J Geograph Syst (2006) 8: 145–164
DOI 10.1007/s10109-006-0026-9

ORIGINAL ARTICLE

Roger White

Pattern based map comparisons

Published online: 6 May 2006

Abstract Map comparison techniques based on a pixel-by-pixel comparison are useful for many purposes, but fail to reveal important aspects of map similarities and differences. In contrast, pattern based map comparison techniques address the question of structural similarity, although with these approaches the comparison problem becomes open ended, since there is an unlimited number of ways to characterise a pattern. Two types of pattern based technique are used here to analyse the test sets of maps. The first, a fuzzy polygon based matching technique, focuses on the meso-scale aspects of pattern. It is based on the areal intersection of land use polygons on the two maps being compared. The areal intersection, areal complement, and polygon size values are fuzzified into an appropriate set of categories, a set of fuzzy inference rules is applied to derive memberships in local matching categories, and finally the local matching category memberships are defuzzified to yield local matching values for each land use polygon on the reference map. The second approach, fractal analysis, captures macro-scale or global qualities of the maps. Two measures are calculated here: the radial dimension and the cluster size—frequency dimension. The polygon matching approach provides only limited insight when applied to the case of the map set representing differences in classification. It proves much more effective when applied to the problem of change detection, revealing areas where the pattern has changed and giving local measures of the degree of change. The approach is particularly useful in the case where there is a considerable degree of random change at the pixel level, as changes in the underlying pattern are extracted from the noise, while pixel based approaches largely detect the noise.

Keywords Polygon matching · Fuzzy comparison · Pattern similarity · Radial dimension · Cluster size frequency dimension

R. White
Department of Geography, Memorial University, St. John's, NL, Canada A1B 3X9
E-mail: roger@mun.ca

1 Introduction

Most techniques relevant to map comparison have their roots in the remote sensing and GIS communities, and the nature of these techniques reflects the interests and problems of those working in those communities. Questions relating to classification, resolution, registration, and error propagation are of fundamental importance, because applications presuppose that these problems have been dealt with reasonably well. In particular, change detection is a type of application that is only possible after these problems have been adequately addressed. However, there is often an implicit assumption that if the data are error free, and classification and registration are perfect, then the change detection problem is trivial: a pixel either has a particular value in both an earlier map and a later one, or it has different values; in the latter case change has been detected. But as the GEOIDE programme recognizes, the problem is more complex than that. The pattern of changes—or more precisely, the changes in pattern—may in many cases be more important than simply the crude magnitude of change. This is especially so in cases where the pattern has some functional significance for the system, as is the case, for example, in cities and ecosystems. But in the case of techniques that are focused on aspects of pattern, the problem becomes open ended, because there may be an unlimited number of ways to characterise a pattern, and among these, the ones that are useful or appropriate can only be determined with reference to some particular point of view or problem. For example, is the convolutedness of polygon edges an important feature? For New Urbanists concerned with urban sprawl and for some ecologists the answer may be yes; for many others, it may be no.

The need for pattern based techniques has only been recognized relatively recently, and the possibilities have only begun to be explored. Landscape ecology is one discipline where the ability to characterize spatial patterns is recognized to be essential, and many of the techniques routinely used in that field have been assembled in the FRAGSTATS software package (McGarigal et al. 2002). The indices provided by FRAGSTATS constitute a useful characterization of some important spatial qualities—specifically those thought to be important to ecological processes. Fragstat is not a map comparison technique as such, although maps may be compared by comparing the values output by FRAGSTATS. Closely related is the fractal analysis now used in a number of fields from physics to geography to economics (Mandelbrot 1983). The various fractal dimensions that may be calculated are highly generalized descriptions of pattern, and, like the FRAGSTATS measures, may be used to compare maps. In particular, two fractal dimensions, the radial dimension and the cluster size frequency dimension, are used in this paper to analyse the pair of maps in pair D2 and then to compare them on the basis of their dimensions.

Image analysis for identification purposes is another area of research with important implications for map comparison. Here the focus is on deciding whether an imaged object is a member of a class (is the figure in this image the letter “T”?), where in many important applications—for example, identification for security purposes—the class holds only one individual. In this latter case the problem is that the image currently in the video camera may be quite different from the stored reference image, because of differences in lighting, perspective, orientation, etc., but it is necessary to determine whether the object imaged is the same—i.e. is it the identical person represented in both images? The techniques developed in attempting to solve this problem are of great relevance in advancing the field of pattern based map comparison. Indeed, some of these techniques were adapted in developing the polygon based fuzzy local matching map comparison technique that is the focus of the analysis in this paper (Power et al. 2001).

2 Polygon based fuzzy local matching map comparison

The polygon based fuzzy local matching technique (Power et al. 2001) was developed primarily to deal with certain problems that arise in modelling urban spatial dynamics, but also with change detection in mind. In particular, the problem context was the calibration and validation of predictive cellular automata based models of land use dynamics. In the calibration process a model is run with initial conditions given by data for the beginning year of the simulation. The final year of the simulation is also one for which data are available. In the case of calibration, the model output for the final year, in particular a predicted land use map, is compared with the actual land use map for that year. On the basis of differences between the two maps, parameters are adjusted to improve the match between the simulated and actual maps. In the case of validation, the procedure is the same, but the comparison of the simulated and actual maps is done with the aim of discovering to what degree and in what ways the model is capturing the processes shaping the actual city. In both cases the situation is asymmetrical in that the actual land use map is the standard to which the simulated map is compared. A simple cell-by-cell comparison and an associated statistic like Kappa are useful up to a point, but something more is required. An urban land use map is highly structured, and the real question is, to what extent is a simulated land use *pattern* similar to the actual *pattern*? As models become more powerful this problem of comparing the maps becomes more acute (White et al. 1997). The fuzzy polygon based matching technique provides an approach to answering that kind of question. It is, therefore, also useful for answering such qualitative change detection questions as, to what degree has the *pattern* of land use changed between 1995 and 2005?

The technique is based on the areal intersection of land use polygons on the two maps being compared, one of which is taken as the reference map, with the overlay of the two maps producing a set of unique polygons. The unique polygons are used to calculate areal intersection and complement ratios for each polygon on the reference map. In addition to the relative degree of overlap of polygons, the polygon size is considered, so that very small polygons can be given little weight. The areal intersection, complement, and polygon size values are fuzzified into an appropriate set of categories; a set of fuzzy inference rules is then applied to derive memberships in fuzzy local matching categories, F; and finally the local matching category memberships are defuzzified to yield local matching values, L, for each unique polygon. In the current implementation the comparison is based only on areal overlap of polygons and on polygon size. But the technique can be extended in a straightforward way to include any relevant polygon qualities, such as shape index, convolutedness of edge, or orientation—i.e. any polygon characteristics that may be relevant in a particular application. The output consists of (1) a comparison map showing the degree to which the test map is similar to the reference map, and (2) a summary measure of similarity, the global fuzzy matching index.

