In a sense, urban planning is the science of spatial distribution with particular focus on place. Place is dependent on the material and spiritual spaces of a region. Although region is traditionally a central theme in spatial planning, planners have always been reticent about the ways in which a region can be built up. In ancient China, there have been studies on the relationship between the urban and the regional. Chronicles of the Han Dynasty in 79–105 refer to elements to consider when selecting city sites: to find out the balance of Yin and Yang, to taste the flavour of spring, to examine the suitability of land, and to build the city state (Ban, 105). Already, geographers have worked on improving the regional structure of hospitals (Godlund, 1961), schools (Yeates, 1963), social administration areas (Massam, 1975), and regional applications (Haggett, Cliff, and Frey, 1977), as well as the more widely known cases of local government boundaries. This chapter will try to explore a quantitative method for region building under urban sprawl conditions. It will include a clarification of the regional concept, regional combinatory, nodal regions, and graphs. Also, this chapter will present the case study of the Sunan (Southern Jiangsu Province), China, with the aim of discussing the methods for regional planning and in order to bring together the explicit techniques that have been used in regional delimitation and in grouping.

1. The regional concept

A region is usually constrained to a particular geographical scope, ranging from the entire planet to a county, township, village factor, school, or even a particular space or place. In other words, the region is everywhere.

1.1 Definition of region

Geographers define region as geographical section on the surface of the Earth, which could be overlapped and should be seamless and cover the entire area. Economists see region as an economic complex in which economic activities take place. Sociologists see region as a social unit which could be classified by ethnic, linguistic, and other characteristics, such as ethnic minority region, Chinese region, English-speaking region, etc. Politicians see region as an administrative
unit which is measureable and hierarchical. For planners, a region is concept that is used to study the variety of physical or non-physical phenomena within a particular area. It refers to a complex that contains a place, a core, a gradient, and an edge.

As Harvey (1969) observes, the region was “sometimes accorded the status of a ‘theoretical entity’ rather like an atom or neutron which could not be precisely observed but whose existence could be inferred from its effect. The areal differentiation of the earth’s surface could thus be ‘explained’ with reference to this theoretical object which governed human spatial organization.” Jacobs (1961) also said, “A region, someone has wryly observed, is an area safely larger than the last one to whose problems we found no solution.”

1.2 Types of regions

Four main types of region are commonly recognized by planners: planning regions, administrative regions, uniform regions, and nodal regions.

Planning regions

When a city is ready to draw up plans, it is usually the first delineation of the urban planning area. It may be defined as areas, contiguous or non-contiguous, delimited on an ad hoc basis for purposes of administration or organization. Planning regions may be overlapping or non-overlapping; they may exhaust the complete study area, or be confined to any part of that area. An objective in the design of planning regions is to maximize overlap between the needs of administration and ‘naturally occurring’ uniform or nodal regions.

Administrative region

Cities can divide their territory into small administrative units – that is, administrative regions of different sizes, based on their political, economic, ethnic, historical, and other differences, in order to establish the appropriate institutions for social management accordingly.

Uniform regions

These may be defined as contiguous areas within which, conditional upon the purpose for which the regional system is being defined, place-to-place variations may be regarded as trivial. More formally, a set of uniform regions may be defined as the arrangement of regional boundaries (cf. analysis of variance). Characteristically, uniform regions are non-overlapping and completely exhaust the space available. Uniform regions are sometimes termed homogeneous or formal regions.

Nodal regions

The nodal regions are defined in terms of bonds or links between pairs of places. Unlike uniform regions, nodal regions may be overlapping and interpenetrating. Nodal regions are sometimes termed functional regions.
1.3 Regional scale

That scale problems have long troubled planners is rather plainly shown in the attempts made to define regions in scale terms. Regions under different scales in planning could be defined using an accessible distance for pedestrians (Table 4.11.1).

1.4 Partitioning and grouping procedures

We can proceed to classify the regions in either of two ways: by logical division or else by grouping.

Interregional (across regions)

Logical division or ‘classification from above’ proceeds by dividing the universal set according to a particular property/attribute. In order to follow this method, we must have prior information on the property being used as an indicator, and thus this approach is sometimes called a deductive one (Figure 4.11.1).

