Chapter 16

Multiple Dimensions Of Settlement Systems: Coping With Complexity

Mike Coombes[1]

This chapter considers some opportunities for improving definitions of urban and rural areas. The complex processes which are reshaping settlement patterns, particularly in more developed countries (MDCs), and creating urban systems make the simple framework of the traditional binary urban-rural divide seem inadequate. A mix of academic literature – but predominantly British and recent – is drawn upon here, together with some major crossnational statistical and policy-related documents (e.g. Decand, 2000).

The first task is to reconsider what the terms urban and rural now mean in the complex reality of MDCs' settlement patterns. Here the term 'settlement' will refer to a single separable builtup area, whether it be a village or a conurbation, defined using United Nations (UN) principles for identifying urban areas (OPCS, 1984). Three main *dimensions* of settlements are identified; these dimensions provide a framework for evaluating different ways to demarcate rural from urban areas, leading on to the development of some more appropriate measures. Attention centers first on options for bricks-and-mortar definitions, then moves on to deal with measures which take account of the wider context within which any rural or urban area is set. Finally, the paper considers some implementation issues posed by the newer methods of definition, and especially the question of how several indicators can be combined in a multidimensional approach to representing settlement patterns.

What Are The Fundamental Issues?

There is now a wide consensus that in MDCs there is an increasingly 'fuzzy' distinction between urban and rural areas. This trend led to a commonly accepted model in which the two categories were seen as the two parts of a continuum, with the difficulty then lying in choosing the best point at which to draw a line between one set of areas and the other. The current position is more complex, so that place A can be more 'urban' than place B in some respects, but more 'rural' than it in

[1] The author is very grateful to his colleague Simon Raybould for implementing the analysis described in Annex 3 and providing the three maps.

others. The distinction between urban and rural areas is arguably now so ‘fuzzy’ that the two categories need to be understood as each grouping a set of areas with a family resemblance among themselves *but* with no single ‘litmus test’ distinction between them and the areas in the other group. From this view, the continuum model is too restrictive because it assumes that one single dimension can provide an adequate basis for distinguishing the two categories satisfactorily.

A key reason for the increasing problems of definition is, of course, that modern ways of living and working have systematically challenged longstanding contrasts between the rural and urban domains. Much of the countryside now shares many of the characteristics previously seen to typify urban areas. A major driving force has been the growing preference for living away from larger cities: this preference might be dated back to royal relocations such as to Versailles, but of course it is now an option open to many. The fundamental process here is of increasing mobility, and a crucial consequence is the growth in flows and linkages between cities, towns and countryside. At a high level of abstraction, cities have been ‘deconstructed’ and portrayed as purely nodes in a ‘space of flows’ (Castells, 1989). This conceptual development has been followed in the policy context by the European Spatial Development Perspective’s suggestion that to counterpose the urban and the rural is now an ‘outdated dualism’ (Nordregio, 1999). Less contentiously, this critique reinforces the emphasis on flows and linkages between cities and countryside.

Yet there is plenty of evidence that key urban-rural differences such as settlement size do indeed influence people’s life chances (e.g. Denham and White, 1998). More generally, Chase-Dunn and Jorgenson (2002) argue that the ‘settlement size distribution – the relative population sizes of the settlements within a region – is an important and easily ascertained aspect of all sedentary social systems’. The distinctions between urban and rural areas are less clear-cut than once they were, but they are still real (Carter, 1990). For example, there are rural policy concerns, particularly about accessibility and the natural environment, which do not take the same form or intensity in urban areas.

Reviewing an emerging consensus nearly 30 years ago, Rees (1970, p.276) summarized the importance of settlement patterns with the following, slightly cautious, claim: ‘The densities at which people live have profound effects on their lives, or so the classical urban ecologist or experimental psychologist would have us believe’. In the USA at the time especially, city size and central area density were seen to be positively correlated, whilst the density of urban neighborhoods was negatively correlated with distance from the city center. These two regularities were reinforced by cities’ average densities being positively correlated with their overall size. In addition, people’s behavior was frequently argued to be shaped by the population density levels where they live.

Many changes to the patterns Rees summarized have led to today’s more complex picture. For instance, many larger cities have lost population, particularly in their inner areas. Urban decline has brought social problems, so that lowered population densities are now often associated with those social and economic problems previously associated with high densities. Also, substantial new towns and other developments often include large populations living at low densities.

The continuing restructuring of advanced economies has also led to an ever-increasing emphasis on flows, both in the form of physical mobility and as electronically-mediated interaction such as teleworking. Increasing flows are often a corollary of increased specialization, and within regions specialization tends to reinforce polynuclear structures where each town and rural area has a mutually interdependent role (Batty, 2001). At the same time, the increasing flows of people across a modern region also results from lifestyle trends, and dual-career households in particular. It is especially important to note that both these micro and macro level trends have transformed patterns of flow *spatially* as well as greatly increasing their number. The traditional model, with a simple pattern of flows centered on a single city, is often now replaced by a reality featuring linear or 'edge' cities and networks of towns, all of which are linked by multidirectional flows including 'reverse' commuting and chained school-shop-work trips. As a result, the crucial step of recognizing the importance of flows in MDCs has to be followed through with a far wider perspective than a simple 1950s-based analysis of commuting into cities. Unless flows in all directions are considered — along with flows other than commuting if possible — then the analysis still presumes that all areas conform to a mid-century model of a monocentric urban system.

This emphasis on flows is part of situating any rural or urban area in its context. Taking the context into account is, as argued elsewhere (Coombes and Raybould, 2001), crucial to recognizing that there are several dimensions to the differences between modern urban and rural areas. Looking at recent trends in MDCs suggested three main *dimensions* which were not substitutable one for another:

- the intensity or concentration of settlements;
- the size of settlements; and
- accessibility to services and other facilities.

The last of these most clearly relates to the earlier discussion of flows, although urban intensity can also be seen to be increasingly characterized by a multitude and diversity of movement. Most pressingly for this paper, these different dimensions will produce different rankings from rural to urban, because place X can be more 'rural' than Y in terms of one dimensions while it is more 'urban' than it in one or both of the other two respects.

