Fractals in urban geography: a theoretical outline and an empirical example

Fractales et géographie urbaine : aperçu théorique et application pratique

Cécile Tannier, CNRS (National Centre for Scientific Research), UMR 6049 ThéMA, Besançon, France
cecile.tannier@univ-fcomte.fr

Denise Pumain, Université Paris I Panthéon-Sorbonne, UMR Géographie-cités, France
pumain@parisgeo.cnrs.fr

Prefatory note
This paper is a follow up of a presentation given for the *68th annual meeting of the Society for American Archaeology*, Symposium "*Fractals in Archaeology*", organised by C. T. Brown and W. J. Stemp, Milwaukee, April, 2003.

Abstract
Recently, fractal theory has become popular in urban geography. Actually, its formalisation is compatible with many characteristics of the urban systems: self-similarity in clustering and fragmentation of spatial patterns at different scales, hierarchical organisation, sinuosity of borders, and non linear dynamics. First, we recall how fractal properties can be related to important features of urban morphology just as easily as to the evolution of urban systems. Second, we briefly review the main trends in the application of fractals to urban issues: the description of urban morphologies (built-up areas, distribution of activities, networks, borders…), the simulation of urban growth and settlement systems analysis. A specific application to the question of urban limits will be presented in detail. Issues of relevance and validation will be discussed, especially regarding the combination of different types of spatial structures.

Résumé
La géométrie fractale est devenue récemment très populaire en géographie urbaine. En effet, son formalisme est en accord avec de nombreuses caractéristiques des systèmes urbains : auto-similarité des formes urbaines à différentes échelles ; organisation spatiale hiérarchique ; sinuosité de la bordure urbaine ; dynamique non linéaire.
Cet article s'attache en premier lieu à rappeler en quoi les propriétés des objets fractals peuvent être rapportées à des caractéristiques majeures tant, de la morphologie urbaine, que de l'évolution des systèmes urbains. En second lieu, les principales tendances concernant l'application des fractales à des questions urbaines sont rapidement évoquées. Enfin, une application spécifique s'intéressant à la question des limites urbaines est présentée. La validité et la pertinence des résultats sont alors discutées, notamment au regard de la combinaison de différents types de structures spatiales.

Fractal geometry was developed and has become popular through the work of the mathematician B. Mandelbrot (*1977*). It deals with mathematical objects which exhibit properties of self-similarity (that is, which present the same type of structure at different scales) and which take intermediary dimensions when compared to Euclidean geometrical objects (for instance, while a straight line has a dimension 1, fractal geometry considers lines which are able to fill a surface such as the Peano curve and whose dimensions take values between 1 and 2).

Such mathematical objects are useful for describing spatial forms which are not regular in the sense of Euclidean geometry but which are characterised by alternate patterns of continuity and fragmentation, or some varying degrees of concentration, and include similar structures at different scales of analysis. Geographers have taken a specific interest in this new concept. One famous example is the question of measuring the length of coastal lines (one of the cases first mentioned by Mandelbrot) and the problem of their generalisation in cartography. But most applications refer to the analysis of spatial distributions which are generated by asymmetrical interaction processes between a centre and its periphery, and which reproduce the same way of alternating free and occupied places at different geographical scales.

In this paper, we focus more precisely on the utility of fractal geometry for urban geography especially when taking a global level of analysis (system of cities or a city considered as a global object, but without developing the analysis of networks within cities). After recalling why it is compatible with some of the major principles of urban theory, we briefly review different ways of applying fractal measures and simulation methods to urban problems. We develop a particular application of fractal measures for studying the structuring of urban space and the limits of built-up areas. Unresolved problems will be discussed as well as the question of the usefulness of fractals for social sciences, especially geography.

1. Concepts in urban geography and fractal theory

By escaping rigid rules of Euclidean Geometry, fractal objects allow the development of useful tools for the description of observed spatial patterns. In the case of urban systems, many properties which have been formalised as major concepts of geographical theory can be related to the framework of fractal geometry. Indeed, the main properties of fractal objects are the same as the properties of urban patterns.

1.1 Heterogeneity of spatial distributions

The traditional approach of the spatial distribution of population and activities in geographical space relies on the concept of density (*Haggett, 2001*). This concept is borrowed from physics and refers to a specific concentration level which is typical of a homogeneous milieu. The measure of the density is particularly well suited for analysing and comparing, for instance, the performance of regional agriculture in given conditions of soil, topography and techniques. When applied to rural population it can be interpreted as a yield (it is the only sociological index which has as a denominator a measure of surface and not of population).
Although widely used, the concept of density is not so well adapted to the description of urban milieu. On one hand, as urban population survival no longer relies on the local resources of their site but on more distant advantages of their situation (for instance, linked to comparative advantages in trading networks), the conceptual meaning of density referring to a direct relationship between the urban population and the occupied surface is not so relevant. On the other hand, from a measurement perspective, towns and cities introduce major discontinuities in statistical landscapes of spatial population distributions, since urban average densities are

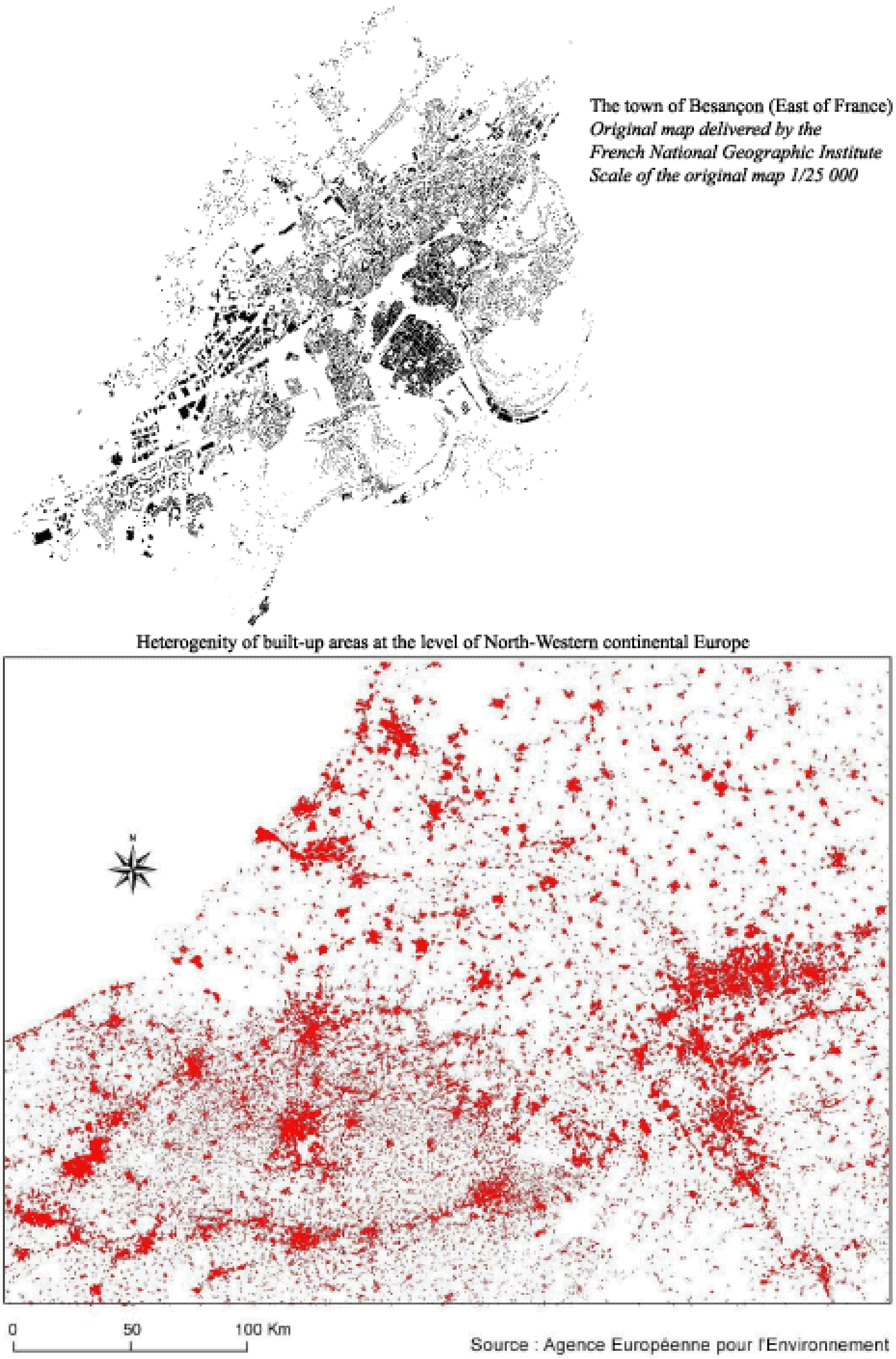
always several times higher than the average surrounding rural densities. Inside towns and cities, there are also major contrasts between urban density levels, linked to the higher rents attached to central or more accessible locations, which give rise to more or less regular heterogeneous patterns of density, generally decreasing from the centre to the periphery and following the land prices gradient.

