HAL open science

Spatial distribution of human population in France: exploring the modifiable areal unit problem using multifractal analysis

François Sémécurbe, Cécile Tannier, Stephane G. Roux

▶ To cite this version:

François Sémécurbe, Cécile Tannier, Stephane G. Roux. Spatial distribution of human population in France: exploring the modifiable areal unit problem using multifractal analysis. Geographical Analysis, 2016, 48 (3), pp.292-313. 10.1111/gean.12099 . hal-01272310

HAL Id: hal-01272310

https://hal.science/hal-01272310v1

Submitted on 28 Jun 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.

SÉMÉCURBE F., TANNIER C., ROUX S. G. (2016), Spatial distribution of human population in France: exploring the Modifiable Areal Unit Problem using multifractal analysis, *Geographical Analysis*, vol. 48, n°3, pp. 292–313.

Spatial distribution of human population in France: exploring the MAUP using multifractal analysis

François Sémécurbe
ThéMA UMR 6049, CNRS-Univ. Bourgogne Franche-Comté
32, rue Mégevand
F-25 030 Besançon Cedex, France
françois.semecurbe@univ-fcomte.fr

Cécile Tannier (corresponding author)
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
cecile.tannier@univ-fcomte.fr

Stéphane Gaëtan Roux
Laboratoire de Physique UMR 5672, CNRS-Ecole Nationale Supérieure de Lyon
46, allée d'Italie
F-69364 Lyon Cedex 07, France
stephane.roux@ens-lyon.fr

Abstract

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: 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 paper, 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.

Spatial distribution of human population in France: exploring the MAUP using multifractal analysis

September 8, 2015

1 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 km^2. In the Île de France region (around Paris), the population density is 1,000 inhabitants per km^2 whereas population density in the Limousin region is 44 inhabitants per km^2. 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 km^2) whereas the subregion Creuse in the Limousin region is almost deserted (22 inhabitants per km^2).

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); Benguigui et al. (2006); Feng and Chen (2010); Shen (2002); Thomas et al. (2008)). 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 et al. (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:

$$\frac{\rho_i}{\rho_{i+1}} = r^{-2-D} \tag{1}$$

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 et al. (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 (Chen and Zhou (2004); Haag (1994)). 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 USA and Great Britain. He considered the multifractal organization of each settlement pattern to be homogeneous within each country. That strong assumption was contested by Vega Orozco et al. (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 paper, 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 paper is organized as follows. Section 2 sets out the mathematical bases of multifractal analysis. Section 3 describes how it applies to geographical analyses. Section 4 reports the analysis results of the spatial distribution of population density in France. Section 5 discusses the results in the light of the centuries-old history of French rural settlements and recent changes in periurban areas. Section 6 concludes the paper.

2 Multifractal analysis: mathematical bases

Multifractal analysis is about studying local variations of a measure (Jaffard et al. (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:

$$\mu(F) = \text{Population contained in F} \tag{2}$$

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 section 2.1). 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 section 2.2). 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 section 2.3).

NB: *although the mathematical definitions and examples are presented in* $\mathbb{R}^2$*, they remain valid and can be generalized for any vector space* $\mathbb{R}^n$*.*

2.1 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. 1-dimensional shapes) if the line is contained in $\mathbb{R}^2$ but not in $\mathbb{R}$. Similarily, it is impossible to measure the area of a surface by covering it with circles (i.e. 2-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.

2.1.1 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 partioning 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:

$$[o_x + m_x\delta, o_x + (m_x + 1)\delta] \times [o_y + m_y\delta, o_y + (m_y + 1)\delta] \tag{3}$$

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:

$$dim_B(F) = \lim_{\delta \to 0} \frac{log(N_\delta(F))}{-log(\delta)} \tag{4}$$

where $N_\delta(F)$ is the number of δ-grid cells that intersect F.

2.1.2 Hausdorff dimension

The problem with the box-counting dimension is that $\lim_{\delta \to 0}$ does not always exist for any set of points F. The Hausdorff dimension can overcome this problem by 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:

$$\mathcal{H}^d(S) = \lim_{\delta \to 0} \inf\{\sum_{i=1}^{+\infty} |U_i|^d : \{U_i\} \text{ is a } \delta\text{-cover of } F \tag{5}$$

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 non-zero:

$$dim_{\mathcal{H}}(F) = \inf\{d : \mathcal{H}^d(F) = 0\} \tag{6}$$

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.