The Challenge Of Definitions

As stated earlier, the categories rural and urban are seen here to represent 'family resemblance' across a variety of characteristics. All rural areas will share *most* of the characteristics which make up a stereotypical contrast with urban areas but, taking any one characteristic separately, a few predominantly rural areas will be more like a typical urban area whilst a minority of urban areas will be more like a typical rural area. This 'fuzzy' distinction implies that methods to delimit urban and rural areas on the ground:

- will be unlikely to generate a clear 'definitive answer' with a single criterion;
- may better meet different purposes by drawing on different approaches, stressing different characteristics; and
- can at best provide a set of definitions suitable for a majority of the very broad range of users of standard urban-rural definitions.

Before discussing alternative ways of identifying rural and urban areas, it is important to establish the criteria against which the alternatives should be assessed. The discussion so far has aimed to sketch out what the categories urban and rural mean in MDCs, and clearly the principal criterion here is that the definition should represent the most important contrasts between the two categories. In practice, this leads back to the three key dimensions of intensity of settlement, settlement size and accessibility.

A first practical issue concerns the 'building block' units of analysis for the definitions. Two of the three key dimensions focus on settlements, but there is an issue over whether these can be taken as given. While it is a valuable starting-point if all settlements have been identified according to the UN builtup area principles, questions do remain to be answered before these definitions can be considered ideal. The first question is whether the detail of UN definitions can be reviewed to keep separate substantial towns which are only tangentially linked by development, thereby avoiding the problem faced by Pumain *et al*. (1991) whose definitions could link together a 'necklace' of towns across distances nearing 100 km. Another is whether settlements should be subdivided so as to help distinguish areas with different levels of access to certain facilities. A parallel question is whether there is an optimal way to partition the area between towns to create the 'building blocks' needed for accessibility analyses. In general, smaller units of analysis lead to superior definitions, but the more important imperative here is avoiding units which group (parts of) larger settlements together with substantial areas of countryside.

As for the definitions which are to be created, it is important that they produce boundaries which separate one area from the next. This may seem obvious, but it was *not* the criterion for the methods in Coombes and Raybould (2001) where the outputs were needed at the level of larger administrative areas (i.e. to measure the *degree* to which such a broad and inevitably mixed area is urban rather than rural). Setting aside several specific policy concerns of that research, two key guidelines identified there are also still relevant in the present context:

- robustness – the data used as input must be consistently reliable across all areas, and
- plausibility – the output values should be largely in accord with prior expectations.

Asking 'where is the most urbanized part of Britain?' provides a suitable rehearsal of these issues. The most plausible candidate is the City of London (the 'square mile' which centers on St. Paul's Cathedral), with its intensive land use and its centrality within Britain's major metropolis. Yet both the two most commonly used measures of urban character – population size and population density – suggest that the City of London is far from the most urban area of the country. The simple reason for the population size measure's result is that the City includes only one square mile (approximately) out of the vastly wider London conurbation and so its total population is inevitably quite low. This can be resolved by shifting analysis to the wider settlement level, although of course then the City as such is no longer identifiable. The density measure's counter-intuitive result highlights three other important issues to be faced when trying to devise better urban-rural measures:

- The units of analysis are critical. Analyzing the City of London in isolation from its neighbors radically underestimates the number of people for whom the City is a core part of their daily lives. An immediate response could be to include incommuters and measure the 'daytime population' (cf. Goodchild and Janelle, 1984), but a more fundamental response is to recognize that adopting 'off-the-shelf' administrative areas was never likely to provide a satisfactory geostatistical unit of analysis (Coombes, 2000).
- Land area is problematic. Density measures treat the measure of land area (the denominator) as of equal importance to the measure of population (the numerator). Yet the land area includes not only residential areas but also industrial and undevelopable areas which are irrelevant to the population. This effect varies markedly between one area and another, due to idiosyncratic boundaries, so land area measures fail the robustness criterion set out above.
- Context cannot sensibly be ignored. A population density measure for the area purely within a boundary such as the City of London's produces a result which would implausibly remain the same if the City was relocated to a remote offshore location! Less obviously, the same principle could be said to apply the whole London conurbation, even though its significance too substantially depends on its place within a wider region.

Population density is all too often used as a 'proxy of first resort' for a more robust measure (cf. Dorling and Atkins, 1995). Its familiarity has led it to being used in an extraordinarily wide range of ways, but it is probably most plausible as an indicator of the intensity of settlement. Its simplicity has encouraged the sort of overinterpretation in which low density areas are often described as 'leafy' or 'peripheral' and high density areas as 'overcrowded' even though the measure does not indicate any of these features either directly or consistently.

This critique of the population density measure does at least draw attention to users' evident preference for easily understood measures and definitions. In particular, highly technical 'black box' methods would have to produce hugely superior results before their benefits would be seen by many to outweigh the disadvantage of their obscurity. This is a problem for the Urbanization Index proposed in Coombes and Raybould (2001) as an indicator of settlement intensity in preference to population density. On the same basis, the fact that the urban-rural dualism has been utilized for so long by so many surely shows that users prize simpler categorizations. Thus any new approach to definition needs to err more towards few – inevitably broader – categories, rather than producing a larger number of categories so as to provide finer distinctions between types of area.

What Can Be Built Upon?

The three key dimensions distinguishing rural from urban areas mentioned above – intensity of settlement, settlement size, and accessibility – pose a problem when simplicity of definitions is a priority. Either one or two of the dimensions must be overlooked or there is increased complexity. Whilst each of the dimensions is fundamentally distinct, intensity and size are both bricks-and-mortar aspects of urbanization, whereas accessibility prioritizes flows and linkages. The need for simplicity here prompts an assessment of the relative merits of settlement size and intensity measures, plus any alternatives, to provide a basic bricks-and-mortar distinction of urban from rural settlements.