Alternative measures for analysing the spatial repartition of a phenomenon are auto-correlation functions and concentration indices. The first method calculates the probability similar elements being located either close to each other (spatial autocorrelation measures *(Odland, 1988, Cliff and Ord, 1973)*) or far away (variograms (*Lajoie, Mathian, 1991*)). Such measures are very useful for studying contagion phenomena characterised by a high probability of close areas having the same characteristics. They are also useful to describe repulsion processes inducing a high probability that if a given area has a given characteristic, this characteristic will be missing for the closest areas.
A second alternative is to study concentration or dispersion phenomena (*e.g.* of a type of retail or industrial activities) by using the classical means of spatial analysis, whether on points or on areas. The spatial analysis indexes measure the deviation from a situation of equi-distribution. They suppose a linear relationship (proportionality) between population and surface. But, such a relationship is not present in most cases: the most populated units are very often smaller (in size) than the less populated ones. Thus, concentration indices give different results according to the geographical scale considered for the calculation. Considering the same scale, they even give different results according to the number of spatial units considered (*Bretagnolle, 1996*).
Thus, density measures and spatial analysis indexes all have the major inconvenient to refer to a homogeneous spatial repartition of elements.

Let us now consider the physical morphology of cities. Urban landscapes have become heterogeneous and fragmented especially since they escaped the enclosure of medieval walls and suburbanisation began to shape their spatial form. Clusters of buildings alternate with empty spaces. Local concentrations may take highly variable levels and forms. When looking at land use maps at any scale, the spatial distribution of urban population or activities appears as intrinsically non-homogeneous: smaller and medium-sized clusters appear in the vicinity of much larger clusters (*figure 1*).
Finally, the fact is that theoretical thinking in architecture and planning mainly refers to objects stemming from Euclidean geometry (as the circle or the square) whereas the emerging urban forms with their irregularities and fragmentation are more often better described by fractal geometry. This results from the polygenic character of most cities, which never reflect a unique and homogeneous concept in their construction. Even the most geometric master plan ends up with unfinished irregular parts or has to become inserted in a different spatial pattern of areas, which are built over the following periods.

Fractal structures share the same property of fundamental heterogeneity. Like a city, or like a set of towns and cities, the distribution of their mass in space is never uniform, neither dense nor diluted. Nevertheless, this fragmented distribution is not purely random, since fractal objects are structured following a central organisation principle, self-similarity throughout the scales, which is a property especially useful for studies in urban geography.

Figure 1: Settlement patterns at two different scales



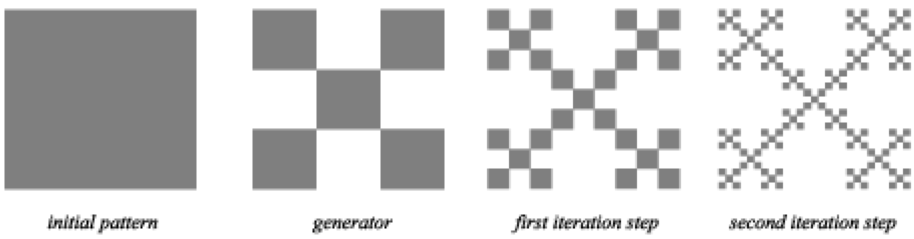
1.2 The centre-periphery pattern and self-similarity

The American geographer Philbrick (*1957*) suggested a systematic description of the structuring of geographical space based on the attractiveness of centres on a surrounding area, of more or less circular shape, at different scales of analysis. A major law of geography is that the intensity of spatial interaction decreases with increasing distance (*Ullman, 1980*). The gravitation model describes the polarisation of the circulation flows around the centres and explains the rather regular spacing of centres for a given type of spatial interaction while a set of similar centres surrounded by their spheres of influence may constitute a homogeneous surface at a higher scale of analysis. For example, a farm-house is point of attraction for the different fields and lands of an agricultural domain, but several farms together make a homogeneous pattern in a village's territory. At a higher level, a market town attracts population and activities from surrounding villages, and a regional capital is a major centre of attraction for several of those elementary farming districts. Because of the very general and dominating character of the centrality principle, which structures spatial patterns whatever the spatial range of interactions, the spatial organisation of geographical space is highly self-similar. *(Arlinghaus & Arlinghaus, 1985)* were the first to mention the fractal as a possible fruitful theoretical framework for interpreting patterns of central places.

Such a nested organisation of centres of different size attracting their periphery (called complementary region) has been formalised in the regular patterns of central place theory by W. Christaller (*1933*). It is linked to an economic explanation based on a series of unequal levels of scarcity or frequency in use (and costs of supplying them on the market) attached to different services and products which are offered to consumers via centres scattered throughout the periphery. Inside towns and cities, the same type of structure has been identified, but it produces different spatial patterns because of stronger differences in land prices and accessibility.

Fractal structures are also characterised by the repetition of the same distribution principle of elements at a multitude of scales. Theoretical fractal forms are built from the iteration of a given pattern of points, curve or surface, at infinity of scales, either by multiplying or by dividing their mass by a fixed quantity at each iteration of the process. But, the same spatial distribution mode does not always mean the same form: that is only the case for theoretical patterns such as Sierpinski carpet or Fournier's dust (*figure 2*). Repetition of the same distribution principle means the repetition of alternating free and occupied places and not necessarily the repetition of the same form. Considering cities, some basic interaction principles involving land prices, accessibility, etc. lead to spatial distributions of elements which seem apparently different, but which are actually similar in terms of the way in which free and occupied places alternate through the scales.

Figure 2: An example of theoretical fractal patterns - The Sierpinski Carpet



The construction principle of this pattern has already been presented in different publications (e.g. *Mandelbrot 1985; Frankhauser 1994*). The above figure shows the first three steps of the construction of such a fractal. The initial pattern is a square of a given length L. This square is reduced by a factor r = 1/3 and N = 5 of these squares of length $l_1 = (r \cdot L) = (1/3\ L)$ are arranged to represent a chess-board. This figure is the generator of the fractal. At the first iteration step each square of length l_1 are replaced by $N = 5$ squares of length $l_2 = (r \cdot l_1) = (1/9\ L)$. The pattern is then made of $N_2 = N^2$ squares of length l_2. The same process is applied to following iteration steps. Thus, for the iteration step n, the length of the squares is equal to $l_n = r^n \cdot L$ and the surface of this object is:

$$A_n = N^n \cdot (l_n)^2 = (5/9)^n \cdot L$$

The fractal dimension is: $D = \underbrace{\log N/\log(1/r)}_{\text{iteration parameters}} = \log 5/\log(1/3) = 1.46$

A result of the self-similarity property of fractals is the regular hierarchical spatial distribution of elements through the scales, which characterises the distribution of central places: self-similarity and heterogeneity (local concentration of elements) lead to centre-periphery patterns.

1.3 Spatial gradients: fractal and non fractal scaling exponents

Self-similarity is a property very often linked with scaling effects, producing regular spatial gradients or hierarchies. A well-known example is the gradient describing the intensity of land use which characterises the internal structure of cities. This gradient was first mathematically described by Clark (*1951*), who formulated an exponential curve for describing the regular decrease in population densities or in land prices from the city centre to the periphery. Density $\rho(r)$ at a distance r from the centre, which has maximal density $\rho(0)$, can be expressed by the following equation:

$$\rho(r) = \rho(0)\, e^{-br} \text{ with } b>0$$

Such a spatial distribution of local densities can also be approximated by a hyperbolic law (*i.e.* an inverse power law)[1]:

$$\rho(r) = \rho(0)\, r^{-a}$$

In principle, the estimated value for a depends upon the size and number of subregions. For instance, in the case of the urban area of Paris, subdivided into arrondissements and communes, the estimated value of a was 2.69 in 1982 and 2.57 in 1990 with the power law. In the case of Lyon, using the exponential model, the b parameter reduces from 0.28 in 1968 to 0.17 in 1990.

In both cases, the absolute value of the parameters a or b measures the rate according to which the density is decreasing over the distance, it is known as an urban density gradient. Both models refer to a non linear but regular distribution of the mass (of population, but it also applies to built-up areas, to rents…) in urban space. The densities are decreasing more quickly than proportionally to the surface when considering more distant outer rings from the city centre. The rapidity of this decrease is however regular and is measured by parameters (b in the exponential model, a in the Pareto model) which have constant value for all the urban structures.

