) # to create maps

radius = c(50, 100, 200, 400, 800) # Distance radii used to perform GWFA

## Transform a shapefile that represents buildings in two dimensions (polygons) into a table that contains
building centroids e with their xy coordinates. ##

buildings = read_sf("filename.shp")

# Transform the initial cartographic projection into Lambert 93
buildings = st_transform(buildings, 2154)

# Calculate the coordinates of the centroid of each building
points = st_coordinates(st_centroid(buildings))
points = data.frame(points)
names(points) = c("x", "y")

## Count the number of neighbours of each point (i.e. building centroid) e within each radius (50, 100, 200, 400,
800). ##

for (r in radius)
{
# Identify the neighbours of each point e
nn = ftNN(points[,1:2], r)

# Count the number of neighbours of each point e
points[, paste0('M_', r)] = sapply(adjacencylist(nn),length) + 1
}

## Apply GWFA to the building centroids e for five distance radii (50, 100, 200, 400, and 800 m). ##

# Calculate the indexes  $M_r(i)$  for each radius  $r$  and each estimation point  $i$  (cf. eq 4)
neighbors = points %>% mutate_at(paste0('M_', radius), function(x){x^-1})
msf = kernelSmoothing(neighbors, 2154, 2000, 8000)
msf = msf %>% mutate_at(paste0('M_', radius), function(x){x^-1})

# Compute the fractal dimension  $D_0$ 
logradius = log2(radius)
X = (logradius - mean(logradius)) / var(logradius) / (length(logradius) - 1)
M_r = msf %>% select(paste0('M_', radius)) %>% st_set_geometry(NULL)
msf$fd_50_800 = as.matrix(log2(M_r)) %*% matrix(X, length(logradius), 1)

## Map the fractal dimension  $D_0$  of each estimation point  $i$  ##

choroLayer(msf, var = "fd_50_800")

```

Appendix B: Computing GWFA using both Python and R

1. Computing code in Python

```

import pandas as pd
import geopandas as gpd
from shapely.geometry import Polygon
from sklearn.neighbors import KDTree
import numpy as np

## Transform a shapefile that represents buildings in two dimensions (polygons) into a table that contains
building centroids e with their XY coordinates ##

building=gpd.read_file("gis.osm_buildings_a_free_1.shp")

# Transform the initial cartographic projection into Lambert 93
building=building.to_crs({'init': 'epsg:2154'})

centroid = building['geometry'].centroid

points=np.column_stack([centroid.x,centroid.y])

## Count the number of neighbours of each point e within each radius r (50, 100, 200, 400, 800) ##
radius=np.array([50,100,200,400,800])

neighbors=np.zeros((points.shape[0],radius.size),dtype=np.float)
tree = KDTree(points)

for i in range(radius.size):
neighbors[:,i]=tree.query_radius(points,r= radius[i],count_only=True,return_distance=False)

output=np.column_stack([points,neighbors])
names=np.hstack(['x','y'], np.core.defchararray.add('r_',np.char.mod('%d', radius)))

output=pd.DataFrame(output,columns=names)
output.to_csv("neighbors.csv", float_format='%0.6f')

```

2. Computing code in R

```

## Apply GWFA to the building centroids e for five distance radii r (50, 100, 200, 400, and 800 m) ##
library(btb)
library(sf)

neighbors=read.csv("neighbors.csv")
neighbors=neighbors[,-1]

carreauxSF <- function(df, iCellSize, sEPSG)
{
  r <- iCellSize / 2
  df$geom <- sprintf("POLYGON ((%i %i, %i %i, %i %i, %i %i))", df$x-r, df$y+r, df$x+r, df$y+r,
df$x+r, df$y-r, df$x-r, df$y-r, df$x-r, df$y+r)
  sdf <- st_as_sf(df, wkt = "geom", crs = as.integer(sEPSG))
  return(sdf)
}

neighbors[,c(-1,-2)]=neighbors[,c(-1,-2)]^-1
liss=kernelSmoothing(dfObservations = neighbors,cellSize = 2000,bandwidth = 8000)
msf=data.frame(liss@.Data)
names(msf)=c("x", "y", "M_50", "M_100", "M_200", "M_400", "M_800")

msf[,c(-1,-2)]=log2(msf[,c(-1,-2)])
msf$df_20_5=-(msf$M_200-msf$M_50)/2
msf$df_80_20=-(msf$M_800-msf$M_200)/2
scale=c(1,2,3,4,5)
x=-(scale-mean(scale))/(var(scale)*4)
msf$df80_5=(as.matrix(msf[,3:7])%*%matrix(x,5,1))

output_8000=carreauxSF(msf,2000,2154)

