WITTEN, T. A. and SANDER, L. M. (1981) 'Diffusion-limited aggregation: a kinetic critical phenomenon', *Physical Review Letters* 47: 1400–3

WOLDENBERG, M. J. (1973) 'An allometric analysis of urban land use in the United States', *Ekistics* 36, 282–90

ZIELINSKI, K. (1979) 'Experimental analysis of eleven models of population density', *Environ. Plan. A* 11, 629–41

is fundamental to the analysis of urban form. In future research, we hope to extract more accurate population and area measures (Martin, 1989) within more generally-defined urban envelopes, to explore less restrictive theoretical models of urban growth, and to devise more accurate ways of measuring related urban forms.

ACKNOWLEDGEMENTS

The authors wish to thank three anonymous referees for their helpful comments, and the Department of the Environment for use of their digital database of urban areas associated with the 1981 Census of Population. However, the views set out here are solely those of the authors. This research was supported by the Economic and Social Research Council under Grants WA504-28-5006 (Batty and Longley) and WA504-28-5001 (Shepherd).

REFERENCES

ALEXANDER, C. (1966) 'The city is not a tree', *Design* February: 46–55

ARLINGHAUS, S. L. (1985) 'Fractals take a central place', *Geografiska Annaler* 67B: 83–8

BATTY, M. (1991) 'Generating urban forms from diffusive growth', *Environ. & Plan. A*, in press

BATTY, M. and LONGLEY, P. A. (1986) 'The fractal simulation of urban structure', *Environ. & Plan. A* 18: 1143–79

BATTY M. and LONGLEY, P. A. (1988) 'The morphology of urban land use', *Environ. & Plan. B* 15: 461–88

BATTY, M., LONGLEY, P. A. and FOTHERINGHAM, A. S. (1989) 'Urban growth and form: scaling, fractal geometry and diffusion-limited aggregation', *Environ. & Plan. A* 21, 1447–72

BENGUIGUI, L. and DAOUD, M. (1991) 'Is the suburban railway system a fractal?, *Geographical Anal.*, in press

CLIFF, A. D. and ORD, J. K. (1981) *Spatial processes* (Pion Press, London)

DOXIADIS, C. A. (1968) *Ekistics: the science of human settlement* (Hutchinson, London)

DUTTON, G. (1973) 'Criteria of growth in urban systems', *Ekistics* 36: 298–306

FEDER, J. (1988) *Fractals* (Plenum Press, New York)

GOODCHILD, M. F. and MARK, D. M. (1987) 'The fractal nature of geographic phenomena', *Ann. Ass. Am. Geogr.* 77: 265–78

GOULD, S. J. (1966) 'Allometry and size in ontogeny and phylogeny', *Biological Rev.* 41: 587–640

HAGGETT, P., CLIFF, A. D. and FREY, A. (1977) *Locational analysis in human geography* (Edward Arnold, London)

KAYE, B. H. (1989) 'Image analysis techniques for characterizing fractal structures', in AVNIR, D. (ed.) *The fractal approach to heterogeneous chemistry: surfaces, colloids, polymers* (John Wiley, New York) pp. 55–66

LONGLEY, P. A. and BATTY, M. (1989) 'Fractal measurement and cartographic line generalisation', *Computers and Geosciences* 15: 167–83

MANDELBROT, B. B. (1967) 'How long is the coast of Britain? Statistical self-similarity and fractional dimension', *Science* 156: 636–8

MANDELBROT, B. B. (1982) *The fractal geometry of nature* (W. H. Freeman, San Francisco)

MARK, D. M. and PEUCKER, T. K. (1978) 'Regression analysis and geographic models', *Canadian Geographer* 22: 51–64

MARTIN, D. J. (1989) 'Mapping population data from zone centroid locations', *Trans. Inst. Br. Geogr.* N.S. 14: 90–7

MULLER, J-C. (1986) 'Fractal dimension and inconsistencies in cartographic line representations', *Cartographic J.* 23: 123–30

MULLER, J-C. (1987) 'Fractal and automated line generalization', *Cartographic J.* 24: 27–34

MUTH, R. (1969) *Cities and housing: the spatial pattern of urban residential land use* (University of Chicago Press, Chicago)

NAROLL, R. S. and BERTALANFFY, L. VON (1956) 'The principle of allometry in biology and the social sciences', *General Systems Yearbook* 1: 76–89

NEWLING, B. E. (1966) 'Urban growth and spatial structure: mathematical models and empirical evidence', *Geographical Rev.* 56: 213–25

NIEMEYER, L., PIETRONERO, L. and WEISMANN, H. J. (1984) 'Fractal dimension of dielectric breakdown', *Physical Review Letters* 52: 1033–6

NORDBECK, S. (1965) 'The law of allometric growth' Michigan inter-university community of mathematical geographers', Discussion paper No. 7, University of Michigan, Ann Arbor, Michigan

NORDBECK, S. (1971) 'Urban allometric growth', *Geografiska Annaler* 53B: 54–67

OPCS (1984) *Key statistics for urban areas* (Office of Population Censuses and Surveys, HMSO, London)

OPENSHAW, S. (1984) 'The modifiable areal unit problem', *Concepts and Techniques in Modern Geography 38* (Environmental Publications, University of East Anglia)

RICHARDSON, L. F. (1961) 'The problem of contiguity: an appendix to "The Statistics of Deadly Quarrels" ', *General Systems Yearbook* 6: 139–87

SHEPHERD, J. and CONGDON, P. (1990) Small town England: an investigation into population change among small and medium-sized urban areas, *Progress in Planning* 33: 1–111

VICSEK, T. (1989) *Fractal growth phenomena* (World Scientific Company, Singapore)

VOSS, R. F. (1985) 'Random fractals: characterization and measurement', in PYNN, R. and SKJELTORP, A. (eds) *Scaling phenomena in disordered systems* (Plenum Press, New York) pp. 1–11

shown in Figure 8. The rationale for the first two functional regionalizations was two-fold: first, to identify whether the settlements within two more broadly-defined urban fields, approximating the sphere of influence of each of the two largest settlements, shared common characteristics; and, secondly, to make a first attempt at identifying common characteristics between them. The results shown in Tables Va to Vc suggest that although the Norwich region appears to generate higher dimensions than the King's Lynn area and the full set of 70 settlements (Table III), no startling differences emerge.

The rationale for separating out the coastal region was to identify how the constraining impact of the sea restricts the shape and form of the settlements. All of the four dimensions – d, D, δ and Δ – will fall in value if the space within which any settlement can grow is restricted. This is an obvious consequence of constraining the geometry and this effect has been clearly demonstrated on simulated urban growth patterns using DLA (Batty, 1990). In fact, this effect can be seen in Table Vc for the DLA dimension associated with the form of the Norfolk coastal settlements. The slightly higher dimension of the envelope–area relation reflects increased concentration of growth upon the inland portion of each of the settlements, although the dimension of the envelope–radius relation is lower, reflecting the restrictions upon the growth field. From Tables IV and V, it is also significant that it is the DLA dimension D which shows the greatest sensitivity to our regionalization varying from 1·603 to 1·976 in contrast to the other three dimensions where the range of variation is much narrower. Strictly speaking, the R^2 statistics given in Tables IV and V should not be compared with one another, nor with those in Table III for the data sets created by successive deletions from the original set of 70 settlements produce statistically different populations. However, the confidence limits can be compared and this suggests that all the analyses produce results consistent with our prior expectations.

CONCLUSIONS

The relationships between urban morphology and the size and spacing of settlements has been a much neglected realm in spatial analysis, and only since the renaissance prompted by the development of fractal geometry has the interest of geographers been rekindled in these questions. It is in this spirit that we have attempted to reappraise the relationship between population density and urban form within a unified theoretical framework. The framework provided by allometric growth would seem to imply constant urban densities over time and space, whilst fractal growth theory based on deterministic or stochastic growth processes such as DLA implies an attenuating effect of distance upon density from a central seed site, consequent upon the manner in which growth comes to fill space. Although we have demonstrated that our theoretical hypotheses are empirically consistent with the urban settlement pattern in Norfolk, we still require much more empirical analysis if our confidence in these conclusions is to be strengthened. Our work represents only a beginning in this quest.

At this stage, there is also a need for further empirical study in order to ascertain how relations between density and form vary according to the size, spacing and urban history of settlements. There is some scope for such analysis using the OPCS urban areas database. Moreover, there are also prospects for developing more coherent settlement classification systems based upon relating this kind of dimensional analysis to functional regionalizations as well as relating this to change data pertaining to the relative growth and decline of settlements within the broader settlement system. Such research might enable the underlying theoretical assumptions of the fractal model to be made more realistic and to incorporate more of the diversity of real growth processes that we know exists through our casual observation. Moreover, such an analysis might permit a broad-brush appraisal of the reactive role of planning policy in the context of such change, as well as permitting controlled analysis of the impact of long-standing urban containment policy instruments such as green belts. And in a more abstract sense, our theoretical framework would permit investigation of the form of functional settlement hierarchies and their relation to the deterministic fractal geometries of central places such as those suggested by Arlinghaus (1985).