2.2 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.

$$\alpha(x) = \lim_{\delta \to 0} \frac{\mu(B(x,r))}{-log(\delta)} \tag{7}$$

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 $density(x,r)$:

$$density(x,r) = \frac{\mu(B(x,r))}{\pi r^2} \simeq \frac{r^{\alpha(x)}}{\pi r^2} \simeq r^{\alpha(x)-2} \tag{8}$$

$\alpha(x) < 2$ In this case, $density(x,r)$ tends to infinity when r tends to 0. As $\alpha(x)$ comes close to 0, $density(x,r)$ 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 behaviour of $density(x,r)$ can not be determined. As r tends to 0, $density(x,r)$ tends either toward a non-zero 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 $F_\alpha = \{x : \alpha(x) = \alpha\}$:

$$f(\alpha) = dim_{\mathcal{H}}(F_\alpha) \tag{9}$$

$f(\alpha(0))$ 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 non-existent in the case of a monofractal

distribution) (Wendt et al. (2007)).

The singularity spectrum $f(\alpha)$ complies with the following property:

$$0 \leq f(\alpha) \leq \alpha \tag{10}$$

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.

Since 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 et al. (2007)). It enables the determination of the singularity spectrum through a Legendre transform of the scaling exponent τ_q:

$$f(\alpha) = \inf_{q \in R} (q\alpha - \tau_q) \tag{11}$$

2.3 Multifractal quantification of the Modifiable Areal Unit Problem

2.3.1 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 upon 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:

$$S_{q,\delta}(\mu) = \sum_{\mu(G_{\delta,i})>0} \mu(G_{\delta,i})^q \tag{12}$$

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:

$$S_{q,\delta}(\mu) \approx \delta^{\tau_q} \tag{13}$$

Here τ_q is a scaling exponent. It can be calculated using the following formula:

$$\tau_q = \lim_{\delta \to 0} \frac{log(S_{q,\delta}(\mu))}{-log(\delta)} \tag{14}$$

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 $log(\delta)$.

2.3.2 Generalized dimensions D_q

On the basis of τ_q, it is possible to define the generalized dimensions D_q:

$$\begin{cases} D_q = \frac{1}{1-q}\tau_q \text{ pour } q \neq 1 \\ D_1 = \lim_{\delta \to 0} \sum_{\mu(G_{\delta,i})>0} \mu(G_{\delta,i}) log(\mu(G_{\delta,i})) \end{cases} \tag{15}$$

D_q with respect to q is called the dimension spectrum. The coefficient q acts as a filter when calculating the generalized dimensions. For $q \to +\infty$, only the cells in which high measures are concentrated are taken into account. Conversely, for $q \to -\infty$ 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 multi-viewpoint (via the coefficient q).

2.3.3 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:

$$\alpha(q) = \frac{d}{dq}(q-1)D_q \tag{16}$$

$$f(\alpha(q)) = \alpha q - (q-1)D_q \tag{17}$$

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, multi-family 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 (Chevillard 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 particulary for image classification. In this paper, we use the singularity spectrum in the same way in order to obtain a typology map of the spatial distribution of human population in France (see section 4.2).

3 Multifractal analysis: practical application

3.1 Data

For this study we used the 200m-grid population 2010 dataset 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 $200m$. One advantage of this dataset 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 $200m$ 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 200m-grid population dataset 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".*

3.2 Estimation of multifractal indexes

3.2.1 Estimation of the generalized dimensions D_q

In practice, the generalized dimensions D_q are defined as:

$$D_q = \lim_{\delta \to 0} \frac{\frac{1}{1-q} log(\sum_{i=1}^{n} \mu(G_{\delta,i})^q)}{-log(\delta)} \tag{18}$$

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.

