The morphology of built-up landscapes in Wallonia (Belgium): a classification using fractal indices

Abstract

The spatial pattern of built-up areas within a NUTS-1 European region (Wallonia in Belgium) is analysed using fractal indices. Methodologically, this paper illustrates the usefulness of fractal indices in measuring built-up morphologies, and also shows that clustering techniques have to be adapted for the non-Euclidean nature of the fractal measurements. An expectation maximisation algorithm (EM) combined with a Bayesian information criterion (BIC) is used . Empirically, we show that fractal indices partition the region into sub-areas that do not correspond to "natural landscapes" but result from the history of urbanisation. Urban sprawl seems to affect most communes, even the remotest villages: traditional (compact, ribbon, etc.) villages are transformed into more complex and heterogeneous shapes. These indices seem to be useful for characterising and understanding the built landscapes, as well as for modelling and planning urban realities.

Keywords: fractal dimension, built-up geometry, pattern analysis peri-urbanisation, Belgium.

1. Introduction

Landscapes result from numerous processes that interact spatially and temporally. Landscapes are often analysed in terms of natural and land-use components. The anthropogenic effect is often discussed in terms of deforestation, ecological and/or land use changes, etc. This paper is limited to one often-neglected component: the spatial arrangement of buildings. By means of fractal indices, we capture the morphology of the built environment (built-up surfaces) within one European NUTS-1 region, measure their diversity/resemblance and see how much recent residential practices have affected the built landscape. In other words, how much urban sprawl has modified the original aspect of the built-up regions.

It is well known that urbanisation modifies the traditional landscape, often making the distinction between urban and rural areas quite unclear (see e.g. Champion, 1989). Cities correspond to very large patches, and suburban landscapes are characterised by a wide variety of land uses, creating complex and diverse landscapes consisting of a highly fragmented mosaic of different forms of land cover and a dense transport infrastructure (see e.g. Antrop, 1997, 2000; Antrop and Van Eetvelde, 2000). In a country such as Belgium (small and highly urbanised), urbanisation affects the remotest villages.

Recent explosive urban growth has led to a particular and interdisciplinary scientific interest in urban peripheries, that is to say mixed areas where urban and rural spaces compete (see Johnson, 2001; Longley and Mesev, 2002; Caruso, 2002; Caruso et al., 2005; Cavailhès et al., 2002, 2004). Indices for defining and measuring sprawl (density, continuity, concentration, clustering, centrality, nuclearity, mixed uses, proximity, etc.) have been widely discussed. Many of the standard metrics are, however, highly correlated and very few take into account the relative location of the buildings (see e.g. Galster et al., 2001; Wolman et al., 2005). In this paper, the potentialities of fractals for measuring the spatial organisation of built surfaces are explored.

Fractals have been used for more than four decades for the description of outlines and surfaces, and have generated a large number of papers in various scientific disciplines (geology, biology, landscape analysis, architecture, physics, remote sensing, etc.) including landscape analysis (see e.g. Milne, 1991; McGarigal and Marks, 1995). By

definition, a fractal is a rough or fragmented geometric shape that can be subdivided into parts, each of which is (at least approximately) a smaller copy of the whole. Fractals are generally self-similar and independent of scale. The use of fractals in urban analysis was mainly developed in the 1990s. Early papers showed that cities can be conceptualised as fractals at several interrelated scales. In many papers, each city is considered individually and studied as a unit; fractality is then measured for that unit at one or several periods of time. The fractal dimension is then considered as a global measure of surface coverage (see e.g. Arlinghaus, 1985; Batty, 2005; Batty and Longley, 1994; Benguigui et al., 2000; Frankhauser, 1994; Goodchild and Mark, 1987; Longley and Mesev, 2000, 2002; MacLennan et al., 1991; Schweitzer and Steinbrick, 1998; Shen, 2002; Wentz, 2001; White and Engelen, 1993).

However, detailed measures of spatial distribution are clearly needed to complement the description of the morphology of an urban area (for examples, see Badariotti, 2005; Batty and Xie, 1996; Benguigui and Czamanski, 2004; Carvalho and Penn, 2004; Frankhauser, 1998; Schweitzer and Steinbrick, 1998; Tannier and Pumain, 2005). Fractal analyses provide synthetic measures of complexity, and thereby allow a numerical characterisation of places. Benguigui and Czamanski (2004) proved that fractality implies that a city possesses a similar structure at several different scales. This suggests the presence of some hidden process that operates at different urban scales and generates similarity.

This paper uses fractals at a local level: it considers the morphology of the built space within each commune (administrative entity), the 2D spatial organisation of the buildings. Theoretically, we suspect that villages created in the same "natural landscape" by the same (or almost the same) process(es) will show similarities in their spatial organisation and thus be fractally similar. The analysis is performed within an entire NUTS-1 region (Wallonia, Belgium), made up of urban, suburban, peri-urban and rural areas. Fractal indices are computed for each commune: the geometry of patches created by a common landscape transformation process should be statistically similar, i.e. their fractal dimensions and their form factors should be alike. The objectives here are to analyse the fractal dimension of all built-up surfaces and their perimeters; to consider to what extent they are alike within a heterogeneous region (Wallonia); to test whether the computed fractal geometry is indicative of a common historical,

geographical and/or planning process; and to test whether different types of built-up landscapes can be detected using appropriate clustering techniques. In other words, we want to propose a map of fractal indices (footprints) defining types of built landscapes in Wallonia.

The paper is organised as follows. Section 2 shows how the concept of fractal dimension fits the objectives. Section 3 explains the classification method used for the fractal dimensions. Section 4 briefly describes the area being studied and the data used; and Section 5 discusses the main results. Our conclusions are reported in the last section.

2. Fractal dimensions of built environments

Let us here intuitively review some general concepts of fractal dimension and explain how they fit our objective. Extensive mathematical formulations can be found in Mandelbrot (1982), Batty and Longley (1994), Frankhauser (1994, 1998), Russ (1994), Mattila (1995), Farina (1998), Lam and de Cola (2002), Falconer (2002), and more recently in Batty (2005, pp. 457–514).

2.1 Fractal hierarchy

By definition, fractal behaviour is associated with a scaling principle that governs how the constituent elements of a structure are distributed in space. The best way to illustrate this property is to look at how a theoretical fractal is constructed by iteration (see Figure 1). Here, the initiator is a square of length l_0 (see Figure 1a) which is then reduced by a factor $r = ¼$ into $N = 8$ elements. These elements are smaller replicates of the initiator with base length $l_1 = r\, l_0$, and they are organised within the area of the initial figure which is now called a generator (Figure 1b). Figure 1c, the first iteration, is obtained by repeating the process. A cluster hierarchy emerges, with large free lanes generated in the first step and smaller lanes separating the elements generated in the next step. The hierarchical aspect becomes obvious as smaller and smaller elements lying closer and closer together are generated in further steps.

Figure 1c shows that the clusters generated in the course of the iteration are distributed in a non-uniform way since the spaces separating the clusters are different. Fractal sets, such as the Fournier dust in Figure 1, can be characterised by their fractal dimension D, which describes the non-uniform distribution of the elements in space: the closer D is to 2, the more homogeneous the structure is (i.e. the more similar the width of the lanes separating clusters – see Figure 5b). For constructed fractals, D may directly be linked to N and r, the parameters of the generator (Mandelbrot, 1982). We should emphasise that D does not depend on the shape of the initial figure, or on the position of the elements in the generator. This is illustrated in Figure 2, where different generators are compared in which N, r and D are the same.

Another feature of Figure 2 is also important: in all the generators illustrated there, the elements are contiguous. Hence, in all the iterations, the fractal consists of one cluster. Such fractals are called Sierpinski carpets (see, for example, Batty, 2005, p. 503; Frankhauser, 1994; Mandelbrot, 1982). In Figure 2c it is unclear that the fractal really consists of one cluster, since the initial figure is a black central square surrounded by white lanes. However, the difference between the logic of the Sierpinski carpet and that of the Fournier dust becomes more obvious when the next iteration is considered (Figure 3). In Figure 3a, the Sierpinski carpet, all the lanes separating the black squares have the same width. This is not the case in Figure 3b, the Fourier dust, where the lanes follow a well-defined hierarchy. Hence the spatial hierarchy is stronger in the Fourier dust and so the fractal dimension is lower. This is confirmed by looking at the parameters of the generator. For both fractals, the number of elements in the generator is $N = 8$, but since the reduction factors, r, are different, the D values are also different: in the Sierpinski carpet $D = 1.89$ (as for the fractals in Figure 2), while for the Fourier dust $D = 1.50$. This example illustrates the obvious link between fractal geometry, dimensions and spatial hierarchy.

As expected, the spatial hierarchy of Fournier dusts is closely linked to the size of the empty lanes. In Sierpinski carpets, this hierarchy reflects the system of unoccupied sites (often called lacunas) as shown in Figure 4. The hierarchy of the lacunas is more obvious in Figure 4a; in Figure 4b, the situation is actually the same, as can be seen by delimiting the empty squares generated at each step. Once again, the dimension

describes the extent to which the occupied sites are spatially distributed in a non-uniform way.

2.2 Fractality and urban patterns

Urbanisation is a complex process to which a large number of public and private decision-makers contribute. The exact processes that control and produce the development of built-up landscapes are still not well understood (see e.g. Bartel, 2000; Van Eetvelde and Antrop, 2005). Several investigations (see references in the introduction) have shown that using fractal measures is a powerful tool for describing the morphology of sprawling urban patterns. This must be considered as a hint that the complex underlying socio-economic interactions generate built-up forms that can be described by means of a few morphological parameters, in particular by fractal dimension(s).

Such observations are well-known in theories of self-organisation and complex systems. They were first introduced in the context of phase-transition phenomena (Haken, 1977). In such systems, parameters that characterise a well defined organisation at a macro-scale are usually called “order parameters”. The use of this concept does not require a detailed knowledge of the underlying micro-processes, such as individual decisions. Hence, authors such as Weidlich and Haag (1983) and Haken (1977) have focused on the possibility of considering macrostructures stemming from socio-economic processes without tackling the underlying micro-processes. Frankhauser (1998) referred to such an approach in showing that the more an urban agglomeration grows, the more the spatial organisation of the built-up surfaces accords with a fractal distribution law. The present paper considers the fractal dimension as an order parameter. This does not exclude the possibility that this kind of macroscopic order may be more or less present in any given pattern.

2.3 Measuring the fractal dimension of urban patterns

Section 2.1 has shown that there is a direct link between N and r (the parameters of the generator) and D (the value of the fractal dimension) in constructed fractals. However, empirical structures do not derive from theoretical iterations. Moreover they are less symmetric than constructed fractals. We might expect that built-up surfaces are spatially

distributed according to the same type of law that holds for constructed fractals, in which case we can interpret urban patterns as random fractals (Frankhauser, 1994; Mandelbrot, 1982). To test this theory, we use cartographic representations of settlements. The black pixels represent built-up surfaces (buildings) and are called "occupied sites". In order to compute the fractal dimension, the cartographic representation is displayed on a computer screen; a rectangular zone is selected within that screen and called the "window". The fractal dimension describes the spatial organisation of the buildings within that window.

Several methods have been developed to test the extent to which the spatial organisation of empirical patterns follows a fractal law and to estimate their fractal dimension(s). Correlation analysis is one of these methods. Comparative investigations have shown that this method is reliable for analysing both surface and boundary dimensions, and for making comparisons, even when there are slight variations in the window's size and position (see e.g. Frankhauser, 2004; Badariotti, 2005). An occupied pixel I is chosen as the counting centre and the number $N_i(\varepsilon)$ of occupied pixels lying within a distance ε from this centre is counted. This number is called the "number of pair correlations". The procedure is repeated for each occupied site i. The mean number $N(\varepsilon)$ of all $N_i(\varepsilon)$ is then computed and the procedure is repeated for other values of ε (see e.g. Vicsek, 1989). If the occupied pixels of an urban pattern follow a fractal distribution law, the relation

$$N(\varepsilon) \propto \varepsilon^{D} \qquad [1]$$

between the mean number $N(\varepsilon)$ and ε is obtained. The *Fractalyse* software (developed by Vuidel, Frankhauser and Tannier and freely downloadable from http://fractalyse.org) is particularly suitable for studying the fractal dimension of built-up patterns and transport networks, and we employed it to analyse our data.