2.1 The comparison map

A major advantage of the fuzzy polygon based approach compared to pixel based Boolean approaches is that the structure of the original reference map remains visible in the comparison map, so that it is relatively easy to relate areas of similarity and dissimilarity to features in the original maps. This facilitates interpretation of the comparison in terms of underlying processes, and in the case where the test map is generated by a model, also facilitates calibration by making it easier to relate regions of poor model performance to local spatial relationships, and the way these may not have been adequately represented in the model. Some pixel based Boolean comparisons, specifically those for individual classes (e.g. those in Fig. 1), also preserve structure, but only for the particular class being compared; spatial relationships with other classes are not visible. Another advantage of the fuzzy polygon based approach, one shared with the fuzzy neighbourhood techniques developed by Hagen, is that the output map shows a measure of *degree* of similarity, rather than simply the binary agree/disagree of Boolean approaches. This helps the user to assess where the differences between two maps are important.

2.2 The global fuzzy matching index

The global fuzzy matching index, g, is calculated as

$$g = (\Sigma_i L_i A_i)/\Sigma_i A_i \tag{1}$$

where L_i = local matching value for polygon i

A_i = area of polygon i

Values of g range between 0 (no similarity) and 1 (the maps are identical). Values of g are generally somewhat higher than values of Kappa calculated for the same pair of maps, but lower than the fraction of cells that match, since Kappa adjusts for the probability that a match will occur simply at random.

Since g contains far less information than the comparison map, it may seem to be of relatively little interest, except as a quick way of determining which of a series of test maps is most similar to a reference map, ignoring the important issue of the *ways* in which it is more similar than the other

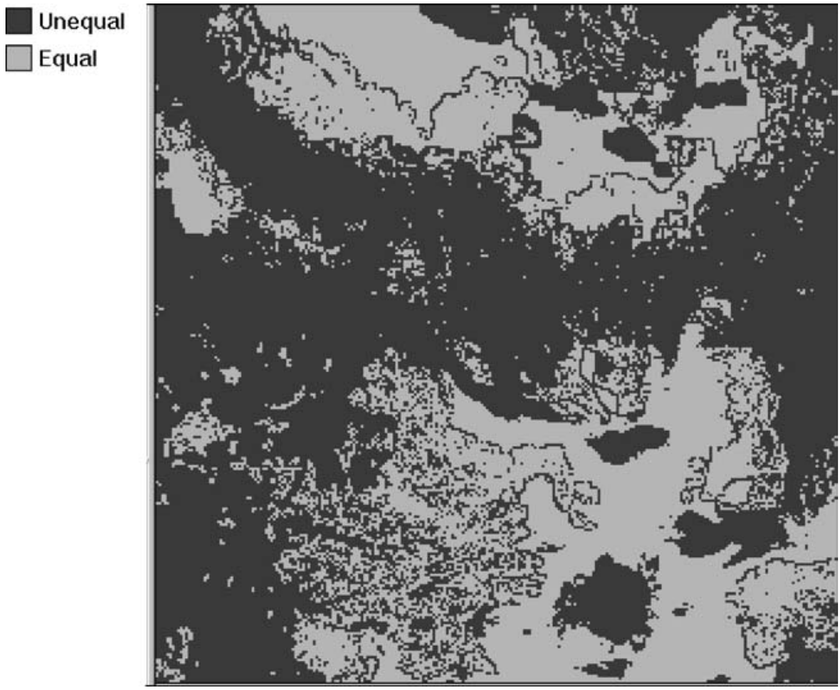

Fig. 1 Pair D1: cell-by-cell comparison

maps—and the ways in which it is *less* similar. But as automatic calibration routines for cellular automata based land use models are further developed, g may become a key measure. An automatic calibration routine works by adjusting parameter values so as to increase the fit between the model output and the reference map of actual land use, where the fit is most conveniently represented by a single number. A statistic like Kappa can be used and gives useful results up to a point, but because Kappa captures the similarity of patterns only indirectly and very imperfectly, the visual results are often disappointing. Using instead a value like g, which reflects the similarity of patterns and can potentially include any number of relevant pattern qualities, may significantly improve the quality of automatic calibrations.

The basic fuzzy polygon based map comparison technique is asymmetrical in that the results depend on which map is chosen as the reference. It has been supplemented by four symmetrical variations developed by Hagen (2002): each map in a pair is taken in turn as the reference map, and then the two results are combined by (1) taking the average of the two values on each cell, (2) taking the product of the two values, (3) taking the minimum of the two, or (4) taking the maximum. All five variations are included in the Map Comparison Kit, developed by Hagen (free download from http://www.riks.nl/mck; Visser 2004), together with several basic cell-by-cell techniques and associated numerical comparisons, also used in this paper. In addition, the Map Comparison Kit includes the fuzzy cell techniques, fuzzy Kappa, and moving window approaches (Hagen 2003) used by Hagen in this issue.

3 Map comparisons: pair D1

A preliminary comparison of the two maps on a cell-by-cell basis is shown in Fig. 1. This comparison is shown to provide a context for assessing the performance of the polygon matching approach. The two primary comparison maps generated by the polygon matching routine—one taking map a as the reference map (Fig. 2), and the other taking map b (Fig. 3)—essentially reproduce the two very different styles of the reference maps. The comparison using map b as the reference suggests (as does the original map) that the spatial resolution is finer in some parts of map b (e.g. the lower left quadrant) than in others (the top third). The fuzzy global matching values for the comparisons are $g = 0.450$ (map a is reference) and $g = 0.454$ (map b is reference). The four symmetrical versions of the polygon matching comparisons do not add much to our understanding of the differences of the two maps; two of these are shown in Fig. 4.

The polygon matching comparison technique does not seem to be very useful when applied directly to this data set. This is not surprising since it was not developed to detect and characterize classification problems. A second order comparison, however—a comparison of the two primary polygon matching comparison maps using the absolute value of the difference of the two maps (Fig. 5)—is perhaps more useful. It shows a high level of agreement in the large area in which the classification is *different* on maps a and b (essentially the large areas dominated by classes 4 and 5), since both underlying or first order comparison maps pick up the disagreement between