st_write(output_8000,"analyse_8000.shp")

```

References

- Andersson C (2008) Ontogeny and ontology in complex systems modeling. In: Albeverio S, Andrey D, Giordano P, Vancheri A (eds) *The dynamics of complex urban systems: an interdisciplinary approach*. Physica-Verlag, Heidelberg, pp 43–58
- Batty M, Longley P (1986) The fractal simulation of urban structure. *Environ Plan A* 18:1143–1179
- Batty M, Longley PA (1987) Urban shapes as fractals. *Area* 19(3):215–221
- Batty M, Xie Y (1996) Preliminary evidence for a theory of the fractal city. *Environ Plan A* 28:1745–1762
- Benguigui L, Czamanski D (2004) Simulation analysis of the fractality of cities. *Geogr Anal* 36(1):69–84
- Benguigui L, Czamanski D, Marinov M, Portugali Y (2000) When and where is a city fractal? *Environ Plan B Plan Des* 27:507–519
- Bloch M (1931) *Les caractères originaux de l'histoire rurale française*. Instituttet for sammenlignende kulturforskning, Oslo
- Brunsdon C, Fotheringham AS, Charlton M (1996) Geographically Weighted Regression: a method for exploring spatial nonstationarity. *Geogr Anal* 28(4):281–298
- Chaudhry O, Mackaness W (2008) Automatic identification of urban settlement boundaries for multiple representation databases. *Comput Environ Urban Syst* 32:95–109
- Chen Y, Wang J (2013) Multifractal characterization of urban form and growth: the case of Beijing. *Environ Plan* 40(5):884–904
- De Keersmaecker M-L, Thomas I, Frankhauser P (2003) Using fractal dimensions for characterizing intra-urban diversity: the example of Brussels. *Geogr Anal* 35(4):310–328
- Demangeon A (1927) *La géographie de l'habitat rural*. *Annales de géographie* 36:1–23
- Diggle P (1985) A kernel method for smoothing point process data. *J Roy Stat Soc Ser C (Appl Stat)* 34(2):138–147
- Dion R (1934) *Essai sur la formation du paysage rural français*. Arrault, Tours
- Feng J, Chen Y (2010) Spatiotemporal evolution of urban form and land-use structure in Hangzhou, China: evidence from fractals. *Environ Plann B Plann Des* 37(5):838–856
- Fotheringham AS, Brunsdon C, Charlton M (2003) *Geographically Weighted Regression: the analysis of spatially varying relationships*. Wiley, Hoboken
- Frankhauser P (1988) Fractal aspects of urban systems. *Sonderforschungsbereich 230 "Natürliche Konstruktionen"* 1:67–76
- Frankhauser P (1998) The fractal approach: a new tool for the spatial analysis of urban agglomerations. *Population* 4:205–240
- Goodchild M (1980) Fractals and the accuracy of geographical measures. *Math Geol* 12(2):85–98
- Grassberger P, Procaccia I (1983) Measuring the strangeness of strange attractors. *Physica D* 9:189–208
- Jenny B, Kelso NV (2007) Designing maps for the colour-vision impaired. *Bull Soc Cartogr* 41:9–12
- Lam N (1990) Description and measurement of Landsat TM images using fractals. *Photogramm Eng Remote Sens* 56(2):187–195
- Lam N, Quattrochi D (1992) On the issues of scale, resolution, and fractal analysis in the mapping sciences. *Prof Geogr* 44(1):88–98
- Le Gléau J, Pumain D, Saint-Julien T (1997) Towns of Europe: to each country its definition. *INSEE Stud* 6:1–14
- Longley P, Batty M (1989) On the fractal measurement of geographical boundaries. *Geogr Anal* 21(1):47–67
- Mancebo F, Salles S (2014) De l'autre côté du miroir. Un périurbain pensé par le rural, pour une périurbanisation modelée par les usages. *Premier plan Le journal d'informations du PUCA (plan urbanisme construction architecture)* 30:4–6
- Mandelbrot B (1982) *The fractal geometry of nature*. Freeman, San Francisco
- Sémécurbe F, Tannier C, Roux SG (2016) Spatial distribution of human population in France: exploring the modifiable areal unit problem using multifractal analysis. *Geogr Anal* 48(3):292–313
- Shen G (2002) Fractal dimension and fractal growth of urbanized areas. *Int J Geogr Inf Sci* 16(5):437–519
- Silverman BW (1978) Choosing the window width when estimating a density. *Biometrika* 65:1–11
- Tannier C, Pumain D (2005) Fractals in urban geography: a theoretical outline and an empirical example. *Cybergeo Eur J Geogr* 307. <https://doi.org/10.4000/cybergeo.3275>
- Tannier C, Thomas I (2013) Defining and characterizing urban boundaries: a fractal analysis of theoretical cities and Belgian cities. *Comput Environ Urban Syst* 41:234–248

- Tannier C, Thomas I, Vuidel G, Frankhauser P (2011) A fractal approach to identifying urban boundaries. *Geogr Anal* 43(2):211–227
- Tél T, Fülöp Á, Vicsek T (1989) Determination of fractal dimensions for geometrical multifractals. *Physica A* 159(2):155–166
- Thomas I, Frankhauser P, Biernacki C (2008) The morphology of built-up landscapes in Wallonia, Belgium: a classification using fractal indices. *Landsc Urban Plan* 84:99–115
- Thomas I, Frankhauser P, Frenay B, Verleysen M (2010) Clustering patterns of urban built-up areas with curves of fractal scaling behaviour. *Environ Plan* 37:942–954
- Vicsek T (1990) Mass multifractals. *Physica A* 168(1):490–497
- White R, Engelen G (1993) Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. *Environ Plan A* 25(8):1175–1199
- White R, Engelen G (1994) Urban systems dynamics and cellular automata: fractal structures between order and chaos. *Chaos, Solitons Fractals* 4(4):563–583

Affiliations

François Sémécurbe¹ · Cécile Tannier¹  · Stéphane G. Roux²

François Sémécurbe
francois.semecurbe@univ-fcomte.fr

Stéphane G. Roux
stephane.roux@ens-lyon.fr

¹ ThéMA, UMR 6049, CNRS-Univ. Bourgogne Franche-Comté, 32, rue Mégevand, 25030 Besançon Cedex, France

² Laboratoire de Physique, UMR 5672, CNRS-Ecole Nationale Supérieure de Lyon, 46, allée d'Italie, 69364 Lyon Cedex 07, France

Annexe C

Principes à la base du *deep learning*

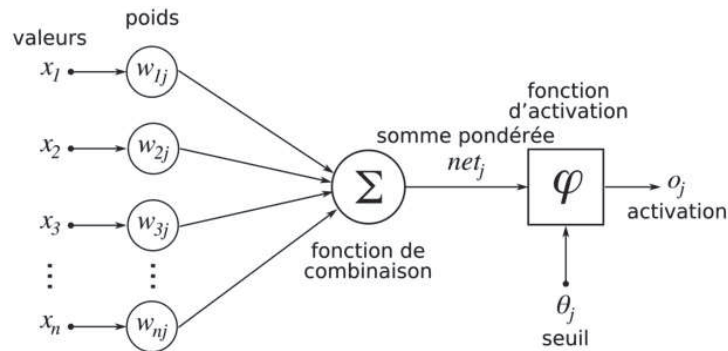
Les réseaux de neurones sont des fonctions qui associent à une entrée donnée, telle qu'un ensemble de caractéristiques décrivant un objet, une réponse, une prédiction, qui peut être une valeur quantitative ou qualitative. Par exemple, si l'on souhaite modéliser le prix des logements (variable quantitative) ou leur diagnostic thermique (variable qualitative), on mobilisera les variables suivantes : coordonnées spatiales des logements, surfaces, type de bâtiment (maison individuelle ou immeuble collectif). La vocation des réseaux de neurones ne se distingue pas de celle de modélisations plus simples telles que les régressions linéaires. Leur principal avantage par rapport à ces dernières consiste en leur capacité à reproduire des comportements non linéaires et d'être ainsi plus à même de reproduire des phénomènes complexes.

La notion de neurone artificiel remonte aux années 1950 avec les travaux de F ROSENBLAT. L'idée de créer des couches de neurones et de les empiler afin de former des réseaux de neurones a émergé plus tard. Le *deep learning*, l'apprentissage profond, désigne les réseaux de neurones composés de nombreuses couches.

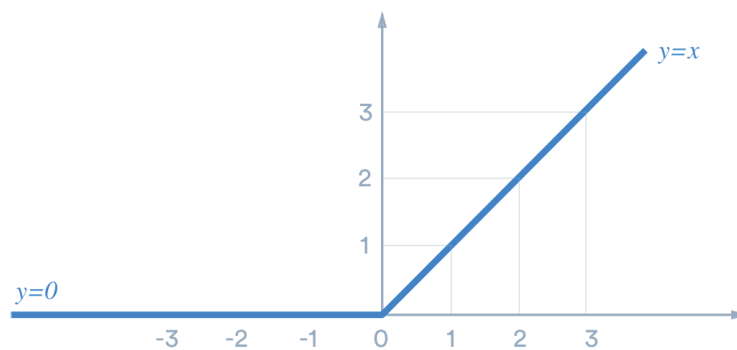