3.2.2 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):

$$\alpha(q) = \lim_{\delta \to 0} \frac{\sum_{i=1}^{n} P_q(\delta, i) log(\mu(G_{\delta,i}))}{-log(\delta)} \tag{19}$$

and

$$f(\alpha(q)) = \lim_{\delta \to 0} \frac{\sum_{i=1}^{n} P_q(\delta, i) log(P_q(\delta, i))}{-log(\delta)} \tag{20}$$

with

$$P_q(\delta, i) = \frac{\mu(G_{\delta,i})^q}{\sum_{i=1}^{n} \mu(G_{\delta,i})^q} \tag{21}$$

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.

3.3 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 (Figures 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.

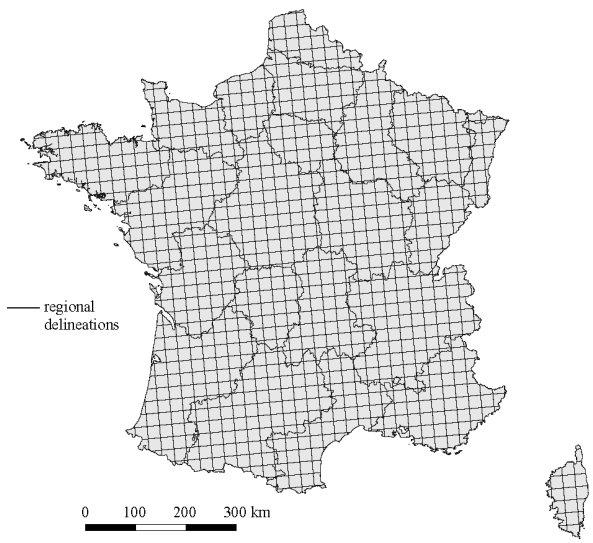


Figure 1: 992 square spatial units regularly dividing the area of mainland France

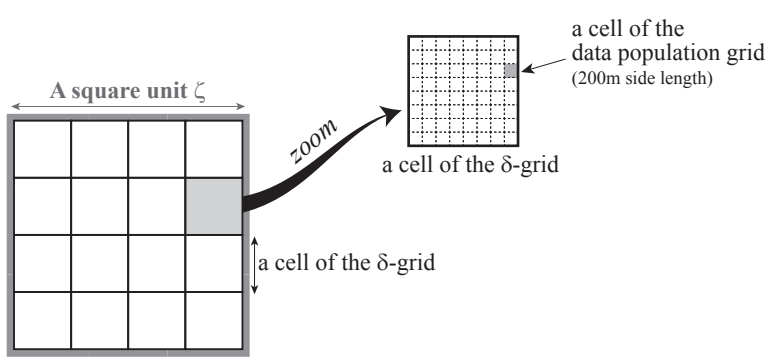


Figure 2: A square spatial unit ζ, δ-grid cells, and 200m-grid population cells

3.3.1 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 $800m$), 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 carried out 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 $6400m$), 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 $800m$ and $6400m$) the multifractal analysis of grid population data interestingly informs us about the effect of urban sprawl on regional settlement patterns.

3.3.2 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 $6400m$, 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 $25km$ side length (Figure 1). The spatial units ζ are therefore almost four times larger than the larger δ-grid cells when the resolution δ is $6400m$.

With a side length of $25km$, 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 $4900km^2$, which corresponds approximately to a diameter of $79km$. This partitioning of metropolitan areas enables us to distinguish the urban centers from their periphery. Moreover, it is possible to 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 et al. (2015), this intermediate scale range is suitable for studying the influence of societies on space.

3.3.3 Choice of the appropriate δ-grid position for calculating entropies for each $q-$dimension

Each grid cell of a δ-grid results from the aggregation by longitude and latitude of k cells of the grid population dataset. 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 $200m$ by $200m$.

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.

3.3.4 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 $\{0, 0.5, 1, 1.5, 2\}$.