Table 16.1 summarizes the most frequently used urban-rural characteristics, within a framework derived from a 30-year-old international overview of urban area definitions (UN, 1969). Table 16.1 also shows that this framework comfortably embraces criteria which were identified in a more recent crossnational comparison of *rural* area definitions (OECD, 1994). The two studies also found that the different criteria were used with similar frequency, which tend to support the earlier view that conventionally they are two sides of the same coin.

Table 16.1. Conventional criteria for defining urban and rural areas

UN (1969) classification of urban criteria *listed in descending order of frequency of use*	OECD (1994) classification of rural criteria *showing the number of countries using them (out of 24)*
population size (of administrative area or settlement)	settlement population size *14* administrative area population outwith conurbations *6* administrative area population size *8*
population (or housing) density	population density *7*
economic activity	agricultural share of workforce *3* in/out commuting ratio *4*
other urban characteristic(s)	centrality or service levels *2*
administrative status	administrative status *5*

Leaving the more heavily-used criteria at the top of the table to the last, this review can quickly dispose of the *administrative status* of areas as not only an unscientific criterion but also an increasingly irrelevant one as many countries reorganize their administrative geographies to group rural and urban areas together. Next comes something of a ‘catch all’ category, with the right-hand side of the table giving the example of *service levels* which, of course, is here more related to accessibility measures than to the bricks-and-mortar approach. The *economic activity* criteria also include an access-related factor in the form of the commuting measure. Agriculture has for some time been far from fundamental to rural areas in many modern economies so it is an unsuitable indicator of urban-rural profile, just as manufacturing used to be concentrated in larger urban areas but now is more often a feature of small towns.

The paper's earlier critique of the *population density* measure guarantees that there will be no deviation here from a recent OECD working party's explicit rejection of density as a basis for defining rural areas. Looking at the settlement intensity dimension more broadly, the question remains whether there is in any case a level of population concentration which is either necessary or sufficient to distinguish rural from urban areas. Whereas a highly compact form is highly evocative of Victorian urban areas, its absence does not make a spaciously planned new town like the UK's Milton Keynes into a rural area, just as a small clustered village's compact form is not enough to make it urban. In other words, intensity *is* a distinct dimension along which high or low values tend to be characteristic of urban and rural areas respectively, but in modern developed countries it is not a robust basis for defining the urban-rural boundary.

This leaves *population size* – which for present purposes can be taken as settlement size as opposed to administrative area size – as the remaining candidate for the bricks-and-mortar urban-rural discriminator. Certainly in everyday terms, towns and cities are urban, whereas villages and countryside are rural. A settlement size criterion, like any solo discriminator, can only be a 'blunt instrument' *even if* the anomalies it produces are relatively few. The specific size threshold above which settlements are deemed to be urban is inevitably a difficult choice. There are only a few countries like Scotland whose urban hierarchy is sparse enough for certain thresholds to readily distinguish towns which are widely accepted as comparable. By contrast, in England a threshold such as 10,000 cuts through the settlement size ranking at a point where there are many very different types of town. A very recent statement by the British government (DTLR, 2002) illustrated this by being unable to choose between 10,000 and 20,000 as the threshold for urban policies.

The difficulty of achieving broad agreement on a settlement size threshold means that the aim becomes choosing the 'least worst' value where the urban-rural boundary should be positioned. The next section of the paper asks whether there are new alternative approaches – emerging since the work underlying the review in Table 16.1 – which might side-step the need to accept such a difficult compromise solution.

How Can Definitions Be Improved?

Just as the challenge of urban-rural definitions has become more complex, so the potential for more complex forms of analysis has greatly increased (e.g. Burrough, 1996). It is no longer reasonable to continue following the earlier approach of relying almost exclusively on a single 'handy' indicator like population density, as if no other indicators were available. Where much more information is available, the challenge becomes choosing from options which can range from satellite imagery to lifestyle datasets. Moreover, data can be accessed at increasingly fine scales, often due to postcoding, whilst Geographic Information System (GIS) techniques make all this information much more manipulable. Nevertheless, new possibilities do not *necessarily* imply that entirely new approaches are needed: there are advantages of continuity in retaining or building upon existing approaches. Given that no single 'definitive' solution will meet all needs in any case, starting entirely afresh would need careful justification.

Before getting too optimistic about new possibilities arising, it is important to keep in mind that the growth in the availability of small area data has not occurred everywhere. This means that the task of finding methods that are both appropriate and applicable crossnationally may actually have been made more difficult rather than less. That said, there now *is* satellite imagery available for most countries, so this can provide an important additional data source for urban-rural definitions. Keeping the focus here on establishing basic urban-rural boundaries, rooted in the bricks-and-mortar tradition, the question then becomes how to synthesize information of different types so as to output improved boundary definitions.

Although there can be no doubt that GIS has greatly eased the process of bringing together diverse datasets, it is equally well known that there has been far less progress in providing new ways of analyzing the collated datasets. In particular, there has been no breakthrough which provides a simple synthesis of multidimensional datasets. The more complex methods which present themselves as alternatives here include:

- a *cluster analysis* of the multivariate database,
- a *synthetic index* approach, and
- certain *rules-based* definitions which work through a series of steps.

There are relatively few examples of cluster analyses used to distinguish urban from rural areas. The most familiar versions are the private sector's geodemographic classifications, but their use of terms like 'suburban' are often unscientific because they have no consistent basis in the data analysis. A particularly good example of a synthetic index is a new British 'town centeredness' index (Hall and Thurstain-Goodwin, 2000), but its sequence of complex geostatistical modeling is not readily explainable to users. A particularly relevant recent example of a rules-based approach is the new Bamford *et al*. (1999) classification of rural areas according to their travel time from larger settlements, although this may not be very transferable to countries where a set of key urban size thresholds will not find ready agreement.

A disadvantage of both the index and the cluster analysis approaches is that they are less able to implement 'everyday' ideas of how to combine criteria. For example, it might be thought that all settlements above a certain size should be recognized as urban areas, together with smaller ones which have other characteristics (e.g. local facilities) similar to those of larger settlements.