C.1 Perceptron multicouche

Le perceptron multicouche est le réseau de neurones artificiel le plus simple. Il est généralement utilisé comme réseau « témoin » afin d'évaluer l'apport de réseaux plus complexes. Le perceptron multicouche est dédié à l'apprentissage supervisé. Afin qu'il soit opérationnel, autrement dit qu'il associe convenablement une entrée à une sortie, on calibre celui-ci à l'aide d'une phase d'apprentissage. On dispose pour cela d'une base de données contenant des objets X qui ont été labellisés y . L'objectif est de généraliser cette relation entre un objet et son label à des objets inconnus mais qui néanmoins ressemblent suffisamment à ceux de la base d'apprentissage pour que l'extrapolation ait un sens.

Les perceptrons multicouches sont composés de neurones élémentaires qui réalisent deux opérations successives : tout d'abord, ils calculent la somme pondérée des entrées x_i , $\sum w_{ij}x_i$ puis cette valeur passe à travers la fonction d'activation ϕ pour produire sa sortie o (fig. C.1.a). Le passage par la fonction d'activation a pour objectif d'introduire de la non-linéarité. Sans cette étape, un réseau de neurones serait une somme de fonctions linéaires, et donc encore une fonction linéaire. La fonction d'activation la plus utilisée est la fonction *ReLU* (Unité de Rectification Linéaire) (fig. C.1.b). Lorsque la somme pondérée est inférieure à θ_j , la sortie du neurone est 0, sinon elle vaut $\sum w_{ij}x_i - \theta_j$.



(a) Opérations au sein d'un neurone artificiel. Source : Wikipédia (Réseau de neurones artificiels) consultée le 07/03/2020.



(b) La fonction d'activation ReLU. Source : Médium consultée le 07/03/2020.

FIGURE C.1 – Schéma de fonctionnement d'un neurone artificiel.

Le perceptron multicouche est constitué de couches de neurones positionnés en série : les entrées d'une couche sont les sorties de la précédente. Le réseau présenté sur la figure C.2 est composé d'une couche d'entrée de trois variables et d'une couche de sortie de quatre neurones qui représentent des classes d'affectation. Chaque neurone de cette dernière couche fournit une probabilité d'appartenance à une classe spécifique, son label. Par exemple, avec ce type de réseau nous pourrions définir une règle d'association entre la longitude, la latitude, la surface d'un bien immobilier et sa performance énergétique dans un référentiel en 4 classes : bon, moyen, mauvais et très mauvais. Entre ces deux couches, le réseau contient deux couches cachées : la première à 4 neurones et la deuxième 3 neurones.

C.2 Étape d'apprentissage : estimation des poids

Tout comme pour la régression linéaire, il faut calibrer les paramètres du réseau, ses poids, afin de rendre pertinentes les prédictions. Mais contrairement à la régression linéaire, il n'existe pas de formule théorique explicite permettant de le faire. La solution proposée par RUMELHART, HINTON et WILLIAMS 1985 consiste à entraîner progressivement le réseau à l'aide d'une base d'apprentissage. L'avantage de cette méthode itérative est qu'elle est adaptée au traitement de données massives. Cette étape d'entraînement s'effectue de la façon suivante :

- Premièrement, on découpe la base d'apprentissage en deux parties. La première X_{train}, y_{train} contient 80% des données et la seconde X_{test}, y_{test} en contient 20%.

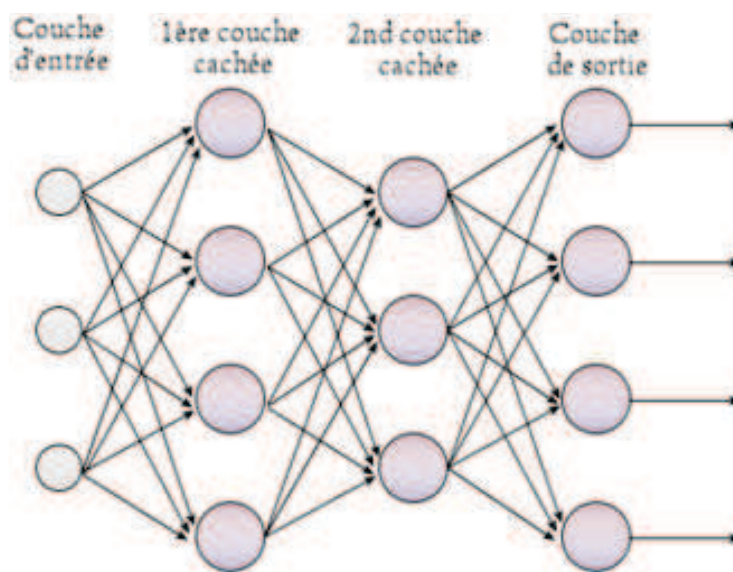


FIGURE C.2 – Perceptron multicouche. Source : Wikipédia (Perceptron multicouche) consultée le 07/03/2020.

- On opère plusieurs cycles d'apprentissage. A chaque cycle, on détermine un échantillon aléatoire contenant plusieurs X_{train} et y_{train} sur lequel on calcule un indicateur de qualité entre les prédictions obtenues et les vrais labels. Cet indicateur de qualité, appelé fonction de perte, indique la façon dont il faut modifier les poids pour améliorer les prédictions.
- Lorsqu'à la fin d'un cycle, la qualité des prédictions n'est pas améliorée, la phase d'apprentissage est terminée et l'on teste la précision du réseau de neurones sur la base de données X_{test} .

La séparation de la base d'apprentissage en deux est indispensable pour éviter le sur-apprentissage, c'est-à-dire le fait que le réseau soit tellement adapté à la base de données d'apprentissage qu'il n'est pas capable de généraliser ses prédictions à de nouveaux objets (fig C.3).

La profondeur du réseau et le nombre de neurones pour les différentes couches sont obtenus empiriquement après plusieurs tests : on conserve le réseau dont la qualité des prédictions est la plus élevée. Il existe plusieurs fonctions de pertes en fonction des sorties du réseau de neurones. Pour une sortie quantitative, on utilise généralement une fonction quadratique $L(y, y_{pred}) = ||y_{pred} - y||^2$ et pour une sortie qualitative, l'entropie croisée.

C.3 Algorithme du gradient descendant

L'algorithme du gradient descendant est utilisé pour minimiser la fonction de perte L à chaque étape. A la fin de l'algorithme, les paramètres obtenus définissent un minimum local.

Supposons que le réseau que nous souhaitons calibrer contienne uniquement un poids w_1 . $L(y, y_{pred})$ est une fonction de w , $L(w)$ dont on peut calculer la dérivée : $\frac{dL}{dw}$. Cette dérivée décrit la pente de la fonction de perte autour de w . Une dérivée positive indique que la fonction est croissante. Dans ce cas, si on choisit un poids légèrement plus grand, on augmente la valeur de la fonction de perte. En revanche, si l'on choisit un poids légèrement

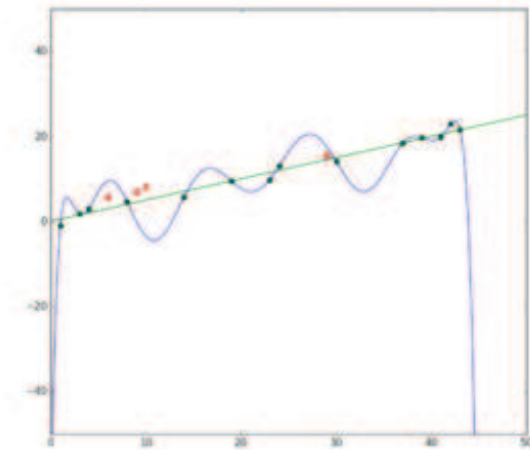


FIGURE C.3 – Sur-apprentissage. En vert, les points de l'ensemble d'apprentissage et une régression linéaire sur ces points. En rouge, les points de l'ensemble de test. Source : Wikipédia (Surapprentissage) consultée le 07/03/2020.

plus petit, on diminue la valeur de la fonction de perte. L'algorithme du gradient descendant formalise cette idée. Supposons que w a été initialisé aléatoirement, on peut construire une suite récurrente w_k de paramètres tels que $L(w_k)$ soit décroissante (fig. C.4) :

$$\begin{aligned} &— w_0 = w \\ &— w_{k+1} = w_k - \alpha \frac{dL}{dw} \end{aligned}$$

En pratique, on arrête la suite lorsque $w_{k+1} - w_k$ devient très petit. Le dernier w_k calculé est la valeur optimisée du paramètre. α est une quantité qui permet de quantifier l'ampleur de la correction à chaque étape (si *alpha* est trop faible alors la convergence est lente, au contraire si *alpha* est trop élevé, on risque d'osciller autour du minimum sans jamais l'atteindre). α est ce que les informaticiens appellent un hyper-paramètre. Ce n'est pas un paramètre de la modélisation mais un paramètre qui permet de l'estimer. Il n'existe pas de bonne valeur *a priori* de α . Le plus souvent, on essaye plusieurs valeurs avant de conserver celle qui donne le plus petit minimum.

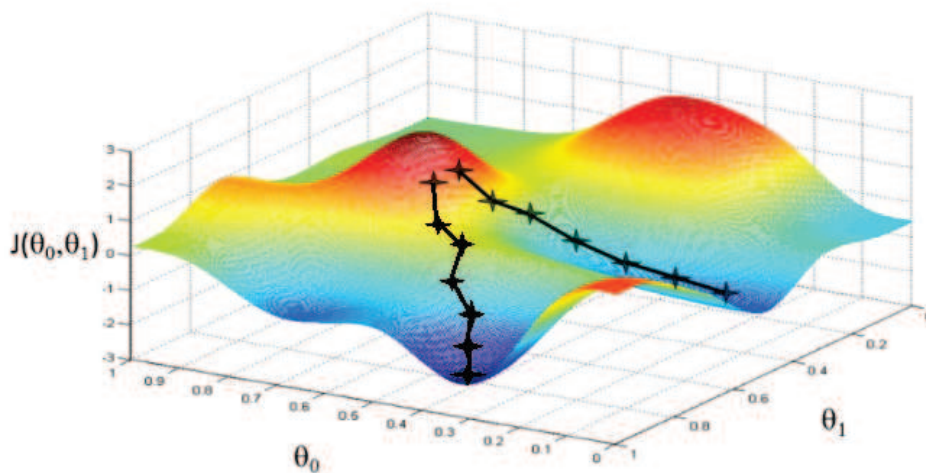


FIGURE C.4 – Algorithme du gradient descendant