4 Results

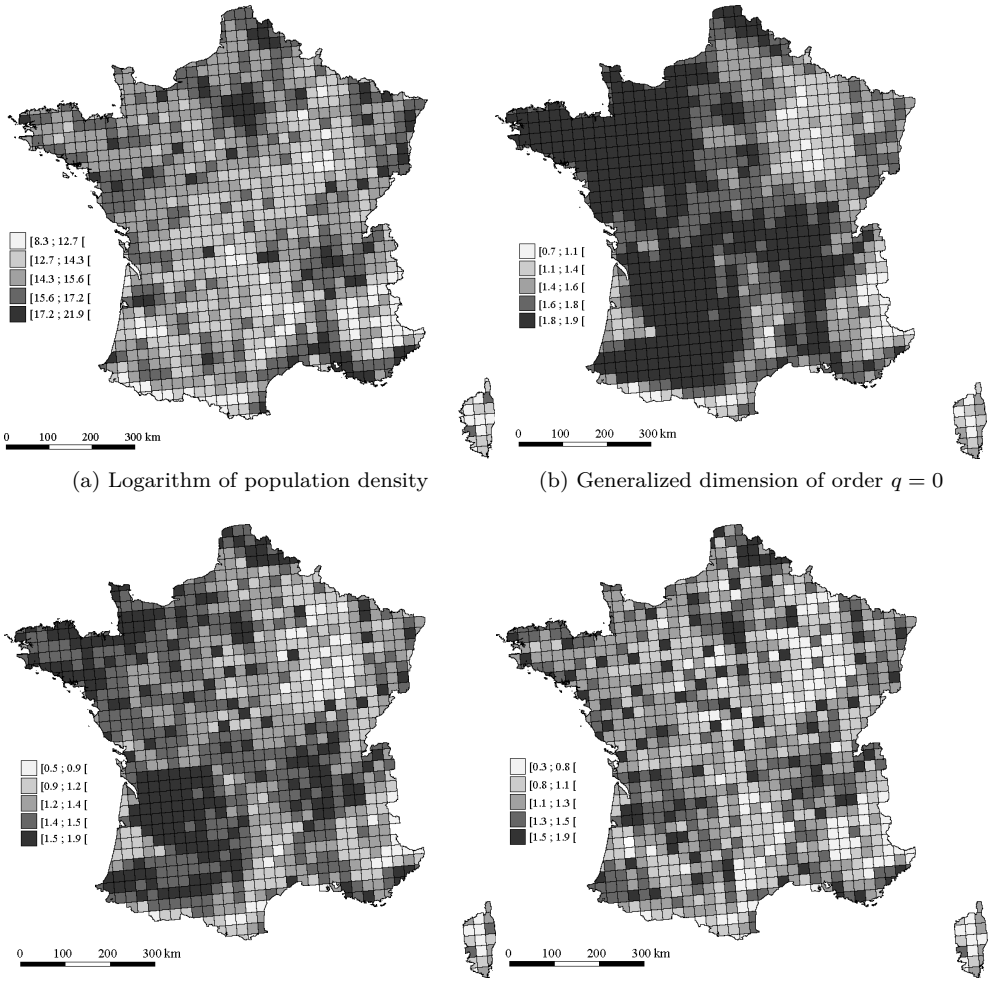
When estimating multifractal indexes, we need to verify the quality of the estimation. As shown in Table 1, R^2 values calculated are satisfactory: more than 90% are higher than 0.96. Only the quality of the estimation of $\alpha_2, f(\alpha_2)$ is slightly lower.

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 1: Distribution of R^2 values obtained by estimating the multifractal indexes

4.1 Maps of generalized dimensions D_q

Figure 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 section 3.1, the quality of data used does not vary spatially. Hence we assume that the revealed spatial structures do not result from cal-



(a) Logarithm of population density

(b) Generalized dimension of order $q = 0$

(c) Generalized dimension of order $q = 1$

(d) Generalized dimension of order $q = 2$

Figure 3: Population densities and corresponding multifractal generalized dimensions in France. *Source: 200m-grid population 2010 dataset provided by the French National Institute of Statistics and Economic Studies (INSEE)*

culation 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 Figure 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.

In contrast, dense urban metropolitan areas (see Figure 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 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, non-urban 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.

Maps of $q-$dimensions yield different and non-redundant 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 non-urban 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 et al. (2007)). By contrast, generalized dimensions measure the variations of population density across scales.

4.2 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 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 $\delta-$neighborhood.