Because this approach depends heavily on theory, it might appear somewhat grandiose, for most empirical research in settlement geography seldom draws directly on theory. What is clear, however, is that measurement prescribes analysis in this work, and that the boundary data we have, only provide very inexact measures of the various individual activity spaces that together define an 'urban' area. As settlements grow and acquire new functions, so the range of land uses which must be incorporated within the urban area also increases. These axioms of central place theory have been almost totally disregarded in research into urban population densities and yet it is clear that this

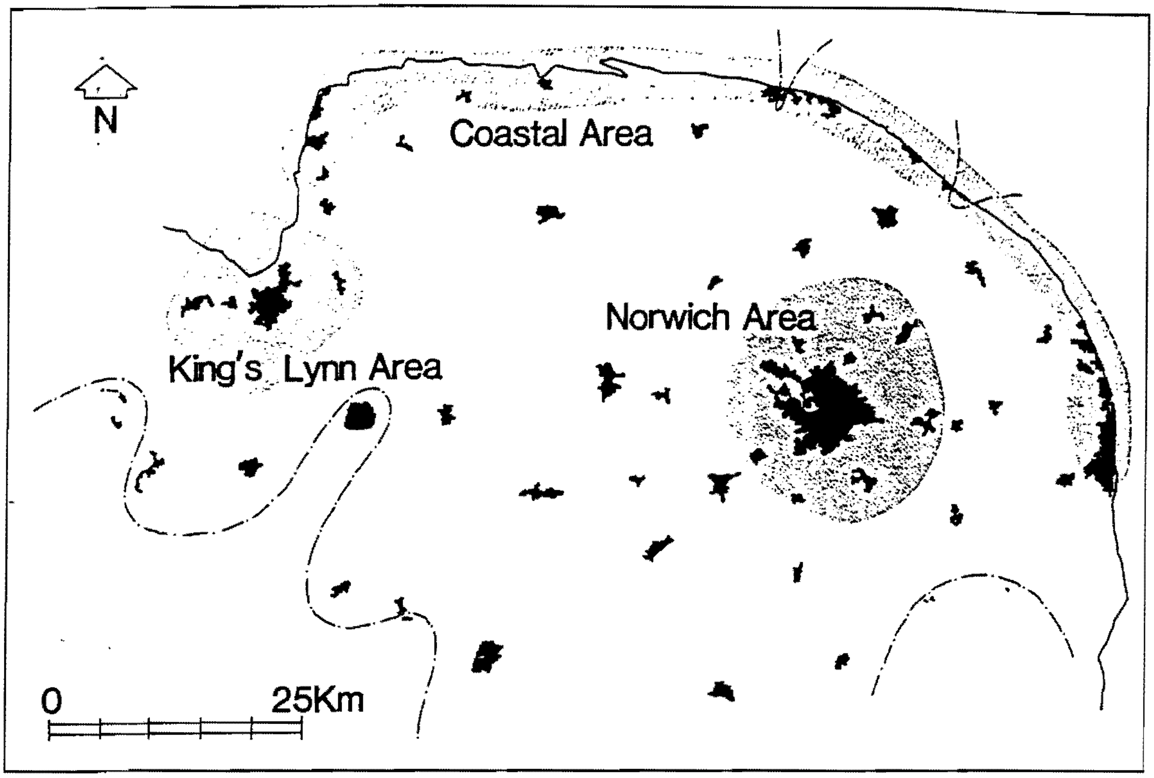

FIGURE 8. Regionalization of the Norfolk settlement pattern

TABLE V. *Estimated dimensions for three subregions of the urban settlement pattern*

	Relationship			
Statistic	*Population–urban area* $d \approx 2$	*Population–radius* $D \approx 1{\cdot}7$	*Envelope–urban area* $\delta \approx 1{\cdot}3$	*Envelope–radius* $\Delta \approx 1{\cdot}2$
[a] Norwich Region				
Slope coefficient	$\alpha = 1{\cdot}037$	$D = 1{\cdot}976$	$\beta = 0{\cdot}601$	$\Delta = 1{\cdot}303$
†Percent of variance R^2	96·3	83·9	86·5	97·2
Dimension	$d = 2{\cdot}074$	$D = 1{\cdot}976$	$\delta = 1{\cdot}202$	$\Delta = 1{\cdot}303$
*95% CI	1·766–2·381	1·319–2·633	0·845–1·560	1·137–1·468
[b] King's Lynn Region				
Slope coefficient	$\alpha = 1{\cdot}009$	$D = 1{\cdot}749$	$\beta = 0{\cdot}623$	$\Delta = 1{\cdot}263$
†Percent of variance R^2	94·0	74·4	90·4	97·6
Dimension	$d = 2{\cdot}018$	$D = 1{\cdot}749$	$\delta = 1{\cdot}246$	$\Delta = 1{\cdot}263$
*95% CI	1·523–2·511	0·802–2·696	0·872–1·622	1·080–1·445
[c] Coastal Region				
Slope coefficient	$\alpha = 1{\cdot}011$	$D = 1{\cdot}635$	$\beta = 0{\cdot}634$	$\Delta = 1{\cdot}029$
†Percent of variance R^2	75·4	72·3	90·8	87·9
Dimension	$d = 2{\cdot}022$	$D = 1{\cdot}635$	$\delta = 1{\cdot}268$	$\Delta = 1{\cdot}029$
*95% CI	1·452–2·591	1·137–2·132	1·069–1·466	0·841–1·217

Notes: *, †, for explanation see Table III

TABLE IV. *Estimated dimensions for the urban settlement excluding the largest towns*

	Relationship			
Statistic	*Population–urban area* $d \approx 2$	*Population–radius* $D \approx 1\cdot7$	*Envelope–urban area* $\delta \approx 1\cdot3$	*Envelope–radius* $\Delta \approx 1\cdot2$
(a) Excluding Norwich				
Slope coefficient	$\alpha = 1\cdot024$	$D = 1\cdot603$	$\beta = 0\cdot624$	$\Delta = 1\cdot125$
†Percent of variance R^2	87·4	71·7	82·6	89·9
Dimension	$d = 2\cdot048$	$D = 1\cdot603$	$\delta = 1\cdot247$	$\Delta = 1\cdot125$
*95% CI	1·858–2·238	1·480–1·849	1·107–1·387	1·033–1·217
(b) Excluding King's Lynn				
Slope coefficient	$\alpha = 1\cdot038$	$D = 1\cdot698$	$\beta = 0\cdot616$	$\Delta = 1\cdot146$
†Percent of variance R^2	89·4	74·5	84·7	90·9
Dimension	$d = 2\cdot075$	$D = 1\cdot698$	$\delta = 1\cdot233$	$\Delta = 1\cdot146$
*95% CI	1·901–2·249	1·455–1·941	1·105–1·361	1·057–1·234
(c) Excluding Norwich and King's Lynn				
Slope coefficient	$\alpha = 1\cdot014$	$D = 1\cdot541$	$\beta = 0\cdot629$	$\Delta = 1\cdot115$
†Percent of variance R^2	85·7	69·3	81·0	89·1
Dimension	$d = 2\cdot029$	$D = 1\cdot541$	$\delta = 1\cdot259$	$\Delta = 1\cdot115$
*95% CI	1·825–2·232	1·288–1·793	1·108–1·409	1·019–1·211

Notes: *, †, for explanation see Table III

Both of the envelope analyses produced high fitting estimates of their dimensions of $\delta = 1\cdot227$ and $\Delta = 1\cdot152$. It is interesting to note that the average dimension of the individual settlement dimensions, computed by applying Richardson's (1961) method to the envelopes of each settlement discussed earlier, was 1·148 and that this compares quite favourably with the value of Δ which is its closest comparator. Finally computing the limits around the values of the slope parameters shows that we can be 95 per cent confident that the true value of the parameter, hence dimension, lies within these limits.

Although these results are encouraging, confirming our initial hypotheses and demonstrating (at least to us) the value of prior theoretical analysis in underpinning such hypotheses, we are also concerned to identify whether or not our results can be disaggregated and generalized to subsets of settlements of different sizes and in different locations. Accordingly, we carried out two further sets of analyses on the data. First, the two largest outlying settlements, representing Norwich and King's Lynn in the graphs of Figure 7, were removed from the data set, first individually and then together. In a statistical sense, this was carried out in order to verify that the high potential leverage effect of these observations was not exerted too strongly against the dominant trend in the data points. In a theoretical sense, this was also important insofar as all of the size and area relations confirm that these two settlements are the most important in the study area, and thus that they might exhibit different relations between density and form. The results of this analysis are shown in Tables IVa to IVc. The R^2 statistics shown there are consistently lower than the corresponding values in Table III, indicating that the major settlements reinforce the general trend in the rest of the data, although all the dimensions in Table IV remain within the 95 per cent confidence limits. With the exception of the envelope-urban area relation, all of the analyses which exclude Norwich and/or King's Lynn produce lower fractal dimensions, suggesting that the global figure is boosted by the particularly tentacular structure of these two settlements.

The second set of disaggregate analyses considered the relations within several subsets of settlements which were defined a priori. Three classes were identified: two regions were delineated around the hinterlands of Norwich and King's Lynn, whilst a third was drawn to embrace all of the settlements along the coast. Settlements which did not clearly fall into any of these categories were omitted. This regionalization is