Par construction, en règle générale, le minimum de la fonction de perte obtenu n'est pas global mais local : en fonction de la valeur choisies pour l'initialisation de l'algorithme, la suite peut converger vers des valeurs optimales distinctes. Pour mettre en œuvre l'algorithme, il faut choisir explicitement une valeur de α . Une valeur de α trop petite implique une convergence très lente et une valeur de α trop grande un phénomène d'oscillation autour du minimum. Le choix du paramètre est déterminé empiriquement après plusieurs essais.

Les réseaux de neurones n'ont pas un seul paramètre mais bien souvent plusieurs milliers contenus dans plusieurs couches successives. Calculer les dérivées partielles¹ de la fonction de perte pour chaque paramètre est dans ce cas très délicate. RUMELHART, HINTON et WILLIAMS 1985 propose une méthode appelée rétropropagation du gradient qui consiste à partir de la couche de sortie et à utiliser la formule des dérivées des fonctions composées pour en déduire de façon itérative les dérivées des poids des couches cachées précédentes.

C.4 Réseaux convolutifs : prise en compte de la spatialité des données

Les réseaux de neurones convolutifs dédiés aux traitements des images sont adaptés à l'étude des phénomènes géographiques : l'information contenue dans une image est généralisée successivement à travers les échelles afin d'en faire émerger ses éléments structurants, conformément à la théorie de la vision proposée par MARR et HILDRETH 1980. De façon imagée, aux échelles les plus fines, le réseau se concentre sur les détails tels que les feuilles, puis il agrègera aux échelles intermédiaires cette information pour proposer une *représentation* des branches qu'il utilise pour *décrire* un arbre comme étant une combinaison de branches feuillues aux échelles les plus grossières.

Quand on utilise le perceptron multicouche sur des images en niveaux de gris (deux dimensions) ou en couleur (3 dimensions avec les composantes spectrales), ceci nécessite d'avoir au préalable transformé les images en vecteur à une dimension, par exemple, en alignant les lignes des images les unes derrière les autres. Les images ainsi transformées perdent donc totalement leur structure spatiale. Les réseaux convolutifs ont été proposés par LECUN et al. 1998 pour traiter explicitement la dimension spatiale des images.

Les réseaux convolutifs traitent l'information spatiale contenue dans une image à l'aide de deux blocs d'opérations répétés successivement : une convolution suivie d'une étape de *pooling*.

Convolution. L'opération de convolution consiste à appliquer un filtre sur l'image. Cette opération se rapproche d'une transformée en ondelettes d'une image, l'ondelette étant un filtre spécifique choisi explicitement par l'utilisateur. Dans le cas des réseaux convolutifs, les poids des filtres sont déterminés lors de la phase d'apprentissage. La convolution consiste à découper l'image en petits morceaux de taille identique à celle du filtre, puis à calculer la somme pondérée des pixels de chaque morceau avec les valeurs correspondantes du filtre².

1. généralisation de la notion de dérivée aux fonctions à plusieurs variables

2. [fr.wikipedia.org/wiki/Noyau_\(traitement_d'image\)](http://fr.wikipedia.org/wiki/Noyau_(traitement_d'image))

Afin de simplifier les notations, nous nous restreignons dorénavant à la présentation d'images en niveau de gris. Sur la figure C.5, nous donnons un exemple de convolution entre l'image I et le filtre K de taille (3×3) . La valeur du carré vert de la convolution $I * K$ est le résultat de l'opération suivante $1 \times 1 + 0 \times 0 + 0 \times 1 + 1 \times 0 + 1 \times 1 + 0 \times 0 + 1 \times 1 + 1 \times 0 + 1 \times 0 = 4$. Plusieurs types de traitement des bords sont possibles. Le plus simple, proposé dans cet exemple, est de se restreindre aux pixels intérieurs de l'image I qui permettent le calcul du filtre.

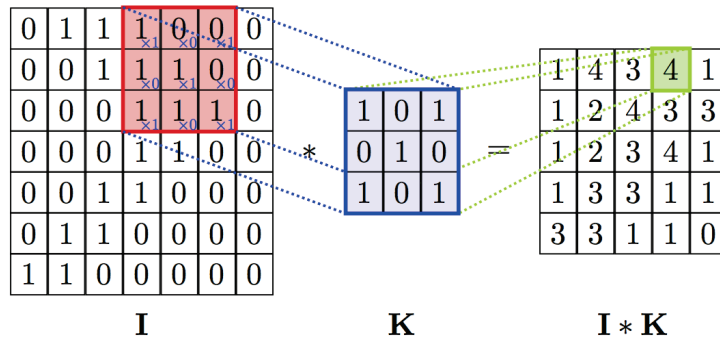


FIGURE C.5 – Convolution

Pooling En sortie de l'opération de convolution on obtient une nouvelle image dont l'information contenue dans deux pixels voisins est très redondante. Afin de supprimer cette information surnuméraire, on applique une opération de *pooling* qui consiste à ne conserver qu'une partie de l'image convoluée : pour chaque sous-matrice de taille (2×2) de l'image convoluée, on garde la valeur du pixel maximum (fig. C.7).

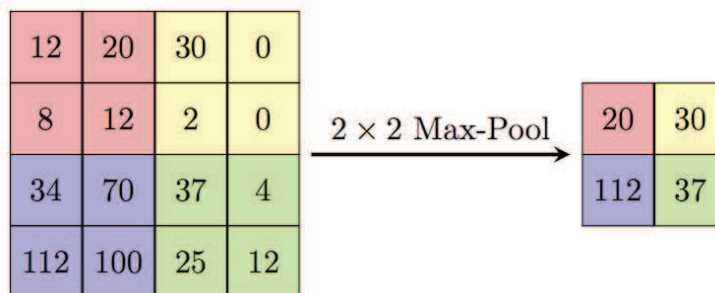


FIGURE C.6 – Pooling

Cette combinaison (couche de convolution suivie d'une couche de *pooling*) est répétée plusieurs fois dans les réseaux convolutifs jusqu'à ce que le niveau d'abstraction soit jugé suffisamment important. Ensuite des couches de perceptrons sont utilisées pour produire la sortie.

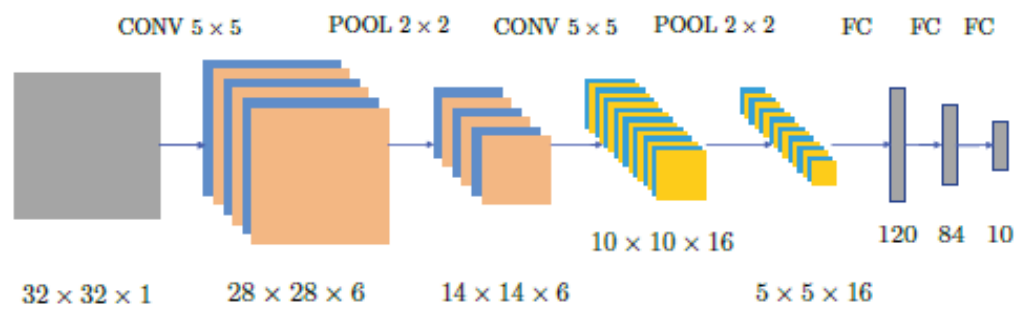


FIGURE C.7 – LeNet est un réseau proposé en 1998 par l'équipe de Yann LeCun pour lire les chiffres contenus dans une image. Il se compose de deux étapes de convolution permettant de passer d'une image de taille 32×32 pixels à une image de 5×5 .

Bibliographie

- ABDISHAKUR (2019). « Multi-label Land Cover Classification with Deep Learning ». In : *Towards Data Science*.
- ABRY, Patrice, Stéphane JAFFARD et Herwig WENDT (2015). « Irregularities and scaling in signal and image processing : multifractal analysis ». In : *Benoit Mandelbrot : A Life in Many Dimensions*. World Scientific, p. 31-116.
- ABRY, Patrice, Herwig WENDT et Stéphane JAFFARD (2013). « When Van Gogh meets Mandelbrot : Multifractal classification of painting's texture ». In : *Signal Processing* 93.3, p. 554-572.
- ABRY, Patrice, Herwig WENDT, Stéphane JAFFARD et Gustavo DIDIER (2019). « Multivariate scale-free temporal dynamics : From spectral (Fourier) to fractal (wavelet) analysis ». In : *Comptes Rendus Physique* 20.5, p. 489-501.
- AGNEW, John (1994). « The territorial trap : the geographical assumptions of international relations theory ». In : *Review of international political economy* 1.1, p. 53-80.
- AGTERBERG, Frederik P (2013). « Fractals and spatial statistics of point patterns ». In : *Journal of Earth Science* 24.1, p. 1-11.
- AKTAS, Nihat (2004). « La «finance comportementale» : un état des lieux ». In : *Reflets et perspectives de la vie économique* 43.2, p. 19-33.
- ALBIZZATI, Colin, Mathilde POULHES et Joyce SULTAN PARRAUD (2017). « Caractérisation des espaces consommés par le bâti en France métropolitaine entre 2005 et 2013 ». In : *Insee Références*.
- ANKERST, Mihael et al. (1999). « OPTICS : ordering points to identify the clustering structure ». In : *ACM Sigmod record*. T. 28. 2. ACM, p. 49-60.
- APPLEBY, Stephen (1996). « Multifractal characterization of the distribution pattern of the human population ». In : *Geographical Analysis* 28.2, p. 147-160.
- ARALDI, Alessandro et Giovanni FUSCO (2018). « Describing retail agglomeration form in urban space-Local fractal dimension for point pattern analysis on network ». In : *Revue Internationale de Géomatique* 28.4, p. 505-532.