The independence of the parameters a and b from the distance to the city centre is one major characteristic which exists in fractal structures too. It corresponds to the mathematical iteration process which is generating them. It is usually summarised by a measure which is called the fractal dimension (see below). Actually Batty and Kim (*1992*) have demonstrated that there is a strict equivalence between the parameter a of the Pareto model and the fractal dimension D, which are linked through the simple relation $D+a=2$. D and a are designed as scaling exponents.

The fractal dimension D of an urban pattern may be obtained by counting the number of built-up elements (or resident population) at several scales and then, by fitting a fractal law. Such a law can be written as following:

$$N(\varepsilon_i) = c + \varepsilon_i^{\,D}$$

where c is a constant, ε_i, the analysis level (*i.e.* the considered distance between the elements) and N, the number of counted elements.

The Pareto model expresses the fact that the largest elements of a statistical distribution are much less numerous than the smallest ones and the parameter a is a measure of the inequality of the distribution of the elements with respect to their number and their size. The Pareto

[1] In social sciences, a hyperbolic law is most often designated as a “Pareto model”, referring to the researcher (Pareto), who had the idea of using a hyperbolic equation for representing the distribution of incomes of a population.

model applied to urban densities is close to a fractal law because it considers a heterogeneous spatial distribution of the elements, just like a fractal law does. But the fractal dimension of a pattern is an indicator of the heterogeneousness of a spatial repartition, at a multitude of scales whereas the Pareto model is non-scalar (or uniscalar).
Hence, some precision is required: even if the density function can be derived from a spatial organisation of a hierarchical nature, the reverse would not be the case. In other words, if a hierarchy is observed, then it is possible to determine a gradient which describes (measures) the change between one level and an other... But the existence of a gradient does not necessarily imply a hierarchical spatial organisation. Indeed, a gradient is a purely descriptive approach including no reference model, and no explanation. Basically, a gradient is the derivative of the incremental change of something.
However, in the case of the Pareto model applied to the urban densities, the formalisation *implicitly* refers to a radioconcentric model of the city. The difference with a fractal distribution of elements is the *explicit* geometrical nature of such a model, which is intrinsically hierarchical.

At another scale of analysis, inverse power laws are also very frequently used for modelling the hierarchical organisation of urban systems. Known as Zipf's *rank size rule*, this model describes the distribution of the number of towns and cities according to their population size as a Pareto function. According to Zipf's notation, the population P_i of a town or a city is inversely related to its rank R_i in the system of cities by the following power law:

$$P_i = K / R_i^{\,a}$$

Zipf's law is obviously like the Pareto model a hyperbolic law, and the same analogy with a fractal distribution can be derived in that case. One of the first papers about fractals in geography (*Arlinghaus, 1985*) suggested that the geometry of central places is a subset of fractal geometry and that an iterative fractal process could generate all possible systems of central places. N. François (*François et alii, 1995*) has demonstrated it for Christaller's models and applied measurements of fractal dimension to the French system of towns and cities.

1.4 Scaling and geographical scales

A clarification has to be made regarding what is called the hierarchical structure of a geographical system. A first meaning of this term is that a collection of geographical objects (sub-systems) are strongly differentiated by their size (which may be measured by the number of smaller elements that each subsystem contains). This scaling effect can be expressed by a statistical distribution following a Pareto law, or measured by a single fractal dimension which can characterise the whole system. A second meaning of a hierarchical system relies on the concept of geographical scale. Geographical objects may be defined as multi-scalar structures, and their relations can be observed meaningfully at different scales of analysis because significant properties appear only at given levels of observation. For instance, an urban system can be conceptualised at three levels: at the individual scale, there are urban actors or agents (as residents, firms, political bodies, pressure groups…); through their interactions, they generate what is called a "town", or a "city", which is a different geographical object, whose aggregated properties cannot simply be derived from the mere addition of individual characteristics. In the same way, interacting towns and cities define at a third level of observation a new type of geographical object known as an "urban network" or "system of cities", which is characterised by new emerging properties (as the hierarchical structure according to Zipf's law and our first definition). In that meaning, even if fractal structures can be observed in both cases, fractal dimensions are not the same: whereas their values are usually comprised between 1 and 2 at the city level (*Batty, Longley, 1994;*

Frankhauser, 1994), they oscillate between 0 and 1 for systems of cities (*François et alii, 1995*). This reflects two different ways of structuring geographical space, for different purposes in terms of location and interaction, the intra-urban organisation of activities on the one hand and inter-urban connections on the other (*Bretagnolle et alii, 2002*).
Also considering only the intra-urban spatial organisation, the combination of different types of fractal behaviours at different scales of analysis can often be observed. In practice, it is not easy to separate the local, more or less random fluctuations around an estimated fractal dimension, and a systematic combination of different processes which can lead to multifractality. Adapted methods are nevertheless likely to improve our understanding of such complex cases.

1.5 Fractal aspects of urban growth

Several aspects of urban growth are in complete agreement with the fractal description of towns. The first and simplest observation that can be made is that, the more a city spreads in surface the more it appears as fragmented and shredded.

The second observation has emerged from studies relating the built-up surface of a set of urban areas to the length of their border (*Batty and Longley, 1994*; *Frankhauser, 1994*). If those areas were simple geometrical objects, their border would be characterised by the dimension 1 and their surface by the dimension 2. But although the observed relation between border and surface was regular, the ratio surface to border was about 1.05, which is in contradiction to Euclidean geometry… But corresponds to fractal geometry.
Such a phenomenon is explained by the very lengthening of the urban border, where it tends towards a complete coverage of the space, close to a plane. It is possible to draw a parallel with observations related to the evolution of the towns. We know that to a specific spatial distribution of the activities corresponds a specific way of people acting on this space. In that sense, the very lengthening of the urban border may partly result from the fact that every person living in a suburban area wants to live close to a green area. Indeed, some examples of urban plans were conceived following the principle that each building should be connected both to the transportation network and to a green area. When implementing this in a fractal manner, the whole population of a city can take advantage of the proximity of the natural areas without spending too much time reaching other more central amenities. This idea that each building is part of the border of the whole urban area exactly corresponds to the fractal geometry of the Sierpinski carpets: because such structures tend to decompose themselves into isolated elements even though forming clusters, the length of their perimeter tends to infinity whereas their surface tends to 0.

Thus, the sinuosity of the urban border provides a way to improve the accessibility of the population to the amenities. But the sinuosity of the urban border is also a property of the urban patterns arising from the behaviour of residents. Residents of an urban area tend to preserve this property by preventing other people settling near to their house and hampering their access to green areas. For that, they may lobby and organise their resistance. These observations support the hypothesis that the interactions between urban planning and self-organising processes lead to fractal cities (*Frankhauser, 1994; Salingaros, 2003).*

More generally, there are obvious analogies between the incremental character of urban evolution and the way fractal forms are generated, through iterative mathematical processes. Batty and Xie (*1996*) relate scaling laws of residential patterns in six American cities to the degree to which space is filled and the rate at which it is filled, by comparing the observed fractal dimensions and the ones resulting from a stochastic process of diffusion (Diffusion Limited Aggregation model). As fractal objects may be generated by non linear dynamic processes, a fruitful research programme is to identify possible social processes leading to different urban forms and to simulate how they may generate fractal patterns or not.

2. Some applications of fractals to urban questions

Applications of fractal geometry in the urban field are now too numerous to be completely reviewed here. We have selected a few which seem representative of the main research currents.

2.1 Description of urban morphologies

The most frequent use of fractal dimension in urban geography has involved measuring the fractal dimensions of urban patterns, aiming at finding new descriptions of the variety of urban morphologies. The morphology of urban patterns is analysed following principles from fractal geometry.

Such analysis relies mainly on the study of the built-up surface of cities and shape and length of their border. Three main sets of results can be obtained:

1) The verification of the hierarchical nature of the spatial structure and the characterisation of this hierarchy;
2) The identification of thresholds in the spatial organisation of the city;
3) The determination of the number of different types of spatial organisation (for instance, connected and weakly hierarchical built-up clusters when considering an analysis window of length from 0 to 200 meters, then non connected and more hierarchical built-up clusters for an analysis window greater than 200 meters). Such results could be related to the multifractality of an urban structure.

The identification of these potential uses of fractal geometry for the analysis of the urban patterns raises two types of questions:

- *Which properties of urban patterns are revealed by the different measures of fractal dimensions?*

  E.g. if the border of a city is characterised by a very high fractal dimension, it means that this border is full of tentacles. Thus, the very extension of such a border allows the access to free spaces (mostly green spaces and roads) for almost all the buildings.