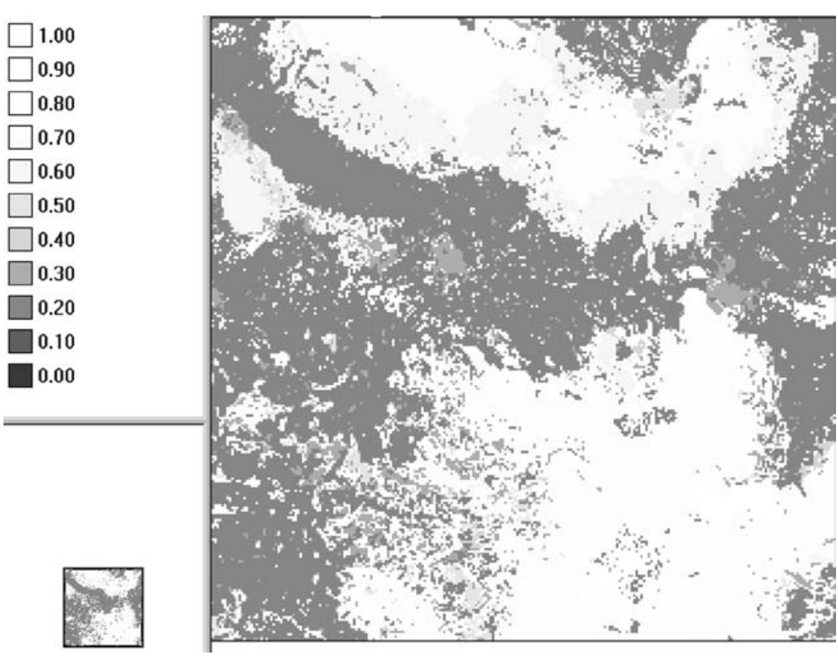

Fig. 2 Pair D1: polygon matching comparison; map a is the reference map

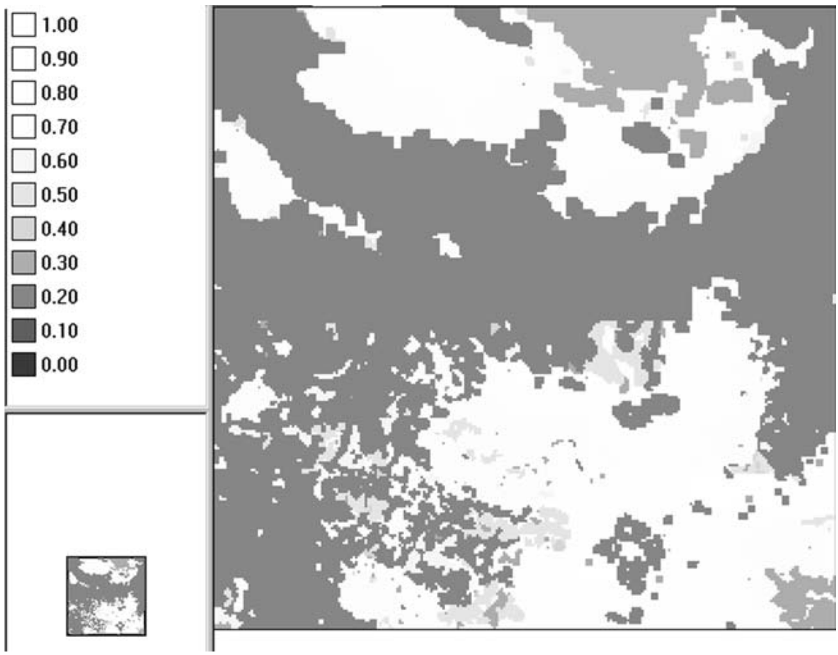

Fig. 3 Pair D1: polygon matching comparison; map b is the reference map

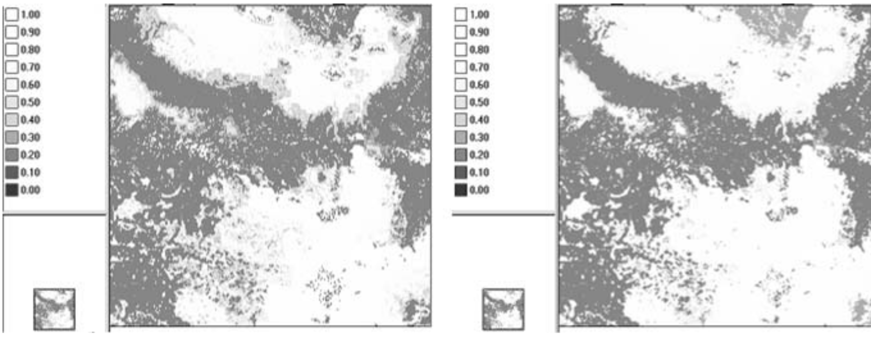

Fig. 4 Pair D1: two symmetrical polygon matching comparisons. *Left*: map produced by averaging the values of the two comparisons using maps a and b, respectively as the references. *Right*: map produced by taking the maximum of the two values from the two reference comparisons

maps a and b. It also gives relatively high agreement levels in areas where the patterns and classes in maps a and b are *similar*, since both first order comparisons pick up the similarity. But the second order comparison shows low levels of agreement between the first order comparisons where boundaries between classes are complex in one case (map a) and simple in the other (map b). The major strip of low level second order agreement follows the northern edge of the large area of classes 4 and 5. This area of confusion and differing classifications does not show up nearly so clearly in either the cell-by-cell (Fig. 1) or fuzzy set comparisons (see paper by Hagen, this issue), and not at all on the first order polygon matching comparisons. The second order comparison also shows some of the one cell wide classification fringes that appear so clearly on the cell-by-cell and fuzzy set comparisons but not in the first order polygon matching comparisons.

4 Map comparisons: pair D2

4.1 Cell-by-cell comparisons

The maps in this set represent a distribution at two different times, and the challenge is to identify or characterise the change. To provide context, a preliminary or baseline analysis is carried out at the cell level. The coincidence matrix (Table 1) shows that there are exactly equal numbers of cells in each of classes 1–4 in map a. The cell-by-cell comparison map (Fig. 6) reveals a curious diagonal running from the upper left corner (row 0, col. 0) to the mid right (row 130, col. 256), with much more change below the diagonal: from map a to map b there is a considerable amount of apparently random "churn" in the items—cells change apparently at random from one

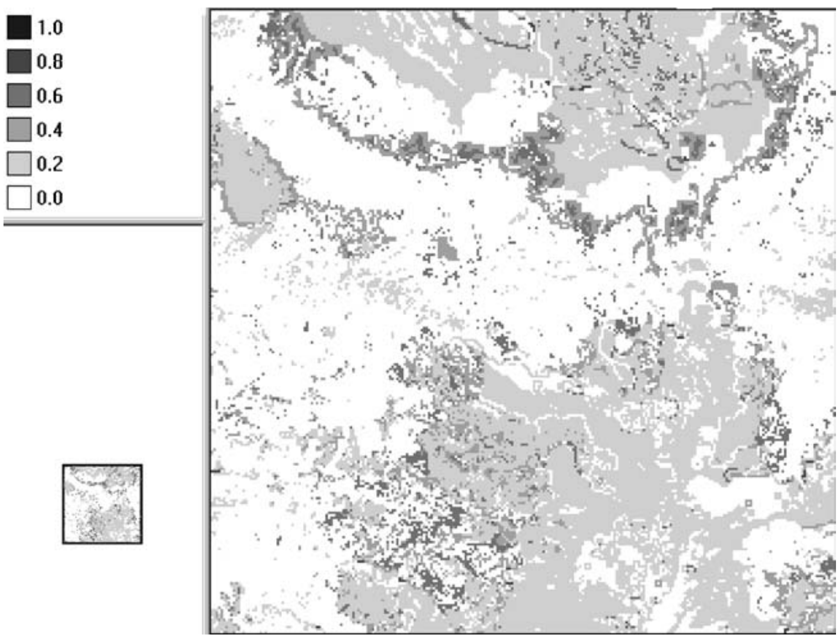

Fig. 5 Pair D1. Comparison of the two polygon matching comparisons in Figs. 2 and 3. The map shows the absolute value of the difference of the two values from the reference comparisons