A $\delta-$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):

$$A_\delta = \{x \in \mathbb{R}^2 : |x - a| \neq \delta \text{ for some } a \in A\} \tag{22}$$

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 $\delta-$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. Figure 6 displays the dendrogram. Classes are defined by cutting branches off 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 (non-urban 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 (Figure 7).

$$\alpha_q(Class) = \frac{\sum_{i \in Class}(\alpha_q(i))}{|Class|} \tag{23}$$

$$f(\alpha_q(Class)) = \frac{\sum_{i \in Class}(f(\alpha_q(i)))}{|Class|} \tag{24}$$

Logically, the map of the multifractal spectra typology (Figure 8) synthesizes the three maps of generalized dimensions displayed in Figure 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 non-urban 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 Figure 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. Figure 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.

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 (Figure 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.

5 Discussion

5.1 Traces of century-old agrarian systems

In the early twentieth century, renowned French geographers and historians analyzed the geography of rural areas (Bloch (1931)), (Dion (1934)), (Demangeon (1927)). 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;
- 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 Figure 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 non-urban 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 agrairian 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.

5.2 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 eveywhere 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 multifrac-

tal 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.

6 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 non-existent. 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.

References

Appleby, S. (1996), 'Multifractal characterization of the distribution pattern of the human population', *Geographical Analysis* **28**(2), 147–160.

Ariza-Villaverde, A. B., Jiménez-Hornero, F. J. and Ravé, E. G. D. (2013), 'Multifractal analysis of axial maps applied to the study of urban morphology', *Computers, Environment and Urban Systems* **38**, 1–10.

Batty, M. and Longley, P. A. (1994), *Fractal cities: a geometry of form and function*, Academic Press.

Batty, M. and Xie, Y. (1996), 'Preliminary evidence for a theory of the fractal city', *Environment and Planning A* **28**, 1745–1762.

Batty, M. and Xie, Y. (1999), 'Self-organized criticality and urban development', *Discrete Dynamics in Nature and Society* **3**(2-3), 109–124.

Benguigui, L., Blumenfeld-Lieberthal, E. and Czamanski, D. (2006), 'The dynamics of the tel aviv morphology', *Environment and planning B* **33**(2), 269.

Benguigui, L., Czamanski, D., Marinov, M. and Portugali, Y. (2000), 'When and where is a city fractal?', *Environment and Planning B: Planning and Design* **27**(4), 507–519.

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 Wang, J. (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 Zhou, Y. (2004), 'Multi-fractal measures of city-size distributions based on the three-parameter zipf model', *Chaos, Solitons & Fractals* **22**(4), 793–805.

Chevillard, L., Castaing, B., Arneodo, A., Lévêque, E., Pinton, J.-F. and Roux, S. G. (2012), 'A phenomenological theory of eulerian and lagrangian velocity fluctuations in turbulent flows', *Comptes Rendus Physique* **13**(9), 899–928.

Chhabra, A. and Jensen, R. (1989), 'Direct determination of the f(α) singularity spectrum', *Phys. Rev. Lett.* **62**, 1327–1330.

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*, Arrault.

Falconer, K. (2013), *Fractal Geometry: Mathematical Foundations and Applications*, Wiley.

Feng, J. and Chen, Y. (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), 838–856.

Frankhauser, P. (1994), *La fractalité des structures urbaines*, Anthropos.

Frankhauser, P. (1998), 'The fractal approach. a new tool for the spatial analysis of urban agglomerations', *Population* **10**(1), 205–240.

Frankhauser, P. (2008), Fractal geometry for measuring and modelling urban patterns, *in* 'The dynamics of complex urban systems', Springer, pp. 213–243.

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 & Fractals* **4**(4), 519–534.

Jaffard, S., Lashermes, B. and Abry, P. (2007), Wavelet leaders in multifractal analysis, *in* T. Qian, M. Vai and Y. Xu, eds, 'Wavelet Analysis and Applications', Applied and Numerical Harmonic Analysis, Birkhäuser Basel, pp. 201–246.

Lam, N. (1990), 'Description and measurement of landsat tm images using fractals', *Photogrammetric engineering and remote sensing* **56**(2), 187–195.

Lam, N. S.-N. and Quattrochi, D. A. (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**, 133–186.

Openshaw, S. (1983), 'The modifiable areal unit problem', *Geo Books* **38**.