Table 4.11.1 Different scales in planning

<table>
<thead>
<tr>
<th>Meters (Log Scale)</th>
<th>City-region</th>
<th>Planning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^0$</td>
<td>House</td>
<td>Urban design</td>
</tr>
<tr>
<td>$10^1$</td>
<td>Block</td>
<td></td>
</tr>
<tr>
<td>$10^2$</td>
<td>Neighbourhood</td>
<td></td>
</tr>
<tr>
<td>$10^3$</td>
<td>City</td>
<td>Urban planning</td>
</tr>
<tr>
<td>$10^4$</td>
<td>Region</td>
<td>Regional planning</td>
</tr>
<tr>
<td>$10^5$</td>
<td>Nation state</td>
<td>Spatial researches</td>
</tr>
<tr>
<td>$10^6$</td>
<td>Continent</td>
<td></td>
</tr>
<tr>
<td>$10^8$</td>
<td>Global</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.11.1 Interregional: main low-income graduates gregarious and spatial diffusion in Beijing. Source: Chaolin Gu, SHENG Mingjie, and Lingqian HU, Spatial and Social Characteristics of the “Ant Tribe” Urban Village in Beijing: Case Study of Tangjialing. Submitting.
Figure 4.11.1 presents an example of how some urban villages, such as Tangjialing, Xiaoyuhec, and Malianwa, are typical gregarious villages; they are also the bordering villages of the Haidian district outside the Fifth Ring Road of Beijing. Tangjialing village was removed in 2011; some other villages such as Xiaoniufang, Shigezhuang and Liulitun, which are adjacent to Tangjialing, and other villages that are much more remote, such as Xiyuan, Fenghuying, Eastern Banbicun, Western Banbicun and Huoying, have become the new gregarious villages for low-income new graduates, but these villages are located in the Changping district of Beijing. Observe, however, that some of these partitions may put them together as a low-income graduate gregarious region but an interregional region which is located in two administrative districts.

Intraregional (within regions)

Figure 4.11.2 shows the population density of China’s Census 2010. According to this map, the sharp division between the two main procedures, partitioning and grouping, may not be maintained. For example, we may be able to partition only part of our universe by logical division; the remaining elements may then be assigned to these cores by a grouping process. This progressive grouping process, shown in Figure 4.11.3, shows population changes in the Beijing-Tianjin-Tangshan conglomeration since the 1990s.

![Figure 4.11.2](image_url)  
*Figure 4.11.2 Population density of China’s Census 2010.*  
Boundary overlap of regions

Although regions can be partitioned or grouped in an unambiguous fashion, in practice the contrary situation is more likely. Figure 4.11.4 shows a great variety in regional definitions. Figure 4.11.4a shows these boundaries. This core area is shaded in Figure 4.11.4b.

2. Regional combinatory analysis

Where a small number of areas have to be assigned to a fixed number of regions, complete enumeration of all possible allocations of areas to regions may be feasible, and the ‘best’ grouping for the purpose at hand chosen. Figure 4.11.5a illustrates such a complete enumeration for this problem.

Gu et al. (2005) argued that the region-building process is based on a weighted combination of time and space distances. The Beijing urban mosaic is illustrated in Figure 4.11.5b.
Figure 4.11.4  Boundary overlap of regions in EU.  

Figure 4.11.5a  Beijing’s social areas, 1998.  
Figure 4.11.5b  An urban mosaic in Beijing, 1998.
3. Nodal regions and graphs

Our discussion of region-building methods has been conducted mainly in terms of formal or homogeneous regions. To a large degree, procedures developed for these can be extended to include nodal regions. Goddard (1970) was able to develop a nodal regional structure for central London from a principal components analysis of dyadic data on taxi flows. The fullest account of dyadic components analysis is given by Berry (1966, pp. 189–237) as part of a massive study of commodity flows for sixty-three commodities moving between thirty-six trade blocks within India. Despite this evidence of similarities between the two main types of regions, there remains a set of techniques which have been developed to tackle the special problems of nodal regions.

The data input for nodal regionalization consists of dyadic scores measuring flows or links between each county and all other counties in the system under analysis. Typically, such scores measure spatial interaction of some kind — for example, number of commuters, migrants, commodities, or telephone calls.

3.1 Primary linkage analysis

The use of graph theory in the interpretation of transport networks has already been clarified. Nystuen and Dacey (1961) have shown how the same type of analysis may be extended to the regionalization of flow data. A study of intercity telephone calls led them to argue that “within the myriad relations existing between cities, the network of largest flows will be the ones outlining the skeleton of the urban organization within the entire region” (Nystuen and Dacey, 1961, p. 7).

Since Nystuen and Dacey built up a regional hierarchy using the dominant outflow from each city, their method has become known as primary linkage analysis. The approach is essentially a simple one. As an illustration, consider Table 4.11.2 which shows a hypothetical matrix.