- ARNEODO, A et al. (1996). « Wavelet analysis of fractals : from the mathematical concepts to experimental reality ». In : *Wavelets : Theory and Application*, p. 349-502.
- ARNEODO, A et S JAFFARD (2004). « L'analyse multi-fractale des signaux ». In : *Journal du CNRS* 107.
- ARNEODO, Alain, Benjamin AUDIT et al. (2008). « Wavelet-based multifractal analysis ». In : *Scholarpedia* 3.3, p. 4103.
- ARNEODO, Alain, Emmanuel BACRY et al. (1995). « Characterizing long-range correlations in DNA sequences from wavelet analysis ». In : *Physical Review Letters* 74.16, p. 3293.
- BACHELIER, Louis (1900). « Théorie de la spéculation ». In : *Annales scientifiques de l'École normale supérieure*. T. 17, p. 21-86.

- BADDELEY, Adrian, Ege RUBAK et Rolf TURNER (2015). *Spatial point patterns : methodology and applications with R*. Chapman et Hall/CRC.
- BADDELEY, Adrian J, Jesper MØLLER et Rasmus WAAGEPETERSEN (2000). « Non- and semi-parametric estimation of interaction in inhomogeneous point patterns ». In : *Statistica Neerlandica* 54.3, p. 329-350.
- BARNESLEY, Michael F (2014). *Fractals everywhere*. Academic press.
- BARTLETT, MS (1964). « 207. Note : A Note on Spatial Pattern ». In : *Biometrics* 20.4, p. 891-892.
- BATTY, Michael (1976). « Entropy in spatial aggregation ». In : *Geographical Analysis* 8.1, p. 1-21.
- (1991). « Cities as fractals : simulating growth and form ». In : *Fractals and chaos*. Springer, p. 43-69.
- BATTY, Michael et Paul A LONGLEY (1994). *Fractal cities : a geometry of form and function*. Academic press.
- BATTY, Michael, Robin MORPHET et al. (2014). « Entropy, complexity, and spatial information ». In : *Journal of geographical systems* 16.4, p. 363-385.
- BECK, Ulrich, Scott LASH et Brian WYNNE (1992). *Risk society : Towards a new modernity*. T. 17. sage.
- BESAG, Julian (1977). « Contribution to the discussion on Dr Ripley's paper ». In : *JR Stat. Soc.* 39, p. 193-195.
- BESAG, Julian et James NEWELL (1991). « The detection of clusters in rare diseases ». In : *Journal of the Royal Statistical Society : Series A (Statistics in Society)* 154.1, p. 143-155.
- BEVAN, Andrew et James CONOLLY (2006). « Multiscalar approaches to settlement pattern analysis ». In : *Confronting scale in archaeology*. Springer, p. 217-234.
- BIRABEN, Jean-Noël et Françoise DUHOURCAU (1974). « La mesure de la population dans l'espace ». In : *Population (french edition)*, p. 113-137.
- BLOCH, Marc (1931). « Les caractères originaux de l'histoire rurale française ». In : *Bulletin de l'Association Guillaume Budé* 33.1, p. 56-62.
- BOLIBOK, Leszek (2008). « Limitations of Ripley K function use in the analysis of spatial patterns of tree stands with heterogeneous structure ». In : *Acta Scientiarum Polonorum Silvarum Colendarum Ratio et Industria Lignaria* 7.1, p. 5-18.
- BOUDA, Martin, Joshua S CAPLAN et James E SAIERS (2016). « Box-counting dimension revisited : presenting an efficient method of minimizing quantization error and an assessment of the self-similarity of structural root systems ». In : *Frontiers in plant science* 7, p. 149.
- BOURDIEU, Pierre (2014). *Ce que parler veut dire : l'économie des échanges linguistiques*. Fayard.
- BOURLARD, Hervé et Yves KAMP (1988). « Auto-association by multilayer perceptrons and singular value decomposition ». In : *Biological cybernetics* 59.4-5, p. 291-294.
- BOUVERESSE, Jacques (1999). *Prodiges Et Vertiges de l'Analogie de l'Abus des Belles-Lettres Dans la Pensée*. Éditions Raisons d'agir.
- BRAS, H.L. et E. TODD (2013). *Le mystère français*. République des idées. La République des Idées.
- BRETAGNOLLE, Anne (2015). « La naissance du périurbain comme catégorie statistique en France. Une perspective internationale ». In : *L'Espace géographique* 44.1, p. 18-37.
- BROMILEY, PA, NA THACKER et E BOUHOVA-THACKER (2004). « Shannon entropy, Renyi entropy, and information ». In : *Statistics and Inf. Series (2004-004)*.

- BRUNSDON, Chris, A FOTHERINGHAM et Martin E CHARLTON (1996). « Geographically weighted regression : a method for exploring spatial nonstationarity ». In : *Geographical analysis* 28.4, p. 281-298.
- BRUNSDON, Chris, AS FOTHERINGHAM et Martin CHARLTON (2002). « Geographically weighted summary statistics—a framework for localised exploratory data analysis ». In : *Computers, Environment and Urban Systems* 26.6, p. 501-524.
- CARUSO, Geoffrey, Mohamed HILAL et Isabelle THOMAS (2017). « Measuring urban forms from inter-building distances : Combining MST graphs with a Local Index of Spatial Association ». In : *Landscape and Urban Planning* 163, p. 80-89.
- CASILLI, Antonio A (2019). *En attendant les robots-Enquête sur le travail du clic*. Le Seuil.
- CHARMES, Eric (2013). « L'artificialisation est-elle vraiment un problème quantitatif ? » In : *Études foncières*.
- (2019). *La revanche des villages-Essai sur la France périurbaine*. Le Seuil.
- CHHABRA, Ashvin et Roderick V JENSEN (1989). « Direct determination of the $f(\alpha)$ singularity spectrum ». In : *Physical Review Letters* 62.12, p. 1327.
- CHOUQUER, Gérard (2000). *L'étude des paysages : essais sur leurs formes et leur histoire*. Editions Errance.
- CHRISTALLER, Walter (1933). *Die zentralen Orte in Süddeutschland : eine ökonomisch-geographische Untersuchung über die Gesetzmässigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen*. (Reprint d. Ausg. Jena 1933).
- CLARK, Philip J et Francis C EVANS (1954). « Distance to nearest neighbor as a measure of spatial relationships in populations ». In : *Ecology* 35.4, p. 445-453.
- CLOUT, Hugh (2005). « Geographers in their ivory tower : Academic geography and popular geography in Paris 1931 ». In : *Geografiska Annaler : Series B, Human Geography* 87.1, p. 15-29.
- CLOUT, Hugh D (2013). *Themes in the historical geography of France*. Elsevier.
- COLAS, René (1945). « Répartition de l'habitat rural ». In : *Bulletin de l'Association de géographes français* 22.169, p. 51-56.
- CORBUSIER, Le (1959). *Les trois établissements humains*. Éditions de minuit.
- COTTINEAU, Clementine et al. (2015). « Paradoxical interpretations of urban scaling laws ». In : *arXiv preprint arXiv :1507.07878*.
- COUCLELIS, Helen (1986). « Artificial intelligence in geography : Conjectures on the shape of things to come ». In : *The professional geographer* 38.1, p. 1-11.
- COURGEAU, Daniel (2000). « Réflexions sur la causalité en sciences sociales ». In : *Revue des politiques sociales et familiales* 60.1, p. 49-60.
- CRESSIE, Noel (1992). « Statistics for spatial data ». In : *Terra Nova* 4.5, p. 613-617.
- CURTIS, Daniel R (2013). « The emergence of concentrated settlements in medieval western Europe : explanatory frameworks in the historiography ». In : *Canadian Journal of History* 48.2, p. 223-251.
- DARONG, Huang et Wang PENG (2012). « Grid-based DBSCAN algorithm with referential parameters ». In : *Physics Procedia* 24, p. 1166-1170.