When investigating real world patterns, it is useful to introduce a generalised fractal law that contains (in addition to D) two other parameters:

$$N = a\,\varepsilon^{D} + c \qquad [2]$$

The prefactor a summarises various deviations from the fractal law (Gouyet, 1996). Such deviations may be due to the presence of huge lacunas (which can exist even in

structures that exhibit strict fractal scaling behaviour) in the selected window. Another reason for deviations may be that the structure being studied consists of n identical fractals. In this case, the number of correlations is simply multiplied by n at each scale, what yields simply to $a = n.$. In real-world structures, we do not expect the distribution of built-up areas to follow a strictly fractal law. Such random deviations will be measured globally by the value of a.

In many cases, the fractal behaviour is constant over a certain range of distances ε_0 to ε_1, which corresponds to a fractal dimension of D_1. However at distance ε_2, the fractal behaviour changes; another fractal behaviour corresponding to a value D_2 is observed. In order to ensure a correct estimate for D_2 it is necessary to introduce a constant c (Frankhauser, 1998). The values of the prefactor a and the constant c are controlled in our empirical analyses, but not thoroughly studied.

2.4 Fractal dimension versus density

The information captured by fractal dimension measured on surfaces has nothing to do with density (see e.g. Batty, 2005; Batty and Kim, 1992; Batty and Xie, 1996; Thomas et al., 2007). This can be illustrated intuitively by an example. Figure 5a shows the second iteration of the fractal shown in Figure 3a, which consists of 64 black squares. In Figure 5b, the same number of squares (64) has been arranged in a uniform way in the same area. Hence the density is exactly the same in both examples. This does not hold for the fractal dimension: in the uniform distribution (Figure 5b) D is equal to 2.00, whereas in Figure 5a $D = 1.89$. When comparing the two subsequent steps, it is obvious that the total black area decreases as the iterations continue. Hence, in fractals, the density changes in subsequent iterations, while D remains the same. Density is a descriptor that is not suitable for fractals.

Spatial structures characterised by $D = 2.0$ correspond to structures where the occupied sites are distributed uniformly. The other limiting case is that of several isolated points, which has a fractal dimension of 0.0. $D = 1.0$ represents a threshold in fractals: when D is smaller than 1.0 the structure comprises a set of unconnected points such as a Fournier dust. D describes the extent to which the set of built-up sites are concentrated in clusters at different scales, or, in other words, the degree to which a set of built-up

sites fills the space in which it is embedded at different scales (see Frankhauser (1994) and Thomas et al. (2007) for further discussion of this point).

2.5 Fractal dimensions of borders

It is also possible to apply the concept of fractal dimensions to urban boundaries. However, when looking at real world patterns such as that illustrated in Figure 6a, urban boundaries do not exist as such: a settlement consists of disconnected houses (or blocks of houses) and hence looks more like a Fournier dust than a Sierpinski carpet. Nevertheless, it is well established that it is possible to identify urban boundaries by using simplified representations of urban patterns. Such boundaries have already been studied by means of fractal geometry (see e.g. Batty and Longley, 1994; Frankhauser, 1994).

In this paper, a procedure based on dilation was used to extract urban boundaries. To this end the buildings represented as black surfaces on the map are dilated in a stepwise procedure. Hence initially separated buildings touch neighbouring ones at a certain step and large clusters appear in the course of the procedure. Boundaries can be seen (Figure 6c) and may be identified as an approximation of the urban border. Investigations have shown that large clusters emerge after just a few dilation steps, when courtyards and small streets are filled in (Figure 6b). This method is explained in more detail by De Keersmaecker et al. (2003) and a mathematical justification was recently given by Frankhauser and Tannier (2005).

The dilated patterns consist of a few clusters corresponding to large settlements, with a large number of small villages, hamlets and isolated farms. Such a hierarchical system can be illustrated theoretically by combining Sierpinski carpets and Fournier dusts as shown in Figure 6d. Sierpinski carpets provide the possibility of modelling tentacular structures with tortuous boundaries (as in Figure 4b), whereas Fournier dusts add the possibility of creating detached elements. Combining these two types of fractal allows a cluster hierarchy to be generated, as shown in Figure 6d. Several tests (not reported here) have shown that the results are quite stable when the boundaries are extracted after different steps in the dilation process. However, the boundaries become progressively smoother and empty intra-urban space is filled up as the dilation process continues,

which may affect the fractal dimension. Hence it is recommended that dilation be limited to a few steps; in this paper it is limited to three.

The fractal dimension of boundaries is denoted D_{Bord}, while that of surfaces is denoted D_{Surf}. For topologically linear objects D_{Bord} characterises the extent to which they cover space in a uniform way. It is obvious that the boundary of the fractal in Figure 6d does not cover space in a uniform way, but it is possible to imagine linear fractals that would do so.

2.6 Borders versus surfaces

The ratio of D_{Bord} to D_{Surf} has been used in empirical studies (e.g. Imre and Bogaert, 2004). Its meaning can be elucidated by using a fractal relationship for both the border and the surface:

$$N_{surf}(\varepsilon) \propto \varepsilon^{D_{surf}} \quad \Rightarrow \quad \varepsilon \propto N_{surf}(\varepsilon)^{\frac{1}{D_{surf}}} \qquad [3]$$

$$N_{bord}(\varepsilon) \propto \varepsilon^{D_{bord}} \quad \Rightarrow \quad N_{bord}(\varepsilon) \propto N_{surf}(\varepsilon)^{\frac{D_{bord}}{D_{surf}}} \qquad [4]$$

Equation [3] shows that, by re-writing the fractal relationship for the surface, we obtain a relationship between the distance parameter ε and the number of elements $N_{Surf}(\varepsilon)$. By inserting this in the fractal relationship for the border, we can link the length of the border $N_{Bord}(\varepsilon)$ directly to the surface occupation $N_{Surf}(\varepsilon)$ by means of a scaling exponent which is simply the ratio D_{Bord} / D_{Surf} (Equation [4]). In a sense, this ratio measures the compactness of a structure. Geometrical objects such as squares or circles have smooth borders and compact uniform inner surfaces without lacunas. Their dimensions are $D_{Bord} = 1$ and $D_{Surf} = 2$ and thus $D_{Bord} / D_{Surf} = 0.5$. This is the minimum value of the ratio. For fractals, the borderlines have higher D values and the surfaces may have values lower than 2. For teragons, $D_{Surf} = 2$ and $1 < D_{Bord} < 2$, and the ratio varies between 0.5 and 1.0. For Sierpinski carpets $D_{Bord} / D_{Surf} = 1$, since the two dimensions are identical and $D_{Bord} = D_{Surf}$. We might expect this to be the maximum value of the ratio. However we will come back to this question when dealing with real-world patterns of built-up areas. Like all ratios, D_{Bord} / D_{Surf} is ambiguous: multiplying both the numerator and the denominator by the same factor does not affect its value.

Hence we obtain the value 1 not only for Sierpinski carpets, but also for straight lines, since $D_{Bord}=D_{Surf}=1$.

The objective now is to see how *D* values obtained for observed built-up structures vary within a region: do places close to each other look alike? As the ratio of *D* values is not sufficient, the following section develops a clustering method.

3. Classifying fractal dimensions

3.1 The problem

In spatial analysis, it is interesting not only to characterise the morphology of each pattern, but also to test whether it is possible to use morphometric parameters to distinguish patterns that are linked to particular historical or geographical topics. For instance, we might expect that settlements which grew up during early periods of industrialisation would show different patterns to sprawling peri-urban villages or villages located in rural regions. Recent observations seem to confirm such a hypothesis: the comparative study of samples of built-up environments shows that different types of urbanisation generate patterns that correspond to typical ranges of fractal dimensions (Badariotti, 2005; Thomas et al., 2007). However this has not been tested using appropriate clustering methods.

For several reasons, statistical clustering procedures may be an interesting alternative to ratios in distinguishing homogeneous and meaningful classes of settlements as described by morphometric indicators. First, such procedures often allow some traditional non-statistical methods, for instance geometric-based methods, to be used and generalised. Moreover, the underlying probabilistic hypotheses suggest a formal context for selecting models, the number of classes in a model being of particular interest. Consequently, we adopted a model-based clustering method to address the classification problem.

3.2 A solution

The aim here is to define clusters of places characterised by similar fractal dimensions of their borders and/or surfaces. In more formal terms, the purpose is to estimate an unknown partition P of a collection of n places $\mathbf{x}=\{x_1,...,x_n\}$ with $x_i=(D_{bord\,i},D_{surf\,i})$ into K clusters. Since the places x_i are continuous observations in $\mathbb{R}^2$, mixtures of bivariate Gaussian distributions may be used. Indeed, such a model provides a powerful and nowadays standard tool for clustering (see references in McLachlan and Peel, 2000). In this context, $x_1,...,x_n$ are assumed to arise independently and identically distributed (i.i.d.) from a mixed distribution with density

$$f(x;\theta)=\sum_{k=1}^{K} p_k \varphi(x;\mu_k,\Sigma_k) \qquad [5]$$

where p_k is the proportion of the k^{th} cluster ($p_1+...+p_K=1$, $p_k>0$) and $\varphi(x;\mu,\Sigma)$ is the bivariate Gaussian distribution with mean μ and variance matrix Σ. The parameter for the whole mixture is denoted by $\theta=(p_1,...,p_K,\mu_1,...,\mu_K,\Sigma_1,...,\Sigma_K)$. As the attentive reader will have noticed, the underlying idea of model-based clustering is to link each cluster to each Gaussian mixture component.

An estimate P' of the partition P may easily be obtained from the estimate θ' of the mixture parameter θ by invoking the maximum a posteriori procedure (which consists of assigning each place x_i to the cluster k having the largest conditional probability that x_i arises from it). Then

$$t_k(x_i;\theta')=\frac{p_k \varphi(x_i;\mu_k',\Sigma_k')}{\sum_{h=1}^{K} p_h \varphi(x_i;\mu_h',\Sigma_h')} \qquad [6]$$

An estimate θ' can be obtained by maximising the log-likelihood given by

$$L(\theta;x)=\sum_{i=1}^{n} \ln(f(x_i;\theta)). \qquad [7]$$

Maximising $L(\theta;$ x) is generally performed using the expectation maximisation (EM) algorithm, which is an iterative procedure well adapted to incomplete data structures (Dempster et al., 1977).

The Gaussian density model leads to an ellipsoidal class with centre μ_k, whose geometric characteristics can be deduced from the eigenvalue decomposition of the variance matrix Σ_k. Following Banfield and Raftery (1993) and Celeux and Govaert (1995), each mixture component variance matrix can be written as

$$\Sigma_k = \lambda_k B_k A_k B_k^t , \qquad [8]$$

where $\lambda_k = \det(\Sigma_k)^{1/d}$, B_k is the matrix of eigenvectors of Σ_k, and A_k is a diagonal matrix such that $\det(A_k)=1$ with the normalised eigenvalues of Σ_k on the diagonal in decreasing order. The parameter λ_k determines the volume of the k^{th} cluster, B_k determines its orientation and A_k its shape. By allowing some of these quantities to vary between clusters, parsimonious and easily interpreted models can be obtained, which are useful in describing various clustering situations. Varying the assumptions concerning the parameters λ_k, B_k and A_k leads to fourteen models of interest, divided into three main families. From the simplest to the most complex these are the spherical family (A_k is the identity matrix), the diagonal family (B_k are permutation matrices and so Σ_k are diagonal matrices), and the general family (other situations). Celeux and Govaert (1995) identify a specific maximisation step of the EM algorithm for each of these 14 models.