- *What reflects these properties in terms of individual behaviours?*

  For instance, the very high number of tentacles of an urban border could mean that everyone has tried to settle as near as possible from a green area and then, that they try to maintain this situation.

Answering these two questions could allow the identification of types of city or urban patterns with well identified properties.

Actually, fractal dimension measures are a good instrument for a global comparison of the morphology of cities: they are more homogeneous in the case of American or Australian cities (fractal dimensions near to 2), more variable for European cities or more generally for very polygenic cities characterised by their high density gradients from the town centre to the periphery (fractal dimensions between 1 and 2, but nearest to 1) (*Frankhauser, 1994; Batty and Longley, 1994*). However the number of comparable measures is not sufficient to obtain a clear classification of the cities of the different parts of the world. Moreover, the results obtained by fractal analysis are highly dependent on the generalisation methods of the maps representing the built-up surfaces that are used for the measurement of fractal dimension.

In addition to static analysis of urban forms, the comparison of the fractal measures over time may throw light on the urban growth process. Studying the evolution of the fractal dimensions of a city in the course of time shows how the urban pattern is progressively self-organising, following a centre to periphery gradient. The structuring of the peripheral areas often occurs a long time after the emergence of the first buildings in the suburbs. A set of fractal analyses of urban patterns across time have shown that urbanised space is increasingly strongly organised around a central cluster. Moreover the urbanisation is accompanied by a self-structuring process which appears in the growing regularity of the curves resulting from the fractal analysis, despite the fragmented morphology of the urban patterns (*Frankhauser, 1998*).

Now, even if fractals are mainly used in urban geography for identifying different forms of cities and of urban growths, some research also tackles the question of the patchwork of intra-urban patterns (*Batty & Xie, 1996; Frankhauser, 1998; Frankhauser & Pumain, 2002*). In that field of research, the analysis recently undertaken by M.L de Keersmaecker, P. Frankhauser et I. Thomas (*2003)* and *(2004*) are particularly interesting. On the basis of statistical analysis of an exploratory nature, they tried to determine if the fractal dimension is a useful index for distinguishing either urban wards (*de Keersmaecker et al., 2003*) or types of peri-urban built-up patterns (*de Keersmaecker et al., 2004*). Indeed, they showed firstly that different fractal dimensions measure complementary aspects of the structure of the urban and peri-urban built-up pattern, secondly that interesting statistical associations can be found between fractal dimensions and the structure of the housing market, the rent, the distance to the city centre, the income of the households as well as some planning rules.

2.2 Simulation of urban spatial dynamics

Analysis and measurement of urban morphologies led to the conception of urban models which simulate urban growth and are able to reproduce the observed properties of the urban spatial patterns. In that field of application, fractals have two different kinds of contributions. They can be used to control the results of simulations: they help to say if the results are realistic or not (*White et alii, 2001; Engelen et alii, 2002*). This is the case for the dynamic model of land use developed by R. White and G. Engelen (*1994*) for Cincinnati. But fractals can also be used as basic principles to generate urban forms.

Indeed, several authors have suggested urban growth models based on fractal rules (*Batty, Longley, 1986; Batty et al., 1989; Markse, Halvin, Stanley, 1995*). Cellular automata are frequently used as simulation tool for modelling urban growth or land use changes, whereas available physical growth models (Eden, DLA: Diffusion Limited Aggregation) could be profitably substituted by more detailed and realistic models of spatial evolution dealing with social processes. As an example, we briefly describe a model developed by E. Bailly (*Bailly, 1999)*. To start with, we have a raster image of an urban pattern made up of two types of pixels: black pixels which represent built-up spaces and white pixels representing non built spaces. An iterative fractal growth model (the DLA model) is applied to the image. At each iteration step, new built-up pixels appear under the constraint that their location is compatible with the fractal nature of the simulated pattern. Other non fractal constraints have been integrated into the model, accelerating, slowing down or preventing the apparition of the built-up areas (rivers, slopes declivity, exposure…). When applied to the town of Marseilles (South of France) in 1930, the pattern simulated by the model presented a global form very similar to the one of Marseilles in 1990's. But locally, the simulated and the real patterns could be very different.

Following the same direction, it would be particularly interesting to provide several models of fractal growth allowing the simulation of urban patterns with well differentiated characteristics. Thus, it could be possible to simulate different conceivable evolutions of an original urban pattern, each of the simulations corresponding to a particular vision of the urbanisation process (*e.g.* urban intensification or sprawl, increasing or decreasing hierarchy…).

Very recently, J. Cavailhès *et al.* (*2004*) also presented the application of a residential location model (standard in urban economics) on a spatial support provided by fractal geometry: on the one hand, a Sierpinski carpet is used to render a nested hierarchy of the rural and urban places within a metropolitan area. On the other hand, households maximise a utility function which portrays the households' taste for variety in urban and rural amenities. Such a modelling uses the fractal approach to replace the Euclidean spatial representation of the city (*i.e.* the "Thünian city") by a fractal one, which is closer to the actual observed reality. A particularly interesting idea developed in the paper is that the "Thünian city" appears as a limit case for the "fractal city".

3. An empirical example: a fractal analysis of the urban pattern of Basle

We develop here in more detail some elements of a study recently undertaken by C. Tannier and B. Reitel[2] in the framework of a contract directed by P. Frankhauser[3] for the French Ministry of the Public Works[4]. It deals mainly with the morphological evolution of the urban area of Basle[5] in the course of last century. The available data are images of the urban pattern at three dates 1882 – 1957 – 1994 (*Appendix 1, 2 and 3*).
The analysis of the images aims to explore the ability of fractal measures to characterise the process of urban sprawl. The ambition is to provide a set of analyses which may be used for comparing the urban realities of a variety of countries by using a unique methodological tool.

3.1 Method of analysis

The basic tool of this application is software called *Fractalyse*[6], which has been developed especially to measure the fractality of cities.
Fractalyse offers different methods to measure the fractal dimension of an image. But, whatever the chosen method, the general principles are always the same:

1) The material source is a raster image of an urban pattern. This image is composed of two types of pixels: black pixels for representing built-up areas and white pixels, which represent non built-up areas (free spaces).
2) The analysis goes step by step following an iteration principle. At each iteration step, the analysis involved counting the number of black pixels (built-up pixels) contained in a counting window. From one step to the next, the size of the counting window is enlarged. By doing that, we artificially change the level of analysis of the image. So, for each analysis we have two elements varying according to the counting step (iteration step) (i):

- the number of counted elements (which is roughly the number of black pixels present in the window) (N)
- the size of either the counting window or the reference element (ε)[7].

[2]Research team *Image et Ville*, Strasbourg, France

[3]Research team *ThéMA*, Besançon, France

[4]Title of the scientific report: *Morphologie des Villes Emergentes en Europe à travers les analyses fractales*, March 2003. The report is downloadable at the following address: http://thema.univ-fcomte.fr/article67.html

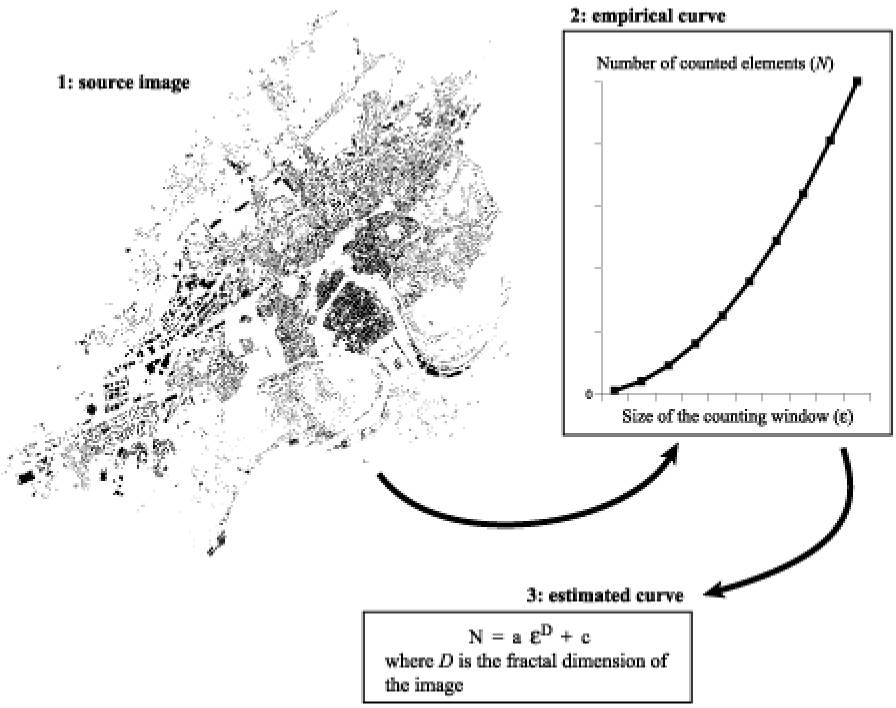
[5] Basle is a frontier urban area which size is about 600 000 inhabitants located over three countries: town centre in Switzerland; extensions in Germany and in France.