class to another. Class-by-class change comparisons (Fig. 7) show that for class 1, the churn seems clustered, with most losses in the transition from map a to map b on the periphery of the clusters ("only in map one"), but with most new cells widely scattered ("only in map two"). There is a net loss of 4,796 class 1 cells. Classes 2 and 3 show scattered (random?) transitions (appearances and disappearances), except in the central regions of clusters of classes 1 and 4, which new cells of classes 2 and 3 largely avoid. Class 4, in the area below the diagonal, loses some cells on the periphery of its clusters and gains them widely throughout the area; the region above the diagonal shows more losses than gains, and these losses occur not just on the periphery of the clusters but from the interior as well. Cells change state preferentially to states with a neighbouring class number (Table 1): for all items, there is a monotonic negative relationship between the difference in class number and the numbers of cells that change from one class to the other. Also, as seen both in Fig. 7 and the coincidence matrix, there is a strong preference for changes from lower to higher class numbers (except, of course for class 4). Class 1 has the largest decline in area, with class 2 having a smaller decline; class 3 increases, but class 4 increases more. The net effect on class 4 is that it remains clustered, but below the diagonal, as the clusters become larger they also become more diffuse. Above the diagonal they become more diffuse as they shrink slightly in total area.

4.2 Polygon matching comparisons

Using map a as the reference, the comparison map (Fig. 8) shows a relative concentration of low similarity areas (similarity = 0.30) in the lower left and also in the lower right areas of the map. These correspond to areas where larger polygons of class 1 were fragmented into many small polygons. The arc of a somewhat higher similarity level (similarity = 0.40) inside these and surrounding the central region of still higher similarity, corresponds to a circular zone of clusters of class 1 that shrank but remained intact. The central area of little change (similarity = 0.60) is essentially the class 4 area. Taking map b as the reference, the comparison (Fig. 9) shows that the centres of the class 1 areas that remained intact show up as white (little or no change: similarity $\geq$ 0.70) rather than as similarity = 0.40. This is not true of the class 4 areas, which grew but also lost on their edges—i.e. the initial areas suffered attrition as well as augmentation, so there was no large area of perfect overlap in either direction of comparison. Also, this map shows a

Table 1 Contingency table: set D2, maps A and B

Cells in map A	Cells in map B				
	State1	State2	State3	State4	Sum map A
State1	8,488	3,057	2,656	2,183	16,384
State2	1,582	8,228	3,605	2,969	16,384
State3	967	2,355	8,977	4,085	16,384
State4	551	1,103	2,778	11,952	16,384
Sum map B	11,588	14,743	18,016	21,189	65,536

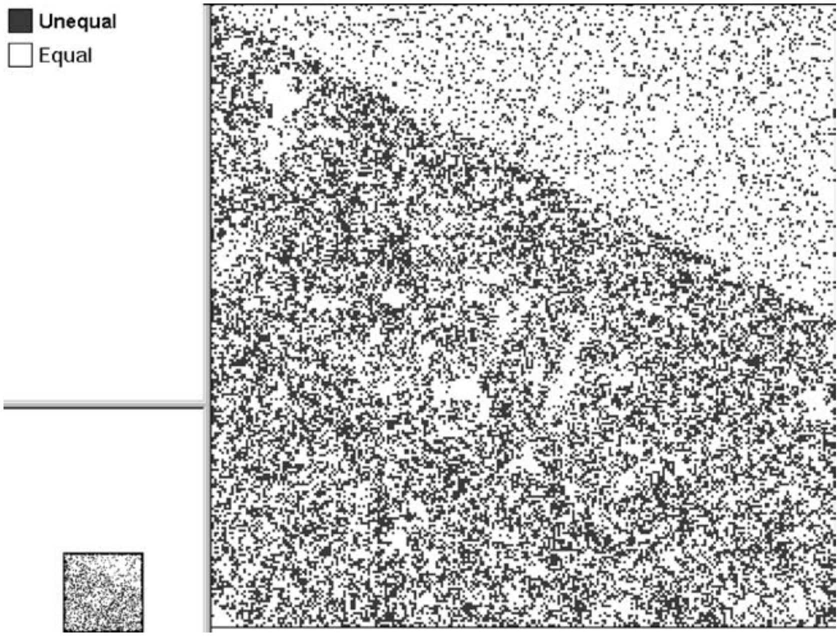

Fig. 6 Pair D2: cell-by-cell comparison

strip, approximately from col. 215 to col. 256 (on the right), above the diagonal, in which there was greater change; using map a as the reference, this strip does not stand out.

Comparison of the two polygon matching comparison maps using the absolute value of the difference of the two maps (Fig. 10) gives a map that shows just a scatter of moderate disagreements, except for a small area centred on 110, 195 that stands out as being in greater disagreement. The strip on the right hand side of the area above the diagonal that showed up as having been subject to greater change in the comparison using map B as the reference stands out conspicuously. In the cell-by-cell comparison this strip does not stand out at all; there seems to be no more change in this area than in any other part of the area above the diagonal. It seems that in this strip, from the standpoint of the later period, the land use pattern changed noticeably, while from the standpoint of the earlier period, the pattern remained essentially the same.

More specifically, the detection of change in the one case and a failure to detect change in the other is due to essentially to the pattern of change in class 4. Class 4 dominates in this strip, and in both maps it is concentrated in several large polygons. In map b, however, a significant number of cells in these large polygons no longer belong to class 4. Thus when map a is the reference, the large polygons of class 4 that remain in map b are compared and found to largely overlap those in a, and so the value of the similarity index over the cells in the map a polygons is high. On the other hand, when map b is the reference, the small polygons that resulted from conversion