Aside from these geometrical features, another important parameter of the k^{th} cluster is its proportion p_k. Two typical situations are generally considered with regard to proportions: it is assumed that they are either equal or free. Combining these alternative assumptions with the 14 previous models produces 28 different models.

It is worth noting that some of these 28 models shed new light on some standard clustering criteria that have been proposed without reference to any statistical model. For instance, Ward's (1963) *K*-means criterion corresponds to the simplest of the models in the spherical family when the mixing proportions are equal. The criterion suggested by Friedman and Rubin (1967) is obtained with the simplest model of the general family (corresponding to the homoscedastic situation). Celeux and Govaert (1992) present a survey of such cases.

There is now a need to select one of the 28 models (denoted by *M*) and *K,* the number of clusters. Schwarz's (1978) Bayesian information criterion (BIC) is generally relevant to

performing such a selection. It is expressed as the following penalisation of the maximum log-likelihood value

$$BIC_{M,K} = L\left(\theta'_{M,K}; x\right) - \frac{v_{M,K}}{2}\ln(n) \qquad [9]$$

where $v_{M,K}$ denotes the number of parameters to estimate (in other words, the dimension of θ) and $\theta'_{M,K}$ denotes the maximum likelihood estimate of θ. The couple (M,K) yielding the largest value of BIC is then retained.

In many situations, practitioners decide to perform the Gaussian clustering procedure on a one to one transformation $\mathbf{g}(\mathbf{x}) = \{g(x_1),...,g(x_n)\}$ of the initial data set instead of on the initial data set **x** itself. The reasons are generally either that the new data set $\mathbf{g}(\mathbf{x})$ either "seems to have a better Gaussian mixture shape" than **x**, or that its unit has a particular meaning for the practitioner. Typically, standard transformations are $g(x)=\exp(x)$ or $g(x)=\ln(x)$. The first transformation expresses data in the same units as $N(\varepsilon)$, which is a more traditional quantity than D for many geographers. This may be sufficient reason to consider such a transformation. To avoid the difficult task of proposing and justifying a particular transformation, the practitioner may use the statistical framework to choose one of the suggested transformations automatically. We describe this interesting and innovative feature below.

It is possible to interpret any transformation as another kind of model which is no longer a Gaussian mixture model (as the previous model M was). Indeed, if the new sample $\mathbf{g}(\mathbf{x})$ arises i.i.d. from the Gaussian mixture density $f(x; \theta)$ then the initial sample **x** arises i.i.d. from a density $f_g(y; \theta)$ which is a transformation of $f(x; \theta)$ and so not necessarily a Gaussian mixture density. Consequently, it is possible to employ the BIC criterion to select this transformation, which can be used for any model. Denoting by J the Jacobian of the transformation g, and by $\theta'_{M,K,g}$ the maximum likelihood estimate obtained with the triplet (M,K,g), we retain the triplet (M,K,g) leading to the largest of the following BIC expressions:

$$BIC_{M,K,g} = L\left(\theta'_{M,K,g}; g(x)\right) - \frac{v_{M,K}}{2}\ln(n) + \ln\left(det(J)\right). \qquad [10]$$

Note that the number of parameters $v_{M,K}$ does not depend on the transformation g. It is also worth noting that in many applications the selection of the transformation is performed in an empirical manner, and that the method proposed here allows an answer to this problem to be found in a formal way. We apply this classification method to observed fractal dimensions below, and show its usefulness in defining built-up morphologies.

4. Data

4.1 Study area

The empirical analysis was carried out on one NUTS-1 region (Wallonia), which encompasses the southern part of Belgium. It is 16,845 sq. km large and has 3.4 million inhabitants. The average density is 200 inhabitants per sq. km with very large differences between sub-areas (see Figure 7). Administratively, Wallonia is divided into 262 communes (townships). The population is mainly concentrated on the former 19th century industrial axis running from the east (Liège) to the west (Mons) and following the rivers Meuse and Sambre. North of this axis, built landscapes are strongly influenced by the presence of Brussels (exhibiting suburbanisation and peri-urbanisation) especially in the Province of Walloon Brabant. In the southernmost part of Wallonia, communes are affected by the presence of the city of Luxembourg (rents are lower in Belgium than in Luxembourg and many Belgian residents commute to Luxembourg). The rest of the territory is less densely inhabited with fewer than 50 inhabitants per sq. km in some communes of the Ardennes. Until now, Walloon landscapes have often been defined in terms of natural characteristics (Antrop, 1997; Dussart, 1957, 1961; Feltz, 2004; Van Eetvelde and Antrop, 2005).

4.2 Database

Buildings are individually identified in the *Plan de Localisation Informatique* (PLI) developed by MRW-DGATLP (2004), which is based on digitised topographical maps at a scale of 1:10,000 (Institut Géographique National) and on the land registry (called *Cadastre*). This database is annually updated and is available by commune. No

information is available about the function of the buildings (residence, public service, industry, service or farm), their height (number of floors), their architectural characteristics, or about their occupants (number of inhabitants or number of jobs within a building). Hence we distinguish only built-up and non built-up sites. One picture (raster) was extracted for each commune, where built-up sites correspond to black pixels. By means of fractal analysis the empirical relation $N(\varepsilon)$ was determined. Let us here remind that $N(\varepsilon)$ is the mean number of black pixels lying within a distance ε one from the other, according to relations [1] and [2]. The software *Fractalyse* estimates the fractal dimension as well as the parameters *a* and *c* by minimising least square deviations between the theoretical fractal law [2] and the empirical relation $N(\varepsilon)$. The quality of adjustment is evaluated by means of the *ratio of correlation* η^2 [1]. For fractal analysis of urban patterns, experience shows, that the values are usually very close to $\eta^2 = 1$ ($\eta^2 > 0.9999$).

Given the scale of analysis adopted here, some cartographic details are already smoothed: small courtyards or very narrow streets are not visible on the map. Hence the original patterns consist of clusters. Taking into account the "semi-aggregated" nature of the database, the patterns are reminiscent of the logic of the hybrid Sierpinski carpet. This is illustrated in Figure 8 where a real pattern is compared to a theoretical one that allows the tentacular aspect of clusters to be combined with highly tortuous boundaries and a hierarchy of clusters. $D_{Surf\text{-}d0}$ is equal to 1.50 in the theoretical fractal pattern and to 1.27 in the real world example.

D values are computed for communes, which are administrative areas. It should be remembered that new communes were defined in Belgium in 1977: former townships were aggregated into larger administrative units. This means that, morphologically, many communes are made up of several villages. This is particularly true in less urbanised areas, where villages have not merged. The fractal dimension is measured here for the entire commune and not for individual villages. This leads to the well-

[1] this index relates the variance of empirical data computed with respect to the general mean value, to the variance of the theoretical function also computed with respect to the general mean value. If the empirical relation suits perfectly to the theoretical one, then $\eta^2 = 1$.

known MAUP (modifiable areal unit problem): we know that *D* values for the communes are not simply an average of the values computed for the villages. Choosing to handle villages separately would, however, be quite tricky in urban and peri-urban areas. Moreover, former villages no longer correspond to functional administrative entities.

As each commune is analysed separately, the size of the window varies. This may lead to slight differences in the values of the indices, but it does not affect the general conclusions (see Thomas et al., 2007). The window is defined as the smallest rectangle including the commune. Fractal dimensions are computed for the original database ($D_{Surf\text{-}d0}$) as well as on slightly dilated patterns ($D_{Surf\text{-}d3}$, $D_{Bord\text{-}d3}$). We expect these dimensions to vary in space and to illustrate the diversity of built environments.

5. Results

5.1 Fractal dimensions

The fractal dimensions computed for the built-up surfaces of the 262 communes are often (but not always) larger than those computed for their borders (see Section 5.2 below), and the observed variation in *D* values is larger for D_{Surf} than for D_{Bord}. *D* values are far from revealing a compact pattern of built-up areas: compactness would give values of $D_{Surf} \approx 2$ and $D_{Bord} \approx 1$, but the history of urbanisation seems to have led to a large variety of morphologies in Wallonia.

Some of the fractal dimensions observed on the original (not dilated) surfaces ($D_{Surf\text{-}d0}$) are smaller than 1.0. This means that the built-up sites are mostly isolated one from the other; this looks like Fournier dusts that consist in a fractal arrangement of isolated sites. In a settlement pattern, such a situation corresponds to small dispersed villages in areas which are not densely built-up and communes mainly comprised of isolated hamlets.

In addition, the significance of these small values is relatively low compared to that of other measurements ($\eta^2 < 0.9999$). Note that when slightly dilated, the fractal dimension for surfaces ($D_{Surf\text{-}d3}$) is always larger than 1.0, and the same is true for

borders ($D_{Bord\text{-}d3}$). Remember that values of D_{Surf} close to 2.0 correspond to perfect homogeneity, while values equal to 1.0 correspond to a line (that is to single-street or ribbon villages).

$D_{Surf\text{-}d3}$ and $D_{Bord\text{-}d3}$ are positively and significantly, but not perfectly, correlated (Pearson correlation = 0.806). For Sierpinski carpets the two dimensions would be equal, but, as we have already seen, this is not true for teragons. Hence, the two dimensions do not correspond to exactly the same realities. This is also true for $D_{Surf\text{-}d0}$ and $D_{Bord\text{-}d3}$ (Pearson correlation =0.642). If we eliminate D values less than 1.0 (due to their low significance levels), then these two correlation coefficients become 0.732 and 0.507 respectively. The morphology of the borders is different than that of the surfaces (although the two measures are certainly not independent): D_{Surf} and D_{Bord} involve different ways of describing the geometry of built-up areas.

5.2 Ratio of dimensions

5.2.1 The paradox of Walloon urbanisation: an example

As mentioned in Section 2.6, it is common practice to compute the ratio of D_{Bord} to D_{Surf}. In this study this ratio is normally distributed, and generally larger than 1.0. This result is mathematically counter-intuitive, since 1.0 is logically the upper limit of the ratio. In our analyses, $D_{Bord\text{-}d3}$ is only smaller than $D_{Surf\text{-}d3}$ in a few communes: the border (perimeter) of the urban patch is less homogeneous than the surface of the built-up area, especially in very densely built-up communes. This surprising result spurred us to analyse several communes in detail. The results from one of them, Le Roeulx, are reported here.

Parameters were estimated for the entire range of distances (1 to 143 pixels) as well as for reduced ranges of distances, selected with respect to the shape of the empirical function $N(\varepsilon)$ i.e. taking into account thresholds in the empirical functions (estimation results see Table 1). Hence the chosen ranges may not be the same for the surface and the border analysis. The quality of the adjustment is again here noted η^2. We know from experience that η^2 values should be higher than 0.9999. This is not the case for the largest range of surface distances. For the reduced ranges, and for the border ranges, the adjustments seem to be correct. The surface prefactors a (see Section 2.3) are higher

than usual in Le Roeulx. However this is not totally inexplicable, as $a > 1$ may indicate an area consisting of several former villages that all have approximately the same scaling behaviour.

In order to verify whether a influences the results, some new tests were performed by fixing the prefactor at $a = 1$ and by estimating only D and c (Table 2). The η^2 values then suggest a lower quality of adjustment. This may be understood as indicating that each value $a \neq 1$ allows the adjustment to be improved. The graphical comparison between the estimated and empirical as (Figure 9) makes this clear. When the Ds are considered we can see that they differ substantially between Table 1 and Table 2, for both surfaces and borders. Moreover in Table 1 D_{bord} is always larger than D_{surf}, whereas the reverse is true in Table 2. As discussed above, it makes more sense mathematically for D_{surf} to be larger than D_{bord}. On this basis, if no other, the second set of results seems more realistic.

5.2.2 An attempt at an explanation

An explanation of these surprising results can be proposed by referring to the special features of Le Roeulx (see Appendix). Figure 10 shows the relative variations, $dN(\varepsilon)/N(\varepsilon)$, by distance for both the border and the surface; it confirms that the relative variations of the border exceed those of the surface. This is why D_{Bord} is greater than D_{Surf}, especially for small distances. The two curves tend to the same limit for large distances. This seems quite reasonable: for large distances, the transient effects discussed in appendix 1 progressively disappear.