Combining criteria in this way can be relatively straightforward within a set of rules – as illustrated by the 'conurbation' definitions of Taylor *et al*. (2000) – but a cluster analysis or index would need the size criterion to be applied as a separate initial step. In effect, they then become part of a rules-based approach in order to create a method in which users can understand the interplay of the various factors.

Another way in which rules-based approaches tend to be different is that they use relatively few criteria, and ones that are recognizable by users. For example, a new Scottish rural area classification (Hope *et al*., 2001), while needing high-level GIS techniques and substantial datasets, uses key criteria of settlement size and drive times that are readily understood. By contrast, cluster analyses and indexing methods frequently synthesize numerous indicators, some of which may be heavily preprocessed or in some other way less recognizable to lay users, and then depend on a combinatory technique that is widely seen as 'black box' to a greater or lesser extent. Whereas users can directly engage in a debate about appropriate drive times, it is hard for them to choose one particular form of principal component analysis, for example, even though these decisions may shape the final outcome more profoundly than the choice of input indicators (Coombes and Wong, 1994). As a result, the previous emphasis on simplicity or transparency of methods leads to a preference for a rules-based approach if several input variables are needed to produce appropriate definitions.

For the present, it seems likely that settlement size is the only plausible candidate to be a sole discriminator between the two types of area, although this will be a 'blunt instrument' and there is unlikely to be ready agreement when it comes to choosing the population size threshold above which settlements will be deemed to be urban. As yet, there is even less consensus likely on the way in which other data, such as satellite imagery, can be brought to bear on the problem so as to reduce the number of anomalous results. This lack of a clear prospect for improvement means that there is insufficient reason to accept the disadvantage of using a less readily understandable procedure. The conclusion here in favor of a rules-based procedure in preference to a cluster analysis or synthetic index approach will become relevant again as attention turns from the bricks-and-mortar aspect of rural-urban distinctions towards issues related to accessibility.

Recognizing The Context

Two types of conclusion can be drawn to summarize the discussion up to this point. The first conclusion is a straightforward reaffirmation of settlement size as the primary – and possibly sole – indicator of urban areas in the bricks-and-mortar sense. At the same time, accessibility-related analyses are needed to consider the broader context of each settlement. No substantial role is seen for a measure of settlement intensity in rural-urban definitions (except that a minimum level of settlement intensity helps to identify the outer boundaries of individual settlements, whose population size can then be measured).

The second type of conclusion is the set of guidelines which have emerged progressively through the discussion to establish types of measurement which will be more 'fit for purpose' in urban-rural definitions. It is timely to restate those which are relevant when seeking a suitable way of mapping the context of each settlement through analyzing the flows which portray the interaction between settlements and their context:

- flows are multidirectional and not necessarily focused on a central city,
- commuting is just one of several types of flow which could be relevant here,
- the boundaries should not cut through settlements, and
- more understandable methods of analysis are much preferred by users of the definitions.

Identifying clusters of flows and interactions around settlements calls for an approach which is at least partly rooted in the models of time geography (Pred, 1984). These models portray local boundaries, if produced by analyzing flows, as the local population's daily pattern of movements. Producing genuinely multi-dimensional definitions calls for analyses which reflect numerous types of interaction, such as migration or access to services, but unfortunately interaction datasets tend to be scarce, and indeed not every advanced country even has information on commuting patterns.

The most well-established approach to defining the surrounding areas with which cities are linked is typified by Metropolitan Statistical Areas (MSAs) in the United States (Spotila, 1999). Viewed from the criteria set out here, this approach has the crucial flaws of presuming an urban-centered structure, and ignoring all flows other than those to the urban core from elsewhere. There are rather fewer examples of boundary definitions which do *not* rigidly presume a certain pattern of flows; most of these are to be found among the range of approaches to labor market area (LMA) definition. A recent review (OECD, 2001) showed that many countries now have LMA definitions, most of which – though not all – rely exclusively on commuting patterns as flow data. The remainder of the discussion here will briefly review ways in which the wider context of urban areas can be identified in three different types of circumstances:

1. where commuting patterns are the only available flow data;
2. where additional flow data can be drawn upon;
3. where no flow data at all is available.

In the first type of circumstance, the need is for a set of consistently-defined LMAs, which are largely *self-contained* in commuting terms with few journeys crossing their boundaries (Goodman, 1970). It is notable that most countries have devised their own LMA definition procedures, although the British software was adopted by ISTAT for their definition of Italian local labor market areas (Sforzi *et al*., 1997) and has also been used in Spain (Casado-Diáz, 1996). A comparative research program applied alternative methods to several countries' commuting datasets, concluding that the British method was the 'best practice' model on which was based Eurostat guidelines for LMA definitions (Eurostat, 1992). Annex 1 (see the end of this chapter) summarizes this procedure, which has been termed the European Regionalization Algorithm (ERA). It can be seen there that flows in all directions – not just those into cities from elsewhere – are equally important to producing optimal LMA boundaries. Figure 16.1 presents LMA boundaries produced by the British software in the West of England: the center of Bristol has quite a large catchment area but the smaller city of Bath nearby emerges from the analysis with a sufficiently self-contained LMA centered on itself.