<table>
<thead>
<tr>
<th>To centre:</th>
<th>A</th>
<th>b</th>
<th>c</th>
<th>D</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
<th>j</th>
<th>k</th>
<th>L</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form centre:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>00</td>
<td>75</td>
<td>15</td>
<td>20</td>
<td>28</td>
<td>02</td>
<td>03</td>
<td>02</td>
<td>01</td>
<td>20</td>
<td>01</td>
<td>00 Satellite</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>69</td>
<td>00</td>
<td>45</td>
<td>50</td>
<td>58</td>
<td>12</td>
<td>20</td>
<td>03</td>
<td>06</td>
<td>35</td>
<td>04</td>
<td>02 Dominant</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>05</td>
<td>51</td>
<td>00</td>
<td>12</td>
<td>40</td>
<td>00</td>
<td>06</td>
<td>01</td>
<td>03</td>
<td>15</td>
<td>00</td>
<td>01 Satellite</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>19</td>
<td>57</td>
<td>14</td>
<td>00</td>
<td>30</td>
<td>07</td>
<td>06</td>
<td>02</td>
<td>11</td>
<td>18</td>
<td>05</td>
<td>01 Satellite</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>07</td>
<td>40</td>
<td>48</td>
<td>26</td>
<td>00</td>
<td>07</td>
<td>10</td>
<td>02</td>
<td>37</td>
<td>39</td>
<td>12</td>
<td>06 Dominant</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>01</td>
<td>06</td>
<td>01</td>
<td>01</td>
<td>10</td>
<td>00</td>
<td>27</td>
<td>01</td>
<td>03</td>
<td>04</td>
<td>02</td>
<td>00 Satellite</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>02</td>
<td>16</td>
<td>03</td>
<td>03</td>
<td>13</td>
<td>31</td>
<td>00</td>
<td>03</td>
<td>18</td>
<td>08</td>
<td>03</td>
<td>01 Dominant</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>00</td>
<td>04</td>
<td>00</td>
<td>01</td>
<td>03</td>
<td>03</td>
<td>06</td>
<td>00</td>
<td>12</td>
<td>38</td>
<td>04</td>
<td>00 Satellite</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>02</td>
<td>28</td>
<td>03</td>
<td>06</td>
<td>43</td>
<td>04</td>
<td>16</td>
<td>12</td>
<td>00</td>
<td>98</td>
<td>13</td>
<td>01 Satellite</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>07</td>
<td>40</td>
<td>10</td>
<td>08</td>
<td>40</td>
<td>05</td>
<td>17</td>
<td>34</td>
<td>98</td>
<td>00</td>
<td>35</td>
<td>12 Dominant</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>01</td>
<td>08</td>
<td>02</td>
<td>01</td>
<td>18</td>
<td>00</td>
<td>06</td>
<td>05</td>
<td>12</td>
<td>30</td>
<td>00</td>
<td>15 Satellite</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>00</td>
<td>02</td>
<td>00</td>
<td>00</td>
<td>07</td>
<td>00</td>
<td>01</td>
<td>00</td>
<td>01</td>
<td>06</td>
<td>12</td>
<td>00 Satellite</td>
<td></td>
</tr>
<tr>
<td>Total:</td>
<td>113</td>
<td>337</td>
<td>141</td>
<td>128</td>
<td>290</td>
<td>071</td>
<td>118</td>
<td>065</td>
<td>202</td>
<td>311</td>
<td>091</td>
<td>039</td>
<td></td>
</tr>
<tr>
<td>Rank order</td>
<td>8</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>10</td>
<td>7</td>
<td>11</td>
<td>4</td>
<td>2</td>
<td>9</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

Source: Nystuen and Dacey, 1961, p. 35.
Urban sprawl and region building

of cities (a,b…) for which the numbers in the matrix measure the flow (e.g., telephone calls) from one city to another. Figure 4.11.6 nodal structure of the matrix shows in Table 4.11.2 in terms of graph theory.

3.2 Multiple linkage analysis

Nystuen and Dacey urged caution in interpreting the primary link as evidence of urban dominance. It is clear that an outflow representing, say, 10 percent of all flows should be regarded as of quite different importance to one of 90 percent. Holmes (1973) found a strong inverse relationship between hierarchical order and traffic dispersion. The information provided by a single dominant link needs to be combined with that from other subdominant links. If, however, we wish to use more than one link from each town or county, then a criterion is needed to separate the ‘significant’ from the ‘insignificant’ flows. A regional application of this link method has been made by Holmes and Haggett (1977) to migration flow data. A number of parallel approaches to multiple link analysis have been put forward, notably that of transaction flow analysis (Brams, 1966; Soja, 1968) and graph hierarchization analysis (Rouget, 1972).