Last but not least, and quite empirically, we should remember that nowadays communes consist of several villages. So two different distance ranges come into play when measuring the distribution of built-up areas: the intra-urban scale referring to the location of buildings and blocks of houses within a village, and the inter-urban scale corresponding to the position of one village with respect to the others. For intermediate distances both phenomena contribute to the shape of the empirical curve $N(\varepsilon)$. Moreover since Walloon planning legislation has traditionally been rather weak, diffuse urbanisation is often observed between the villages. This can also influence the results.

Additional explanations can put forward. Firstly, the scale of the analysis: measurements are made at the level of the commune, and by definition communes differ in shape and size. This certainly leads to MAUP, but fractals – by their very nature – should not be too sensitive to this problem. Secondly, in Wallonia there are both urbanised and rural communes. The ratio of D_{Bord} to D_{Surf} is smaller than 1.0 in very densely populated communes (cities). In these areas the morphology is totally different (see De Keersmaecker et al., 2004). And thirdly, as in many other places throughout the world, the administrative limits of the commune do not correspond to the real functional borders of the urban agglomerations: in some rare cases the town is smaller than the commune; in many others the city extends beyond its administrative limits. The border of the urban area is thus not included in the commune.

This discussion suggests that fractal dimensions seem to vary sufficiently to characterise the morphology of the built-up surfaces, although errors and the scatter in the data suggest that no perfect relationship exists. There are several reasons why D_{Bord} may exceed D_{Surf} in some places. Taking into account the reliability of the parameter estimation, we treat the estimated D-values as morphological order parameters, according to our descriptive approach. We will now map the D values in order to consider their spatial variation (Section 5.3) and then cluster the communes by means of several D values (Section 5.4).

5.3 Mapping fractal dimensions

The surface dimensions observed here are rather low compared to those observed in France and Germany (see e.g. Frankhauser, 1994, 2004). This means that built-up areas are less uniformly distributed in Wallonia than elsewhere, which may be explained by the rather weak control of peri-urbanisation in Belgium. Mapping fractal dimensions enables the spatial variation to be analysed visually. Univariate choropleth maps were drawn for the entire region, natural breaks being used to define exploratory classes of values.

Figure 11 shows the fractal dimension computed for the undilated built-up surfaces. It clearly reveals the urbanisation structure: $D_{Surf\text{-}d0}$ is high (> 1.58) in Walloon cities along the 19th century Sambre–Meuse industrial axis and also in some communes in the central northern part of the region due to the peri-urbanisation process along the

southern edge of Brussels (see communes such as Waterloo and Wavre–Ottignes–Rixensart which could be considered as emerging edge-cities). This shows that in these areas the built-up surface is more uniformly distributed than in other areas. Indeed in the former industrialised communes (Charleroi, Liège, La Louvière, etc.), we can assume that the initially empty interstitial sites have been progressively filled up with time. This also holds for the areas south of Brussels where peri-urbanisation is very significant. $D_{Surf\text{-}d0}$ values between 1.57 and 1.39 characterise peri-urban communes and the western part of the region (Hainaut, with its specific industrial history). The southern part of Wallonia (Province of Luxembourg) is less urbanised, and two types of morphologies seem to coexist in accordance with traditional landscape/geographical analyses (Dussart, 1957; 1961; Feltz, 2004): compact ($D_{Surf\text{-}d0}$ = 1.23 to 1.39) and elongated villages ($D_{Surf\text{-}d0}$ = 1.01 to 1.23). In these elongated villages, tentacular concentrations of buildings along transportation axes lead to patterns with a high degree of contrast: almost empty corridors exist between the different built-up branches. Finally, values smaller than 1.0 are observed in some communes in the southern part of Wallonia. These communes are composed of isolated settlements (villages, hamlets and large farms) that are highly dispersed and consist of detached splashes rather than large clusters. In this respect they resemble Fournier dusts.

At this stage of the analysis, the map suggests a typology of $D_{Surf\text{-}d0}$ that clearly reveals the history of urbanisation as well as differences in morphology: it seems likely that the shape of the built-up surfaces within the Walloon communes is the result not only of physical characteristics (soil, relief, initial vegetation, afforestation, etc.), but also of human activities and land use (suburbanisation, etc.). Hence the older and more densely populated communes are more uniformly built up and have high values of D_{Surf}; medium values occur in communes with more tentacular patterns, which were more recently affected by peri-urbanisation; and more rural areas are dominated by small isolated "splashes" (D_{Surf} <1.0). This corresponds to the different stages of peri-urbanisation: firstly, isolated detached houses are built which correspond to "leapfrogging"; then "tentacular" evolution occurs where interstitial spaces are filled in along transportation axes; and finally the large empty lanes between the former axes are developed and a fairly uniform pattern emerges. This result accords with Benguigui and Czamanski's (2004) analysis of Tel Aviv.

The spatial distribution of the fractal dimension of dilated borders is shown in Figure 12. As expected, the patter here differs slightly from the preceding one. Some densely populated communes (e.g. Waterloo and Liège) have relatively low $D_{Bord\text{-}d3}$ values, while the western communes (Hainaut) (which do not, on average, have high population densities) have large values of $D_{Bord\text{-}d3}$. This shows that in some highly urbanised areas the composition of the town is such that no large non-built-up areas remain. The remaining small non-built-up spaces (streets and small squares) have disappeared after three dilation steps. Another consequence is that the built-up space has become more compact. Hence such patterns show characteristics similar to teragons or even circles: the surface tends to be uniform and the dimension is high, whereas the borders are quite smooth (e.g. Liège). If the dimension of surfaces is low and that of borders high, the inner fragmentation is high: the interstitial spaces between the buildings have obviously not disappeared with dilation and hence we find borders of clusters within the settlement. In the southern part of the Walloon region, low values are also observed for the borders. This corresponds to the non-uniform distribution of built-up surfaces (i.e. to tentacular morphologies).

Comparative investigations of agglomerations in other European countries have shown that D_{Surf} in peri-urban areas lies between 1.60 and 1.87 and D_{Bord} between 1.20 and 1.55 (Frankhauser, 2004). In general, we can conclude that surface dimensions of about 1.7 and border dimensions of about 1.4 – 1.5 guarantee a good articulation of urbanised and green areas. Green areas are easily accessible, but the pattern nevertheless preserves a certain degree of compactness. Compared to these observations, the Walloon case should be considered as an extreme example. It appears that local densification could occur without real loss of quality of life.

5.4 Classifying fractal dimensions

The 262 communes can now be classified in terms $D_{Surf\text{-}d0}$, $D_{Surf\text{-}d3}$ and $D_{Bord\text{-}d3}$ to identify similar built-up landscapes using the methodology described in Section 3.2.

As well as the initial data set, the two transformations $g(x)=\ln(x)$ and $g(x)=\exp(x)$ were considered. The EM algorithm was run for 150 iterations from 5 random starting parameters for each combination: Gaussian model M × transformation g × number of clusters K. Only the trial with the starting parameters leading to the largest likelihood

was retained. The 28 Gaussian models described in Section 3.2, with three types of transformations (no transformation, logarithm and exponential), and 12 different numbers of clusters (*K* varying from 1 to 12) were considered. Whichever pair of variables were used, the results were similar: the data were spatially well discriminated. The first pair ($D_{Surf\text{-}d0}$; $D_{Bord\text{-}d3}$) enabled the communes that are alike in terms of the texture of their surfaces and borders to be clustered; the second ($D_{Surf\text{-}d3}$; $D_{Bord\text{-}d3}$) enabled a classification replacing their simple ratio. The results were quite similar. For the sake of clarity, only one set of results ($D_{Surf\text{-}d0}$ and $D_{Bord\text{-}d3}$) is discussed below.

The optimum number of classes (BIC criterion) with the exponential transformation is two (which means that the clusters are Gaussian when using the unit of $N(\varepsilon)$). Using a Gaussian model with different mixing proportions, the same volume and shape but different orientations between clusters, one cluster includes all the urban communes (cities) while the other covers all the other communes (peri-urban, rural). The built-up landscape of Wallonia is hence very clear cut: urban and non-urban.

If we force the classification procedure to adopt the standard Ward methodology (which corresponds to the simplest model *M*: the same proportions and the same volume in all the clusters, plus the spherical family; for a early reference on the Ward methodology, see for instance Forgy, 1965), the BIC criterion retains five classes and the exponential transformation. In fact, to follow the standard Ward methodology exactly, it would be necessary to optimise the completed likelihood instead of the likelihood (see Celeux and Govaert, 1992). The completed likelihood differs from the likelihood in that the partition appears explicitly in its expression. Thus, unlike the likelihood (which only has to be maximised on the mixture parameters (see Equation [7])), the completed likelihood has to be maximised both on the mixture parameters and on the partition itself. Nevertheless, the results of these two procedures are expected to be very close (Celeux and Govaert, 1992). If the number of classes is forced to six, the *BIC* criterion retains both Ward's model and the exponential transformation. The results for six classes of communes are illustrated in Figure 13. The map reveals strong effects of contiguity: communes close to each other look alike in terms of fractal dimensions. Clusters are, however, spread out, all over the region.

Table 3 gives the mean values of $D_{Surf\text{-}d0}$ and $D_{Bord\text{-}d3}$ for each class, and an example of a commune in each cluster is illustrated in Figure 14. Since the exponential transformation has been retained, the mean values correspond to the logarithm of the centre of each class; this enables the centre of each cluster to be expressed in the unit *D*. Table 3 lists the three communes which are closest to the centre of each class (Mahalanobis distance) and shows that both fractal dimensions (borders and surfaces) have to be considered.

On average, we can say that the morphology of the built-up surfaces in Wallonia is strongly influenced by the history of urbanisation and the underlying processes: the history of the urban network, the 19th century urbanisation leading to the Sambre-Meuse industrial axis, and the 20th century suburbanisation spreading from Brussels and other cities (including Luxembourg). *D* reveals the dispersion and concentration processes but also enables compact and elongated/spread patterns of building to be differentiated.

Some remnants of traditional rural landscapes are to be found in the Ardennes, where two types coexist (east–west partition), so corroborating a traditional description of the habitat (Dussart, 1961; Feltz, 2004). For some communes, the value of D_{Bord} exceeds that of D_{Surf}. These exceptions do not, however, lead to absurd results when the *D*-values are classified. Rather, they seem to indicate that these communes exhibit particularities which are correctly identified by the classification procedure.

Recent peri-urbanisation has generated patterns with low surface dimensions and high border dimensions. These patterns are rather fragmented. Such observations could be useful when thinking about how to manage urban sprawl. The dream of a compact city is no longer realistic: we cannot ignore the fact that a large number of households prefer to live in semi-rural areas offering peace and quiet with pleasant landscapes, while still within reach of urban amenities. In some sense the morphological properties of peri-urban patterns may be seen as resulting from such a demand.

These results confirm that centre–periphery structures seem to dominate the Walloon space and that suburban areas are characterised by a wide variety of land uses, creating complex and diverse landscapes consisting of a highly fragmented mosaic of different forms of land cover (see Antrop, 1997; Paquette and Domon, 2001). The fractal

measurements used here are obviously suitable for characterising the morphology of built-up settlements: this has been demonstrated both theoretically and empirically.

We end up with a well defined spatial structure showing a clear relationship between morphology and historical/geographical/planning contexts, even if nowadays sprawl reduces contrasts between rural and urban areas. The major effect of sprawl is here due to Brussels which lies out of the studied area; southern periurbanisation of Brussels started 40 years ago in several waves mainly due to the construction of transportation axes. This sprawl also occurs more locally around the regional urban centres.