[6]This software has been developed by Gilles Vuidel in the frame of the contractual work for the French Ministry of the Public Works. If you want more information about *Fractalyse*, please consult the website of the research team ThéMA: http://thema.univ-fcomte.fr, heading *"Research teams" -> "City, mobility, territory".*

[7]Series of measures of different sizes ε_i are an analogy to the length l_n of the elements in the constructed fractals.

3) Then, we obtain a series of points that can be represented on a Cartesian graph. The Y-axis corresponds to the number of counted elements (N) and the X-axis corresponds to the size of the counting window or to the size of the reference element ε, with ε increasing from step to step (*figure 3*).

Figure 3: How to calculate the fractal dimension of an image



4) Mathematically, the series of points is a curve (named the empirical curve). The next stage of the analysis is to fit this empirical curve with another one, the estimated curve. If the empirical curve follows a fractal law, the estimated curve has the form of a power law (parabolic or hyperbolic).

$$N = \varepsilon^{D} \text{ or } N = \varepsilon^{-D}$$

A non linear regression is used to find the power law which best fits the empirical curve[8]. Because an image is not a pure fractal (it is not a continuous function but a discrete and finite one), it is only possible to approximate the fractal law. It explains why we do not estimate directly the fractal law $N = \varepsilon^{D}$ but a generalisation of it $N = a\varepsilon^{D} + c$. The quality of the estimation is quantified using a correlation coefficient. If the fit between the two curves (empirical and estimated ones) is bad, two conclusions are possible: either the pattern under study is not of a fractal nature or it is of a multi-fractal nature. In the second case, the empirical curve has to be divided into several portions, each of them corresponding to a different estimated curve (*i.e.* according to the considered portion of curve, the non linear regression gives different values for the three parameters a and D and c).

5) The exponent D of the estimated curve is the fractal dimension.
The parameter c corresponds to the point of origin on the Y-axis. Its absolute value may be very high. The parameter a is called the "pre-factor of shape". It gives a synthetic indication of the local deviations from the estimated fractal law. In the case of a mathematical fractal structure a should be equal to 1. In some cases a is equal to 0.5 or 3. If

[8] D is often estimated by using a double logaritmic representation of the power law but here it has been chosen to minimise the least square deviations by means of a non-linear regression.

its value goes over 10 or beyond 0.1 the fractality of the structure under study is not confirmed.

We may here emphasise that the estimations of the fractal dimension of a structure result from an empirical process. Indeed, it is possible to obtain a great variety of estimations of the fractal dimension stemming from a unique empirical curve. Different methodological choices lead to different estimations of the fractal dimension. This has to be taken into account when analysing the results.

For studying the morphological evolution of the urban area of Basle we used two types of methodological approaches which provide complementary insights on the fractality of the urban patterns. The first method is the calculation of the fractal dimension of the images by using the correlation analysis. The second one is based on an iterative transformation of the images (step by step dilation) and a representation of some information about the transformed images on a two-dimension graph for each step of the iteration. This second approach provides no calculation of fractal dimension, but results from a multi-scalar reasoning on a typical fractal nature.

- *Correlation analysis*

Each point of the image is surrounded with a small squared window. The number of occupied points inside each window is enumerated. This allows the mean number of points per window of that given size to be calculated. The same operation is applied for windows of increasing sizes.
The X-axis of the graph represents the size of the side of the counting window $\varepsilon = (2i+1)$. The Y-axis represents the mean number of counted points per window.
(Because the theory underlying the correlation analysis considers the simultaneous presence of two points at a certain distance, *i.e.* the mean distance between a pair of built-up pixels, the correlation dimension is a second order fractal dimension. In a multi-fractal theoretical framework, this correlation dimension should be extended to a series of three, four or more points).
In the case of Basle, we applied the correlation analysis to the built-up surface of the area (*appendix 1, 2 and 3*) as well as to its border line (*appendix 4*). It is interesting to estimate not only the global fractal dimension of each image, but also the fractal dimensions for several portions of the empirical curves[9]. Actually, whereas the fractality of a structure is clear when the adjustment between the empirical curve and the estimated curve is good, a structure is characterised by the combination of different types of fractal behaviour when the fit between the two curves remains good after having segmented the curve into several portions.

- *Step by step dilation and extraction of information about each dilated image*

The principle of the dilation is to surround each occupied point with a black border, the size of which increases at each step of iteration. At the beginning (non dilated image), the reference element (also called "structuring element") is the pixel. During the first dilation, each pixel is surrounded by a border of one pixel width. Then, the reference element is a square of 3^2 pixels size. At the second iteration step, each pixel is surrounded by a border of two pixels width. The structuring element is then a square of 5^2 pixels size. And so on… As the size of the squares gradually increases, the details smaller than the size of the structuring element are overlooked. Thus, we gradually obtain an approximation of the original form.
In the case of Basle, we applied a step by step dilatation to the three original images and we extracted two types of information:

- the total length of the border of each dilated image,
- the number of clusters of built-up pixels at each dilation step.

[9] Curves of scaling behaviour are used for identifying relevant thresholds and thus, distinguishing different segments of curves.

Then our study is based on two types of results: two-dimension graphs and fractal dimension values. The graphs represent either the evolution of the length of the border of the built-up area at each step of the dilation, or the evolution of the number of clusters of built-up pixels through the dilations. The fractal dimensions result from the correlation analysis of the border of the built-up area and from the correlation analysis of the built-up surface of the urban area.

3.2 Evolution of the border of the urban area

- *Correlation analysis applied to the border*

In 1882, the fractal dimension is nearest to 1 than to 2 and reveals that the border of the urban area was on the whole not very tortuous at that time. In addition, the high fluctuations of the fractal dimensions when changing the limits of the zone under study (*i.e.* the bounds of the estimation) characterise the diversity in shape of the border at the local level (*table 1*).

Table 1 - *Fractal correlation dimensions - Borders of the urban area*

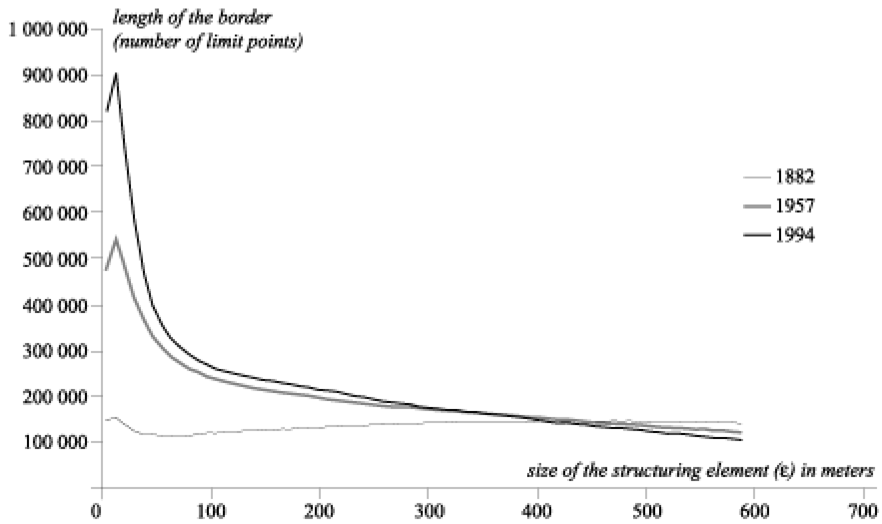
dates	*Bounds of the estimation*					
	1 - 1069	**1 – 57**	**57 - 183**	**183 - 541**	**541 - 757**	**757 - 1069**
	Total curve.	1 to 240 m.	240 to 775 m.	775 to 2300 m.	2300 to 3200 m.	3200 to 4520 m.
1882	1.3	1.6	1.1	1.5	1.2	0.8
	1 - 1069	**1 – 83**	**83 - 449**	**449 - 1069**		
	Total curve.	1 to 350 m.	350 to 1900 m.	1900 to 4520 m.		
1957	1.7	1.8	1.6	1.7		
	1 - 1069	**1 – 83**	**83 - 449**	**449 - 589**	**589 - 1069**	
	Total curve.	1 to 350 m.	350 to 1900 m.	1900 to 2500 m.	2500 to 4520 m.	
1994	1.7	1.8	1.6	1.7	1.9	

In comparison, the border in 1957 appears more tortuous (higher fractal dimensions, close to 1.7) but also more homogeneous through scales (weak variations of the dimension when considering different bounds of estimation). The spatial extension of the urban area happened mostly in the valleys and along the main transportation axis (tramways and railway). Thus, the border has become tentacular and covers more space than in 1882.
Between 1957 and 1994 this trend was only slightly reinforced, which explains that the fractal dimensions are very similar at the two dates. The general form of the border in 1994 is very close to the one in 1957 in a general context of a higher consumption of space. The only difference is the estimation of the fractal dimension of 1.9 for a radius of the correlation larger than 2 500m. The border has become so tortuous, that it covers the space just as a surface does. It indicates a more pronounced urban sprawl in 1994 than in 1957.