4. The case study of regional building of the Sunan region

This case employs the use of satellite image data in 1984, 1991, 2000, and 2005 to obtain the urban cluster sprawl in the Sunan (Southern Jiangsu Province). A fractal dimension calculation (the concept of fractal dimension is described ahead) is used for urban cluster analysis but not for the inner urban analysis as demonstrated by the literature, which, together with the compactness index, shows in the Sunan case study that the urban cluster is becoming a more and more homogenous and compact region. An alternative measure is spatial autocorrelation based on the urban sprawl intensity, which is used for analysis of urban cluster pattern and the hot/cold spot detection of urban sprawl (please see Matos, Chapter 4.6, and Haining, Chapter 4.2, this volume, for examples of these methodologies).
4.1 Data

Measures of areal coverage and spatial distribution are both needed to describe the morphology of an urban area adequately (Schweitzer and Steinbrink, 1998). A satellite image offers the historical footprint of human activities in a certain time, which makes the source of the data derivation comparable (see figure 4.11.7).

4.2 Methodology

Fractal dimension

Fractal dimensions are a good instrument for a global comparison of the morphology of cities (Tannier and Pumain, 2005). So in this section, we give more details about fractals through the description of the methods of determination of their fractal dimension, including radius dimension, grid dimension, correlation dimension, and boundary dimension. The first three dimensions can be assigned to the counting method, which was implemented by Fractalyse, developed by Gilles Vuidel. The fourth may be calculated by a regression function available in most statistic software.

Figure 4.11.7a The urban morphology of Sunan in different years.
a. Grid method

Theoretically, the grid dimension ($D_g$) ranges from 0 to 2, which shows a uniformity degree of urban distribution in a certain area. If $D_g$ is equal to 0, it shows that all towns are concentrated on one point – that is, there is only one town in the area, which is generally impossible in reality; if $D_g$ is equal to 1, it shows towns uniformly distribute along a line – for example, a railway, a river, or a coastal line; and if $D_g$ is equal to 2, it shows the spatial distribution of towns is completely uniform. In general, $D_g$ ranges from 1 to 2; with the value of $D_g$ increasing, there is more and more uniform spatial distribution of towns. In this study, we choose the counting centres, respectively, at the barycenter (310, 335), Changzhou (113, 229), Wuxi (253, 322), and Suzhou (388, 441).

b. Radius method

This method refers to a specific point known as the counting centre and gives the law of distribution of the occupied sites around this point. The radius dimension ($D_r$) indicates that the attenuation features of spatial distribution depart from the centre of the analysis window to its periphery. If $D_r$ is less than 2, it shows that spatial distribution of the town attenuates in density from the centre to its periphery; if $D_r$ is equal to 2, it shows the spatial distribution of the town system element is uniform in the direction of radius; and if $D_r$ is more than 2, it shows that the spatial distribution of the town system element increases in density from the centre to its periphery. For a more convenient analysis and comparison, the counting centres are used as those in the grid method.

c. Correlation method

Each point of the image is surrounded with a small squared window. The number of occupied points inside each window is enumerated. This allows the mean number of points per window of that given size to be calculated. The same operation is applied for windows of increasing sizes. In principle it is possible to choose any shape for the window, such as circle, hexagon, etc. However, since pixels are square-like, the choice of a square helps to avoid rounding errors. Like $D_g$, $D_c$ also shows a uniformity degree of urban distribution in a certain area; however, it gives more detailed results about the distribution of occupied points than the grid dimension. Generally, the correlation dimension ($D_c$) is in the range of 0 to 2. If $D_c$ is closer to 2, it shows all towns distribute more uniformly; and if $D_c$ is closer to 0, it shows one premier city has been formed.

d. Boundary method (or area-perimeter method)

If those urban surfaces were simple geometrical objects, their borders would be characterized by the dimension 1 and their surfaces by the dimension 2. Although the observed relation between borders and surfaces was regular, the ratio of surface to border was about 1.05 (Tannier and Pumain, 2005), which seems to be in contradiction to Euclidean geometry but seems to confirm the fractal geometry. For each urban polygon, its perimeter $P$ is related to the area $A$ of the same polygon by the basic fractal relationship (Johnson et al., 1995):

$$ P = kAD/2 $$

(3)
where $D$ is the fractal area-perimeter dimension ($D_a$) and $k$ is the constant of proportionality. Equation (3) can be transformed logarithmically:

$$\ln A = \frac{2}{D_a} = \ln P + \epsilon$$

(4)

where $\epsilon$ is the intercept (constant) for linear regression. We employed ARCGIS 9.0 for $A$ and $P$; the statistic software SPSS 11.0 was employed for regression to acquire the value of $D_a$. Similarly, the quality of the estimation is quantified using a correlation coefficient. In general, $D_a$ is in the range of 1 to 2.