Obviously, fractal dimensions are suitable for characterizing the form of the built-up patterns and for defining orientations for urban development. However, considering the spatial distribution of built-up areas does not provide information about how "free spaces" between the buildings is effectively used. It can however be argued that the uniform distributions characterizing city centres as well as some individual housing areas can never offer a diversified range of green amenities. If homogenously built-up residential areas offer an individual garden for each house, they provide no public space for events, which necessitate large 'green areas' at a different spatial scale.

Accessibility to the green amenities and hence traffic is also to be considered (see on going research of Frankhauser et al., 2007). Let us also add here that in a fractal structure, the penetration of built-up and non built-up spaces (green, empty) is multi-scale. Such a spatial organisation offers the chance of having zones of high concentration near urban amenities (shopping, services), e.g. in the vicinity of public transportation networks, but also more diluted spaces. It then becomes possible to maintain a social mix by means of a higher local variety of densely and less densely built-up (populated) zones and also to preserve huge empty zones in the neighbourhood of urbanized areas, which may be imagined as natural reserves, agricultural zones, etc. (see Frankhauser, 2004).

6. Conclusion

We have analysed the spatial arrangement of the built elements in all the communes of Wallonia, Belgium, using fractal dimension(s) measured on their surfaces and borders.

Wallonia turned out to have particular values which were, from a strictly mathematical point of view, counterintuitive. Possible explanations have been given.

Despite the many limitations in the data, fractal indices partition the region into clear-cut sub-areas that do not match the physical landscape: the structures observed correspond more closely to the history of urbanisation. It seems likely that the shape of the built-up areas within Wallonia is the result not only of physical characteristics (soil, relief, initial vegetation, afforestation, etc.) but also of human activities and land-use (such as suburbanisation). Nowadays, centre–periphery structures seem to dominate the region. Peri-urbanisation has affected most communes and traditional village structures (compact, ribbon, etc.) are rare, as they are progressively transformed into more uniform clusters of buildings which loose the tentacular aspect often observed in the first phase of urbanisation.

Indices of fractal dimension represent the morphology of the built-up environment as well as its extent. They are – on average – quite independent of the appearance of the built environment. Hence suburban landscapes are characterised by a wide variety of land uses, creating complex and diverse landscapes consisting of a highly fragmented mosaic of different forms of land cover and a dense transport infrastructure.

This paper has introduced a coherent method of identifying classes of settlements with respect to their fractal dimensions. A specificity of the clustering method is its reliance on statistical hypotheses, in particular on the bivariate Gaussian model-based hypothesis. In this context, the estimation of the partition is performed through maximisation of the likelihood of the mixture parameter through the EM algorithm. The BIC criterion allows a model to be selected from among many proposals. Models include the number of clusters, the geometrical feature of the Gaussian class and the possible transformation of the axis. Note that this last kind of model is rarely selected in a formal way in applications, and the specific procedure to perform this selection contributes to the originality of the present paper.

The fractal dimension measured on surfaces gives different information from that measured on borders. Fractal dimensions are of value in studying built-up landscapes as they provide a theoretical basis for the observed features of self-similarity and scale-independence. They also provide a quantitative measure of the roughness or complexity

that could help in the interpretation of maps of built-up areas, and assist the explanation of processes creating built-up landscapes. Linking surface dimensions and perimeter dimensions revealed counterintuitive results, for which a heuristic explanation has been found and verified. Nevertheless, more multidisciplinary research on other regions is required to decide whether the relationships observed here reveal fractal properties of the shapes or whether they are an artefact.

Fractal measures serve to characterise the spatial organisation of urban patterns with unequivocal values. They can be used to measure to what extent the built-up area is distributed in a uniform or a varied way within an urban pattern. When interpreting the results it becomes obvious that a fractal approach to urban patterns helps to improve our knowledge of their spatial organisation, regardless of the extent to which they were planned. Obviously multi-scale pattern organisation is an interesting way of managing the consequences of the new peripheral lifestyle which tends to have good access to different kinds of urban and rural amenities, while simultaneously reducing the risks of a diffuse sprawl which tends to reduce the quality of the environment and generate more and more traffic flow. This empirical work can also help to inform the choices made in urban simulations (see Cavailhès et al., 2002 and 2004).

Acknowledgements (after acceptance)

References

Antrop, M., 1997. The concept of traditional landscapes as a base for landscape evaluation and planning: the example of Flanders Region. Landscape Urban Plann. 38 (1/2), 105–117.

Antrop, M., 2000. Background concepts for integrated landscape analysis. Agr., Ecosyst. Environ. 77, 17–28.

Antrop, M., Van Eetvelde, V., 2000. Holistic aspects of suburban landscapes: visual image interpretation and landscape metrics. Landscape Urban Plann. 50, 43–58.

Arlinghaus, S., 1985. Fractals take a central place. Geogr. Ann. B 67 (2), 83–88.

Banfield, J., Raftery, A., 1993. Model-based Gaussian and non-Gaussian clustering, Biometrics 49, 803–821.

Badariotti, D., 2005. Des fractales pour l'urbanisme ? Quelques pistes de réflexion à partir de l'exemple de Stasbourg-Kehl. Cahiers de Géographie du Québec 49 (137), 133–156.

Bartel, A., 2000. Analysis of landscape patterns: towards a 'top down' indicator for the evaluation of land use. Ecol. Model. 130, 87–94.

Batty, M., 2005. Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals, Cambridge, Massachusetts, The MIT Press,

Batty, M., Kim, S., 1992. Form follows function: reformulating urban population density functions. Urban Stud. 29 (7), 1043–1070.

Batty, M., Longley, P., 1994. Fractal Cities: A Geometry of Form and Function, London: Academic Press, 394 pp.

Batty, M., Xie, Y., 1996. Preliminary evidence for a theory of the fractal city. Environ. Plann. 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. Plann. B 27 (4), 507–519.

Caruso, G., 2002. La diversité des formes de périurbanisation. In Perrier-Cornet, P. (Ed.) Repenser les campagnes. Datar, Editions de l'Aube, 67–99.

Caruso, G., Rounsevell, M., Cojocaru, G., 2005. Exploring a spatio-dynamic neighbourhood-based model of residential behaviour in the Brussels peri-urban area. Int. J. Geogr. Inf. Sci. 19 (2), 103–123.

Carvalho, R., Penn, A., 2004. Scaling and universality in the micro-structure of urban space. Physica A, 539–547.

Cavailhès, J., Frankhauser, P., Peeters, D., Thomas, I., 2002. Aménités urbaines et périurbaines dans une aire métropolitaine de forme fractale. Revue d'Economie Rurale et Urbaine 5, 729–760

Cavailhès, J., Frankhauser, P., Peeters, D., Thomas, I., 2004. Where Alonso meets Sierpinski: an urban economic model of a fractal metropolitan area. Environ. Plann. A 36 (8), 1471–1498.

Celeux, G., Govaert, G., 1992. A classification EM algorithm for clustering and two stochastic versions. Comput. Stat. Data An. 14 (3), 315–332.

Celeux, G., Govaert, G., 1995. Gaussian parsimonious clustering models. Pattern Recogn. 28 (5), 781–793.

Champion, A., 1989. Counter-urbanisation, Edward Arnold, London.

De Keersmaecker, M.-L., Frankhauser, P., Thomas, I., 2003. Using fractal dimensions to characterize intra-urban diversity: the example of Brussels. Geogr. An. 35 (4), 310–328.

De Keersmaecker, M.-L., Frankhauser, P., Thomas, I., 2004. Analyse de la réalité fractale périurbaine: l'exemple de Bruxelles. L'Espace Géographique 3, 219–240.

Dempster, A., Laird, N., Rubin, D., 1977. Maximum likelihood from incomplete data via the EM algorithm (with discussion), J. Roy. Stat. Soc. B 39, 1–38.

Dussart, F., 1957. Geographie der ländlichen Siedlungsformen in Belgien und Luxemburg. Geographische Rundschau IX, pp. 12–18.

Dussart, F., 1961. Types de dessins parcellaires et leur répartition en Belgique. Bulletin de la SOBEG XXX, pp. 21–65.

Falconer, K., 2003. Fractal Geometry: Mathematical Foundations and Applications, John Wiley & Sons, Ltd, West Sussex.

Farina, A., 1998. Principles and Methods in Landscape Ecology, Chapman & Hall, London.

Feltz, C., 2004. Les territoires paysagers de Wallonie, Etudes et Documents, CPDT 4, 68 pp.

Forgy, E. W., 1965. Cluster analysis of multivariate data: efficiency vs interpretability of classifications, Biometrics 21, 768–769.

Frankhauser, P., 1994. La fractalité des structures urbaines, Collection Villes, Anthropos, Paris.

Frankhauser, P., 1998. The fractal approach: a new tool for the spatial analysis of urban agglomerations. In Population: An English Selection, special issue New Methodological Approaches in Social Sciences, 205–240.

Frankhauser, P., 2004. Comparing the morphology of urban patterns in Europe. A fractal approach. In: Borsdorf, A. and Zembri P. A. (eds.). European Cities: Insights on Outskirts. Report Cost Action 10 Urban Civil Engineering, Brussels 2, 79–105.

Frankhauser, P., Tannier, C., 2005. A multi-scale morphological approach for delimiting urban areas. CUPUM 05: Computers in Urban Planning and Urban Management, 9th Conference organised by the CASA, University College of London, 29/6/05–1/7/05, London.

Frankhauser P., Tannier C., Vuidel G., Houot H. (2007). Fractal analysis of urbanisation. Accessibility analyses and multi-scales simulations (in French) Submitted to Transport Policy and paper presented at Paper presented WCTR (Berkeley) in June 2007.

Friedman, H., Rubin, J., 1967. On some invariant criteria for grouping data. J. Am. Stat. Assoc. 62, 1159–1178.

Galster, G., Hanson, R., Ratcliffe, M., Wolman H., Coleman S., Freihage, J., 2001. Wrestling sprawl to the ground: defining and measuring an elusive concept. Hous. Policy Debate 12 (4), 681–717.

Goodchild, M., Mark, D., 1987. The fractal nature of geographic phenomena. Ann. Assoc. Am. Geogr. 77, 265–278.

Gouyet, J.-F., 1996. Physics and Fractal Structures, Masson and Springer, Paris, 234pp.

Haken, H., 1977. Synergetics: An Introduction, Springer, Berlin.

Imre, A., Bogaert, J., 2004. The fractal dimension as a measure of the quality of habitats. Acta Biotheor. 52, 41–56.

Johnson, M., 2001. Environmental impacts of urban sprawl: a survey of the literature and proposed research agenda. Environ. Plann. A 33, 717–735

Lam, N., de Cola, L., 2002. Fractals in Geography, The Backburn Press, Caldwell, New Jersey, 308 pp.

Longley, P., Mesev, V., 2000. On the measurement and generalisation of urban form. Environ. Plann. A 32 (3), 471–488.

Longley, P., Mesev, V., 2002. Measurement of density gradients and space-filling in urban systems. Pap. Reg. Sci. 81, 1–28.

MacLennan, M., Fotheringham S., Batty, M., Longley, P., 1991. Fractal geometry and spatial phenomena: a bibliography. NCGIA Report, 91–1.

Mandelbrot, B., 1982. The Fractal Geometry of Nature, Freeman, San Francisco.

Mattila, P., 1995. Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability, Cambridge University Press, Cambridge.

McGarigal, K., Marks, B.J., 1995. FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure. Gen. Tech. Report PNW-GTR-351, USDA Forest Service, Pacific Northwest Research Station, Portland, OR.

McLachlan, G., Peel, D., 2000. Finite Mixture Models, Wiley, New York.

MRW-DGATLP, 2004. Plan de Localisation Informatique. Ministère de la Région Walonne.

Milne, B.T,. 1991. Lessons from applying fractal models to landscape patterns. In Turner, M.G. and R.H. Gardner, (eds.). Quantitative Methods in Landscape Ecology. Springer-Verlag, New York, pp. 199–235.

Paquette, S., Domon, G., 2001. Rural domestic landscape changes: a survey of the residential practices of local and migrant populations. Landscape Research 26 (4), 367–395.