- *Evolution of the length of the urban border through the dilations*

On figure 4 we have plotted the number of counted elements in ordinate (number of points belonging to the limit of urbanised area which appear in the counting window) and on the X-axis the size of the dilation. The first point on the X-axis is 4.23 m. and corresponds to the initial size of the non-dilated pixel. For this value of 4.23 on the X-axis, the corresponding value on the Y-axis is the total length of the border of the non-dilated image of the urban area. The total length of the initial border varies greatly between 1882 (150 362 limit points), 1957 (477 686 limit points) and 1994 (819 700 limit points). The first dilation step is characterised by an extension of the border for each of the three curves: the clusters, which were initially constituted by isolated buildings, grow bigger; their perimeter grows longer too without enough fusion of clusters happening to decrease the total length of the border. Clear differences may be observed between the shape of the curves of 1957 and 1994 on the one hand, and the shape of the curve of 1882 on the other hand. But, the differences dwindle in the course of the dilations.

Figure 4: Evolution of the length of the urban border with the dilations



The curve of 1882 indicates first a decrease in the length of the border, for ε values comprised between 12 and 40 m, because inside the city the built-up units are aggregated at the next steps of the analysis, whereas for longer distances this process is compensated by the rejoining of further settlements in the outskirts, which tend to elongate the total border.

The curves of 1957 and 1994 are more similar. The general morphology of the whole urban area, although it was expanding, did not change much between these two dates. As early as the second step of dilatation, many built-up elements are aggregated and the length of the border sharply decreases, while above the 85 m threshold, the buildings are more distant from each other and do not aggregate so rapidly.

This type of analysis could be used for comparing sprawling processes for different cities. The longer the initial border, the less compact is a town. A steep curve slope indicates that numerous settlements are close enough for aggregating at further steps of the analysis and coins therefore urban sprawl. A variety of shapes of curves could be related to different types of urban growth.

3.3 Evolution of the built-up structure of the urban area

- *Correlation analysis of the built-up surface of the area*

On table 2, fractal dimensions are higher in 1957 and 1994 than in 1880, which reveals on the whole that the repartition of built-up areas have become more homogeneous over time.

In 1882, the built-up area is highly contrasted. The computed fractal dimensions decrease sharply for the highest values of ε (between 2 300 and 4 520 m), revealing that the spatial organisation becomes like a Fournier's dust (*d* value is below 1). This corresponds to the numerous villages which are distant from each other.

Fractal dimensions in 1957 and 1994 are higher (closer to 2) and keep similar values for different estimation intervals, which mean that the built-up area has become more homogeneous.

Table 2 - *Fractal correlation dimensions – Built-up surface of the urban area*

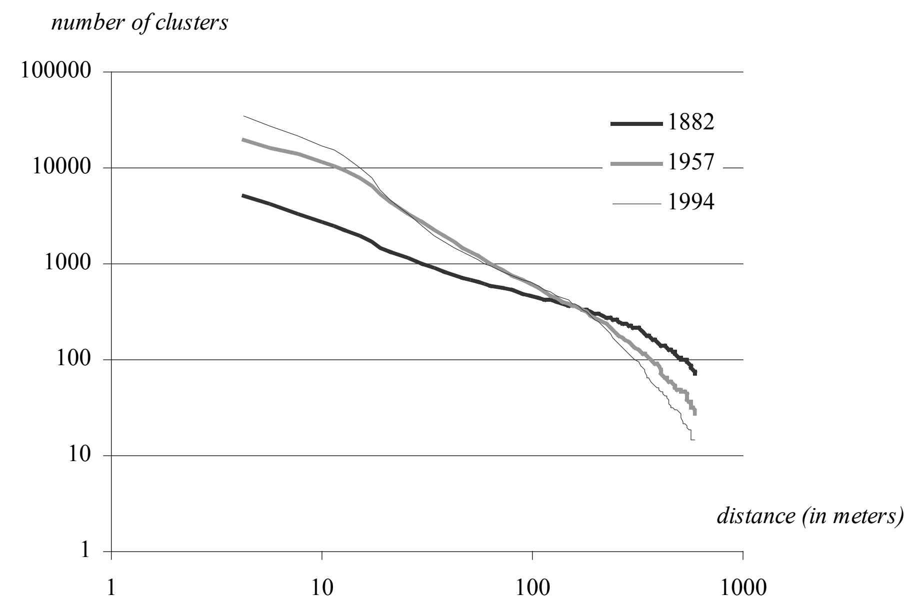
Bounds of the estimation

dates		**1 - 1069**	**1 - 39**	**39 - 159**	**159 - 343**	**343 - 541**	**541 - 1069**
		Total curve.	1 to 165 m.	165 to 675 m.	675 to 1450 m.	1450 to 2300 m.	2300 to 4520 m.
	1882	1.1	1.6	1.4	1.5	1.3	0.5
		1 - 1069	**1 - 83**	**83 - 623**	**623 - 1069**		
		Total curve	1 to 350 m.	350 to 2650 m.	2650 to 4520 m.		
	1957	1.7	1.8	1.7	1.6		
		1 - 1069	**1 - 83**	**83 - 623**	**623 - 891**	**891 - 1069**	
		Total curve.	1 to 350 m.	350 to 2650 m.	2650 to 3770 m.	3770 to 4520 m.	
	1994	1.7	1.8	1.7	1.8	1.7	

- *Evolution of the number of clusters of built-up pixels through the dilations*

On figure 5, an intermediary result helps us to understand the fractal description. The number of clusters varies according to the steps of dilation. At the beginning, it is much lower in 1882 (5103 clusters) than in 1994 (34 250 clusters), while the number of clusters in 1957 was in between (19 710). This corresponds to the number of non contiguous buildings which has increased in the recent periods, following a growing trend to urban sprawl. For the three dates, a sharp decrease in the number of clusters can be observed after the first steps of dilation, with slightly different thresholds corresponding to the mean size of neighbourhoods at the time. The slowing down of the decreasing curve is less pronounced for the more recent periods, due to a larger fraction of space being occupied by non compact built-up zones.

Figure 5: Number of clusters of built-up pixels at each step of the dilation



On the whole, urban sprawl coincides with a large number of built-up sectors (non connected buildings) being enumerated at the first step of the analysis, followed by a sharp decrease of this number during further steps of dilatation. This description is in accordance with the observations made about the border of the urban area.

3.4 Concluding remarks

Fractal analysis as applied to the Basle agglomeration throws a promising light on the evolution of the urban structure of the city. It shows that the general form of the agglomeration was already shaped in 1957, the consecutive evolution being merely a space filling process around the existing built-up cores. Considering tables 1 and 2, it appears that fractal dimensions of the border and of the built-up area are similar in 1957 and 1994, while results are more different in the case of 1882. The relationship between surface and border changed over time. The results obtained should now be interpreted thoroughly in order to identify the substantive meaning of the identified thresholds as well as the substantive meaning of the intersection of the curves which appeared.

From a general point of view, urban sprawl mainly involves the homogenisation of the built-up texture and an increasing sinuosity of the border, which also becomes less contrasted in design. But it seems useful here to sum up the morphological properties of urban patterns which can be identified through the analysis presented and which manifest themselves in the existence of urban sprawl:

- great number of built-up clusters at the initial step of dilation: the space is highly covered with housing; this coverage is locally rather homogeneous; the urban pattern is rather weakly compact; built-up clusters are rather close to one another;
- at the end of dilations, only a relatively small number of built-up clusters remains;
- at the end of dilations, only a small number of lacunas internal to the clusters remains;
- in the course of dilations, emergence of a great number of lacunas when emerge big clusters;
- the initial total border of the urban area is particularly long;
- the curve representing the evolution of the length of the border through dilations is characterised by a steep negative gradient.