**Compactness index**

We employ the indicator compactness derived from “landscape metrics” (for a detailed description of landscape metrics, please Reis, Silva, and Pinho, Chapter 4.4, this volume) to quantitatively measure the urban form. Compactness not only measures the patch shape for the individual patch but also considers the dispersion degree of the landscape. The compactness index ($CI$) defined by Li and Yeh (2004) is

$$CI = \frac{\sum P_i / p_i}{N} = \frac{\sum 2S_i / \pi p_i}{N}$$

(5)

where $S_i$ and $p_i$ are the area and perimeter of patch (here, urban area) $i$, $P_i$ is the perimeter of a circle with the area of $S_i$, and $N$ the total number of patches. According to this definition, the compact patch with the round shape will have the high value. In order to minimize the bias caused by the numerous small compact patches rather than the large complex ones, Li and Yeh (2004) also revised the compactness index as follows:

$$CI' = \frac{CI}{N} = \frac{\sum 2S_i / \pi p_i}{N^2}$$

(6)

**Sprawl intensity**

In addition to static analysis of urban forms by using a fractal dimension at some specific time and dynamic analysis by the evolution of the fractal dimensions in the course of time, it is necessary to select a dynamic index for showing the urban and urban cluster growths more directly. So we employ the sprawl intensity index ($SII$) as follows:

$$SII = \frac{A_s}{A_t \times \Delta t} \times 100$$

(7)

where $A_t$ (in $m^2$) is the total area within the administrative town boundary, and $A_s$ (in $m^2$ / year) the urban sprawl area of each of the towns along some direction or directions in its corresponding administrative boundary during the period time $\Delta t$ (in year). In China, the city or town is the most basic administrative boundary, which usually is merged or split partially or wholly according to its economic development in that time. As a consequence, there is a probability that the
boundary is not the same in different time periods. Here, we take the administrative boundaries in 1991 for the basic calculation and analysis unit in the course of time.

**Spatial autocorrelation**

Some standard global and new local spatial statistics, including the Moran I (Cliff and Ord, 1981), Getis–Ord G (Getis and Ord, 1992), and Local Indicators of Spatial Association (LISA) (Anselin, 1995), can be employed to detect the sprawl pattern of urban cluster (Ma et al., 2006). They start from the assumption of a randomized distribution of spatial pattern. Or the spatial pattern or form for the spatial dependence is derived from the data only without a preconceived theoretical notion. In this study, the global and local Moran I were carried out by GeoDa 0.9.5-i (Beta) developed by Luc Anselin; the global and local G statistics were calculated by Spatial Statistics Tools in ArcGIS 9.0.

a. Global Moran I

The Moran I is defined by

$$I = \frac{n}{S_0} \sum_{i} \sum_{j \neq i} w_{ij} (x_i - \bar{x})(x_j - \bar{x})$$

where $n$ is the number of observations, $x_i$ and $x_j$ denote the observed value (of sprawl intensity in this study) at location $i$ and $j$, respectively, $\bar{x}$ is the average of the $\{x_i\}$ over the $n$ locations, $w_{ij}$ is a symmetric binary spatial weight matrix $(n \times n)$ defined as 1 if location $i$ is contiguous to location $j$ or location $i$ and $j$ are within a certain distance $d$ and 0 otherwise, and $S_0$ is the sum of all the elements from $w_{ij}$.

The value of Moran I ranges from -1 to 1. The Moran I is significant and positive when the observed value of locations within a certain distance or their contiguous locations tends to be similar, negative when it tends to be dissimilar, and approximately zero when the observed values are arranged randomly and independently over space.

b. Global Getis–Ord G

The Getis–Ord G is defined by

$$G(d) = \frac{\sum \sum w_{ij}(d)x_i x_j}{\sum \sum x_i x_j}$$

where the symbols have the same meaning as in equation (4). For ease of interpretation, a standardized form of $G(d)$ can be defined as

$$Z(G) = \frac{G - E(G)}{\sqrt{Var(G)}}$$
where $E(G)$ is the mathematical expectation of $G$ and $\text{Var}(G)$ variance of $G$. If $G$ is more than $E(G)$ and $Z(G)$ is significant, the observations are clustered by relatively large values; if $G$ is less than $E(G)$ and $Z(G)$ is significant, the observations are clustered by relatively small values; and if $G$ is close to $E(G)$, the observations are randomly distributed over space.