Russ, J., 1994. Fractal Surfaces, Plenum Press, New York.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.

Schweitzer, F., Steinbrick, J., 1998. Estimation of megacity growth: simple rules versus complex phenomena. Appl. Geogr. 18 (1), 69–81.

Shen, G., 2002. Fractal dimension and fractal growth of urbanized areas. Int. J. Geogr. Inf. Sci. 16 (5), 519–437.

Tannier, C., Pumain, D., 2005. Fractals in urban geography: a theoretical outline and an empirical example. Cybergeo, Article 307.

Thomas, I., Frankhauser, P., De Keersmaecker, M.-L. (2007). Fractal dimension versus density of the built-up surfaces in the periphery of Brussels. Papers in Regional Science, 86:2, 287-307. .

Van Eetvelde, V., Antrop, M., 2005. The significance of landscape relic zones in relation to soil conditions, settlement pattern and territories in Flanders. Landscape Urban Plann. 70, 127–141.

Vicsek, T., 1989. Fractal Growth Phenomena, World Scientific, Singapore, 355 pp.

Ward, J., 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58, 236–244.

Weidlich, W., Haag, G., 1983. Concepts and Models of a Quantitative Sociology, Springer, Berlin.

Wentz, E., 2001. A shape definition for geographic applications based on edge, elongation and perforation. Geogr. Anal. 32 (2), 95–112.

White, R., Engelen, G., 1993. Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. Environ. Plann. A 25, 1175–199.

Wolman, H., Galster, G., Hanson, R., Ratcliffe, M., Furdell K., Sarzynski, A., 2005. The fundamental challenge in measuring sprawl: which land should be considered? Prof. Geogr. 57 (1), 94–105.

APPENDIX $D_{Bord} > D_{Surf}$

The fractal law in Equation [2] shows that the fractal dimension depends on the distance range considered. By replacing D by $D(\varepsilon)$ we get

$$N(\varepsilon) = a\varepsilon^{D(\varepsilon)} + c \qquad [11]$$

The derivative of $N(\varepsilon)$ can now computed with respect to ε in order to get information about the variation in the number of pair correlations $N(\varepsilon)$ and the variations of $D(\varepsilon)$. In this context it is convenient to introduce a new variable $N'(\varepsilon) = N(\varepsilon) - c$ which yields $dN'(\varepsilon) = dN(\varepsilon)$. Moreover, since $D(\varepsilon)$ is no longer a constant, we ought to rewrite the power law as an exponential function:

$$N'(\varepsilon) = a\varepsilon^{D(\varepsilon)} = ae^{D(\varepsilon)\cdot\ln\varepsilon} \qquad [12]$$

In order to compute the derivative of $N'(\varepsilon)$ it is convenient to take the logarithm of the previous relationship

$$\log N'(\varepsilon) = \log a + D(\varepsilon)\ln\varepsilon \qquad [13]$$

which yields

$$\frac{d\log N'(\varepsilon)}{d\varepsilon} = \frac{D(\varepsilon)}{\varepsilon} + \ln\varepsilon\frac{dD(\varepsilon)}{d\varepsilon}$$

We now use the identity $d\log N'(\varepsilon) \equiv \frac{dN'(\varepsilon)}{N'(\varepsilon)}$ and we multiply both sides by $d\varepsilon$. This gives

$$\frac{dN'(\varepsilon)}{N'} = D(\varepsilon)\frac{d\varepsilon}{\varepsilon} + \ln\varepsilon\cdot dD(\varepsilon) \qquad [14]$$

This relationship links the relative variation of the number of correlations $N(\varepsilon)$ with the relative variation of the distance ε. It shows that the second term vanishes when the dimension remains constant, but it also generates an increase in the ratio $dN(\varepsilon)/N(\varepsilon)$ when $D(\varepsilon)$ changes (see also Frankhauser, 1998).

$N_{Surf}(\varepsilon)$ and $N_{Bord}(\varepsilon)$ are then computed for a slightly dilated pattern. At very small distances, the inner surfaces of the clusters are mainly filled up. The correlation analysis will hence detect a relatively high number of pair correlations. The borders are now

rather smooth due to the scale of analysis, the characteristics of the database (see Sections 2.4 and 4) and dilation. Hence, when analysing the borders, the number of correlations $N_{Bord}(\varepsilon)$ remains low i.e. close to that observed for a simple line. The situation is similar to that of a set of compact smooth objects resembling Euclidean objects and we expect that, for a given distance, $N_{Surf}(\varepsilon) >> N_{Bord}(\varepsilon)$. However as the distance parameter,ε, increases, the tentacular character of the patterns comes into play: the border lines become rather tortuous and $N_{Bord}(\varepsilon)$ increases rapidly. This means that the fractal dimension changes, as well as the second term of Equation [14]. Since the tentacles are rather filigree, the number of pair correlation $N_{Surf}(\varepsilon)$ now increases less rapidly than before. Thus, for this range of variations, we expect $dN_{Bord}(\varepsilon) \approx dN_{Surf}(\varepsilon)$. If we now assume not only that $N_{Surf}(\varepsilon) >> N_{Bord}(\varepsilon)$ but that $N_{Surf}(\varepsilon) - c_{Surf} >> N_{Bord}(\varepsilon) - c_{Bord}$, then we obtain

[15]

which yields $D_{Bord} > D_{Surf}$. In any case, according to Equation [14], a large change in dimensions increases the ratio $\frac{dN'_{bord}}{N'_{bord}(\varepsilon)} > \frac{dN'_{surf}}{N'_{surf}(\varepsilon)}$.

Referring to Equation [12], let us call the c-value corresponding to the fractal law of surfaces c_{Surf}, and that of borders c_{Bord}. In all cases, we observe $N_{Surf}(\varepsilon) >> N_{Bord}(\varepsilon)$ and $c_{Surf} < 0$. c_{Bord} can be positive or negative, but $|c_{Surf}| > |c_{Bord}|$ always holds. Hence we always obtain $N_{Surf}(\varepsilon) - c_{Surf} >> N_{Bord}(\varepsilon) - c_{Bord}$.

TABLES

Table 1: Estimated values of *a* and *D* for different distance ranges in Le Roeulx

Surface					Border				
Distance range		$D_{Surf\text{-}d3}$	*a*	η^2	Distance range		$D_{Bord\text{-}d3}$	*a*	η^2
Smallest	*Largest*				*Smallest*	*Largest*			
1	143	1.447	3.84	0.999836	1	143	1.667	0.21	0.999949
21	57	1.525	2.69	1.000000	57	91	1.771	0.13	0.999995
21	99	1.500	3.02	0.999984	21	57	1.701	0.17	0.999967
57	103	1.407	4.84	0.999999	91	143	1.503	0.49	0.999990

Note: Surface and border dimension are estimated after 3 steps of dilation.

Table 2*:* Estimated *D* values in Le Roeulx using the simplified data model

Surface				Border			
Distance range		$D_{Surf\text{-}d3}$	η^2	Distance range		$D_{Bord\text{-}d3}$	η^2
Smallest	*Largest*			*Smallest*	*Largest*		
1	143	1.717	0.996232	1	143	1.350	0.993683
21	57	1.756	0.999196	57	91	1.352	0.999312
21	99	1.733	0.998467	21	57	1.295	0.997422
57	103	1.720	0.994160	91	143	1.375	0.999929

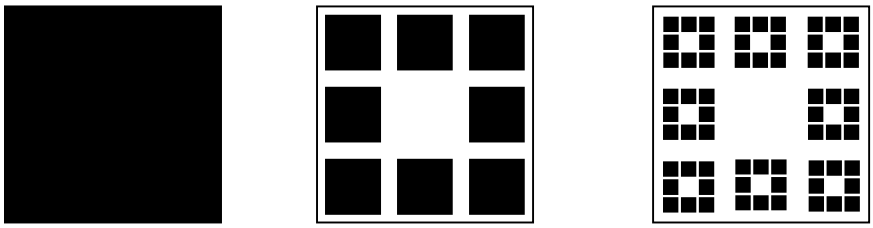
Note: Surface and border dimensions are estimated after three steps of dilation; dataset and distance ranges are the same as in Table 1. In this simplified model *a* is fixed at 1.

Table 3: Results of the fractal classification of communes

Cluster	$D_{Surf\text{-}d0}$	$D_{bord\text{-}d3}$	*n*	Three most representative communes	Typology
1	1.37	1.65	44	Brugelette, *Heron*, Nandrin	Peri-urban I and small cities
2	0.92	1.50	40	*Lierneux*, Havelange, Merbes-le-C.	Rural I: compact isolated hamlets
3	1.50	1.76	49	Pepinster, Saint-Georges, *Blegny*	Peri-urban II and eastern part (Hainaut)
4	1.11	1.59	47	Erquelinnes, Baelen, *Rendeux*	Rural II: hamlets with a linear structure
5	1.68	1.70	40	Ottignies, Châtelet, *Chaudfontaine*	Urban (homogeneous, fully urbanised communes)
6	1.25	1.63	42	Gesves, Jalhay, *Ciney*	Rural III: rural communes with hamlets and one (small) city centre

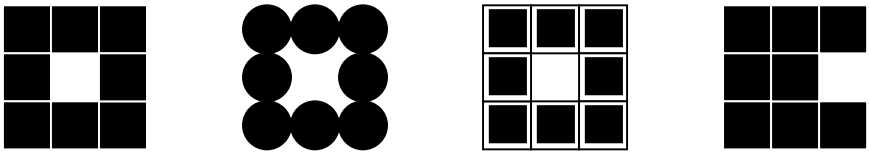
n: number of communes in the class;
names in italics indicate the communes illustrated in Figure 14

FIGURES



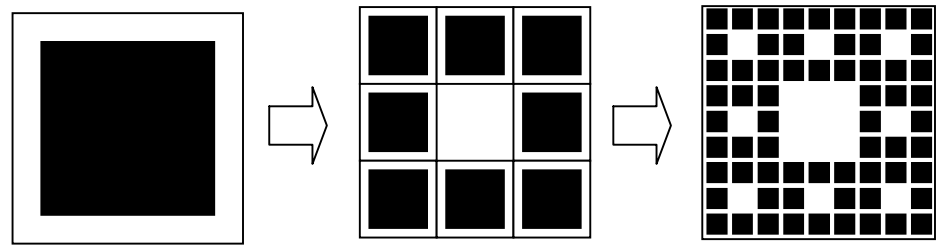
(a) (b) (c)

Figure 1: First steps in generating a Fournier dust



(a) **(b)** **(c)** **(d)**

Figure 2: Same dimension, D, for different initiators ((a), (b) and (c)) or positions of the elements in the generator ((a) and (d)) of a Sierpinski carpet



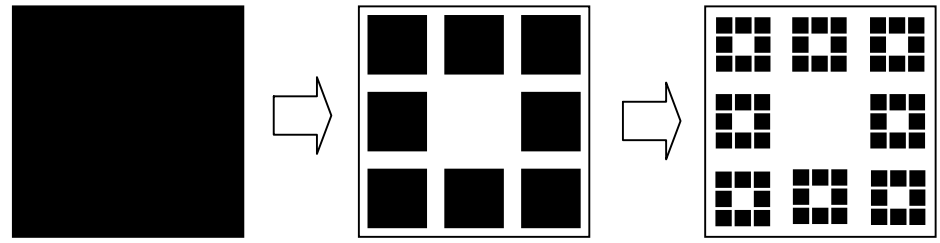


Figure 3: Two generators with identical *N*s, but different reduction factors, *r*, forming a Sierpinski carpet (a) or a Fournier dust (b), depending on the initiator

Note: For the sake of clarity, the lines delimiting the elements have been omitted from the RHS of Figure 3(a).