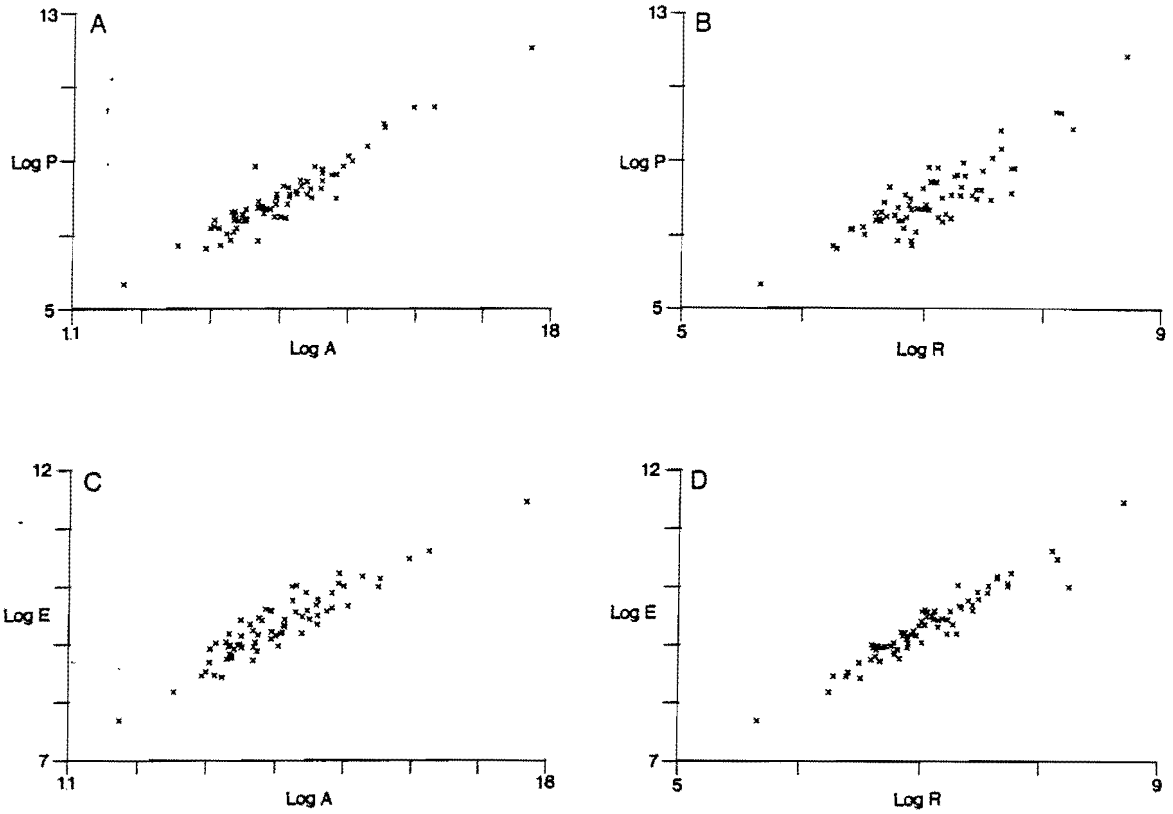

FIGURE 7a–d. Allometric and DLA relations for the 70 urban settlements

TABLE III. *Estimated dimensions for 70 urban settlements*

	Relationship			
Statistic	*Population–urban area* $d \approx 2$	*Population–radius* $D \approx 1{\cdot}7$	*Envelope–urban area* $\delta \approx 1{\cdot}3$	*Envelope–radius* $\Delta \approx 1{\cdot}2$
Slope coefficient	$\alpha = 1{\cdot}043$	$D = 1{\cdot}738$	$\beta = 0{\cdot}613$	$\Delta = 1{\cdot}152$
†Percent of variance R^2	90·3	76·1	85·7	91·5
Dimension	$d = 2{\cdot}085$	$D = 1{\cdot}738$	$\delta = 1{\cdot}227$	$\Delta = 1{\cdot}152$
*95% CI	1·919–2·250	1·502–1·975	1·105–1·348	1·067–1·237

†The R^2 statistic is the coefficient of determination (unadjusted for degrees of freedom) which gives the percentage of the covariation explained by the relationship

*The 95 per cent confidence interval (CI) for the slope coefficient is computed as plus or minus the appropriate percentage point on the '*t*' distribution multiplied by the standard deviation of the slope. In cases where the dimension is twice the slope, we have also doubled the limits of the interval

settlements would have been problematic had it been against the general trend in the rest of the data. The dimension estimated from the DLA population–'radius' analysis is very close to that of a classic DLA structure with $D = 1{\cdot}738$ and this is an encouraging result, particularly in view of the crudity of our approximation to settlement radius. However the level of overall statistical fit is lower with only 76 per cent of the variance explained and high potential leverage effects can again be detected from Figure 7b.

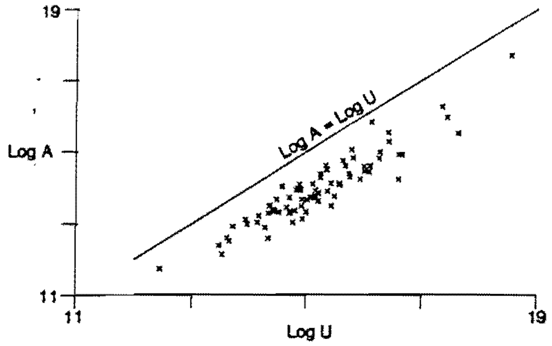

FIGURE 6. The relation between urban area and urban field

EMPIRICAL ALLOMETRY AND FRACTAL GROWTH IN NORFOLK

Central to the assessment of urban shape and form is the notion that the growth of urban areas is fuelled by the functions that each area performs in relation to the rest of the urban system. As we noted earlier, established thinking on the nature of urban densities has paid scant attention either to the spacing or contiguity of settlements or to the relationship between population growth and boundary shape. However, the development of analogies between growth through diffusion-limited aggregation (DLA) and the growth of urban areas offers some prospect for understanding how urban forms and densities evolve within a clearly-specified pattern, whilst investigation of envelope–area relations may reveal how growth occurs at the margins of settlements. Thus both may be seen to complement those more established allometric approaches which reduce form to a simple area measure; hence our approach may contribute towards a more sensitive and comprehensive treatment of population size and form.

Our present empirical analysis is restricted in the degree to which the artifacts of urban growth can be clearly identified. We have already defined the set of urban area data $\{A\}$ through the digitized envelope data $\{E\}$ in the OPCS urban areas data set, and population $\{P\}$ is also a part of this data. However, with respect to our DLA analogies, we do not have data on the field area U or the radius R $(\sim \sqrt{U})$. In the absence of information as to where the historical 'seed' of each settlement is likely to lie, we can calculate a crude approximation to its radius, using 'Feret's Diameter' (F) shown in Figure 2d for Norwich; this enables us to devise a rudimentary 'field' for each of the settlements, and a 'radius' R which is taken as F/2. A further problem is that the rate of urban growth is likely to be uneven at different places around our envelopes and it remains to be seen whether any signals attributable to characteristic growth patterns might be detectable from aggregate measures of the structure and character of the entire set of boundaries. To provide some indication of the way the urban area data set $\{A\}$ relates to the calculated field areas $\{U\}$, Figure 6 illustrates that the relationship between built-up area and field across the range of settlement sizes is strong but erratic, although there is a high positive correlation ($R^2 = 0{\cdot}857$) as might be expected. What Figure 6 does show however is that urban fields are everywhere much larger than urban areas, thus indicating that none of the settlements in the data set are circular and compact, and that all must be irregular, hence possibly dendritic and thus fractal.

In our empirical analysis of the Norfolk data set, we will examine the four sets of relations identified previously. These are: the population-urban area relation, P-A, based on equation 12 in accordance with established allometric analysis; the population–'radius' relation, P-R, based on equation 13 in analogy with urban forms generated by DLA; the envelope-area relation, E-A, based on equation 16 which enables us to identify whether there is any detectable evidence that boundaries are characteristic of growth processes; and the envelope–'radius' relation, E-R, based on equation 17 to identify whether the boundaries of the settlements can be related to fractal growth. Figure 7 illustrates each of these relations for the 70 settlements based on logarithmic transforms of the data as implied by equations 18–21 which are also shown in Table II. We have fitted regression lines to the scatters shown in Figure 7 and the results are given in Table III.

These results generally confirm our a priori expectations. The dimension d of the allometric population-urban area relationship is 2·085, close enough (at conventional confidence limits) to our hypothesized value of 2 to suggest that density is more or less constant with settlement size. Our analysis was carried out for a smaller range of settlement size than previous analyses, and the implication of this finding is to reinforce the simple scaling hypothesis based on a population–area relation found by Wöldenberg (1973) and Dutton (1973), rather than the area–volume hypothesis argued by Nordbeck (1971). The R^2 statistic suggests a high global goodness-of-fit, and the parameter a, hence d, is well above 95 per cent confidence limits, although the high degree of potential leverage exerted by the three largest

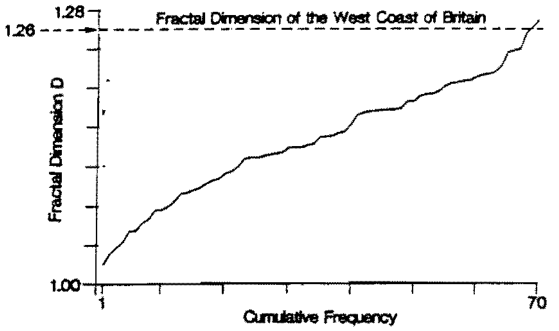

FIGURE 5. The cumulative distribution of fractal dimensions

curve at different scales and calculating their associated lengths. Our algorithm entails measurement of the boundary envelope of each area at a range of successively finer scales, which yield correspondingly increased length measurements as more and more detail on the base curve is picked up. The range of scaled measurements obtained for each parcel was set at between half the mean digitizing intensity for that parcel and one-half of the so-called 'Feret's Diameter', which is the maximum spanning distance between any two points on the digitized base curve (Kaye, 1989), and shown earlier in Figure 2d for Norwich. Regression analysis is then performed on the paired envelope–scale length points to establish whether the envelope is indeed fractal from the value of its fractal dimension. In our previous work (Longley and Batty, 1989), we have found that the structured walk method is the most reliable and robust procedure for computing such dimensions although there is enormous variation in the values of the dimension given using different methods of estimation. This is an important issue but it is outside the scope of this paper and will not be discussed further here.

The motivation for computing the fractal dimension associated with the envelope of each settlement using Richardson's (1961) method is based on the assumption that if the set of settlements are generated by a single process of the kind associated with our theoretical model, then the range of dimensions will be narrow, and this will increase our confidence that there is one single dimension for the whole set of settlements. In Figure 5, we show their cumulative frequency, also indicating the fractal dimension of the west coast of Britain for comparison. Richardson (1961) estimated the fractal dimension of this coastline to be 1·25, and Mandelbrot (1967) suggested that an idealized model of such a coastline might be the Koch snowflake curve which has a fractal dimension of $D = \log(4)/\log(3) \approx 1{\cdot}262$. The mean value for our settlements is rather lower at 1·148 with a standard deviation of 0·059; this would appear to reflect the less intricate nature of man-made boundaries. The range of dimensions is fairly narrow and this does indeed support the theory that a single process may be operating in generating their growth. Moreover, these dimensions do not have high correlations with any of the measures of settlement size and area, namely P, E, A, or U, which we will use in estimating the allometric and fractal relationship in equations 18–21. In terms of R^2, the highest value is 0·130 between D and E, and if logs are taken, this value increases to 0·295, again between D and log E.