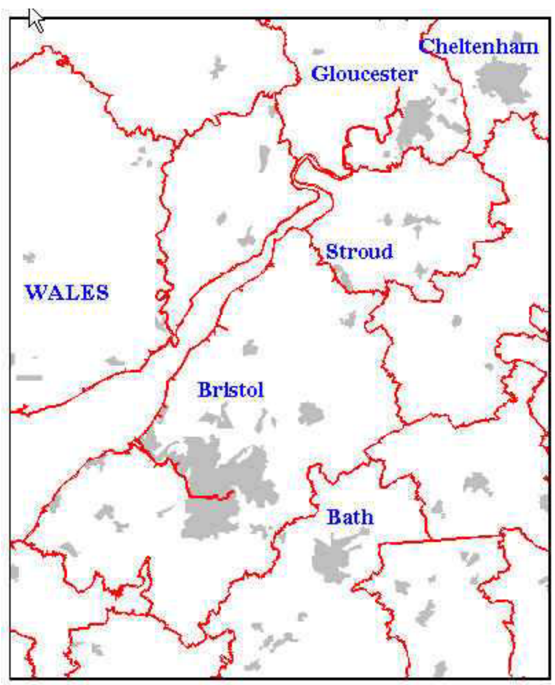

Figure 16.1 Boundaries defined by ONS and Coombes (1998)

Turning to the second type of circumstance, it has already been stressed that commuting patterns alone can be no more than a very partial proxy for a more rounded assessment of flows between rural and urban areas. For example, Claval (1987) called for an examination of numerous facets to the linkages which make up a region in practice. Until recently there has been no method for combining the evidence which data on different types of flow can provide towards such a multidimensional portrayal of regions. Coombes (2000) has now demonstrated an appropriate method, the starting point for which is viewing each strand of evidence as part of a ‘fuzzy’ picture, with the task of the analysis then being to distil the recurring patterns which make up the underlying structure.

The methodological innovation hinges on splitting the procedure into two phases. Phase 1 involves compiling numerous analyses from numerous datasets, and then in phase 2 the results from these analyses are collated within a synthezing analysis. Phase 2 centers on creating *synthetic data* for which the input is the range of evidence provided by phase 1. Described in more detail in Annex 2, the approach involves layering sets of phase 1 boundaries on top of each other and counting the number of layers in which there is no boundary between each pair of areas. This provides an assessment of the ‘strength of evidence’ that two areas should be grouped together, on the grounds that they are linked together by most of the input classifications. Figure 16.2 shows – for the West of England again – the areas which are linked by high synthetic data values. This synthetic dataset is then analogous to a flow matrix, because it represents the level of connectedness of pairs of areas, and so can be analyzed with ERA to produce boundaries grouping rural and urban areas which have a high level of connectedness.

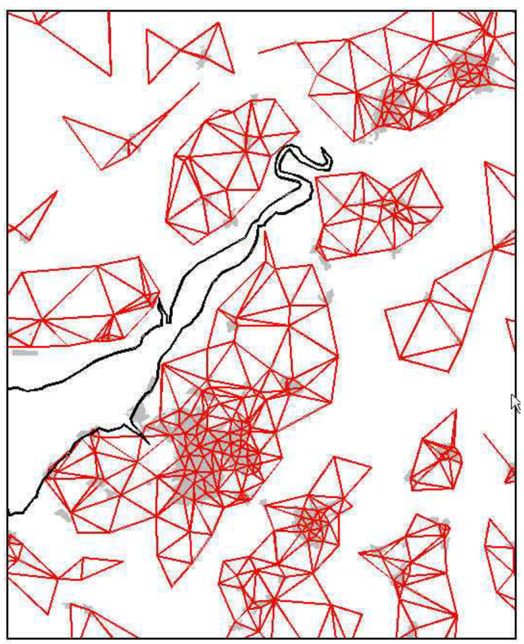

Figure 16.2 Information from the analysis of Coombes (2000), with lines connecting pairs of areas with synthetic data values which are at least half of the maximum possible value in that dataset

In the third type of circumstance – where there are *no* flow datasets to draw upon – how could patterns of linkages be represented in the form of boundaries?

There are, in fact, a number of analysis techniques which can estimate patterns of linkages given some relevant information. A fairly simple, yet plausible, example can be outlined by estimating the pattern of commuting where no flow data is available *but* the distribution of both jobs and employed residents is known. For some time now, spatial interaction models have been able to estimate the pattern of flows which could link where the workforce live with where the jobs are located. Such a matrix of estimated flows can then be analyzed by a method such as ERA to create hypothetical LMAs. A critical problem of spatial interaction models is that they are undeniably 'black box' in their workings. Glover and Openshaw (1995) had to devise a much simplified 'front end' so as to encourage use of a web-based version. Although this ameliorates the problem slightly, a more fundamental issue remains in that few users can understand how the results were produced. A much simpler option has been devised here, providing a very straightforward method which should be readily understood and easily transferable (Annex 3). Requiring only the distributions of jobs and residents as input, plus a centroid for each 'building block' area, the method produces very plausible estimates of commuting patterns. Figure 16.3 covers the West of England again, and shows that the method estimates the main cities' catchment areas well. Indeed, it even hypothesizes a flow across the Severn estuary at exactly the point where the Severn bridge exists, despite the analysis having no input data on transport infrastructure.

The form of analysis outlined in Annex 3 – and more complex spatial interaction modes – do still need additional data to that which is available from nearly every national census. For example, the estimated commuting patterns here required the location of jobs to be known. A different form of modeling is needed if the data available at a detailed scale relates to the resident population only. In such situations there is even greater reliance on the key principle underlying spatial interaction modeling, *viz*. that the probability of interaction declines with increased distance.

In the GIS era, this distance-deterrence principle may be made more sensitive to local circumstances than by simply assuming that a straight-line distance between two places represents the difficulty of traveling between them (Higgs and White, 2000). There are often now measures of actual route lengths, allowing the estimation of point-to-point travel times based on the speed of travel along different route segments. All this sophistication, of course, depends on the availability of these extra datasets, and here the interest in such models is restricted to circumstances where data availability is extremely limited.

Mention has already been made of analyses such as that of Bamford *et al.* (1999) that measure the relative ease with which more rural areas can access urban centers. These analyses could readily identify the areas with a level of accessibility above a predefined level. Given the necessary route data, this approach can draw boundaries round areas within X minutes' drive-time of major centers. Without this information, the approach is limited to simpler analyses which would then appear on maps as purely hypothetical (e.g. circular catchment areas). In both forms of application, the crucial first step in the process is choosing the centers whose catchment areas are to be modeled. This step draws attention to the model's presumption of monocentric regions, in which the only interaction considered radiate to/from the center. A final point to remember is that, as with most models, key thresholds need to be chosen and these will often not command consensus due to their great influence on the resulting boundaries. This problem is likely to be substantial for these models because users may expect differentiated results – e.g. 'within X minutes of a city of size 1 *or* Y minutes of a city of size 2 *or* ...' – which would require a consensus on the choice of numerous thresholds of both city size and accessibility level.