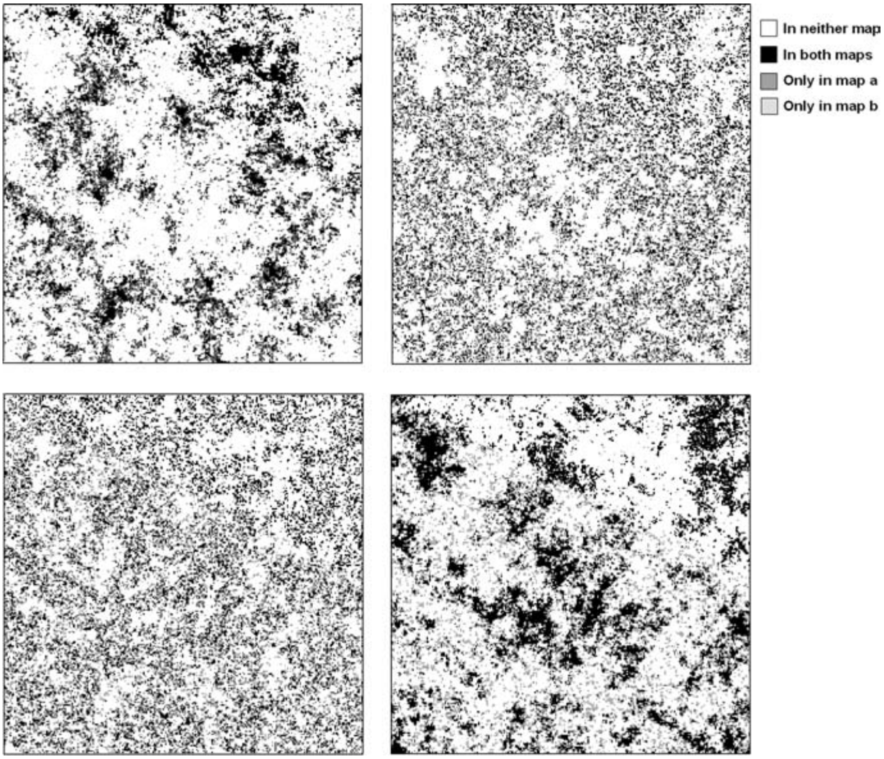

Fig. 7 Pair D2: cell-by-cell comparisons for each of classes 1–4. (In the original colour version of this figure, the upper diagonal feature appears prominently)

from class 4 to other classes do not have any overlap with polygons of equivalent classes in map a and so the (pre-fuzzified) similarity values are zero. In other words, from the viewpoint of map a we see large polygons of class 4 on both maps, though they differ in detail; while from the viewpoint of map b, we see that the numerous small polygons of other than class 4 that are imbedded in the large class 4 areas are largely absent in map a. This is the same phenomenon that was observed in the second order comparison in set 1.

In general, the polygon matching maps show the areas where the *pattern* has changed. In this instance, with the patterns remaining qualitatively similar in spite of much churn, the global matching ($g = 0.557$) is quite a bit higher than the fuzzy Kappa ($K = 0.365$). This is the strength of the polygon matching technique—in the presence of churn or random fluctuations in the class membership of individual cells, it is not blinded by the random fluctuations, but reveals changes in underlying patterns. As noted above, for many applications pattern similarity rather than cell-by-cell agreement is the most important quality. This is especially true of predictive or forecasting models, where patterns can often be predicted but high levels of cell-by-cell agreement may only represent over fitting of the model.