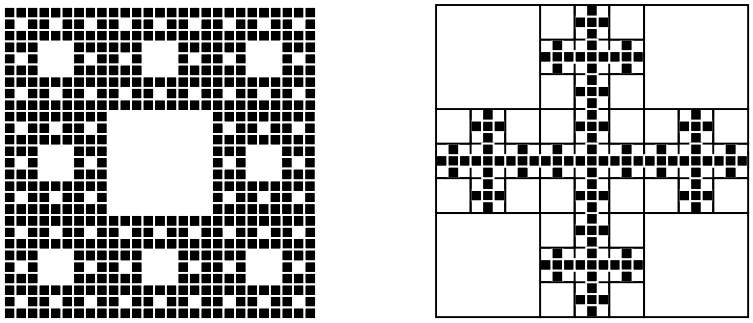
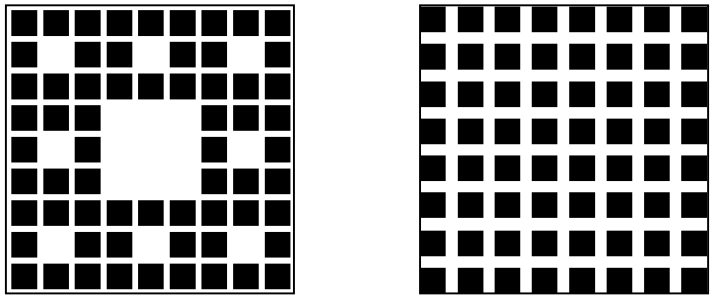


Figure 4: Two Sierpinski carpets in the second iteration



(a) (b)

Figure 5: Two figures, each composed of 64 black squares, having the same density but different fractal dimensions (D = 1.89 in (a) and D = 2.00 in (b))

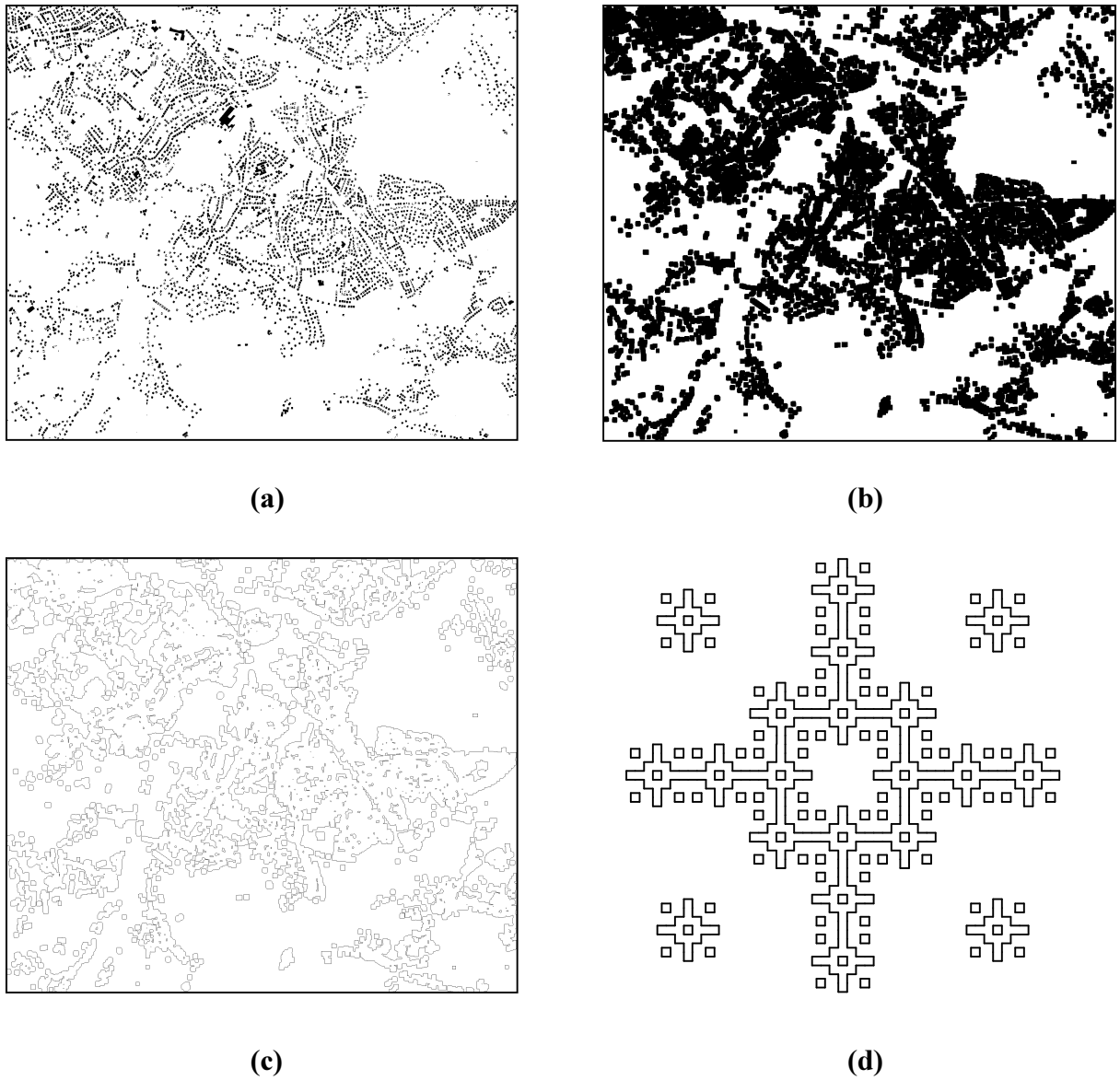


Figure 6: Extracting boundaries by dilation. Figure 6a shows the original urban pattern, Figure 6b the corresponding dilated structure (3 steps), Figure 6c the extracted boundary and Figure 6d a theoretical fractal with similar features to the observed fractal in (c)

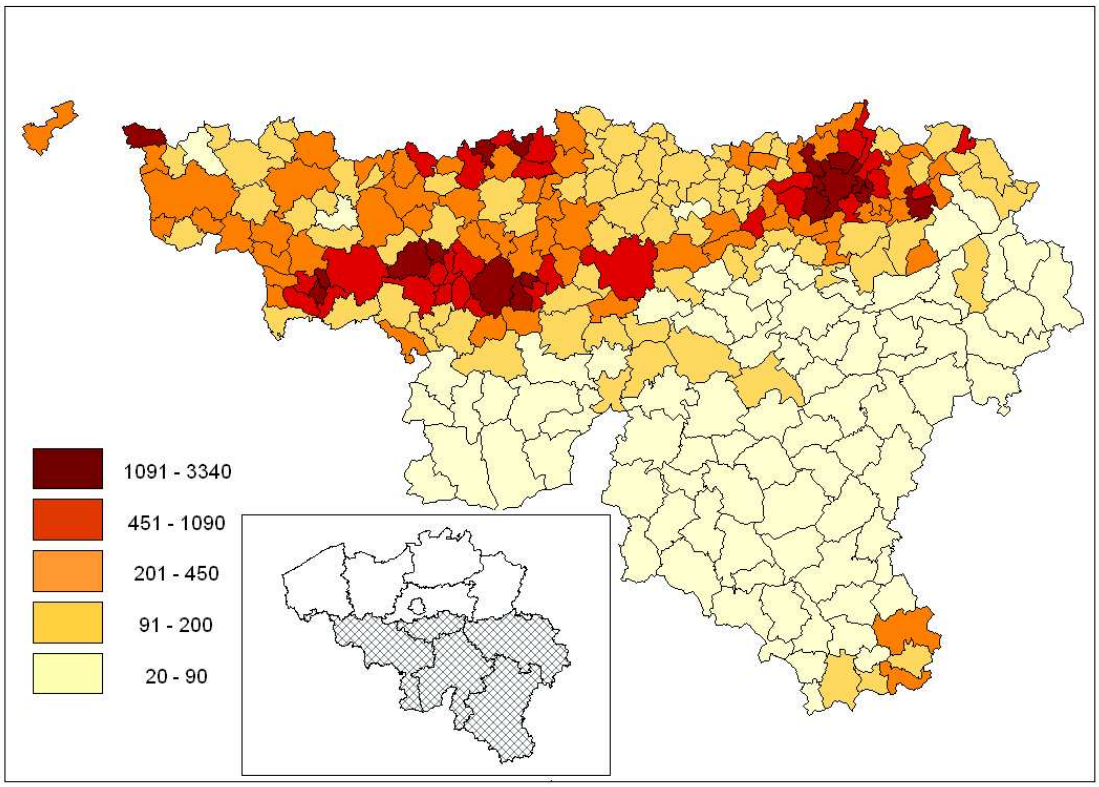
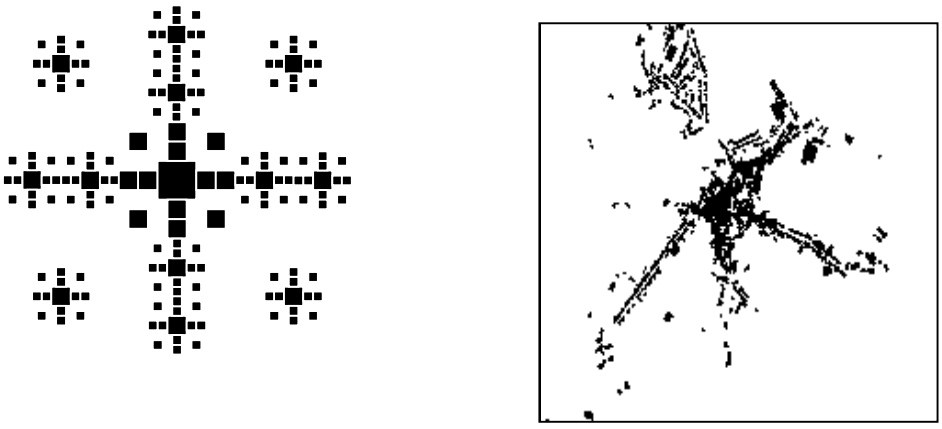
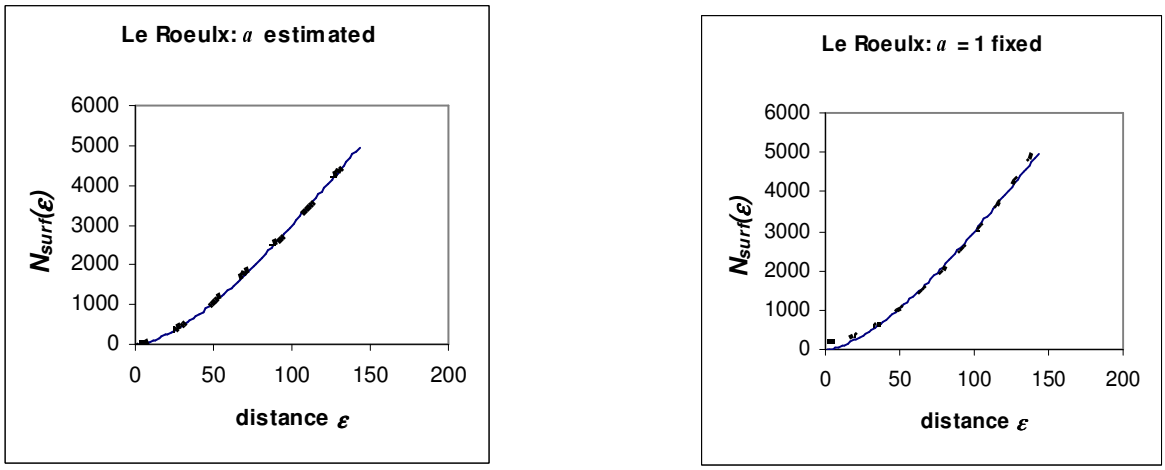


Figure 7: Population density in Wallonia



(a) **(b)**

Figure 8: A hybrid multifractal Sierpinski carpet (a) and a Walloon settlement pattern (b)



(a) (b)

Figure 9: Estimated and empirical results for Le Roeulx with (a) estimated and (b) fixed values of a_{Surf}