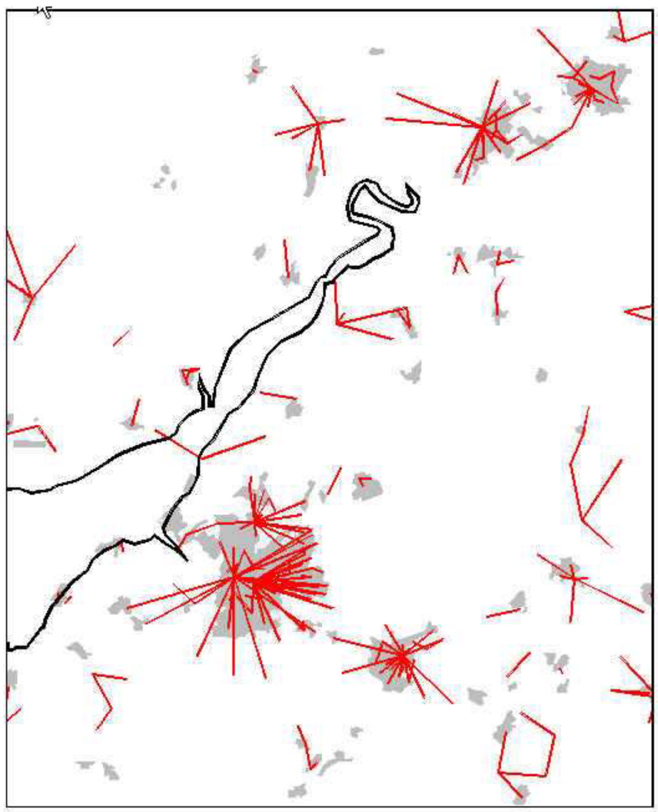

Figure 16.3 A line connects any pair of adjacent areas which step 3.2 ???? processed by canceling out surplus jobs in one area with surplus work-seeking residents in the other

Implementation Issues

Drawing clear conclusions here is scarcely possible because many different alternatives have been discussed, not least to cope with the great variation in data availability that has to be faced when considering new international standard definitions. Emphasis has been placed repeatedly on flows and linkages, but data coverage of these patterns is acutely limited. All the same, the single main conclusion here has been the need to combine a bricks-and-mortar definition of each urban settlement with a mapping of its context. One implementation issue which is provoked by that conclusion is how these two elements of rural-urban definitions should be combined. In tackling this issue, the earlier discussion about users' preference for simplicity becomes an important constraint. A simpler approach will have the additional advantage that it may reduce the number of thresholds to be agreed, thereby avoiding the type of problem noted above in which several different hinterland sizes had to be chosen to match several different size bands of city.

Table 16.2 compares ways of implementing the call here for urban area definitions to be set in context by representing their interactions with nearby areas. The rows in the table each reflect a different level of data availability. Thus the top row is the 'ideal' situation in which there is more than one dataset about patterns of relevant interaction, whereas the second row considers options if there is just one such dataset (e.g. commuting patterns, or just possibly migration flows). The third row provides a reminder that certain types of network – whether fixed networks like roads or service networks like bus services – may be useful as a 'proxy' for data on actual patterns of interaction. The last row assumes that no relevant data is available, apart from the map locations needed to estimate distances between places.

Table 16.2. Representing the linkage of urban and rural areas: analyzing interaction, with differing levels of data availability

	Linkage to:	
Linkage shown by	Foci which the relevant data show to attract flows	Foci which are predetermined in the analysis
Multiple forms of interaction	Localities in Britain (Coombes, 2000)	City Regions in Britain (Coombes, 2000)
Commuting (or possibly migration)	Local labor markets (e.g. Casado-Diáz, 1996)	Metropolitan areas (e.g. Spolita, 1999)
Potential for, or signs of, interaction (as shown by networks)	Bus service hinterlands (Green, 1950)	Road network accessibility (e.g. Bamford *et al*., 1999)
Probability of interaction (as shown by relative distances)	See Annex 3	Basic 'gravity' models (Glover and Openshaw, 1995)

Table 16.2 also has two columns. The left-hand one covers those analyses in which there is no preselection of key 'foci' for the pattern of flows (nb. this type of analysis may also handle the multidirectional flows which are increasingly common in polynuclear regions). The right-hand column is for analyses that

predetermine which places are the foci for the interaction patterns analyzed (e.g. the central cities which are identified first in most metropolitan area definitions, so that the only commuting flows then examined are those of noncore residents who work in these cores).

The principles outlined in this chapter indicate that the analyses in the left-hand column of Table 16.2 are preferable to those in the right-hand one and, data permitting, those higher in the table are to be preferred to those lower down. This interpretation is consistent with the approach of Frey and Speare (1995) who argued that methods of labor market analyses in Europe (e.g. Casado-Diáz, 1996) were preferable to metropolitan area definitions in the USA (not least because the latter only analyze the small subset of commuters who travel to work in the predefined urban cores). Although the table is mainly a summary of material already discussed earlier in this chapter, there has been no previous reference to Green (1950). The latter's use of bus service information to identify urban centers and their hinterlands may be highly relevant in some less developed countries where public transport still dominates mobility patterns, and where Census information may well not include information on the pattern of commuting. Another useful reminder provided by the table is that patterns of local migration can offer a valuable alternative, or supplement, to the more familiar commuting data for these types of analysis. Finally the table draws attention to the role of spatial interaction models in situations where very little data can be found.

In most cases, the wider area within which the urban area is set will be defined as a labor market area (LMA). The next step is to differentiate these LMAs in an appropriate way. Population size is the one universally-available measure, but the categorization could be in terms of either the whole LMA or its largest urban area. The development of polynuclear regions suggests that the whole LMA is the more appropriate metric. This leads to the outline recommendation that the eventual classification uses two population-size measures:

- settlement size (with a threshold providing the simplest rural-urban categorization); and
- size of the labor market area (or nearest-equivalent available boundary) in which the settlement is situated.