These dimensional measurements are not directly comparable with the other measurements reported below due to the fact that our subsequent analysis is based on computing fractal dimensions using the set of 70 settlements as observations of scale change, not scale changes derived by aggregating curves for individual settlements. However, the dimensions reported here are likely to have the same order of magnitude as those we will compute in the next sections for the envelope–area and envelope–field relations and these, as we argued earlier, will be less than those which we will compute from the population–area and population–field relations. This is a consequence of the different means by which the urban boundary is represented as an envelope rather than a perimeter, and strikes at the heart of the argument as to which 'development' should be included in analyses of urban density relations. The urban envelopes which make up the OPCS database each include urban areas which nevertheless have zero population density through space occupied by industrial, commercial or educational land uses, by transport infrastructure or by public open space. By contrast, fine resolution raster representations of urban areas maintain 'holes' of unoccupied land within the outermost urban boundary. This explains why analysis of vectorized urban envelopes are likely to yield lower fractal dimensions, although the measurements will remain internally consistent between settlements. Moreover, when we examine the distribution of the individual fractal dimensions computed here, there is no real evidence of any spatial patterning, suggesting that boundary geometry alone is not a sufficiently strong criterion to enable classification of urban form.

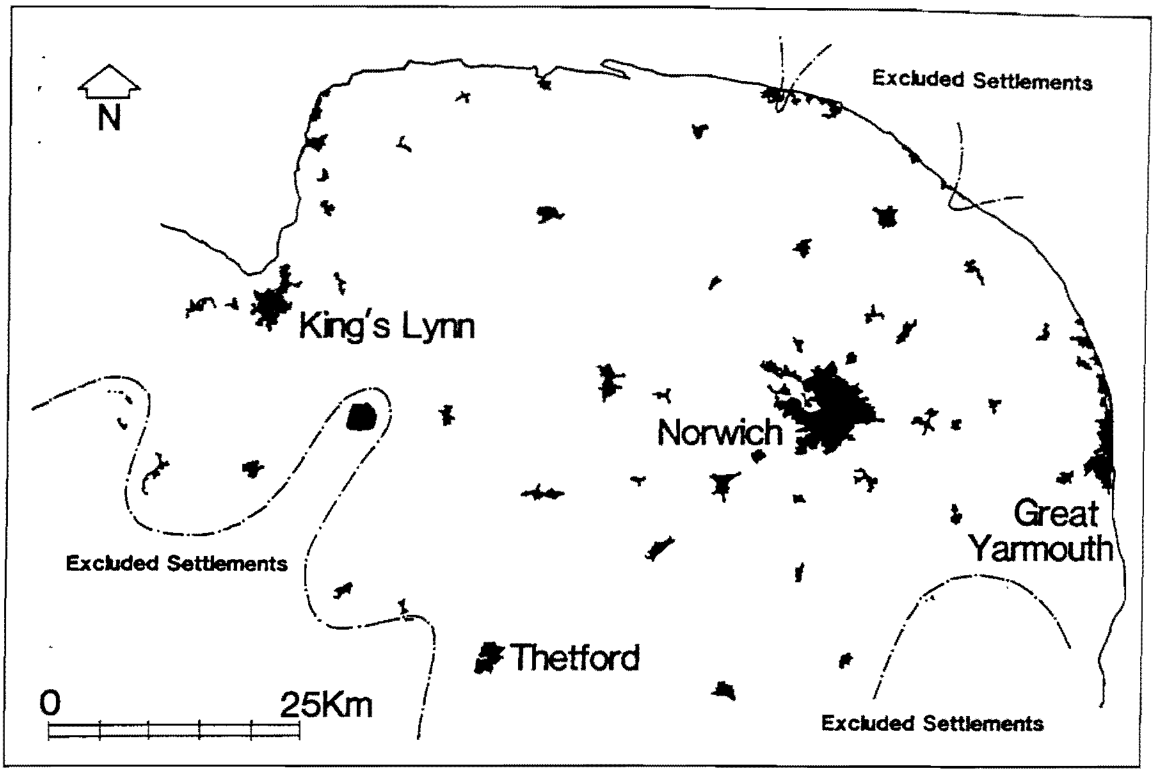

FIGURE 3. The pattern of urban settlement in Norfolk

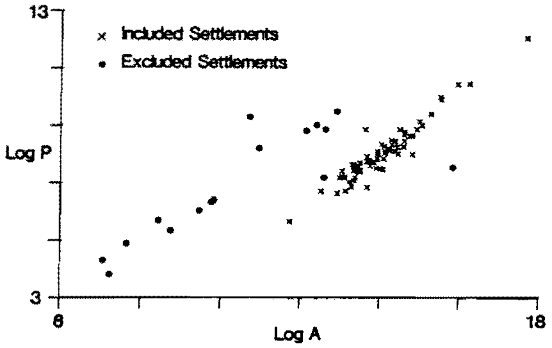

FIGURE 4. The relationship between population and urban area for the entire urban settlement system

pairs of coordinates, making the computation of their fractal dimension using Richardson's (1961) method unreliable, thus further justifying their exclusion.

Our theory posits that population and area co-vary in a systematic way, and thus our preliminary analysis began by checking this. Figure 4 illustrates the relationship between population and urban area for the entire Norfolk settlement system. All 86 settlements are shown here, and the sixteen which were excluded are shown by the open circles. It is clear from Figure 4 that the excluded settlements are amongst the smallest in the system, and with such small settlements – hamlets, and villages – casual observation suggests that their form is dominated by their transport network rather than any neighbourhood differentiation which we assumed in our theoretical model. Moreover, fourteen of the excluded settlements seem to form a different linear relation from the bulk of the remaining settlements while Marham Airfield is a clear outlier which must be excluded on a priori grounds. We must stress that these settlements were excluded at the start of our analysis, and although we computed relationships for the full set of settlements, these will not be reported as they play no part in the subsequent analysis.

We initially computed fractal dimensions for each of these 70 settlements. Calculation of such dimensions is now an established diagnostic for identifying the structure and character of digitized curves (Muller, 1986; 1987). The fractal dimensions of each individual settlement were computed using a 'structured walk' algorithm akin to Richardson's (1961) 'dividers' method based on spanning each digitized

The causes of these discrepancies and sources of possible measurement errors are increasingly understood (e.g. Openshaw, 1984) and the routine innovation of digital databases holds the prospect of greater precision in the delineation of urban areas and monitoring of the areal impacts of change (Shepherd and Congdon, 1990). But nevertheless, there remains cause for concern that even in the data-rich environment of the 1990s, the effects of different measurements of areal units will go undetected in spatial analysis. Moreover, there exist acute definitional difficulties with respect to what is and what is not unambiguously 'urban', and the distance threshold beyond which outlying urban parcels should be classified as physically (and possibly, by extension, functionally) separate from main urban areas. Our own investigations using comparable boundary data recorded at different spatial scales and based upon slightly different digitizing criteria suggest that areal discrepancies of the order of 20 to 30 per cent are likely to be quite common for most settlement sizes. Taken together, this makes it difficult to assess precisely how marginal increments in population lead to changed boundaries of urban forms through the process of accretion, and there is a clear need to develop stronger links between measurement and theory in this context. More generally, vector- and raster-based representations of urban areas are likely to exhibit quite different forms, and in our previous work (Batty *et al.*, 1989; Longley and Batty, 1989), we have attempted to generate insights into the characteristic dimensions which are created using these two different forms of representation.

In this context however, all our data are represented in the vector mode. The data source used in this study is the Office of Population Censuses and Surveys (OPCS) urban areas database (OPCS, 1984) in which urban areas are defined as follows: land on which permanent structures are situated; transport corridors (roads, railways and canals) which have built-up sites on one or both sides or which link built-up sites which are less than 50 metres apart; transport features such as railway yards, motorway service areas, car parks as well as operational airfields and airports; mineral workings and quarries, and any area completely surrounded by built-up sites. The areas were identified using the 1981 1:10 560 Ordnance Survey series in conjunction with 1981 Population Census Enumeration District (ED) base maps. These maps were used to ascertain which areas of urban land contained four or more EDs, and on this basis, these qualified as urban areas. Population figures from EDs which had 50 per cent or more of their population within an urban area were included in the population total for that area. Further general information and details of the treatment of small areas of population and discontiguous urban land can be found in OPCS (1984). These boundaries were then reduced to the 1: 50 000 scale and computer digitized to an accuracy of 0·5 mm permitting inaccuracies of up to 25 m on the ground. The boundaries were digitized manually, in point mode with a weed tolerance of 2·54 mm. Three sources of error thus exist in the digital data, namely, error created by the transfer of the urban areas between the two map scales, digitizing errors of up to 25 m on the ground, and original map error in the two map scales used. Our empirical case study uses data for the County of Norfolk which have been extracted from these sources.

The data comprises 86 distinct urban settlements from populations as small as 45 to the major county town of Norwich which has about 186 000 people. The pattern and form of these urban settlements are shown in Figure 3. We have already alluded to the difficulty of defining and adhering to definitions of urban land which are both unambiguous and appropriate to any specific task, and it is likely that the original decision by OPCS to include some of the smallest settlements was in practice an arbitrary one. We might thus anticipate that the population and area of these smallest settlements would not closely correspond to any empirical regularities extant elsewhere in the data set, as a result of disproportionate errors in the measurement of their populations and bounding envelopes. Settlements whose form is dominated by transport infrastructure are also likely to be 'unusual' in both geometrical and population terms, and such settlements will be primarily, but not exclusively, small in size and scale.