- DE KEERSMAECKER, Marie-Laurence, Pierre FRANKHAUSER et Isabelle THOMAS (2003). « Using fractal dimensions for characterizing intra-urban diversity : The example of Brussels ». In : *Geographical analysis* 35.4, p. 310-328.
- (2004). « Dimensions fractales et réalités périurbaines. L'exemple du Sud de Bruxelles ». In : *L'Espace géographique* 33.3, p. 219-240.

- DEHOUE, Danièle (2014). « La notion de fractale en anthropologie ». In : *Ethnographiques.org* 29.
- DEMANGEON, Albert (1927). « La géographie de l'habitat rural ». In : *Annales de géographie*. T. 36. 199. JSTOR, p. 1-23.
- (1939). « Types de peuplement rural en France ». In : *Annales de Géographie*. T. 48. 271. JSTOR, p. 1-21.
- DEPRAZ, Samuel (2019). « Les gilets jaunes refont la géographie de la France ». In : *Libération*. Consultée le 03/05/2020.
- DESCOLA, Philippe (2005). *Par-delà nature et culture*. Éditions Gallimard.
- (2006). « La fabrique des images ». In : *Anthropologie et sociétés* 30.3, p. 167-182.
- DESROSIÈRES, Alain (2014). *Prouver et gouverner : une analyse politique des statistiques publiques*. La découverte.
- DESSALLES, Jean-Louis (2019). *Des intelligences très artificielles*. Odile Jacob.
- DIGGLE, Peter (1985). « A kernel method for smoothing point process data ». In : *Journal of the Royal Statistical Society : Series C (Applied Statistics)* 34.2, p. 138-147.
- DIGGLE, Peter J (2013). *Statistical analysis of spatial and spatio-temporal point patterns*. Chapman et Hall/CRC.
- DION, Roger (1934). *Essai sur la formation du paysage rural français*. Tours, France : Arrault.
- DONNAY, Jean-Paul, Mike J BARNSELY et Paul A LONGLEY (2000). *Remote sensing and urban analysis : GISDATA 9*. CRC Press.
- DOUILLARD, Arthur (2018). « Object Detection with Deep Learning on Aerial Imagery ». In : *Medium*.
- DUBEY, Abhimanyu et al. (août 2016). « Deep Learning the City : Quantifying Urban Perception At A Global Scale ». In : t. 9905. DOI : 10.1007/978-3-319-46448-0_12.
- DUBOIS, Jacques et Jean CHALINE (2006). *Le monde des Fractales : la géométrie cachée de la nature*. Ellipses.
- DUNCAN, Otis Dudley (1957). « The measurement of population distribution ». In : *Population Studies* 11.1, p. 27-45.
- DURKHEIM, Emile (1897). *Le suicide : étude de sociologie*. Alcan.
- ENCARNAÇÃO, Sara et al. (2012). « Fractal cartography of urban areas ». In : *Scientific Reports* 2, p. 527.
- ESTÈBE, Philippe (2016). *L'égalité des territoires, une passion française*. Presses universitaires de France.
- ESTER, Martin et al. (1996). « A density-based algorithm for discovering clusters in large spatial databases with noise. » In : *Kdd*. T. 96. 34, p. 226-231.
- FALCONER, Kenneth (2013). *Fractal geometry : mathematical foundations and applications*. Wiley.
- FEDER, Jens (2013). *Fractals*. Springer Science & Business Media.
- FLEURY, Vincent et Yun Sun LIMET (1997). « L'escroquerie Sokal-Bricmont ». In : *Libération*.
- FLORINDO, Joao Batista et al. (2012). « A comparative study on multiscale fractal dimension descriptors ». In : *Pattern Recognition Letters* 33.6, p. 798-806.
- FORRIEZ, Maxime et Laurent NOTTALE (2010). « Lois d'échelle et transitions fractal-non fractal en géographie ». In : *L'Espace géographique* 39.2, p. 97-112.
- FRANKHAUSER, Pierre (1994). « La fractalite des structures urbains. Pairs : Collection Villes ». In : *Anthropos*.

- (2015). « From fractal urban pattern analysis to fractal urban planning concepts ». In : *Computational approaches for urban environments*. Springer, p. 13-48.
- (2018). « Fractales, tissus urbains et aménagement – quelques précisions ». In : *L'Espace géographique*.
- FUSCO, Giovanni, Matteo CAGLIONI et Alessandro ARALDI (2017). « Street network morphology and retail locations. Application to the city of Nice (France) ». In : *Plurimondi* 17.
- GABRIEL, Edith et al. (2016). « Adapted kriging to predict the intensity of partially observed point process data ». In : *Spatial Statistics*.
- GAN, Li, Dong LI et Shunfeng SONG (2006). « Is the Zipf law spurious in explaining city-size distributions ? » In : *Economics Letters* 92.2, p. 256-262.
- GETIS, Arthur et Janet FRANKLIN (1987). « Second-order neighborhood analysis of mapped point patterns ». In : *Ecology* 68.3, p. 473-477.
- GRAPS, Amara (1995). « An introduction to wavelets ». In : *IEEE computational science and engineering* 2.2, p. 50-61.
- GRASSBERGER, Peter et Itamar PROCACCIA (1983). « Measuring the strangeness of strange attractors ». In : *Physica D : Nonlinear Phenomena* 9.1-2, p. 189-208.
- GREKOUSIS, George (2019). « Artificial neural networks and deep learning in urban geography : A systematic review and meta-analysis ». In : *Computers, Environment and Urban Systems* 74, p. 244-256.
- GROSSMANN, Alexander et Jean MORLET (1984). « Decomposition of Hardy functions into square integrable wavelets of constant shape ». In : *SIAM journal on mathematical analysis* 15.4, p. 723-736.
- GUO, Xifeng et al. (2017). « Deep clustering with convolutional autoencoders ». In : *International conference on neural information processing*. Springer, p. 373-382.
- GUY, Debord (1967). *La société du spectacle*. Buchet/Chastel.
- HAGGETT, Peter, Andrew David CLIFF et Allan FREY (1977). *Locational analysis in human geography*. T. 2. Arnold London.
- HARRIS, Paul et al. (2015). « Enhancements to a geographically weighted principal component analysis in the context of an application to an environmental data set ». In : *Geographical Analysis* 47.2, p. 146-172.
- HIMPENS, Stéphanie, François SÉMÉCURBE et Mathilde POULHES (2019). *Etude morphologique des parcelles cadastrales*. Les Rencontres de Théo Quant.
- HOANG, Lê Nguyễn et Rachid GUERRAOUI (2018). « Deep Learning Works in Practice. But Does it Work in Theory ? » In : *arXiv preprint arXiv :1801.10437*.
- HOANG), Science4all (Lé Nguyen (2018). *Deep learning et fractales*. Youtube. URL : <https://www.youtube.com/watch?v=61JoMM1uJ8E>.
- IVANOV, Plamen Ch et al. (1999). « Multifractality in human heartbeat dynamics ». In : *Nature* 399.6735, p. 461-465.
- JAFFARD, Stéphane (1989). « Exposants de Hölder en des points donnés et coefficients d'ondelettes ». In : *CR Acad. Sci. Paris* 308.1, p. 79-81.
- JAFFARD, Stéphane et al. (2019). « Multivariate multifractal analysis ». In : *Applied and Computational Harmonic Analysis* 46.3, p. 653-663.
- JANOWICZ, Krzysztof et al. (2020). *GeoAI : spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond*.
- JIANG, Zhi-Qiang et al. (2019). « Multifractal analysis of financial markets : a review ». In : *Reports on Progress in Physics* 82.12, p. 125901.

- JOSSELIN, Didier (1995). « La Déprise agricole en zone de montagne : vers un outil d'aide à la modélisation couplant systèmes d'induction et d'information géographique ». Thèse de doct. Grenoble 1.
- JULIA, Luc (2019). *L'intelligence artificielle n'existe pas*. First.