Implementing such a classification doubtless faces numerous hurdles, ranging from huge issues such as uneven data availability to the more technocratic (e.g. coping with settlements straddling more than one LMA). These issues lie in the future, beyond the major uncertainty of establishing whether this approach is reasonably acceptable to rural and urban experts, data supplies and users.

Concluding Comment

Underlying this chapter is the thesis that the new GIS-facilitated flexibility *increases* the importance of the decision over which areas are best used for data analysis, because the rapid reduction in barriers to data availability increases the need to identify the most 'fit for purpose' urban and rural definitions. The identification of these areas is far from deterministic: a number of alternatives can be shown to meet different requirements. There are few easy answers when trying to respond to new opportunities and user needs whilst adhering to the ethos summarized by the Fundamental Principles of Official Statistics (agreed in 1991/2 by the United Nations Economic Commission for Europe). In practice, these Principles involve a tension between trying to satisfy all the citizens' entitlement to public information, and aiming to retain the public's trust in official statistics by discouraging inappropriate usage of statistics. An additional tension for this paper has been between the drive to innovate – stimulated by acute dissatisfaction with existing classifications and aided by new geostatistical opportunities – and the belief that users prefer simple and stable classifications. In this light, it may be less surprising that a chapter opening with a discussion of new multidimensional complexity ends recommending a two-dimensional classification which builds on the longstanding measure of settlement size.

Annex 1

The basic features of the ERA method derive from an earlier algorithm (ONS and Coombes, 1998), but that multi-step method has been simplified to remove numerous steps which required the setting of applications-specific parameters (e.g. 'select areas which score higher then x on parameter y'). Reducing the procedure to its basic components makes it much more transferable, in that it can be used with different datasets and for different purposes without long periods of experimentation to determine optimal settings for each parameter. At its simplest, the ERA procedure can be described in two steps:

1. consider all input areas to be potential output regions, calculating their values on the statistical objective function set (e.g. level of self-containment), and ranking them accordingly
2. disassemble regions which fail to meet the statistical criteria (starting with the region furthest from meeting the objective function) and group the constituent areas individually with whichever other region they share most flows.

In technical terms, ERA's tendency to be self-optimizing is due to three factors:

1. The software can cope with very different sized 'building block' because in its early stages it groups them into localized clusters without trying to get 'everywhere right first time' and then it gradually iterates towards a more optimal solution.
2. The groupings are not constrained by contiguity, which allows the procedure to always choose the linkage that maximizes the grouping's integration; in fact very few areas are eventually grouped non-contiguously because the groupings are driven by patterns which reflect people's reluctance to travel longer distances than are essential.
3. The procedure is not rigidly hierarchical, in that two areas which are grouped together at an early stage may then later be disassembled and grouped into separate areas.

Annex 2

The essential basis for the synthetic data is the understanding that a set of boundaries is a classification in which each building block area is allocated to one and only one region. Thus each analysis undertaken in phase one produces a classification of all the N building block areas from which the final definition are to be composed. For each area, the initial information in each classification is the region number to which that area is assigned. The crucial step is then to transform this information into binary data in a matrix by taking each pair of areas and identifying whether they are ('1'), or are not ('0'), classified into the same region. In this way, each classification becomes expressed as a binary matrix of N*N cells (although the matrix is in fact symmetrical, so only half of it is needed). For example, if area B was in the same region as area C but in a different region to area D then the cell BC would take the value 1 while cell BD would be 0 (and cell CD would also take the value 0).

The crucial benefit from re-expressing each separate classification in this binary form is that these matrices can then be cumulated to produce the synthetic data needed. For example, if the results from the phase 1 analyses were collated in this way, the value in each cell of the final synthetic data matrix would vary from 0 (the value for any pair of areas which were not in the same region according to any of those analyses) up to 3 (the value for any pair of areas which all three analyses had put in the same region).

Annex 3

The method can be seen as a development from the analysis by Small and Song (1992), who attempted to identify 'wasteful' commuting by creating a hypothetical pattern of commuting which would minimize work trip travelling. The analysis here makes a similar assumption in favour of filling each job with the nearest people available to work. The method can be briefly outlined in a few steps, as below:

1. Assume jobs in each building block area are filled by the same area's residents so far as possible.
2. Identify each area as having surplus jobs *or* surplus work-seeking residents *or* a balance with all its jobs and residents allocated.
3. Proceed through all area-area pairs, in ascending order of distance apart:
    - 3.1 STOP if the next pair are too far apart for commuting to be plausible, otherwise proceed.
    - 3.2 *if* one of that pair of wards has surplus jobs and the other has surplus residents, fill as many of the surplus jobs as possible with these residents and update the areas' surplus/balance status; whether or not such a change has been made, return to step 3.1.

References (to move to composite bibliography)

Bamford, E., Dunne, L. and Hugo, G. (1999), 'Accessibility/Remoteness Index of Australia (ARIA)', *Department of Health and Aged Care Occasional Papers Series No. 6,* Commonwealth Department of Health and Aged Care, Canberra.

Batty, M. (2001), 'Polynucleated Urban Landscapes', *Urban Studies*, Vol. 38, pp. 635-55.

Burrough, P.A. (1996), 'Natural Objects with Indeterminate Boundaries' pp 3-28 in P.A. Burrough and A.U. Frank, (eds), *Geographic Objects with Indeterminate Boundaries*, London: Taylor and Francis.

Carter, H. (1990), *Urban and Rural Britain*, Harlow: Longman.

Casado-Diáz, J.M. (1996), 'Mercados Laborales Locales. Análisis Preliminar del Caso Valenciano', *Revista de Estudios Regionales*, Vol. 45, pp. 129-55.

Castells, M. (1989), *The Informational City*, Oxford: Blackwell.