Orozco, C. D. V., Golay, J. and Kanevski, M. (2015), 'Multifractal portrayal of the swiss population', *Cybergeo: European Journal of Geography* (714). **URL:** *http://cybergeo.revues.org/26829*

Seuront, L. (2009), *Fractals and Multifractals in Ecology and Aquatic Science*, CRC Press.

Shen, G. (2002), 'Fractal dimension and fractal growth of urbanized areas', *International Journal of Geographical Information Science* **16**(5), 419–437.

Stanley, H., Amaral, L. N., Goldberger, A., Havlin, S., Ivanov, P. C. and Peng, C.-K. (1999), 'Statistical physics and physiology: monofractal and multifractal approaches', *Physica A: Statistical Mechanics and its Applications* **270**(1), 309–324.

Tannier, C. and Pumain, D. (2005), 'Fractals in Urban Geography: a theoretical outline and an empirical example', *Cybergeo* (307), http://cybergeo.revues.org/3275.

Tannier, C. and Thomas, I. (2013), 'Defining and characterizing urban boundaries: A fractal analysis of theoretical cities and belgian cities', *Computers, Environment and Urban Systems* **41**(0), 234 – 248.

Tannier, C., Vuidel, G., Houot, H. and Frankhauser, P. (2012), 'Spatial accessibility to amenities in fractal and non fractal urban patterns', *Environment and Planning B: Planning and Design* **39**(5), 801–819.

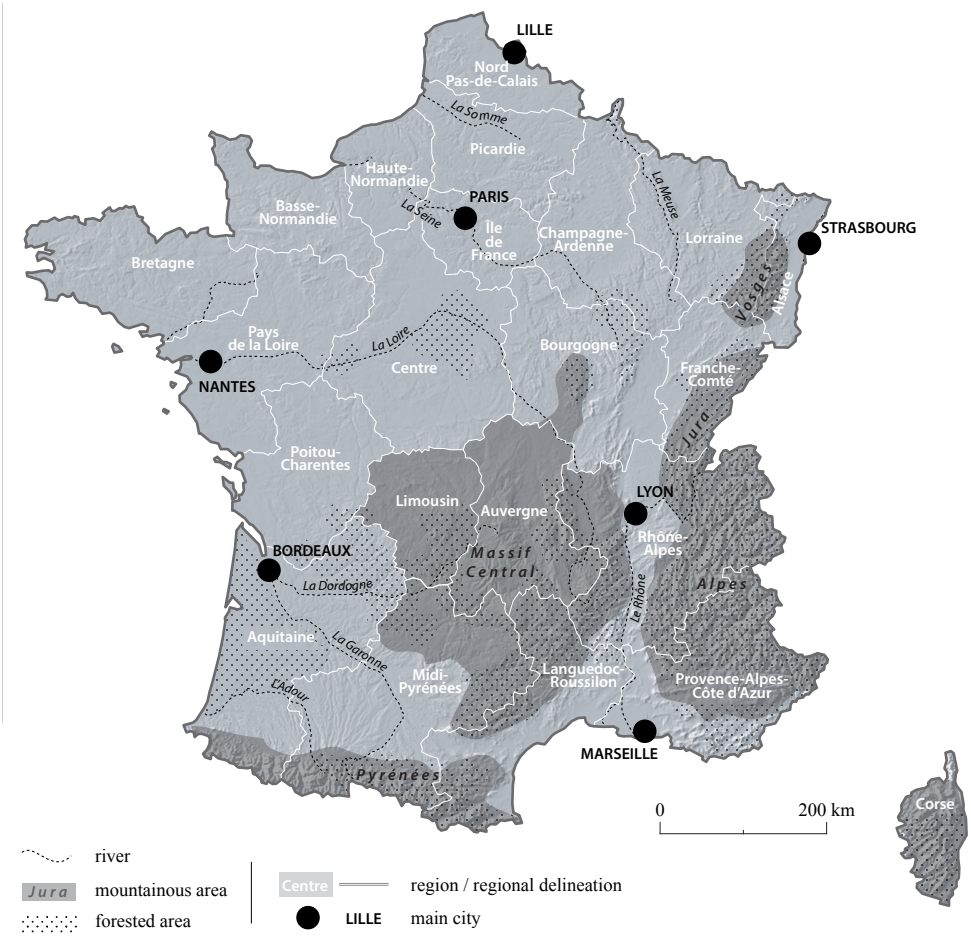
Thomas, I., Frankhauser, P. and Biernacki, C. (2008), 'The morphology of built-up landscapes in wallonia (belgium): A classification using fractal indices', *Landscape and Urban Planning* **84**(2), 99–115.