Each of the two aforementioned statistics gives only a single value to show a whole spatial pattern for observations, so we cannot know about the spatial variance at each of locations.

**LOCAL MORAN I**

The local Moran statistic for each observation $i$ is defined as

$$I_i = \sum w_{ij}Z_iZ_j$$

(11)

where the observations $Z_i$ and $Z_j$ are in standardized form (with mean of zero and variance of one). The spatial weight $w_{ij}$ are in row-standardized form. So, $I_i$ is a product of $Z_i$ and the average of the observations in the surrounding locations. The value of $I_i$ unlike that of global Moran $I$, is tightly related with the observations, and not confined to the range of -1 to 1.

With a significant level (such as p-value less than 0.05), a positive $I_i$ and a positive $Z_i$ indicate that a high observation value at location $i$ is associated with relatively high values at its surrounding locations, viz. high-high value cluster (HH); a positive $I_i$ and a minus $Z_i$ indicate that a low observation value at location $i$ is associated with relatively low values at its surrounding locations, viz. low-low value cluster (LL); a minus $I_i$ and a positive $Z_i$ indicate that the observation value at location $i$ is much more than those at its surrounding locations, viz. high-low value cluster (HL); and a minus $I_i$ and a minus $Z_i$ indicate that the observation value at location $i$ is much less than those at its surrounding locations, viz. low-high value cluster (LH).

c. Local Getis-Ord G

The global Getis-Ord G may not easily distinguish the presence of negative spatial association from spatial clustering, which is often defined as either high-rate or low-rate spatial clustering. The global G has not been evaluated extensively, especially for low-value clustering. It is critical to interpret the local $G$ according to the degree of the global $G$ (Ord and Getis 2001). The local $G$ (including $G_i$ and $G_i^*$) is used to test the deviation of a local pattern from the average values of observations. The spatial statistic $G_i(d)$ and $G_i^*(d)$ can be defined as

$$G_i(d) = \frac{\sum_{j, x_i \neq x_j} w_{ij}(d)x_j}{\sum_{j, x_i \neq x_j} x_j}$$

$$G_i^*(d) = \frac{\sum_{j, x_i \neq x_j} w_{ij}(d)x_j}{\sum_{j, x_i \neq x_j} x_j}$$

(12)

where the symbols are the same as before. For ease of interpretation, a standardized form of $G_i(d)$ in Ord and Getis (1994) can be defined as

$$Z(G_i) = \frac{G_i - E(G_i)}{\sqrt{\text{Var}(G_i)}}$$

$$Z(G_i^*) = \frac{G_i^* - E(G_i^*)}{\sqrt{\text{Var}(G_i^*)}}$$

(13)

where $E(G_i)$ is the mathematical expectation of $G_i$ and $\text{Var}(G_i)$ is the variance; and $E(G_i^*)$ is the mathematical expectation of $G_i^*$ and $\text{Var}(G_i^*)$ is the variance.
Urban sprawl and region building

A significant and positive $Z(G)$ or $Z(G^{*})$ indicates that the location $i$ is surrounded by relatively large values, whereas a significant and negative $Z(G)$ or $Z(G^{*})$ indicates that the location $i$ is surrounded by relatively small values. So the local $G$ statistics can be used to identify spatial agglomerative patterns with high-value clusters or low-value clusters.

4.3 Results

General situation

In the period of 1984 to 2000, the urban area became linearly larger and larger from about 230 km$^2$ in 1984 to 750 km$^2$ in 2000, and then it suddenly speeds up to about 2800 km$^2$ in 2005, and about 900 km$^2$ is covered by various development zones, including industry development zones and economic and technical development zones. The total urban area of Sunan in 1991 is 2.33 times that in 1984, in 2000 is 1.57 times that in 1991 and 3.64 times that in 1984, and in 2005 is 3.41 times that in 2000, 5.34 times that in 1991, and 12.42 times that in 1984. The relationship between the total urban area and the total urban population agrees greatly with a positive exponential function (Figure 4.11.7). And the urban area growth is much faster than the urban population growth, which means that the urban growth is land-enclosed.