Note: Dashed line – estimated results; continuous line – empirical results

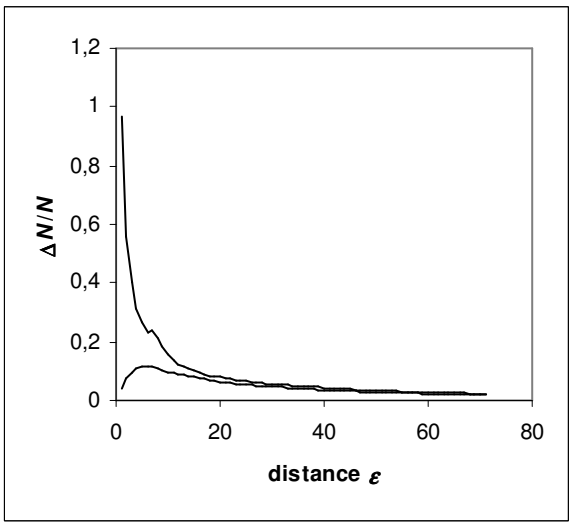


Figure 10: Relative variations ($dN(\varepsilon)/N(\varepsilon)$) for the border (upper curve) and for the surface (lower curve)

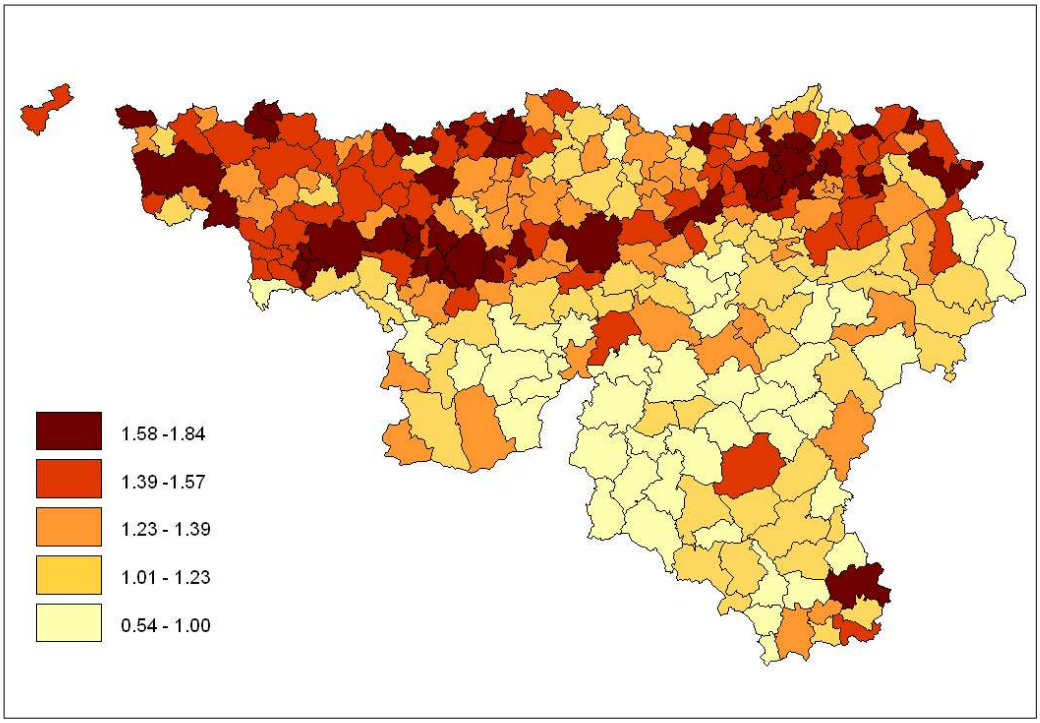


Figure 11: Fractal dimensions of the undilated surfaces ($D_{Surf\text{-}d0}$)

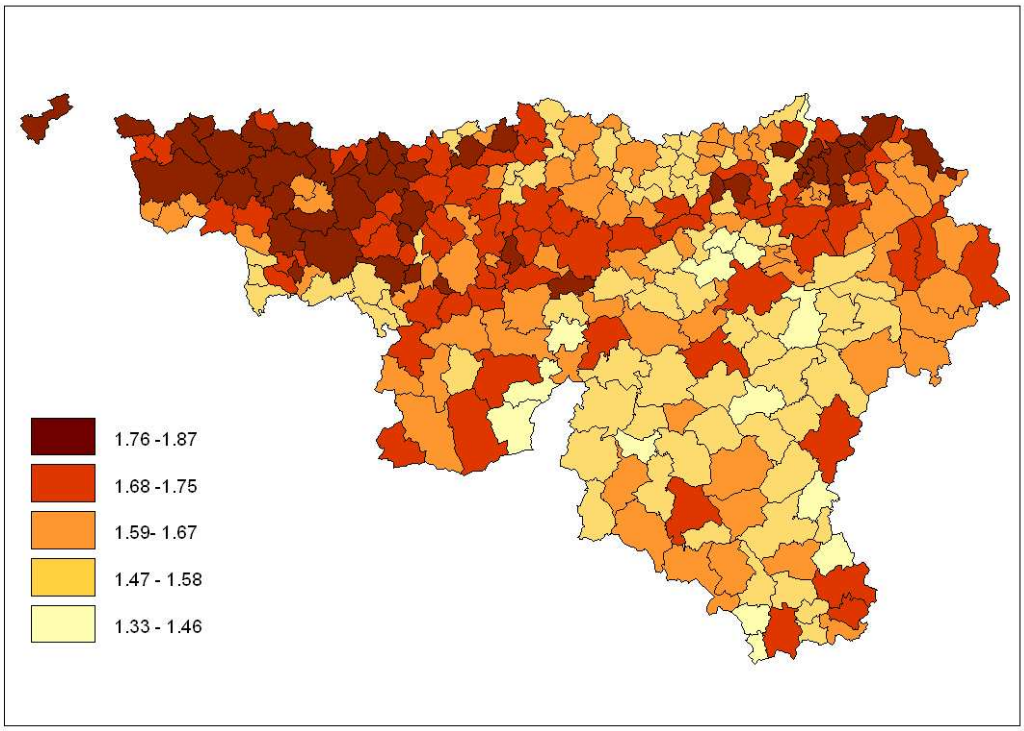


Figure 12: Fractal dimensions of the dilated borders ($D_{Bord\text{-}d3}$)

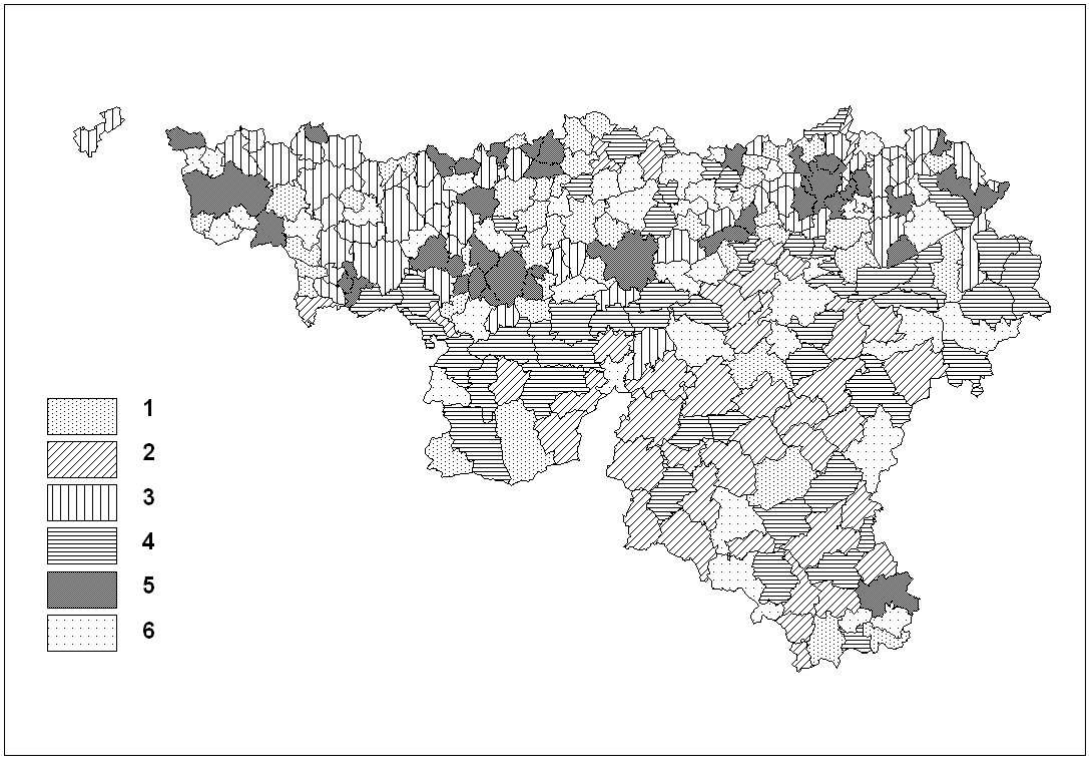
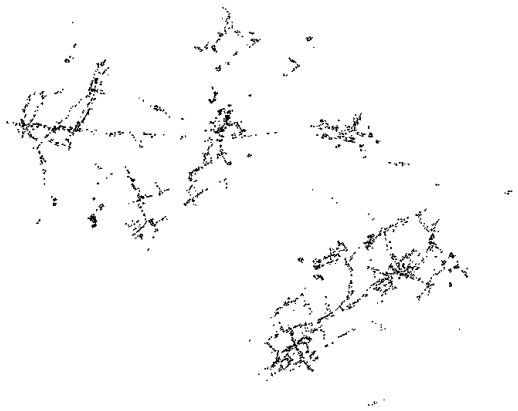
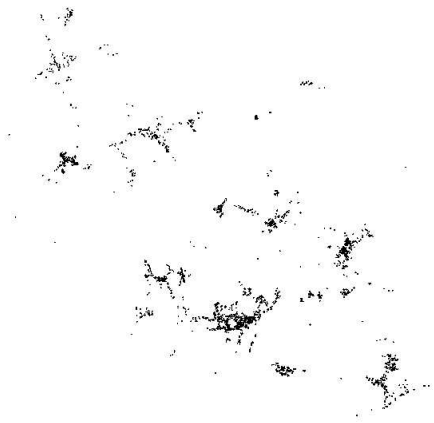


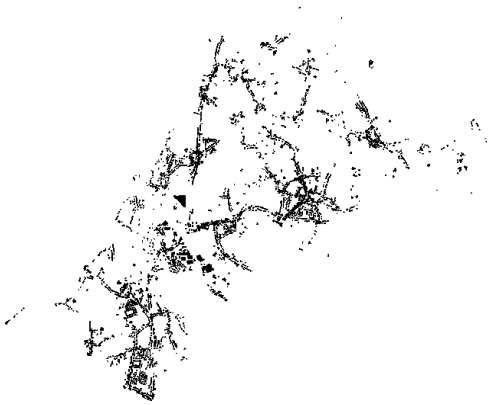
Figure 13: Classification of the 262 Walloon communes in terms of *D* values (see Table 3 for the definitions of the clusters)



Cluster 1: Héron



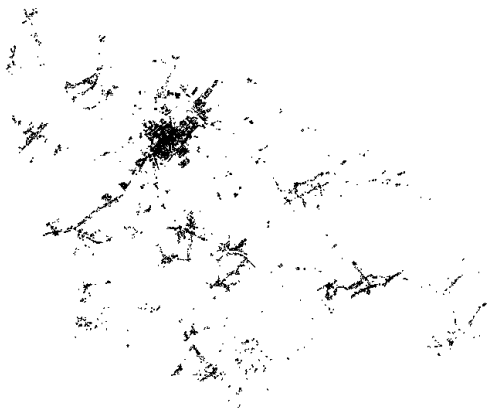
Cluster 2: Lierneux



Cluster 3: Blégny

Cluster 4: Rendeux

Cluster 5: Chaudfontaine



Cluster 6: Ciney

Figure 14: Cluster analysis results: one example of each type of cluster (not to scale)

Source: Plan de Localisation Informatique (MRW-DGATLP, 2004)