We immediately reduced our set of 86 settlements to 70. Thirteen of these settlements were cut quite arbitrarily by the administrative County boundary, and are thus not representative either numerically or geometrically of their related settlement forms. We also removed two coastal settlements which were apparently subject to digitizing errors in that their boundaries criss-crossed in a quite ludicrous fashion. Marham Airfield was excluded because its low population could in no way be judged to be representative of its large land area given over to runways etc., and because its form could not be seen as being consistent with the sorts of urban growth processes we were exploring. In fact, apart from the Airfield, all the settlements we excluded had fewer than 55 digitized

TABLE II. *The basic relationships of settlement size, shape and dimension*

Relationship	*Variables*	*Dimension & predicted value*	*Untransformed equation*	*No.*	*Log-transformed equation*	*No.*
Allometric	Population P Area A	d (≈ 2·0)	$P = aA^{\alpha} = aA^{d/2}$	(12)	$\log P = \log a + \alpha \log A$	(18)
	Length of envelope E Area A	δ (≈ 1·3)	$E = bA^{\beta} = bA^{\delta/2}$	(16)	$\log E = \log b + \beta \log A$	(20)
DLA	Population P Field radius R	D (≈ 1·7)	$P = gR^{D}$	(13)	$\log P = \log g + D \log R$	(19)
	Length of envelope E Field radius R	Δ (≈ 1·2)	$E = hR^{\Delta}$	(17)	$\log E = \log h + \Delta \log R$	(21)

mean value of $\delta = 1{\cdot}296$ (Batty and Longley, 1988). We have also computed the dimensions of the urban envelopes of the town of Cardiff across different scales using variants of Richardson's (1961) 'walking by dividers' method, and there we found that the dimension varied between 1·172 and 1·308 (Longley and Batty, 1989). This also suggests that δ and Δ will have values less than D and d.

Pulling all these threads together, we will hypothesize that the four dimensions associated with the four scaling relationships given in equations 12, 13, 16 and 17 should be ordered as $1 < \Delta < \delta < D < d$, where Δ, $\delta \approx 1{\cdot}26$, $D \approx 1{\cdot}71$, and $d \approx 2$. The constants associated with these four relationships can be estimated from regressions of their log-linearized forms. We will refer to these relationships as being of allometric or DLA (diffusion-limited aggregation) type, involving independent variables of occupied area or urban field. Table II summarizes these relations and for completeness, the log-linearized forms of equations 12, 13, 16 and 17 are given as

$$\log P = \log a + \alpha \log A, \; (d = 2\alpha), \qquad (18)$$

$$\log P = \log g + D \log R, \qquad (19)$$

$$\log E = \log b + \beta \log A, \; (\delta = 2\beta), \text{ and} \qquad (20)$$

$$\log E = \log h + \Delta \log R. \qquad (21)$$

Equations 18 to 21 will be those whose parameters will be estimated in the sequel and used to establish the consistency between the form of the urban settlement system in Norfolk and the theoretical and simulated DLA models outlined in this and the previous sections.

DATA REPRESENTATION AND INITIAL ANALYSIS OF THE NORFOLK DATA

We have already focused upon some of the difficulties of measuring the relationship between the size and form of urban settlements. Early work on the size relations within settlement systems was necessarily restricted by the quality of the measures of the precise areal extent and population size of constituent areas. Naroll and Bertalanffy (1956) attributed much of the variation in international urban–rural population ratios to differing national definitions of 'urbanity' and the differing areal extent of data collection units which together comprise urban areas. Newling (1966) encountered problems of the changing areal basis of data collection in his study of the evolution of intra-urban population density gradients over time. And as we have noted, Woldenberg (1973) obtained some quite radically different estimates of population–size relations in his cross-sectional study of the US settlement system, depending upon his use of one or other of two atlases to obtain his urban area measurements. In the face of such vagaries and inconsistencies, it is scarcely surprising that the nature of the empirical relationship between size and spatial form remains obscure. We have already begun to clarify some of these issues in earlier sections and our empirical analysis which follows is designed to cast further light on these questions.

dimensions which can range over the interval from $D = 1$ to $D = 2$, associated respectively with linear to completely compact clusters (Niemeyer *et al.*, 1984; Batty, 1990). In fact, a related goal of this research which is beyond the scope of this paper, is to explore the extent to which D might vary across a range of settlement sizes.

Finally, it is worth noting that equations 12 and 13 also imply density relations for the areas in question. Dividing equation 12 by area A gives

$$\rho_A = \frac{P}{A} = aA^{\alpha-1}, \qquad (14)$$

and 13 by R^2 ($\sim U$)

$$\rho_B = \frac{P}{R^2} = gR^{D-2}. \qquad (15)$$

If $\alpha = 1$, then equation 14 gives a constant density $\rho_A = a$, while if $1 < D < 2$, equation 15 yields a decreasing density with increasing field size as expected. In the sequel, we will not present any estimates of these density relations because they are equivalent to the basic size-area equations in 12 and 13, although we will examine some of these densities empirically.

Relationships between the length E of the bounding envelope of urban development and the area A and field radius R will also be explored here. It is important to note that the bounding envelope is not the perimeter of the cluster in that any undeveloped interior of the cluster is not detected by the envelope (see Figures 2a and 2b). In fact, as the envelope defines the outer edge of the cluster, it is likely to be smoother and less circuitous than the perimeter, and this suggests that any measure of the fractal dimension of such a line is likely to be less than the fractal dimension of the cluster. In the case of the urban area A, we can relate the envelope to the assumed radius $r \sim A^{1/2}$ of occupied area, giving

$$E = bA^{\beta} = bA^{\delta/2}, \qquad (16)$$

while for the field radius, a similar relation is postulated

$$E = hR^{\Delta}, \qquad (17)$$

where β, hence δ in equation 16 and Δ in equation 17 can be regarded as 'dimensions' with b and h as constants of proportionality.

Our scaling model also predicts values for δ and Δ in the theoretical case shown in Figure 1c. The occupied area A_k of the growing fractal is proportional to the number of occupied units N_k, and T_k, the perimeter, is a function of N_k as shown in equation 10. Thus where the envelope is defined by the perimeter, the value of δ in equation 16 would be 2. In short, E scales directly with occupied area, or $E = bA$. This of course is the same analysis as already applied to equation 12. In the case of the field radius R, where $T_k \propto N_k$, the envelope E scales as R^{Δ} with $\Delta = D$ which is the same as implied by equation 13. At the other extreme, if the envelope scales with L_k, not N_k, this implies that the perimeter is a simple function of the linear scale, and $\delta = \Delta = 1$; these represent lower bounds on the scaling constants. In fact, in defining an envelope for the fractal in Figure 1c, such an envelope would clearly be more circuitous than the linear scale L_k but not as circuitous as the actual perimeter. Thus it is likely that $1 < \Delta, \delta < D$. Moreover, as the occupied area is likely to be less than the field area, then it is likely that $\Delta < \delta$.

Before we state the relationships which we seek to validate empirically, it is worth noting the theoretical bounds within which our analysis will take place. From Figure 1, it is clear that our theoretical model enables us to consider how a continuum of forms might be measured, from urban development which is linear across the plane to that which is a completely compact circle. In the case of the completely compact cluster, its occupied area and its field are coincident with $P = aA = bU \sim \pi R^2$ and with $d = D = 2$. The growing zone at the edge of the cluster is the same as the perimeter and this is defined as the derivative of P with respect to radius R, that is $dP/dR \sim \pi R$. The envelope is also the perimeter in this case with $E = gA^{1/2} = hR \sim \pi R$ and $\delta = \Delta = 1$. In the case where the cluster is linear $P = gA^{1/2} = bU^{1/2} \sim \pi R$ and $d = D = 1$, while the derivative of P does not provide the formula for the perimeter, just the growing zone which is always a point of zero dimension, implying in this case that $\delta = \Delta = 0$. In the case of a real urban cluster which does not completely fill its available space, area, perimeter and envelope can be approximated by space-filling lines which suggest that all the dimensions of significance – δ, Δ, D and d – will be between 1 and 2. The only examples we are aware of where the dimensions of urban envelopes have been estimated is in our own work; we have computed fractal dimensions for a series of land use parcels in the town of Swindon using perimeter–area relations similar to that in equation 16 and this has yielded dimensions in the range $\delta = 1{\cdot}243$ to $1{\cdot}478$ with a

We will examine two types of relationship between these variables, first relating population P to area A and to field radius R, second relating the length of the envelope E to these same variables. These types of relationship are used in the study of allometry or 'relative size' (Gould, 1966), and by relating size and length to area, this enables us to explore questions of density. In this way, we can relate our work to the literature on urban allometry (Dutton, 1973) as well as to our own previous work on fractal geometry (Batty and Longley, 1986; Longley and Batty, 1989; Batty *et al.*, 1989).

The classic allometric relation we will begin with involves the relationship between population size P and occupied area A which we can write as

$$P = aA^{\alpha} = aA^{d/2}. \qquad (12)$$

a is a constant of proportionality and α is a scaling constant. In equation 12, we have also written α as d/2 where d can be interpreted as a 'dimension' of the occupied area, scaling the 'radius' r of such an area ($r \sim A^{1/2}$) to population. The use of this convention will become clear in the sequel when all the scaling parameters have been introduced. There is obviously a strong relationship between population and area although the precise form of the scaling is problematic. Nordbeck (1965; 1971) suggests that the scaling constant α should be 3/2 using the argument that population growth takes place in three dimensions; thus if $r \sim A^{1/2}$ is taken as the linear size of area, then $P \sim ar^3 \sim aA^{3/2}$. This hypothesis is borne out in an analysis of the urban population of Sweden in 1960 and 1965 (Nordbeck, 1971). Results from urban density theory also suggest that as cities get bigger, their average density increases but the empirical evidence on this is mixed and is much complicated by the definitions of urban area used (Muth, 1969). However, Woldenberg (1973) shows quite unequivocally that $\alpha \approx 1$ from an analysis of two large population–area data sets for American cities.