The global fractal radius dimensions (GFRD) centred at the barycenter (310, 335), with a maximum effective circle radius range of 1 to 619 pixels containing almost all the towns, show on the whole that the repartition of urban areas has been becoming more homogeneous over time, except in 1991 with a value below 1 (Figure 4.11.8a), revealing that the spatial organization became like a Fournier's dust in 1991. A similar tendency with big and many oscillations (meaning different reliable local dimensions with many estimated intervals) is illustrated by scaling behaviour curves in 1984, 1991, and 2000 (Figure 4.11.8), revealing similar heterogeneous spatial organization of urban surfaces, and a strong dilution existed in the radius range of 150 to 230 pixels. With urban sprawling, the curve in 2005 was obviously different from others, revealing more homogenous repartition and more compactness especially in the radius range of 1 to 400 pixels. The GFRD’s centre in different cities will present results accordingly; due to space limitations those results aren’t going to be presented in this chapter (see Figure 4.11.9).

Figure 4.11.7b The total urban areas in different years and the relationship with the corresponding total urban population.
In 1984, the global fractal correlation dimension (GFCD) was close to 1 (Figure 4.11.10a), and there also were big fluctuations of dimensions from about 0.55 to 1.43 for different estimation intervals (i.e. the local fractal correlation dimension, LFCD) (Figure 4.11.10b), showing heterogeneous spatial organization like Fournier’s dust, especially in the \( e \) ranges of 23 to 58, 58 to 69, 69 to 86, 118 to 130, and 130 to 154 pixels because of LFCD less than 1 with correlation coefficients more than 0.999. This corresponds to the towns which are distant from each other. The GFCD was higher (less than 1.3) in 1991 and 2000 but kept similar values for different estimation intervals, and the most difference was in 1991 when there still were some LFCD less than 1 with \( e \) ranges of 28 to 90 and 112 to 126 pixels. The fractal correlation dimension in 2005 was highest (more than 1.5) in the different estimation intervals, which means that the urban area has been becoming more homogeneous.

Figure 4.11.11 illustrates that \( CT' \) is becoming higher and higher over time from 1984 to 2005. It reveals on the whole that the connection between towns is more and more compacted, which validates the analysis resulting from fractal dimension. Different from the implication of the fractal dimensions and the revised compactness indexes – that is, the urban surface becoming more and more homogenous and compact – the comparison of fractal boundary dimensions.
Figure 4.11.9 Scaling behaviour in the five years, respectively, centred at: (a) barycenter, (b) Suzhou, (c) Wuxi, and (d) Changzhou; y-axis for $\alpha$ (dimensionless), x-axis for $\varepsilon$ (in pixel).

Figure 4.11.10 Fractal correlation dimension, left y-axis and solid black rectangle for dimension (dimensionless), right y-axis and hollow white rectangle for correlation coefficient (dimensionless), x-axes for time (in year); (b) the scaling behaviour, y-axis for $\alpha$ (dimensionless), x-axis for $\varepsilon$ (in pixel).

shows on the whole that the outlines of urbanized surfaces are unstable and irregular. It reveals that the urban outline is out of order to some extent from 1984 to 2005, maybe because there is a lack of a continuous urban planning schema over a long time even though in some periods some planning exists.
Means of $SII$ of all the towns are, respectively, 0.54 in the periods from 1984 to 1991, 0.34 from 1991 to 2000, and 3.65 from 2000 to 2005, showing that the sprawl intensity increases sharply in the new century and is about 6.8 times in the initial stage of reform and opening-up period after going through a transitional stage from 1991 to 2000.

The sprawl is clustered to a certain extent on the whole in each of the three periods, revealed by the global Moran $I$ of $SII$, whose spatial weights were constructed based on contiguities, respectively, from polygon boundary files ($I=0.427$ in 1984–1991, 0.176 in 1991–2000, 0.294 in 2000–2005), from the average nearest neighbours with a threshold distance of 5,000 m (the nearest neighbour observed mean distance is about 4910 m, calculated by the spatial statistics tools in ARCGIS 9.0) and from the nearest neighbours with a threshold distance mean distance of 10,000 m (Table 4.11.3). On the whole, the clustered degree in the period of 1984–1991 is highest, then in 2000–2005, and the lowest in 1991–2000, which shows that: (1) in 1984–1991, urban sprawl intensity was very heterogeneous and maybe the prominent sprawl happened only
Urban sprawl and region building

Table 4.11.3  Global I of SII whose spatial weights are constructed based on contiguities from the nearest neighbours

<table>
<thead>
<tr>
<th></th>
<th>Threshold distance = 5000 m</th>
<th>Threshold distance = 10000 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>I(d)</td>
<td>1.620</td>
<td>0.200</td>
</tr>
<tr>
<td>E(d)</td>
<td>−0.005</td>
<td>−0.005</td>
</tr>
<tr>
<td>Z Score</td>
<td>15.317</td>
<td>1.878</td>
</tr>
</tbody>
</table>