Now, the objective of further research is to better understand the time evolution of the relation between the length of the border, the number of clusters and the number of lacunas. Such an objective could be attained mainly through systematic comparisons with other urban areas.

4 Discussion: what are fractals useful for?

It is not so easy to assess the main benefits of the use of fractals in geography and more generally for social sciences. Below, we briefly review a list of remaining questions for urban geography which could be solved by intensifying comparative research.

The reference to fractals is relatively recent in geographical literature, the first appeared less than twenty years ago, and probably deeper insights will be gained as studies become more numerous and more systematic. The main advantage of fractal geometry is to provide a model of reference which seems more adapted than Euclidean geometry to the description of spatial forms created by societies: features of heterogeneity, self-similarity and hierarchy are included from the very beginning in fractal structures. When comparing observed spatial patterns to Euclidean geometry, these properties appear as major deviations and anomalies specifying social systems, whereas direct comparison to fractal models may reveal specific features which have not been noticed yet. Another very important although not yet fully explored property of fractals is their relation to underlying non linear generative mechanisms. The design and use in simulation of models which would explicitly connect individual behaviour or micro processes to the emergence of fractal morphologies at upper levels of observation would greatly improve our understanding of the genesis of such forms and allow a more systematic exploration of their stability, limits and rationales.

However, one can enumerate a few of the many questions which remain partially or totally unsolved at the moment.

- What would be an index of the fractality of cities? We know that because of its homogeneity, a perfectly compact city is not fractal, neither are suburbs which would be homogeneously scattered. In between, how should be the variations in the degree of fractality interpreted?
- Fractal dimensions can be compared but they are very concise summaries of entire urban structures which may differ in other ways while exhibiting the same fractal dimension. Urban fractal properties are not well enough known up until now to derive a truly consistent interpretation of measured values from a proper theory.
- A large variety of measures should help to determine if fractality is better explained by relating it either to different schools in urbanism (different ways of conceiving urban shapes) or to successive steps in the urbanisation process.
- Are there any relationships between urban quality of life and the degree of fractality of urban morphology? Would it be more relevant for policies, instead of distinguishing between urban compactness and sprawl, to differentiate between fractal and non fractal cities?

POSTSCRIPT

Fractals in archaeology

As suggested by applications to urban geography, fractals can be used in archaeology as well, for the study of spatial structures of many artefacts, including buildings, networks and land use. This could offer precious references, since conditions of spatial interaction were very different from nowadays but perhaps more similar between different cultures in ancient times, especially in terms of speed and spatial range and consecutively in, for instance, possible extension, hierarchy and differentiation of cities. One major problem of course is to get a good cartography of the supposed fractal structures and to be able to compare them at a given and well identified level of resolution. But in turn, comparison of spatial structures of previous eras with those of today could help to identify the social processes which are behind their morphogenesis. This could suggest the terms of a co-operative research between our disciplines.

Acknowledgement

We would like to acknowledge Richard Stephenson, British geographer and colleague at the University of Franche-Comté, for his carefully considered and valued comments.

References

Arlinghaus S.L. 1985, "Fractals take a central place", *Geografiska Annaler*, 67B, 83-88.

Bailly E. 1996, "Position de recherche sur une méthode de détermination d'un contour urbain", *Cybergeo*, 10, 6 p. (http://www.cybergeo.presse.fr)

Batty M., Longley P., 1986, "The fractal simulation of urban structure", *Environment and Planning A*, 18, 1143-1179.

Batty M., Longley P. Fotheringham S., 1989, "Urban growth and form: scaling, fractal geometry, and diffusion-limited aggregation", *Environment and Planning A*, 21, 1447-1472, 1989

Batty M., Kim K.S. 1992, Form follows function: reformulating urban population density function. *Urban Studies*, 29, 7.

Batty M., Longley P. 1994, *Fractal Cities*, London, Academic Press, 1994

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

Bretagnolle A. 1996, Étude des indices de concentration d'une population. *L'Espace Géographique*, vol. 2, 145-157.

Bretagnolle A., Paulus F., Pumain D., 2002, Time and space scales for measuring urban growth. *Cybergeo,* 219, 12 p.

Cavailhès J., Frankhauser P., Peeters D., Thomas I. 2004, Where Alonso meets Sierpinski: an urban economic model of fractal metropolitan area. *Environment and Planning A*, 36, 1471-1498.

Christaller W., 1933. *Die Zentralen Orte in Süddeutschland*. Iena, Fischer.

Clark C., 1951. Urban population densities*, Journal of the Royal Statistical Association*, 114, 490-496.

A. D. Cliff, J. K. Ord, 1973. *Spatial autocorrelation*. London, Pion.

De Keersmaecker M.L., Frankhauser P., Thomas I., 2003, “Using fractal dimensions for characterizing intra-urban diversity – The example of Brussels”, *Geographical Analysis*, 35, 4, 310-328

De Keersmaecker M.L., Frankhauser P., Thomas I., 2004, “Dimensions fractales et réalités périurbaines – L'exemple du sud de Bruxelles”, *L'Espace Géographique*, vol. 3, 219-240.

Engelen G., White R., Uljee I., 2002. *The MURBANDY and MOLAND models for Dublin*. Report of the Research Institute for Knowledge Systems, Maastricht.

François N., Frankhauser P., Pumain D. 1995, "Villes, densité et fractalité", *Les Annales de la Recherche Urbaine*, 67, 55-64.

Frankhauser P., 1994. *La fractalité des structures urbaines*, Paris, Anthropos, coll. Villes, 291p.

Frankhauser P., 1998. "The fractal approach. A new tool for the spatial analysis of urban agglomerations", *Population: an English Selection*, Special issue *New methodological Approaches in the Social Sciences*, 205-240.

Frankhauser P., Pumain D. 2002, Fractales et géographie, in Sanders L. (ed.) *Modèles en analyse spatiale*. Paris, Hermès, 301-329.

Haggett P., 2001, *Geography: A Global Synthesis.* Prentice Hall, New York. 833 pages.

Lajoie G., Mathian H. 1991, “Application of variograms to urban geography”, *Spatial analysis and Population dynamics*, D. Pumain (ed.), John Libbey-INED, pp. 295-310.

Le Bras H. 1996, *Le peuplement de l'Europe*, Paris, la Documentation française.

Mandelbrot B. 1977, *The fractal geometry of nature*. Freeman, San Francisco.

Markse A., Havlin S., Stanley H.E. 1995, “Modelling urban growth pattern”, *Nature*, 377.

Odland J. 1988, “Spatial autocorrelation”, *Scientific Geography Series*, vol.9, G.I. Thrall editor, USA, 87p.

Philbrick A.K. 1957, "Principles of areal functional organisation in regional human geography", *Economic Geography*, 229-336.

Pumain D., Moriconi-Ebrard F. 1997, City Size distributions and metropolisation. *Geojournal*, 43 :4, 307-314.

Salingros N., 2003. *Connecting the Fractal City*, Keynote speech, 5th Biennal of towns and town planners in Europe, Barcelona, April 2003.