- KANEVSKI, Mikhail et Mário G PEREIRA (2017). « Local fractality : The case of forest fires in Portugal ». In : *Physica A : Statistical Mechanics and its Applications* 479, p. 400-410.
- KAPLAN, David H et Steven R HOLLOWAY (2001). « Scaling ethnic segregation : causal processes and contingent outcomes in Chinese residential patterns ». In : *GeoJournal* 53.1, p. 59-70.
- KARAMI, Amin et Ronnie JOHANSSON (2014). « Choosing DBSCAN parameters automatically using differential evolution ». In : *International Journal of Computer Applications* 91.7.
- KENNEDY, William J et James E GENTLE (2018). *Statistical computing*. Routledge.
- KESTENER, Pierre et al. (2010). « Characterizing complexity in solar magnetogram data using a wavelet-based segmentation method ». In : *The Astrophysical Journal* 717.2, p. 995.
- LAMBERT, Anne (2015). *" Tous propriétaires !". L'envers du décor pavillonnaire : L'envers du décor pavillonnaire*. Le Seuil.
- LAMBERT, E. (2018). *Le parc planétaire : la fabrication de l'environnement suburbain*. La Collection Critiques & cités. Éditions L'Oeil d'or.
- LANG, Gabriel et Eric MARCON (2013). « Testing randomness of spatial point patterns with the Ripley statistic ». In : *ESAIM : Probability and Statistics* 17, p. 767-788.
- LANG, Gabriel, Eric MARCON et Florence PUECH (2015). « Distance-based measures of spatial concentration : introducing a relative density function ». In : *The Annals of Regional Science*, p. 1-23.
- LATOUR, Bruno (1984). *Les Microbes : guerre et paix suivi de Irréductions*, La Découverte.
- LAVIGNE, Cédric (2003). « De nouveaux objets d'histoire agraire pour en finir avec le bocage et l'openfield ». In : *Études rurales* 167-168, p. 133-185.
- LAW, Stephen et Mateo Torres NEIRA (2019). « An unsupervised approach to geographical knowledge discovery using street level and street network images ». In : *GeoAI 2019*.
- LE BRAS, Hervé (1996). *Le peuplement de l'Europe*. La Documentation française.
- LE GOFF, Jacques (2014). *Faut-il vraiment découper l'histoire en tranches ?* Le Seuil.
- LE MÉE, René (1971). « Population agglomérée, population éparse au début du xix e siècle ». In : *Annales de Démographie historique*. JSTOR, p. 455-510.
- LECUN, Yann et al. (1998). « Gradient-based learning applied to document recognition ». In : *Proceedings of the IEEE* 86.11, p. 2278-2324.
- LEDFORD, Heidi (2019). « Millions of black people affected by racial bias in health-care algorithms. » In : *Nature* 574.7780, p. 608.
- LEFÈVRE, Marguerite A (1926). *L'habitat rural en Belgique*. H. Vaillant-Carmanne.
- LEPRINCE-RINGUET, Daphne (2020). « For self-driving cars, winter is coming ». In : *ZDnet*.
- LESNE, Annick et Michel LAGUÈS (2015). *Invariance d'échelle. Des changements d'états à la turbulence*. Belin.
- LIU, Qiliang et al. (2015). « Modeling the effect of scale on clustering of spatial points ». In : *Computers, Environment and Urban Systems* 52, p. 81-92.
- LOONIS, Vincent et Marie-Pierre DE BELLEFON (2018). *Manuel d'analyse spatiale - Théorie et mise en oeuvre pratique avec R*.

- LOUF, Rémi et Marc BARTHELEMY (2014). « Scaling : lost in the smog ». In : *Environment and Planning B : Planning and Design* 41.5, p. 767-769.
- LUSSAULT, Michel (2018). *À propos de la justice spatiale. Un concept utile contre les discriminations ?* Rapp. tech. Conseil national d'évaluation du système scolaire.
- MANDELBROT, Benoit et Richard L HUDSON (2007). *The Misbehavior of Markets : A fractal view of financial turbulence*. Basic books.
- MANDELBROT, Benoit B (1975). « Les Objets Fractals : Forme ». In : *Hasard et Dimension, Flammarion, Paris*.
- MARCON, Eric et Florence PUECH (2017). « A typology of distance-based measures of spatial concentration ». In : *Regional Science and Urban Economics* 62, p. 56-67.
- MARR, David et Ellen HILDRETH (1980). « Theory of edge detection ». In : *Proceedings of the Royal Society of London. Series B. Biological Sciences* 207.1167, p. 187-217.
- MATHIAN, Hélène et Lena SANDERS (2004). « Statistique spatiale ». In : *Hypergeo*. Consultée le 03/05/2020.
- MAUDUIT, V. et al. (2020). « DCGAN for the synthesis of multivariate multifractal textures : How do we know it works ? » In : *ENS de Lyon*.
- MEDVEDKOV, Yu V (1967). « The concept of entropy in settlement pattern analysis ». In : *Papers of the Regional Science Association*. T. 18. 1. Springer, p. 165-168.
- MOOSAVI, Vahid (2017). « Urban morphology meets deep learning : Exploring urban forms in one million cities, town and villages across the planet ». In : *arXiv preprint arXiv :1709.02939*.
- MURCIO, Roberto et al. (2015). « Multifractal to monofractal evolution of the London street network ». In : *Physical Review E* 92.6, p. 062130.
- MUZY, J. -F., D. SORNETTE et al. (2000). *Multifractal returns and Hierarchical Portfolio Theory*. arXiv : cond-mat/0008069 [cond-mat.stat-mech].
- MUZY, Jean-François, Emmanuel BACRY et Alain ARNEODO (1994). « The multifractal formalism revisited with wavelets ». In : *International Journal of Bifurcation and Chaos* 4.02, p. 245-302.
- NORDBECK, Stig (1971). « Urban allometric growth ». In : *Geografiska Annaler : Series B, Human Geography* 53.1, p. 54-67.
- O'KEEFFE, Tadhg (2000). « Reflections on the Dispersed-Nucleated Paradigm in Medieval Settlement Archaeology ». In : *Ruralia III. Pamatky Archeologicke* 104.
- OPENSHAW, Stan (1984a). « Ecological fallacies and the analysis of areal census data ». In : *Environment and planning A* 16.1, p. 17-31.
- (1984b). « The Modifiable Areal Unit problem. CATMOG 38. Norwich ». In : *Geo-Abstracts*.
- OPENSHAW, Stan et Christine OPENSHAW (1997). *Artificial Intelligence in Geography*. 1st. USA : John Wiley Sons, Inc.
- ORAIN, Olivier (2004). *La géographie française face à la notion d'échelle*.
- PALMER, Michael W (1988). « Fractal geometry : a tool for describing spatial patterns of plant communities ». In : *Vegetatio* 75.1-2, p. 91-102.
- PAVON-DOMINGUEZ, P et al. (2017). « Fractal and multifractal characterization of the scaling geometry of an urban bus-transport network. » In : *Computers, Environment and Urban Systems* 64, p. 229-238.
- PIERRE, Bourdieu (2001). *Science de la science et réflexivité*. Raisons d'Agir.
- PINÇON, Michel et Monique PINÇON-CHARLOT (1994). « De l'espace social à l'espace urbain. Utilité d'une métaphore ». In : *Les annales de la recherche urbaine*. T. 64. 1. Centre de Recherche d'Urbanisme, p. 51-53.

- POLONY, Natacha (2014). *Ce pays qu'on abat*. Place des éditeurs.
- POOLE, Ben et al. (2016). « Exponential expressivity in deep neural networks through transient chaos ». In : *Advances in neural information processing systems*, p. 3360-3368.
- PUMAIN, Denise (2007). « Lois d'échelle et mesure des inégalités en géographie ». In : *Revue européenne des sciences sociales. European Journal of Social Sciences XLV-138*, p. 55-65.