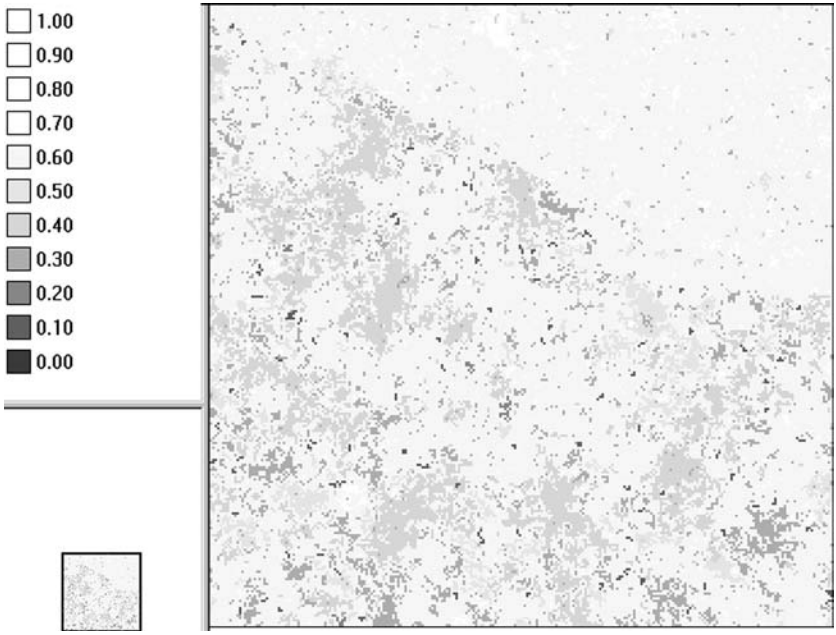

Fig. 8 Pair D2: polygon matching comparison; map a is the reference map

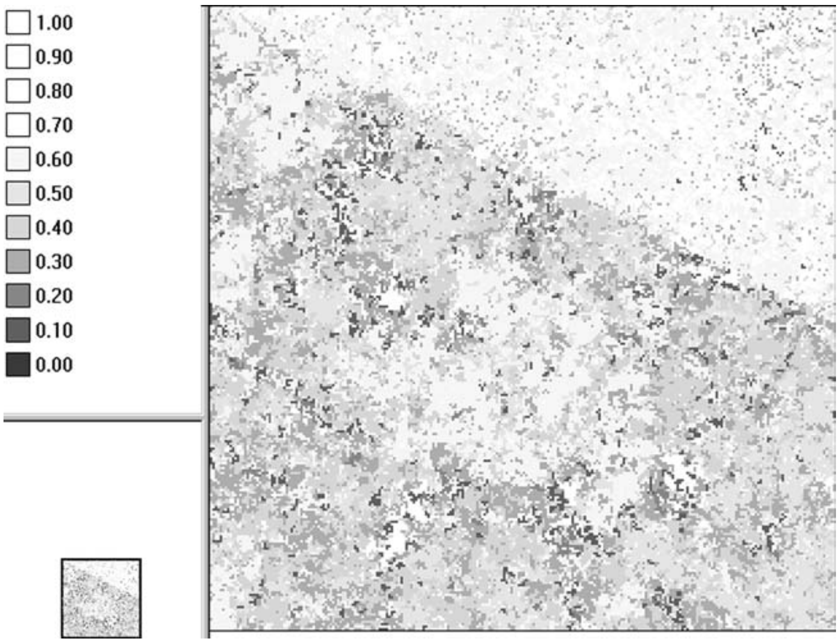

Fig. 9 Pair D2: polygon matching comparison; map b is the reference map

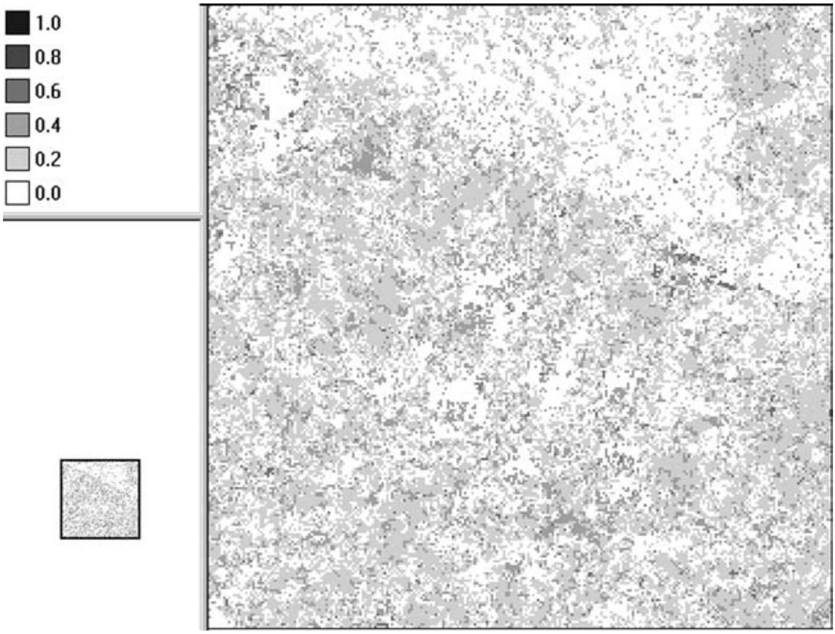

Fig. 10 Pair D2. Comparison of the two polygon matching comparisons in Figs. 8 and 9. The map shows the absolute value of the difference of the two values from the reference comparisons

Symmetrical versions of the polygon matching technique (Figs. 11, 12)—i.e. versions that do not take one of the maps as the reference map against which to measure the other—do not reveal as much when applied to this pair of images, with the possible exception of the maximum value map (Fig. 12), which tends to reveal the features of both the asymmetrical comparison maps. For example, the unchanged cores of the class 1 areas appear as white (high agreement) as they do on the map b reference comparison, but also the concentric bands of disagreement around the central class 4 areas show up. The fuzzy global matching values for these four symmetrical versions are as follows: product, $g = 0.321$; minimum, $g = 0.486$; average, $g = 0.556$; maximum, $g = 0.627$.

4.3 Fractal analysis

The polygon matching approach focuses on pattern at the local or regional scales. But it is also useful to look at global descriptions of pattern. Fractal dimensions constitute one class of global measures that have been found to be useful in characterising self organized structures (Gouyet 1996). Self organized structures like urban settlements, but also like many—perhaps most—natural structures, show scaling properties—i.e. they are fractals (Batty and Longley 1994; Frankhauser 1994, 2000; White and Engelen 1993).

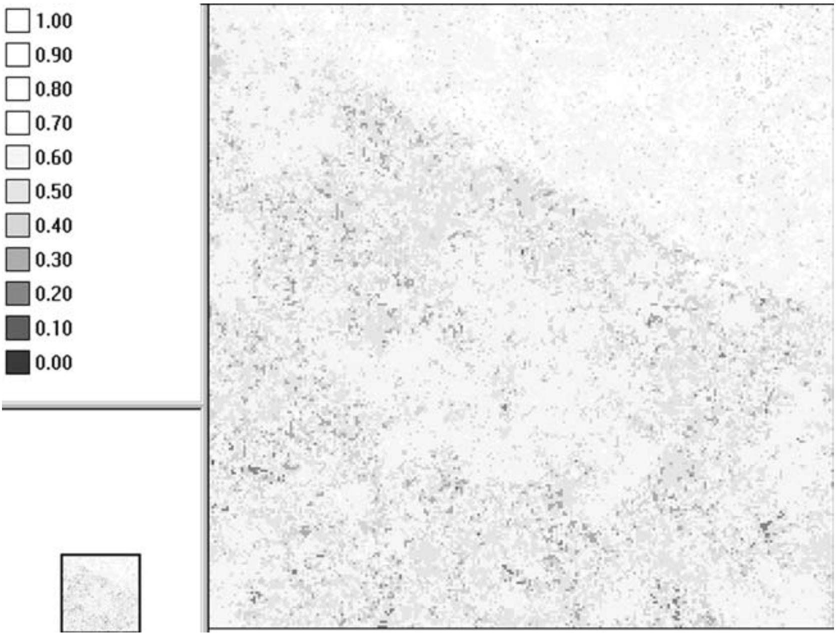

Fig. 11 Pair D2: Symmetrical polygon matching comparison: average map

Two fractal measures are used here: the cluster size frequency relationship and the radial dimension.

4.4 Cluster size frequency relationship

The cluster size frequency relationship is established as the log-log plot of frequency against cluster size. It shows the rate at which clusters become more numerous as they become smaller. The measure is only quasi-spatial, since cluster size is itself a spatial property, but the location of the clusters is ignored. Nevertheless, the proportion of clusters of various sizes is an important characteristic of map pattern, and has functional significance as well, since it reflects the characteristics of the underlying processes that generate the pattern. For example, a distribution with too few small clusters is one in which the organizing process is not generating sufficient new clusters. The cluster size frequency relationship has proven to be very useful in calibrating and validating dynamic models of the spatial structure of cities and regions (White and Engelen 1993).

The cluster size frequency relationships for all four classes in map pair D2 are log linear (an example is shown in Fig. 13), with representative values for the slope in the neighbourhood of 1.5. This is a value that can be observed in actual human landscapes; but since such relationships are so widespread (e.g. most classes in the set 1 maps also have log linear cluster size frequency distributions), in the absence of any context that would make these results significant or useful, they only indicate that the two maps are not blatantly unrealistic images—although they were generated artificially.

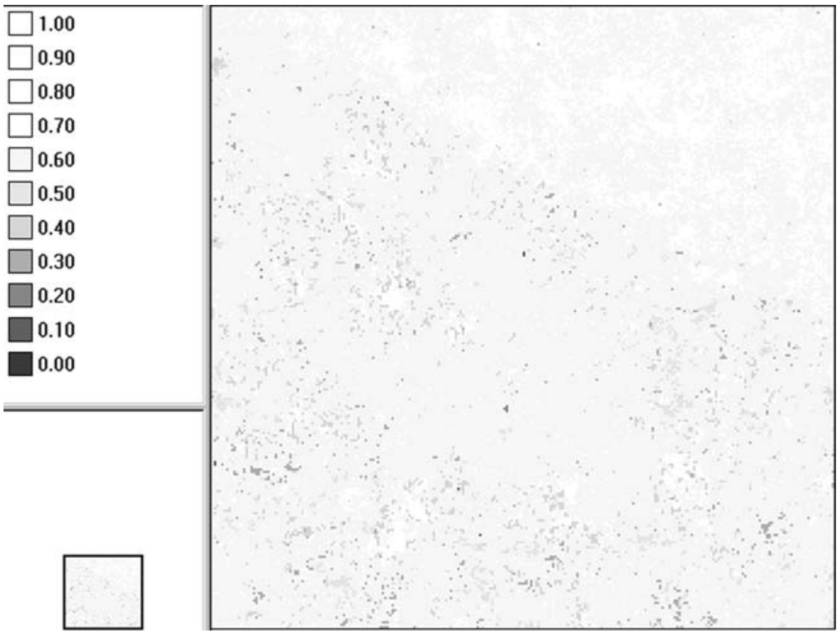

Fig. 12 Pair D2: Symmetrical polygon matching comparison: maximum value map

4.5 Radial dimension

Some processes, for example diffusion limited aggregation (DLA) and urban growth, generate structures that grow outward from a nucleating centre. These structures typically have radial scaling properties, and these are reflected in a log linear plot of area against radius: if the area of the structure (as measured by pixel count) scales as less than the square of the radius, then the structure is a fractal. Cities, for example, scale with an exponent of approximately 1.90–1.95 in an inner, fully urbanized zone, and with an