In the case of the scaling model introduced earlier, it is clear that the area occupied by the units of development N_k varies as the development itself. For the growing fractal, the area of each occupied cell is ξ^2, thus the total area is $A_k = N_k\xi^2$. In short, the population density $N_k/A_k = \xi^{-2}$ is constant regardless of scale or the stage reached in the growth process. In fact, this is an assumption of the model. If we equate population P with N_k and area A with A_k, then we might expect the empirical relation between P and A to be of the simplest kind – perfect scaling – with both the theoretical model and much empirical evidence suggesting that $\alpha \approx 1$ and $d \approx 2$.

With respect to the urban field, the scaling between P and U is more complicated. As cities grow, their field becomes correspondingly larger, growing at a more than proportionate rate, and in the case of very large cities, the urban field is often considered to be global. This implies that as cities grow, their field density P/U always decreases. It is more appropriate in this analysis to represent the field area U in terms of its 'radius' $R \sim U^{1/2}$. Thus the field relationship can be stated as

$$P = gR^{D} = gU^{D/2}. \qquad (13)$$

g is a constant of proportionality and D is the scaling constant. For most cities, this constant will be less than 2 but greater than 1 as is the case in Figure 2b and to anticipate our analysis, D in fact is a fractal dimension, a measure of the extent to which P fills its available field. Only recently has there been any empirical work at all in measuring this relationship – research by two of the authors (Batty *et al.*, 1989; Batty, 1990) has yielded fractal dimensions D for real cities varying between 1·5 and 1·8 – and it is one of the main purposes of this paper to provide such measures for a system of urban settlements.

One of the basic relationships in the theoretical model of the previous section is equivalent to the field relationship in equation 13. Equations 3 for the static cluster and 9 for the growing cluster scale in the same manner as equation 13. For the growing cluster, the theoretical relation is $N_k = L_k^{D}\xi^{-D}$ where it is clear that L_k, the linear measure associated with the scale k, has the same role as the radius R. In the theoretical model, $D \approx 1{\cdot}465$ and it would be a simple matter to extend this model to cases where more of the available space were filled, increasing D towards its maximum bound of 2 where all space is occupied. The scaling model is a deterministic version of a stochastic model which has been widely applied in the theoretical physics of particle clusters; stochastic processes based on the constrained diffusion of particles around seed sites have generated a class of models known as diffusion-limited aggregation or DLA models (Vicsek, 1989; Witten and Sander, 1981). In these models, sparse aggregates similar to those shown in Figure 1 are grown in a continually changing potential field, and these have a remarkably consistent, perhaps universal, fractal dimension of $D \approx 1{\cdot}71$. A generalization of the DLA model has been developed which generates clusters with fractal

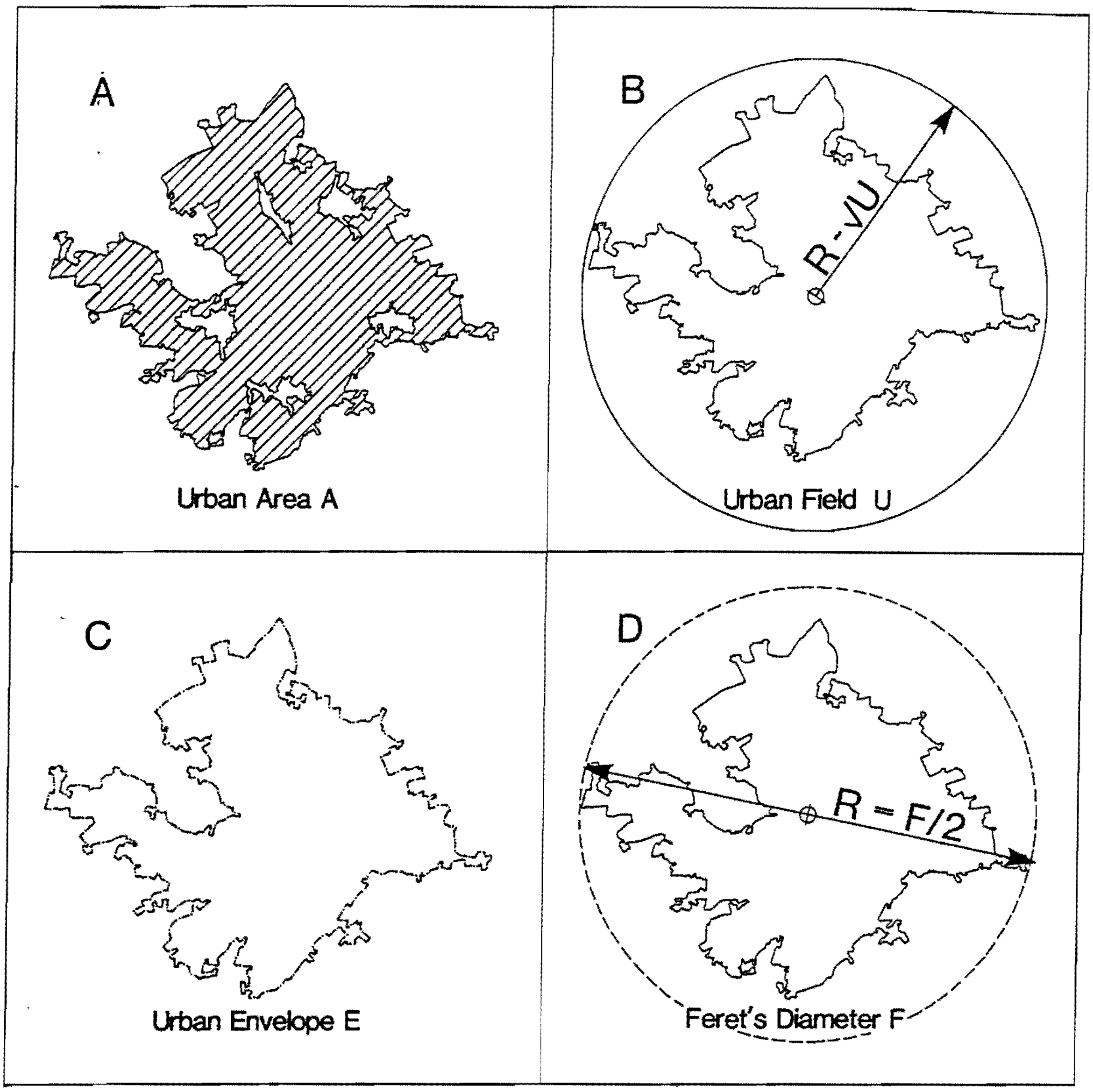

FIGURE 2. Definitions of urban area, field, envelope and radius

Thirdly, there is another variable of interest which relates area A to field size U and this is the urban envelope E defined as the length of the boundary which marks the greatest extent of the built-up area. To provide some meaning to these concepts, we have illustrated their spatial definition using the example of the largest town from our data set, Norwich; these definitions are shown in Figure 2.

Figure 2a shows the built-up urban area whose extent A is indicated by the cross hatch and it is this area that accommodates the population P. The urban field is shown in Figure 2b, and this is the bounding circle based on the centre of the cluster which is marked by the maximum radius R and which contains the whole cluster. The area of the cluster is given as $U = \pi R^2$ and $U > A$. The urban envelope is shown in Figure 2c, its length E being a measure of both the size and the shape of the cluster. In Figure 2d, the maximum spanning distance across the cluster, known as 'Feret's diameter' (Kaye, 1989), is shown. The length of this span is defined as F and this will be used later in estimating and approximating the radius R where the centre of the cluster shown in Figure 2b is unknown.

TABLE I. *Measurements of scale change in the theoretical urban cluster*

Measure of scale	*Equation*	*Scale*			
		k = 0	k = 1	k = 2	k = 3
No. of units	$N_k = 5^k$	1	5	25	125
No. of subdivisions	$n_k = 3^k = L_k$	1	3	9	27
Length of unit	$\xi_k = 3^{-k}$	1	1/3	1/9	1/27
Relative perimeter	$T_k = 5^k 3^{-k}$	1	5/3	25/9	125/27
Perimeter of Figure 1b	$T_k = 4\sqrt{2}(5/3)^k$	5·657	9·428	15·713	26·189
Perimeter of Figure 1c	$T_k = 5^k/27$	1/27	5/27	25/27	125/27

be discussed in the sequel when we equate these measures with population and units of area or distance.

At any scale, it is possible to compute the fractal dimension D from equations 3 or 5. From equations 2 and 3, and using logarithms (which are to the base e throughout this paper), we get

$$D = \log \frac{N_k}{n_k} = \frac{\log N_k}{\log L - \log \xi_k}, \qquad (6)$$

while from equations 2 and 5

$$D = 1 + \frac{\log T_k - \log z - \log L}{\log n_k},$$

$$= \frac{\log T_k - \log z - \log \xi_k}{\log L - \log \xi_k}. \qquad (7)$$

Equations 6 and 7 are equivalent and can be simplified if the arbitrary constants z and L are set to unity.

For the hypothetical case in Figure 1a, it is clear that at each level of resolution k, $N_k = 5^k$ and $n_k = 3^k$. Assuming that $L = 1$ and $z = 1$, then $\xi_k = 3^{-k}$ and $T_k = 5^k 3^{-k} = (5/3)^k$. These calculations for the four scales $k = 0$, 1, 2, and 3 are shown in Table I. It is also clear that the fractal dimension D is constant across scales. From equation 6 for any scale k, $D = \log(5^k/3^k) = \log(5)/\log(3) \approx 1{\cdot}465$. Equation 7 however illustrates that D has a lower bound of 1 which would result if there was no increase in the perimeter of the cluster with scale k and an upper bound of 2 where the perimeter T_k increases at the same rate as the number of subdivisions n_k.

Table I also shows that the linear perimeter of the fractal increases from 1 to 125/27 over the four scales and it is clear that as $k \rightarrow \infty$, the perimeter T_k will continue to increase. This is a simple demonstration of the 'length of coastline conundrum' articulated in recent times by Richardson (1961), and used by Mandelbrot (1967) in his early expositions of fractal geometry. In Figure 1b, we represent the structure of the same cluster at each level of resolution by a connected line which 'fills' the occupied space through spanning. If we were to trace out a perimeter around this curve, we would count each diagonal span of the basic grid unit twice, there also being two such spans (diagonals) of each square. The length of each diagonal at scale k is $\sqrt{2}/3^k$ where $L = 1$ and as in this case where there are four spans, each grid square has a perimeter of length $4\sqrt{2}/3^k$. Using equation 5 where we now assume $z = 4$ and $L = \sqrt{2}$, $T_k = 4\sqrt{2}(5/3)^k$ which we have also shown in Table I.