Thomas, I., Frankhauser, P. and De Keersmaecker, M.-L. (2007), 'Fractal dimension versus density of built-up surfaces in the periphery of brussels', *Papers in Regional Science* **86**(2), 287–308.

Wendt, H., Abry, P., Jaffard, S., Ji, H. and Shen, Z. (2009), Wavelet leader multifractal analysis for texture classification, *in* 'Image Processing (ICIP), 2009 16th IEEE International Conference on', IEEE, pp. 3829–3832.

Wendt, H., Abry, P., Roux, Stéphane, G. and Jaffard, S. (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 Engelen, G. (1994), 'Urban systems dynamics and cellular automata: fractal structures between order and chaos', *Chaos, Solitons & Fractals* **4**(4), 563–583.



Sources: BD ALTI® 1000m, GEOFLA®, SCAN1000®, IGN ; BDCARTHAGE®, Sandre

Figure 4: Mountains, forested areas, rivers, regions, and main cities in mainland France

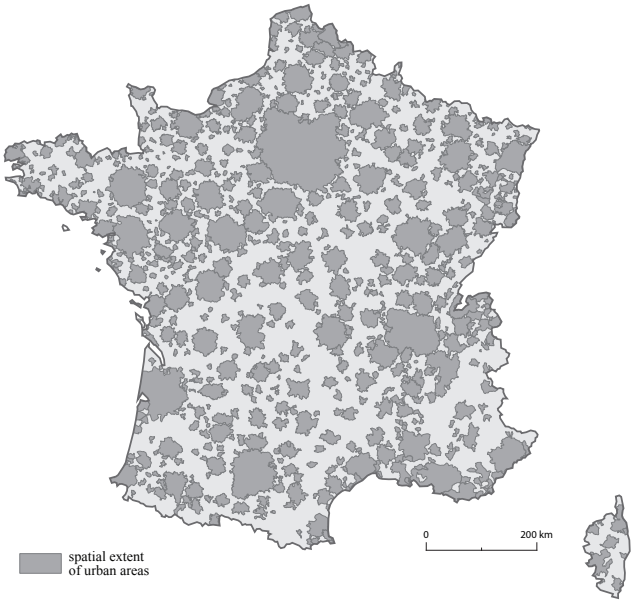


Figure 5: Delineation of French *aires urbaines* (metropolitan areas). In France, an *aire urbaine* encompasses a densely built *unité urbaine* -which is akin to a U.S. urban area- and its commuter belt. *Source: French National Institute of Statistics and Economic Studies (INSEE) 2010*

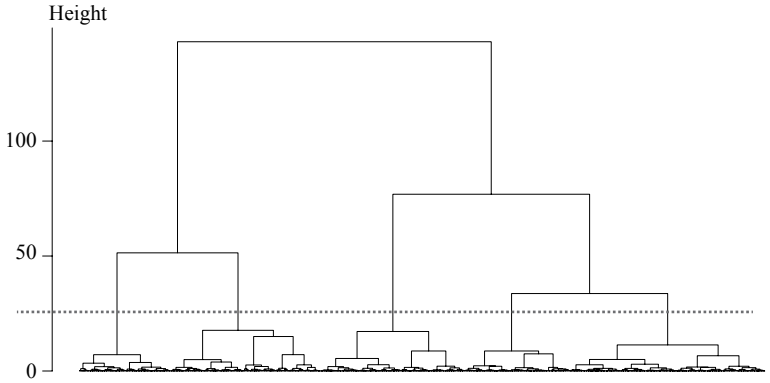


Figure 6: Dendrogram resulting from the classification of spatial units according to their multifractal spectrum