Chase-Dunn, C. and Jorgenson, A. (2002), 'Settlement Systems: Past and Present', *GAWC Research Bulletin 73*, http://www.lboro.ac.uk/gawc/rb/rb/73.html.

Claval, P. (1987), 'The Region as a Geographical, Economic and Cultural Concept', *International Social Science Journal*, Vol. 39, pp. 159-72.

Coombes, M.G. (2000), 'Defining Locality Boundaries with Synthetic Data', *Environment and Planning A*, Vol. 32, pp. 1499-1518.

Coombes, M.G. and Raybould, S. (2001), 'Public Policy and Population Distribution: Developing Appropriate Indicators of Settlement Patterns', *Government and Policy*, Vol. 19, pp. 223-48.

Coombes, M.G. and Wong, Y.L.C. (1994), 'Methodological Steps in the Development of Multivariate Indexes for Urban and Regional Policy Analysis', *Environment and Planning A*, Vol. 26, pp. 1297-1316.

Decand, G. (2000), 'Regional Policy and Statistics – A Long Story', *Sigma*, 1/2000**,** 6-8.

Denham, C. and White, I. (1998), 'Differences in Urban and Rural Britain', *Population Trends*, Vol. 91, pp. 23-34.

Department for Transport, Local Government and the Regions (2002), *Urban Policy Evaluation Strategy Consultation Document.* http://www.urban.dtlr.gov.uk/consult/evalstrat/index.htm

Dorling, D. and Atkins, D. (1995), 'Population Density, Change and Concentration in Great Britain 1971, 1981 and 1991', *OPCS Studies on Medical and Population Subjects 58,* London: HMSO.

Eurostat (1992), *Study on Employment Zones*, Eurostat (E/LOC/20): Luxembourg.

Frey, W. and Speare, A. (1995), 'Metropolitan Areas as Functional Communities', in D.C. Dahmann and J.D. Fitzsimmons (eds) *Metropolitan and Non-metropolitan Areas: New Approaches to Geographical Definition*, US Bureau of the Census (Population Division Working Paper 12), Washington DC, pp. 139-90.

Glover, T. and Openshaw, S. (1995), The AGW Spatial Interaction Modelling Workstation, http://www.geog.leeds.ac.uk/people/s.openshaw/w3agw.html.

Goodchild, M.F. and Janelle, D. (1984), 'The City Around the Clock: Space-Time Patterns of Urban Ecological Structure', *Environment and Planning A*, Vol. 10, pp. 1273-85.

Goodman, J.F.B. (1970), 'The Definition and Analysis of Local Labour Markets: Some Empirical Problems', *British Journal of Industrial Relations*, Vol. 8, pp. 179-96.

Green F.H.W. (1950), 'Urban Hinterlands in England and Wales: An Analysis of Bus Services', *Geographical Journal,* Vol. 116.

Hall, S. and Thurstain-Goodwin M. (2000), 'Geographical Information – Policy Driving, Policy Driven – Providing Statistics for the UK's Town Centres', *Statistical Journal of the United Nations Economic Commission for Europe*, Vol. 17, pp. 125-32.

Higgs, G. and White, S. (2000), 'Alternatives to Census-Based Indicators of Social Deprivation in Rural Communities', *Progress in Planning*, Vol. 53, pp. 1-81.

Hope, S., Braunholtz, S., Playfair, A., Dudleston, A., Ingram, D., Martin, C. and Sawyer, B. (2001), '*Scotland's People: Results From the 1998 Scottish Household Survey*', *(Volume 1),* http://www.scotland.gov.uk/shs/docs/rep99-v1.pdf.

Nordregio (1999), *European Spatial Development Perspective: Towards Balanced and Sustainable Development of the Territory of the European Union*, Luxembourg: European Commission.

OECD (1994), *Creating Rural Indicators for Shaping Territorial Policy*, Paris: OECD.

OECD (2001), '*Functional Regions: a Summary of Definitions and Usage in OECD Countries*', Paris: OECD.

Office for National Statistics and Coombes, M.G. (1998), *1991-Based Travel-to-Work Areas*, London: Office for National Statistics, http://curdsweb1.ncl.ac.uk/files/4753maindoc.pdf

Office for Population Censuses and Surveys (1984), *Key Statistics for Urban Areas*, London: HMSO.

Pred, A. (1984), 'Place as Historically Contingent Process: Structuration and the Time-Geography of Becoming Places', *Annals of the Association of American Geographers*, Vol. 74, pp. 279-97.

Pumain, R., Saint-Julien, T., Cattan, N. and Rozenblat, C. (1991), 'The Statistical Concept of the Town in Europe', *Report 0673002 by the Network for Urban Research in the European Community,* Luxembourg: Eurostat.

Rees, P.H. (1970), 'The Axioms of Intra-Urban Structure and Growth', in B.J.L. Berry and F.E. Horton (eds), *Geographical Perspectives on Urban Systems*, Englewood Cliffs NJ: Prentice-Hall, pp 276-88.

Sforzi, F., Openshaw, S. and Wymer, C. (1997), 'Le Procedura di Identificazione dei Sistemi Locali del Lavoro', in F. Sforzi (ed.), *I Sistemi Locali del Lavoro 1991*, Rome: ISTAT, pp. 235-242.

Small, K.A. and Song, S. (1992), '"Wasteful" Commuting: A Resolution', *Journal of Political Economy*, Vol. 100, pp. 888-898.

Spotila, J.T. (1999), 'Recommendations from the Metropolitan Area Standards Review Committee to the Office of Management and Budget concerning changes to the standards for defining metropolitan areas', *Federal Register*, Vol. 64, pp. 56628-44 (October 20).

Taylor, D., Bozeat, N., Parkinson, M. and Belil, M. (2000), *The Urban Audit:Ttowards the Benchmarking of Quality of Life in 58 European Cities, (Volume 3: The Urban Audit Manual)*, Luxembourg: European Communities.

UN (1969), *Growth of the World's Urban and Rural Population 1920-2000*, New York: UN Department of Economic and Social Affairs, (Population Studies 44).