The same theory can be used to generate a growing cluster such as that shown in Figure 1c (Voss, 1985). Let us now define a linear length scale L_k which is the total length of one side of the growing cluster at scale k. Then assuming a basic unit of development of linear length ξ, the total number of subdivisions of L_k is given as

$$n_k = \frac{L_k}{\xi} \qquad (8)$$

The number of occupied basic units of development at scale k is also called N_k and from Figure 1c, it is clear that the rate of increase of this unit scales with n_k in the same manner as in Figure 1a; that is from equation 8

$$N_k = n_k{}^D = L_k{}^D \xi^{-D} \qquad (9)$$

which is of the same form as equation 3. The perimeter of the growing fractal is, in terms of the length scale ξ,

$$T_k = zN_k\xi \qquad (10)$$

which can also be represented as

$$T_k = zL_k n_k^{(D-1)} = zL_k^D \zeta^{(1-D)} \qquad (11)$$

From equations 9 and 11, the fractal dimension D is clearly the same as that computed for the static cluster while equation 10 shows that the perimeter scales at the same rate as the number of units of development N_k. For purposes of demonstration, if we assume that the basic unit $\zeta = 1/27$, then N_k in equation 9 and L_k from equation 8 vary as N_k in equation 3 and n_k in equation 2 respectively. T_k from equations 10 or 11 can also be computed for this growing fractal and this is shown as the last row in Table I. Once the growing fractal reaches $k = 3$, it is the same size as the static fractal, hence its perimeter is the same as is its number of occupied units. In the case of the static fractal, N_k is the number of units detected at each scale in contrast to the growing fractal where N_k is the number generated by time k. For the growing fractal, scale and time are thus synonymous. Finally although we have not shown this, Figure 1c could be represented as a connected line spanning the growing cluster as was done for the static cluster in Figure 1b; the perimeter of this curve associated with the growing fractal would then be proportional to T_k as given in the last row of Table I.

THE MATHEMATICS OF SIZE, SHAPE AND DIMENSION

Before we discuss the relevance of the scaling model to the form of settlements at different sizes, some of the limitations posed by its adoption should be clarified. The power functions which scale the spatial resolution level given by n_k to the units of development N_k in equations 3 and 9, and to the perimeters T_k in equations 5 and 11 are the only functions uniquely determined by the growth process shown in Figure 1 which is based on the replication of self-similar forms. This is clearly demonstrated by Mandelbrot (1982) but this does not imply that all power functions can be generated by fractal growth processes. Although a growth process may be consistent with a given power function, there are many other processes which could give rise to the same functional form. If we can demonstrate that the size of settlements can be predicted using similar power functions, this will increase our confidence in the process outlined above which we consider to be a plausible baseline model for urban growth and form. This is not a proof that the fractal model applies, for power functions are consistent with other growth processes, for example, those based on self-affine relationships which imply different power functions for different sizes of settlement. In this sense then, our baseline model is quite restrictive and only acts as a starting point in research.

We are also making the assumption that settlements of different sizes represent the operation of the same process at different stages of growth, that is, that the largest settlements in the system are simply larger versions of the smaller ones. In assuming this, we are working in the same spirit as many researchers working with fractal geometry (Feder, 1988; Vicsek, 1989). Moreover, we also equate the number of units of development N_k with the number of units of population, and as each unit of development N_k is located in the same amount of local space, then this means that the density of units is the same regardless of the size of the cluster. Although we will test for this assumption, it is less restrictive than might appear at first sight for it excludes surrounding undeveloped space, and we will pick this up in the sequel in our distinctions between urban area and urban field. In our empirical study below, we minimize reliance on this assumption by restricting our model to a relatively homogeneous pattern of settlement falling within a restricted size range and serving a primarily agricultural region. As this area has been relatively isolated from urban growth based on industrial manufacturing and related urban services, it appears tenable to hold that the smallest settlements have the potential to grow into the largest settlements in this region. The case study based on the pattern of settlement in the English County of Norfolk in the region of East Anglia, reflects these characteristics.

As suggested earlier, the two basic measures of size which we will use are population and area. Our task will be to seek relationships between these variables, first by identifying how these variables might best be defined, and secondly, by exploring how the scaling model of the previous section might be used to illuminate the postulated relations. Associated with the population P of any urban cluster, there might be several definitions of area. We will use two distinctive measures here: first there is the occupied area called A which can loosely be defined as the built-up or developed area. Secondly, there is the urban field whose area U can be defined as the hinterland immediately associated with the greatest radial extent of the cluster, that is the immediate circle of area within which growth has already taken place.

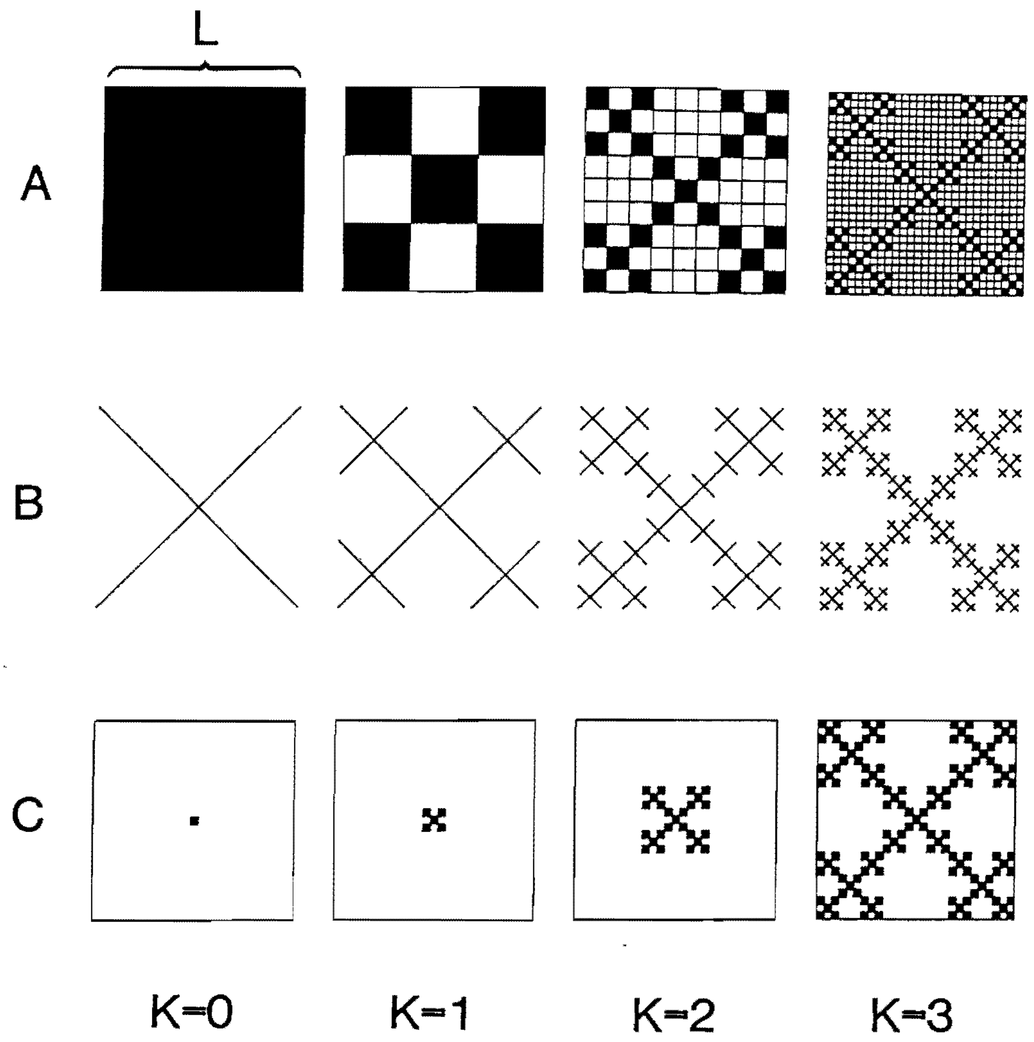

FIGURE 1. A theoretical model for measuring and generating urban growth and form

$$N_k = n_k^D = L^D \xi_k^{-D}, \qquad (3)$$

where D is a scaling constant greater than 1 but less than 2. In fact, D is a dimension which can be interpreted in this case as a measure of the extent to which the urban cluster fills the space available. If the whole space were to be filled at each scale k, then $D = 2$, while if only a line of squares across the space were filled, then $D = 1$ (Mandelbrot, 1982).

Associated with the increase in the number of units of the urban cluster is the absolute amount of space they occupy. One measure of this is the perimeter of the occupied units which in terms of the number of grid squares, N_k, in our example in Figure 1a, is four times the actual length of each side of each square multiplied by the total number of occupied squares, N_k. Thus a general measure of perimeter, T_k, can be given as

$$T_k = zN_k\xi_k, \qquad (4)$$

where z is the constant which scales the length ξ_k to the perimeter of each basic unit. Using equations 1 to 3 in 4 gives

$$T_k = zn_k^D L n_k^{-1} = zLn_k^{(D-1)}$$
$$= zL^D \xi_k^{-D} \xi_k = zL^D \xi_k^{(1-D)}. \qquad (5)$$

Equations 3 and 5 determine basic measures of size which we will use in the empirical investigation to shape might have more practical relevance in the representation of digital data.