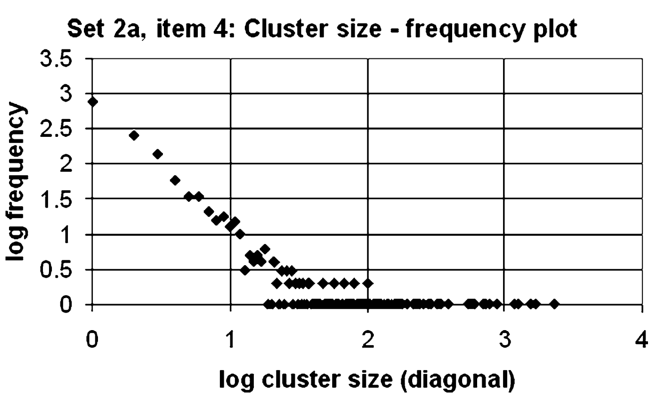

Fig. 13 Pair D2, map a, class 4: cluster size–frequency relationship

exponent of approximately 1.0 in an outer, urbanizing zone, with a clear kink in the area-radius plot at the transition point (Frankhauser 1991, 1994; White and Engelen 1993; White et al. 2001).

An area-radius plot of class 4 in map a, with the radius centred on row 150, col. 135, in the central cluster of class 4, shows a clear scaling relationship with a kink (Fig. 14). The inner zone, with an exponent or scaling factor of $n_i = 1.77$ extends out to a radius of 45, while the outer zone, with a scaling factor of $n_o = 1.55$, extends from 50 to 85 (both area and radii are measured in pixels). These relationships are highly linear ($R^2 = 0.997$ and $R^2 = 0.999$, respectively). Beyond a radius of 90 pixels the slope increases again, as other major class 4 clusters are encountered (Fig. 15). Repeating the analysis for map b (Fig. 16), we find that the inner zone extends out to a radius of 35, with $n_i = 1.85$ ($R^2 = 0.997$), and the outer zone extends from 40 to 150, with $n_o = 1.51$ ($R^2 = 0.999$); beyond a radius of 150 the area-radius relationship becomes non-linear because of the boundary effect of the corners of the map. The fact that the kink point shifts inward shows that the dense core of the class 4 cluster is shrinking, while the dramatic increase of the width of the outer zone shows that the fringe of the cluster has essentially merged with neighbouring class 4 clusters. Combining items 3 and 4 and performing the same analysis raises the inner zone exponent to more realistic values (map a, $n_1 = 1.86$; map b, $n_1 = 1.96$), but also raises the outer zone values (map a, $n_2 = 1.43$; map b, $n_2 = 1.71$) (Figs. 17, 18). So the radial dimension is useful in detecting and characterizing this particular change in pattern.

While the radial dimensions do not permit us actually to infer the process that generated the central cluster on map a and then changed it into the one observed on map b, the particular values of the radial dimension do rule out both DLA and urban growth as the processes involved. While classical DLA could have generated the inner radial dimension observed for the central cluster in either one of the maps, the classical process does not produce a kinked area-radius relationship. Nor could a simple DLA change the cluster observed on one map into the cluster on the other. As for the urban growth process, radial dimensions have been determined for a number of urban

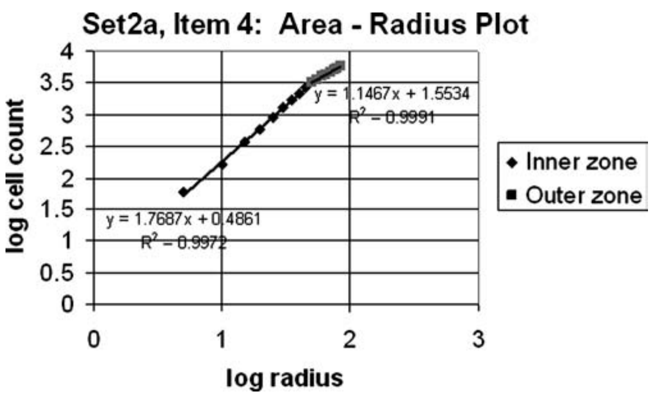

Fig. 14 Pair D2, map a, class 4: area–radius relationship of the central class 4 cluster

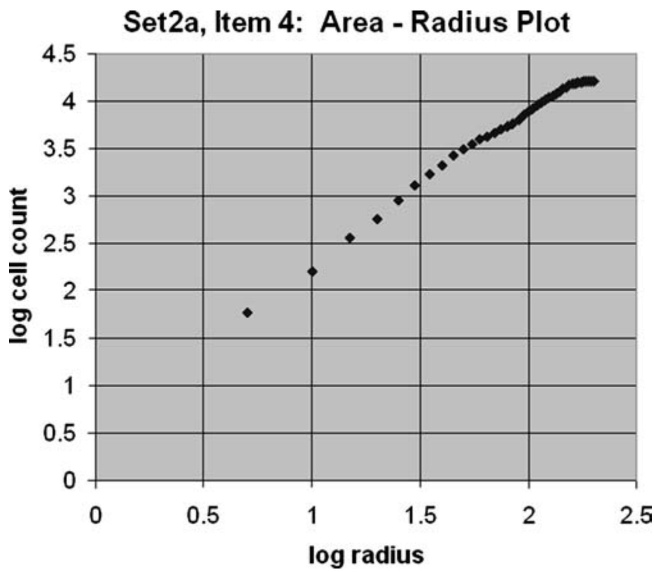

Fig. 15 Pair D2, map a, class 4: complete area–radius relationship

areas around the world (Frankhauser 1994; White and Engelen 1993), as well as for individual land uses within urban areas (White et al. 2001). In every case cities have been found to be characterized by a kinked area-radius relationship like that observed in the data set here. However, the actual values are quite consistent, with typical values of $1.90 < n_i < 1.95$ and $1.0 < n_o < 1.2$, while with one exception the values observed in the present data fall well outside these ranges. Furthermore, in the present data set, the kink point shifts inward from the first to the second time period, a phenomenon which is never observed in actual cities.