- (2010). « Commentaire : Transition fractal-non fractal et objets géographiques ». In : *L'Espace géographique* 39.2, p. 113-116.
- RAGHAVENDRA, BS et D Narayana DUTT (2010). « Signal characterization using fractal dimension ». In : *Fractals* 18.03, p. 287-292.
- RAGHU, Maithra et al. (2017). « On the expressive power of deep neural networks ». In : *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, p. 2847-2854.
- RAMADE, F. (2007). *Ode pavillonnaire*. Filigranes.
- RATZEL, Friedrich (1891). *Anthropogeographie : Die geographische Verbeitung des Menschen*. T. 2. J. Engelhorn.
- RAZEMON, Olivier (2017). *Comment la France a tué ses villes*. Diagonales. Rue de l'échiquier.
- REARDON, Sean F, Stephen A MATTHEWS et al. (2008). « The geographic scale of metropolitan racial segregation ». In : *Demography* 45.3, p. 489-514.
- REARDON, Sean F et David O'SULLIVAN (2004). « Measures of spatial segregation ». In : *Sociological methodology* 34.1, p. 121-162.
- RIPLEY, Brian D (1977). « Modelling spatial patterns ». In : *Journal of the Royal Statistical Society : Series B (Methodological)* 39.2, p. 172-192.
- (1997). « Can Statistical Theory Help Us Use Neural Networks Better? » In : *Interface* 97.
- ROSENBLAT, F (1958). « The perceptron : A probabilistic model for information storage and organization in the brain ». In : *Psychological Review* 65.6, p. 386-408.
- ROTHSTEIN, R. (2017). *The Color of Law : A Forgotten History of How Our Government Segregated America*. Liveright.
- RUMELHART, David E, Geoffrey E HINTON et Ronald J WILLIAMS (1985). *Learning internal representations by error propagation*. Rapp. tech. California Univ San Diego La Jolla Inst for Cognitive Science.
- S. FOWLER, Christopher (2016). « Segregation as a multiscalar phenomenon and its implications for neighborhood-scale research : The case of South Seattle 1990–2010 ». In : *Urban geography* 37.1, p. 1-25.
- SALAT, Hadrien et al. (2018). « Uncovering inequality through multifractality of land prices : 1912 and contemporary Kyoto ». In : *PloS one* 13.4, e0196737.
- SAMUEL, Arthur L (1959). « Some studies in machine learning using the game of checkers ». In : *IBM Journal of research and development* 3.3, p. 210-229.
- SANDERS, Lena (2011). *Géographie quantitative et analyse spatiale : quelles formes de scientificités ?*
- SAPOVAL, Bernard (1997). *Universalités et fractales : jeux d'enfant ou délits d'initié ?* Flammarion.
- SÉMÉCURBE, François, Cécile TANNIER et Stéphane G ROUX (2016). « Spatial Distribution of Human Population in France : Exploring the Modifiable Areal Unit Problem Using Multifractal Analysis ». In : *Geographical Analysis* 48.3, p. 292-313.

- (2019). « Exploring the deviations from scale-invariance of spatial distributions of buildings using a Geographically Weighted Fractal Analysis. An application to twenty French middle-size metropolitan areas ». In :
- SHANNON, Claude E (1948). « A mathematical theory of communication ». In : *The Bell system technical journal* 27.3, p. 379-423.
- SILVERMAN, Bernard W (1978). « Weak and strong uniform consistency of the kernel estimate of a density and its derivatives ». In : *The Annals of Statistics*, p. 177-184.
- SOKAL, Alan et Jean BRICMONT (1997). *Impostures intellectuelles*. Odile Jacob.
- STERLING, Theodore D (1959). « Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa ». In : *Journal of the American statistical association* 54.285, p. 30-34.
- TANNIER, Cécile (2017). « Analyse et simulation de la concentration et de la dispersion des implantations humaines de l'échelle micro-locale à l'échelle régionale-Modèles multi-échelles et trans-échelles ». Habilitation à diriger des recherches.
- (2018). « A propos des modèles fractals en géographie urbaine et en aménagement : réfutation de l'esthétique et de la norme universelle ». In : *Cybergeog*.
- (2020). *Effets de la distance et dépendance d'échelles dans les modèles géographiques de villes et de territoires*.
- TANNIER, Cécile, Marc BOURGEOIS et Maxime FRÉMOND (2017). *Simulation de scénarios fractals d'urbanisation et des mobilités quotidiennes résultantes*.
- TANNIER, Cécile et Denise PUMAIN (2005). « Fractals in urban geography : a theoretical outline and an empirical example ». In : *Cybergeog : European Journal of Geography*.
- TANNIER, Cécile, Isabelle THOMAS et al. (2011). « A Fractal Approach to Identifying Urban Boundaries. » In : *Geographical Analysis* 43.2, p. 211-227.
- TANNIER, Cécile, Gilles VUIDEL et al. (2010). « Simulation fractale d'urbanisation. MUP-city, un modèle multi-échelle pour localiser de nouvelles implantations résidentielles ». In :
- TÉL, Tamás, Ágnes FÜLÖP et Tamás VICSEK (1989). « Determination of fractal dimensions for geometrical multifractals ». In : *Physica A : Statistical Mechanics and its Applications* 159.2, p. 155-166.
- THOMAS, Isabelle et al. (2010). « Clustering patterns of urban built-up areas with curves of fractal scaling behaviour ». In : *Environment and Planning B : Planning and Design* 37.5, p. 942-954.
- VAN ANDEL, Tjeerd H, Curtis N RUNNELS et Kevin O POPE (1986). « Five thousands years of land use and abuse in the southern Argolid, Greece ». In : *Hesperia : The Journal of the American School of Classical Studies at Athens* 55.1, p. 103-128.
- VICSEK, Tamás (1990). « Mass multifractals ». In : *Physica A : Statistical Mechanics and its Applications* 168.1, p. 490-497.
- (1992). *Fractal growth phenomena*. World scientific.
- VIRILIO, Paul (1984). *L'espace critique : essai*. Christian Bourgois éditeur.
- WALTER, Christian (2008). « La représentation brownienne du risque : une faute morale collective ? » In : *Finance Bien Commun* 2, p. 137-144.
- WARD, Lawrence M et Priscilla E GREENWOOD (2007). « 1/f noise ». In : *Scholarpedia* 2.12, p. 1537.
- WATTEAUX, Magali (2005). « Sous le bocage, le parcellaire... » In : *Études rurales* 175-176, p. 53-80.
- (2012). « Le bocage. Un paysage rural à la lumière des études archéologiques et archéogéographiques ». In : *Archéopages. Archéologie et société* 34, p. 64-72.

- WENDT, Herwig (2008). « Contributions of Wavelet Leaders and Bootstrap to Multifractal Analysis : Images, Estimation Performance, Dependence Structure and Vanishing Moments. Confidence Intervals and Hypothesis Tests. » Thèse de doct.
- WEST, Geoffrey (2017). « Scale : The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life ». In : *Organisms, Cities, Economies, and Companies*.
- ZHANG, Zhonghao et al. (2014). « Spatial point pattern analysis of human settlements and geographical associations in eastern coastal China—A case study ». In : *International journal of environmental research and public health* 11.3, p. 2818-2833.
- ZOU, Zilong et al. (2016). « Trabecular bone texture classification using wavelet leaders ». In : *Medical Imaging 2016 : Biomedical Applications in Molecular, Structural, and Functional Imaging*. T. 9788. International Society for Optics et Photonics, 97880E.