In this paper, we will work towards a consistent theory of urban growth and form in a system of urban settlements, combining allometric relationships and fractal geometries. We will illustrate our theory with data on the size, shape and spacing of urban settlements in the County of Norfolk in the English region of East Anglia. We first propose a scale theory of urban growth and form which casts allometric relations in the framework of fractal geometry, and we are then in a position to develop an appropriate mathematics linking size, shape and dimension. After this, we will examine the data for our case study, first briefly reviewing the principal means by which urban shapes and areas are represented through boundaries or 'envelopes', and we will investigate some characteristics of the urban settlement pattern in Norfolk before beginning our major analysis. We will then apply the various scaling relationships which we consider of major importance in linking size to shape through dimension to the urban settlement system in Norfolk, showing how the hypotheses which we will set out in the next two sections, are consistent with the data. Finally we will propose a number of new directions and extensions to this work which we hope to consider in future research.

A SCALE THEORY OF URBAN GROWTH AND FORM

In attempting to measure and interpret salient characteristics of urban form relative to size and shape, it is essential that we have in mind some theoretical baseline against which empirical results can be compared. Moreover, we require a theory applicable to growing cities which enables growth through replication of some basic unit. We postulate that cities grow by the accretion of new neighbourhoods, around which urban development is clustered, neighbourhoods in turn forming districts, districts forming sectors and so on. The theory we will propose utilizes this concept to generate a hierarchy of clusters having the property of self-similarity across a range of scales.

The scale theory we present is immediately applicable to the growth of a single settlement in which a hierarchy of neighbourhoods is strictly determined by the basic unit of development. However, from our casual observation of settlement patterns which range from the smallest hamlets to the largest cities, we would argue that the basic unit of development is very similar in very different sizes of town or city. In this sense, we assume that large cities are simply small cities with a greater level of hierarchical differentiation but with the same basic constituents of urban form at the neighbourhood level. This is implied in a very wide literature (see, for example, Alexander, 1966; Doxiadis, 1968), and it enables us to use the scale theory not only to detect urban form within particular cities, but to explore a system of settlements of different sizes. Our assumption then is that the form of urban settlements is self-similar not only within the settlement in question but between settlements of different sizes.

We will present the theory using an hypothetical example of urban form in which growth and development is located in regular clusters on a square lattice at different scales as shown in Figure 1. A square lattice is not essential for this exposition – we could use hexagonal or other forms of regular tessellation of the plane – but a square grid provides a clearer and more familiar example. Although the main purpose of our theory involves generating growth, it is easier to begin with its application to an existing urban structure such as that portrayed in Figure 1a. Let us assume that the city has an overall linear dimension of L units and an areal dimension of L^2 as shown at scale $k = 0$ in Figure 1a, and that the changes in scale from $k = 0$ to $k = 1$ and so on, represent increasing resolution in observing and detecting the form of the cluster in question. Figure 1a represents this resolution with respect to the occupied areas of the hypothetical city form while Figure 1b represents the same cluster as points joined by a connected line which spans the occupied areas in question.

At each scale or level of resolution k, the basic linear unit of the grid (the side of each grid square in Figure 1a), ξ_k, is given by

$$\xi_k = \frac{L}{n_k}, \qquad (1)$$

where n_k is the number of subdivisions into which L is divided. From equation 1, n_k can be written as

$$n_k = \frac{L}{\xi_k}. \qquad (2)$$

As the scale of resolution in Figure 1a increases, it is clear that the total number of units of development (solid grid squares), N_k, increases at a faster rate than n_k but not as fast as the total number of grid squares n_k^2. Using equations 1 and 2, it is clear that N_k and n_k might be related as

Notwithstanding these elaborate models of the size and spacing of urban settlements, we have made very little progress on models of their size and shape, shape and area, and through this, their density which provides the crucial link to the urban economic theory of the city. Moreover, processes of urban growth and change are reflected in the shape of settlements, and in the way these processes are moulded by physical and planning constraints. In short, it is our contention that the size and spacing of settlements is influenced by the physical geometry of urban form, and it is therefore of fundamental concern that we explore this relationship between form and process. In this quest, our research will be informed by two broad lines of inquiry: first, the renaissance in morphology or the geometry of form which is based on the emergence of a geometry of the irregular – *fractal geometry*, as its foremost proponent Benoit Mandelbrot (1982) has called it; and secondly, *allometry*, or the study of relative size as it has been developed for biological and human systems (Gould, 1966).

Fractal geometry deals with forms which at first sight seem to defy any geometrical order but on further scrutiny, reveal the same degree of irregularity or 'disorder' on many scales. Such sameness in disorder, and its associated degree of irregularity is captured through the notion of fractional or *fractal* dimension which links this new geometry of the irregular to our traditional conceptions of Euclidean geometry. In the last decade, the theory of fractal geometry has been used to synthesize many hitherto unrelated morphological measurements in a wide variety of fields and although its formal application to geography is comparatively recent, the promise of its application has already been revealed in studies as diverse as central place theory (Arlinghaus, 1985), the measurement, classification and simulation of boundaries and surfaces such as coastlines, urban edges, and terrain, river systems and other dendritic forms (Batty and Longley, 1986; Longley and Batty, 1989), as well as climatic change and related forms of turbulence (Goodchild and Mark, 1987).

In this paper, we will draw on one of the most rapidly developing areas of fractal research, that concerned with the growth of far-from-equilibrium clusters which have an underlying dendritic or tree-like form. These forms are generated by constrained diffusion from fixed growth poles or seed sites, being termed diffusion-limited aggregation (or DLA) processes (Vicsek, 1989; Witten and Sander, 1981). So far these types of model have been applied to urban forms by Benguigui and Daoud (1991) and by two of the present authors (Batty *et al.*, 1989), but the line of research we will promote here also relates to more traditional areas of fractals-based research concerned with the measurement of boundaries and edges.

Our second direction to this research has been central to much previous work in the analysis of urban form. This is the concept of *allometry*, which is used 'to designate the differences in proportions correlated with changes in absolute magnitude of the total organism or of the specific parts under investigation' (Gould, 1966, p. 588). More commonly, the term is used to describe scaling relations between two 'size' measures of an organism or system under study (Mark and Peucker, 1978). Applied to urban form, allometric studies have been concerned with the relationship between urban area and population size, linked through the concept of density, this relationship being monitored both over time (Dutton, 1973) and across space (Nordbeck, 1971). However, as we shall see, the a priori conceptions of density adopted in these studies are open to question and, to anticipate our conclusions, a rethinking of the conventional wisdom about urban density gradients is an important precursor to our further understanding of urban growth mechanisms. In this, we will forge a clear and unequivocal link between the evolution of the shape of urban areas through fractal geometry and their size through urban allometry.

A related theme in this paper concerns the measurement of size and shape, area and density. In particular, we will make a central distinction between the concepts of the urban area and the urban field, focusing upon the need to relate the particular measurement in question to the purpose of the analysis. It is already very clear to us in reviewing the literature on the measurement of urban population and urban density that conventional practice is obscure; definitions of both area and population differ between studies, some based on restricted built-up areas, some depending on original census tracts or groups of these into sectors, rings and so on (Zielinski, 1979). Our confidence, therefore, in previous empirical estimates of allometric and other scaling relationships in urban studies is low. A related theme but one which is of different import involves the representation of spatial shape and area in computer models and information systems which are concerned with spatial manipulation, analysis and display. As our research concerns ways in which geometry might be simulated, we would ultimately hope that our models of size and

The size, shape and dimension of urban settlements

PAUL A. LONGLEY
Lecturer in Planning, Department of City and Regional Planning, University of Wales, Cardiff, P.O. Box 906, Colum Drive, Cardiff CF1 3YN

MICHAEL BATTY
Professor of Geography, National Center for Geographic Information and Analysis, State University of New York at Buffalo, 105 Wilkeson Quad, Buffalo, New York 14261, USA

JOHN SHEPHERD
Reader in Geography, Department of Geography, Birkbeck College, University of London, 7–15 Gresse Street, London W1P 1PA

Revised MS received 1 November, 1990

ABSTRACT
In this paper, we propose a scale theory of urban form and growth which enables us to consistently explain and estimate relationships between urban population size, area, field and boundary length for a system of settlements. Our approach is based on a synthesis of allometry and fractal growth theory, and the associated relationships are uniquely specified by dimensional parameters whose values vary from 1 to 2, from the line to the plane. The theory assumes that the form of settlements is tentacular and that the population density of these forms is constant with respect to their size. After the theory has been presented, four relationships – two allometric, relating populations and boundaries (or envelopes) to urban areas, and two fractal, relating the same variables to the urban field size – are estimated for some 70 settlements which compose the urban system in the English County of Norfolk. The hypothesized values of the dimensions characterizing these four relationships are confirmed by regression estimates and these results are given further strength when the same relations are re-estimated for various subsets of settlements in the Norfolk urban system. We conclude that the geometric form of the settlements system is consistent with the model we have adopted, that population density is constant at all scales, and that urban boundaries have a degree of irregularity measured by a fractal dimension similar to that conventionally assumed for coastlines. Finally, we suggest directions for further research.

KEY WORDS: Scale, Fractal, Dimension, Allometry, Urban morphology

---

INTRODUCTION

Our understanding of the growth and evolution of urban settlement systems largely rests upon the edifice of central place theory and its elaboration and empirical testing through spatial statistics. Stochastic processes which govern the evolution of the size distribution of settlements in the urban hierarchy have been widely studied; convincing models exist which explain the spacing and patterning of urban clusters and there have been useful attempts at linking these to theories of spatial innovation and diffusion. Yet although the arguments which comprise this classical location theory are reasonably consistent and complete (Haggett *et al.*, 1977), it has proved difficult to link these ideas to the more detailed geometry of settlement patterns and to questions of population and other forms of urban density. The geometry of urban settlements, insofar as it exists, is based on the idealized hexagonal tessellations of the plane associated with traditional central place theory and the statistics of point patterns (Cliff and Ord, 1981).

CHAPTER 25

THE SIZE, SHAPE AND DIMENSION OF URBAN SETTLEMENTS

(P Longley, M Batty and J Shepherd)

Transactions, Institute of British Geographers, 1991, NS 16(1), 75-94