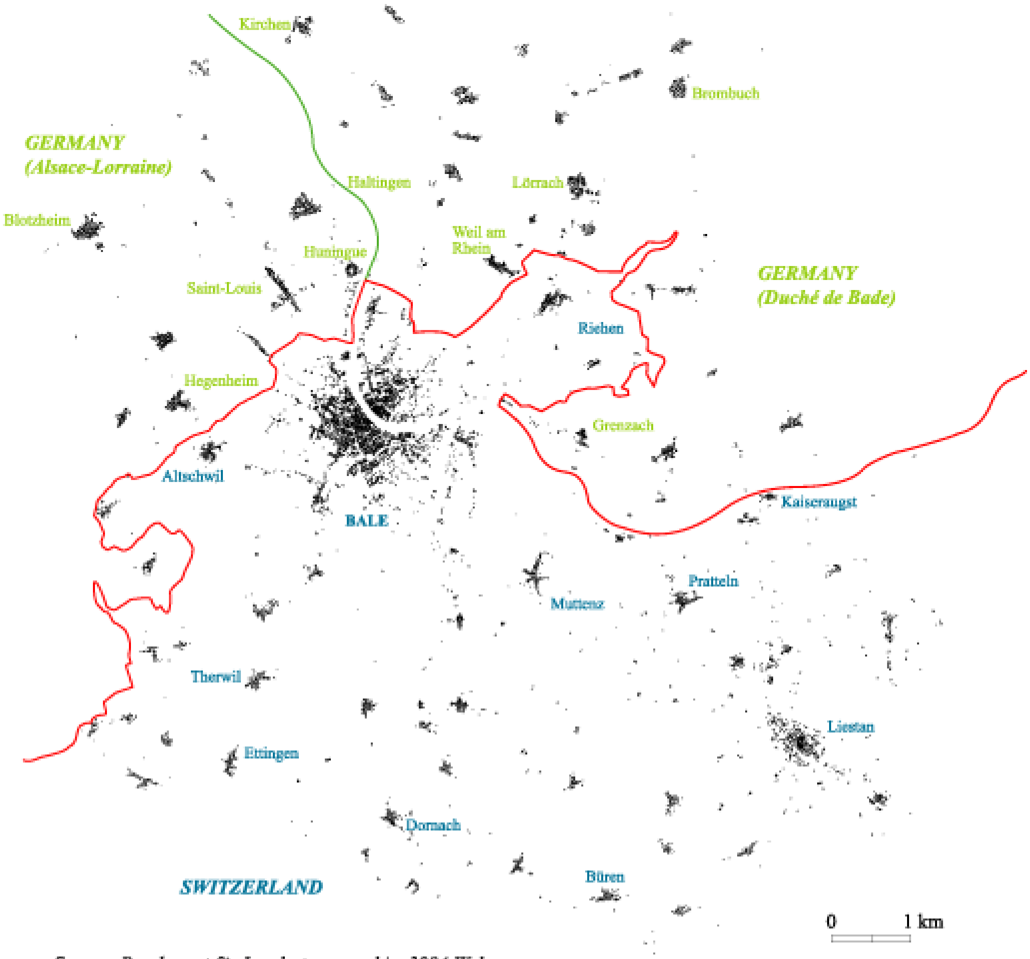
Tobler W. 1979, "Cellular geography", *Philosophy in geography*, Eds Gale S., Olsson G., Dordrecht, Reidel, 279-386.

Ullman E.L. 1980, *Geography as spatial interaction*, Seattle, University of Washington Press (ed. by R.R. Boyce)

White R., Engelen G., 1994, "Urban Systems Dynamics and Cellular Automata: Fractal Structures between Order and Chaos", *Chaos, Solitons and Fractals*, 4, 4, 563-583.

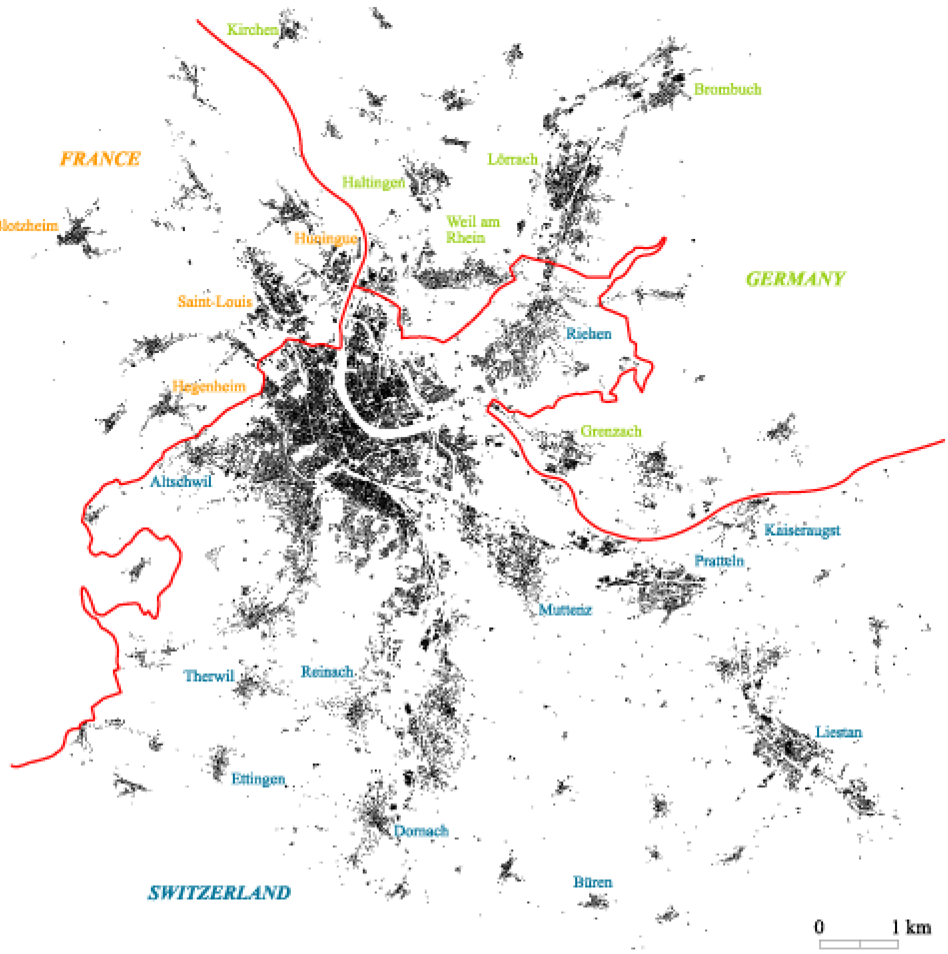
White R., Luo W., Hatna E., 2001, Fractal Structures in Land Use Patterns of European Cities: form and Process, *12th European Colloquium on Quantitative and Theoretical Geography,* Saint-Valéry-en-Caux, France, 11 p.

Appendix 1: The urban area of Basle in 1880



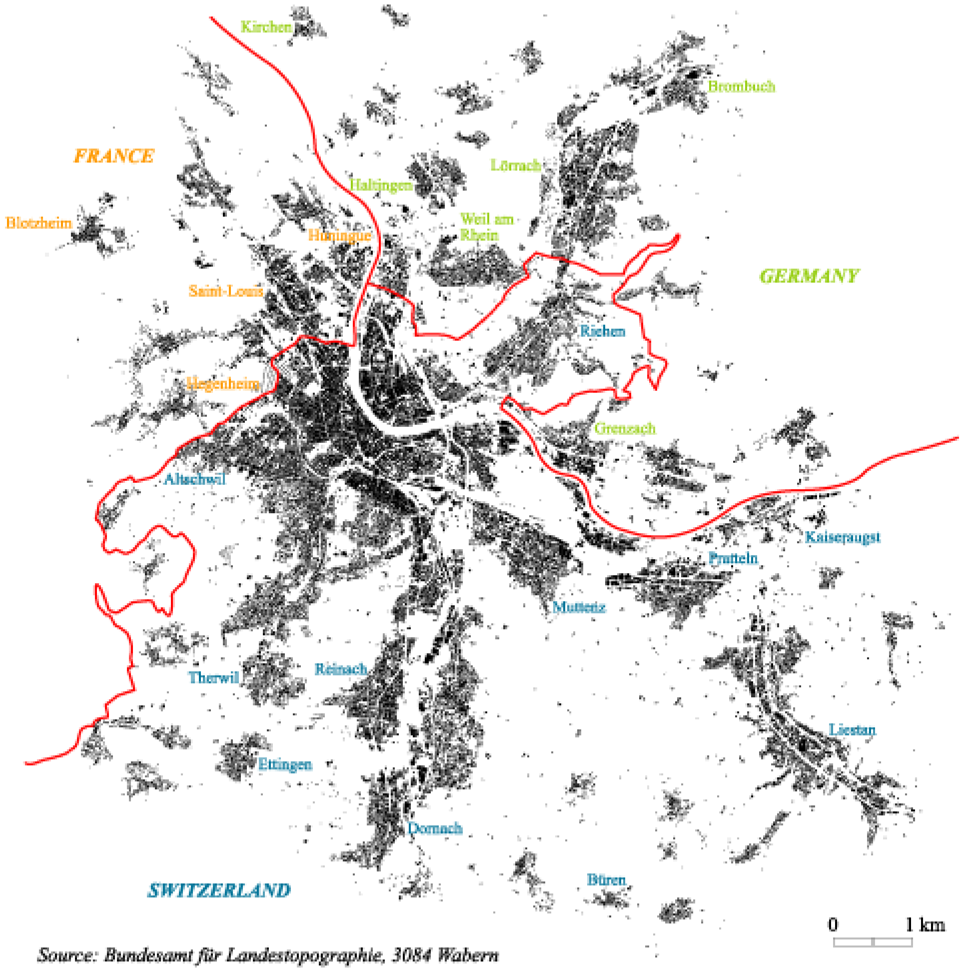
Source: Bundesamt für Landestopographie, 3084 Wabern

Appendix 2: The urban area of Basle in 1957



Source: Bundesamt für Landestopographie, 3084 Wabern

Appendix 3: The urban area of Basle in 1994



Appendix 4: A part of the non dilated border of the urban pattern of Basle in 1880