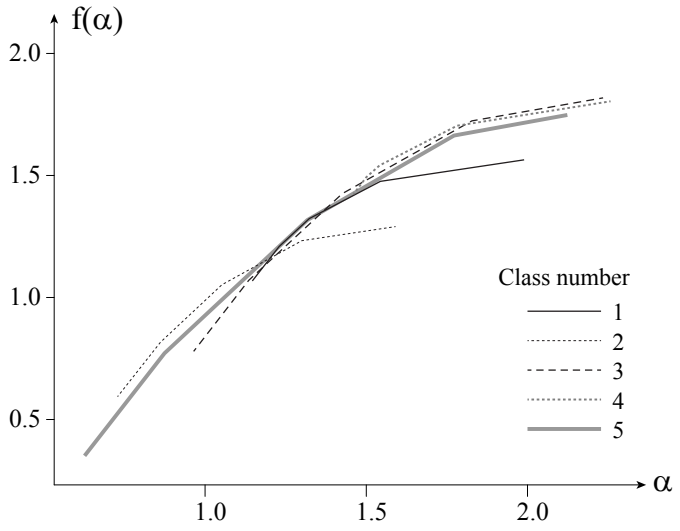


Figure 7: Average multifractal spectrum characterizing each class of the HAC

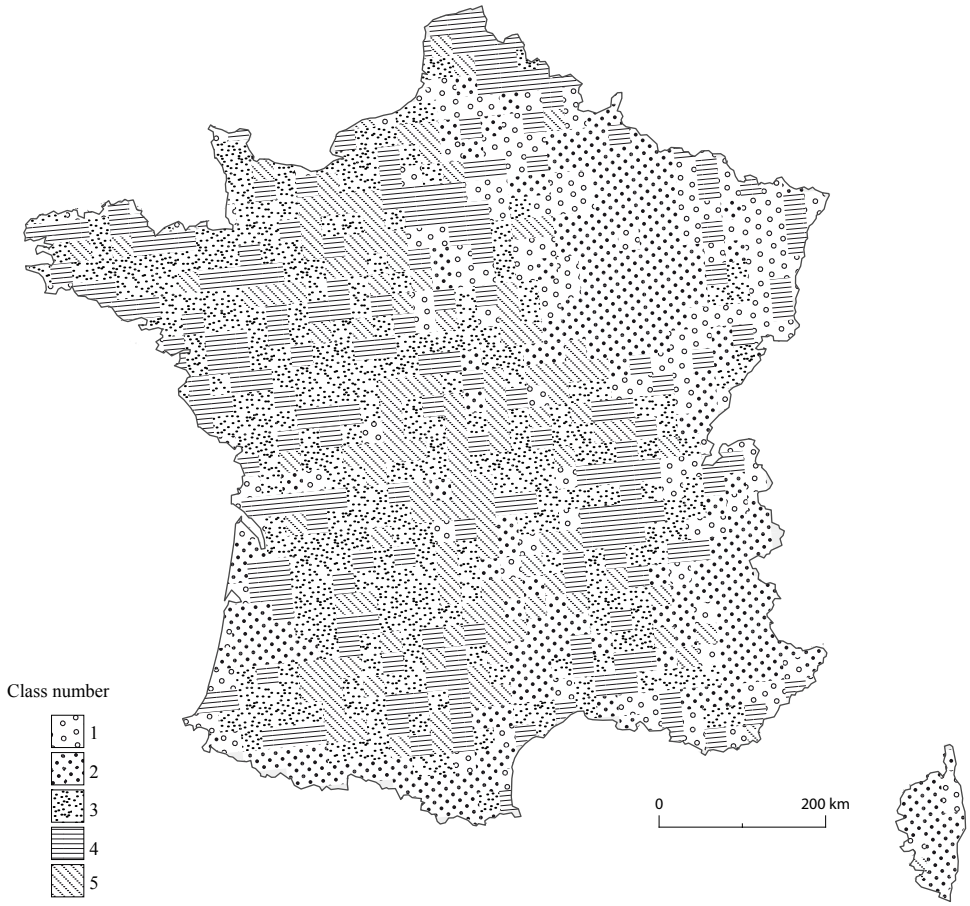


Figure 8: Map of the spatial units classified according to their multifractal spectrum