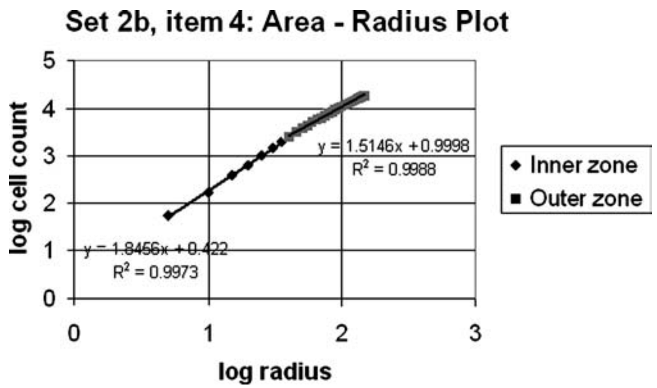

Fig. 16 Pair D2, map b, class 4: area–radius relationship

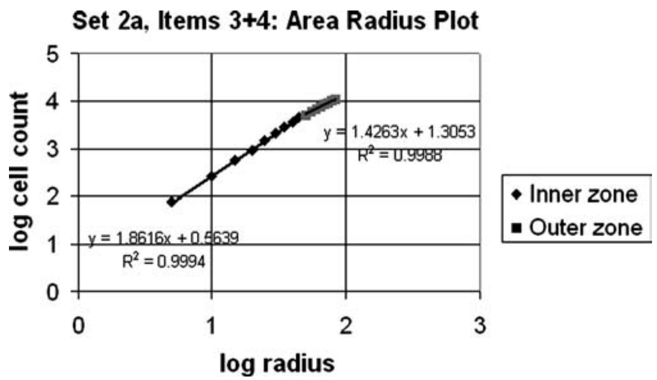

Fig. 17 Pair D2, map a, classes 3 and 4 combined: area–radius relationship

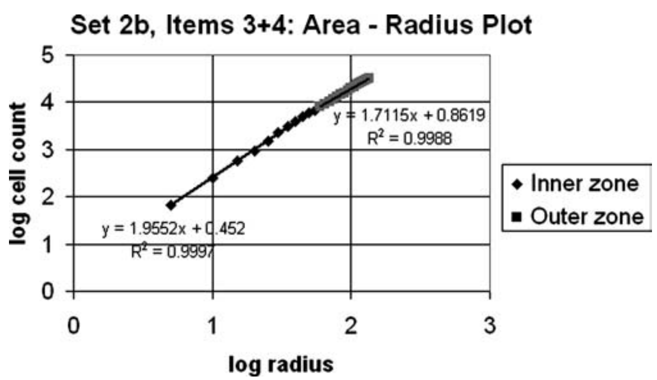

Fig. 18 Pair D2, map b, classes 3 and 4 combined: area–radius relationship

In general then, where radially structured clusters are a significant feature of a map, radial dimension analysis can be used to characterise the pattern; and applied to a series of maps, it can also yield a measure of pattern change.

5 Discussion

On the basis of the analyses performed on the two data sets provided to the participants in this workshop, it would seem that the fuzzy polygon matching map comparison technique has little to offer that is useful in analysing map pairs representing differences in classification, as in pair D1. However, in the context of change detection, the technique seems quite useful, since one of the key concerns in detecting change is to detect changes in *pattern*. This is important in many ecological applications, but also in applications involving the human landscape of cities and regions, as well as the spatial structure of socio-economic systems. Where maps exhibit fractal qualities, the various kinds of fractal dimensions that can be extracted, including the two employed here, can also be useful, although these are highly abstract measures of pattern. However, the importance of pattern is always relative to a particular context, so without knowing the problem context, it is impossible to know what aspects of pattern it is useful to look for or analyse. Ignorance of context is the principal limitation of the present analysis.

Since the early days of GIS and digital remote sensing, the focus has been on data: data acquisition; data accuracy; and data processing, transformation, and presentation. The enormous investments made in this effort have paid off: geographers now work in a data rich environment. Progress in other fields of geography has not been as spectacular. Most conspicuously there is still very little in the way of a formal theory of the dynamics and evolution of geographical systems. In other words, we can now detect change, but cannot say much about what it means in a deeper theoretical sense. What does seem certain is that as in most self-organising complex systems, the processes generating change cannot be rigorously inferred by analysing the data, because in these systems the processes generating the future states are subject to bifurcations, so that their future is open. Thus while change may imply the existence of a process generating the change, the techniques of inferential statistics are not sufficient to make the implicit explicit. Currently the favoured approach is to hypothesise a process, model it, and then see whether it produces *patterns* of change that are similar to what is observed.

In real research situations we almost always know what kind of process is generating the changes we are detecting in the map comparisons, whether or not we have a good model of it. And the type of process operating, along with the purpose of the analysis, together determine the particular map comparison and analysis techniques that are appropriate. For example, if we know that cities are characterised by fractal cluster size frequency relationships and know that these relationships are stable over long time periods, then when we model the urban growth process, we would certainly want to see that the model produces similar, stable relationships. Or, if we are ecologists or New Urbanist planners, perimeter length of the patches or clusters is a significant characteristic, both scientifically and in terms of practical applications. In this case, whatever the processes operating in the real system, the purpose of the analysis would direct us to use measures of perimeter. This is true whether we are testing the adequacy of the model as a vehicle for scientific explanation or whether we are applying it for planning or policy purposes.

The general conclusion is that for problems concerning data quality, like classification problems, the various pixel or cell based approaches to map comparison are most appropriate. For applications in the theory and modelling of geographical systems, however, pattern based map comparison techniques are most useful.

References

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

Frankhauser P (1991) Aspects fractals des structures urbaines. L'Espace Geographique 1:45–69
Frankhauser P (1994) La Fractilité des structures urbaines. Economica, Paris
Frankhauser P (ed) (2000) Morphologie des ''Villes Émergentes'' en Europe à travers les analyses fractales. (Research Report published by CNRS laboratory THEMA) Université de Franche-Comté, Besançon
Gouyet J-F (1996) Physics and fractal structures. Springer, Berlin Heidelberg New York
Hagen A (2002) Comparison of maps containing nominal data. (Report commissioned by the National Institute for Public Health and the Environment (RIVM)) Research Institute for Knowledge Systems, Maastricht
Hagen A (2003) A fuzzy set approach to assessing similarity of categorical maps. Int J Geogr Inf Sci 17:235–249
Mandelbrot B (1983) The fractal geometry of nature. W H Freeman, New York
McGarigal K, Cushman SA, Neel MC, Ene E (2002) FRAGSTATS: spatial pattern analysis program for categorical maps. (Software available at: http://www.umass.edu/landeco/research/fragstats/fragstats.html)
Power C, Simms A, White R (2001) Hierarchical fuzzy pattern matching for the regional comparison of land use maps. Int J Geogr Inf Sci 15:77–100
Visser H (ed) (2004) The map comparison kit: methods, software and applications. (Report 550002005/2004) RIVM, Bilthoven
White R, Engelen G (1993) Fractal urban land use patterns: a cellular automata approach. Environ Plan A 25:1175–1199
White R, Engelen G, Uljee I (1997) The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. Environ Plan B 24:323–343
White R, Luo W, Hatna E (2001) Fractal structures in land use patterns of european cities: form and process. (Paper presented at the 12th European Colloquium on Theoretical and Quantitative Geography, St. Valary-en-Caux, France)