Table 4.11.4  Global G of SII whose spatial weights are constructed based on contiguities from the nearest neighbours

<table>
<thead>
<tr>
<th></th>
<th>Threshold distance = 5000 m</th>
<th>Threshold distance = 10000 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>G(d) (×10^6)</td>
<td>9.234</td>
<td>2.165</td>
</tr>
<tr>
<td>E(d) (×10^6)</td>
<td>1.034</td>
<td>1.034</td>
</tr>
<tr>
<td>Z Score</td>
<td>17.053</td>
<td>3.475</td>
</tr>
</tbody>
</table>

around several towns or cities; (2) in 1991–2000, urban sprawl intensity became a little homogeneous; and (3) in 2000–2005, the towns grew in homogeneously again; however, the prominent sprawl happened around more towns or cities.

However, the clustered patterns indicating urban sprawl were different, which was confirmed by the local Moran I of SII. The local Moran I scatter maps of SII show that (Figure 4.11.12):

1. There are obvious HH, HL, LH, and LL clusters with similar spatial distribution but different sizes in the three periods; some towns subjected to Changzhou, Wuxi, Suzhou were classified into the HH cluster in each of the three periods, and some towns along the Yangtze River always were in the HH cluster region; additionally, HH cluster regions were transformed gradually from the city cores to the suburbs in the periods from 1984–1991 to 2000–2005.
2. In 1984–1991, HH clusters were focused on towns subjected to the three cities of Changzhou, Wuxi, and Suzhou and several towns along the Yangtze River, around which existed several HL clusters; a majority of towns were classified into the LL cluster. It reveals in this period that urban fast growth was centralized mainly in the three big cities.
3. In 1991–2000, the HH cluster was extended in spatial distribution; there were more and more HH cluster towns along the Yangtze River, and towns subjected to Kunshan were classified into the HH cluster but the LL cluster in 1984–1991.
4. In 2000–2005, the HH clusters were joined contiguously into a zonal region from Wuxi, through Suzhou and Kunshan, to Taicang; additionally, the suburbs were in the HH cluster, like in 1991–2000.
5. Wholly and generally, the HH cluster was pattern-pointed at the initial stage, and then they were transformed into the periphery or enlarged into a large region but still pattern-pointed; with a rapidly developing economy, more HH clusters emerged and some of them were joined together into a zonal region.
Figure 4.11.12  Moran I scatter map.
Figure 4.11.13 Spatial distribution of the local Getis-Ord G; 1 denotes Huning railway, 2 Huning expressway, 3 Subei railway, 4 Xicheng expressway, 5 Yanjiang expressway, and 6 Sujia-hang expressway.
In order to uncover different clustered patterns more deeply and formally, the hot/cold spot analysis technique was employed to calculate the global Getis-Ord $G$ values. The global Getis-Ord $G$ value, together with the $E$ value and $Z$ score (Table 4.11.4), shows that the cluster revealed by the global Moran I is a high cluster, which is more significant in 1984–1991 than in the other two periods. So the hot spot of urban sprawl was very concentrative in 1984–1991, and then was dispersed gradually. In order to uncover the spatial distribution of the hot spots and their transformation, the local Getis-Ord $G$ values were calculated also.

Figure 4.11.13 shows from the view of sprawl intensity that: (1) in 1984–1991, there were four hot spots, which were concentrated, respectively, at the four cities – that is, Changzhou, Wuxi, Suzhou, and Jiangyin; the former three are connected directly by the Huning (from Nanjing to Shanghai) railway and the fourth lies along the Yangtze River; (2) in 1991–2000, the hot spots respectively located at Changzhou and Jiangyin still existed but dispersed into a big connected patch; the hot spot located at Suzhou was enlarged and a new one grew up at Kunshan; the one located at Wuxi became weaker; (3) in 2000–2005, the hot spot located at Changzhou still existed but diminished, and so did the big connected patch; the one located at Wuxi was enhanced again; noticeably, a zonal hot spot had grown up from Wuxi, through Suzhou, to Kunshan along Huning railway and Huning expressway; additionally, a new one emerged at Tai-cang along the Yangtze River; and (4) wholly and generally, the hot spots of urban sprawl were concentrated mainly at big cities at the initial stage, where the urban sprawl were self-governed and they didn’t have a strong influence on one another; and then the hot spots were spread to their surrounding towns gradually, or they were joined with other hot spots into a big connected patch; with economic and social development, the hot spots were spread and dispersed continuously and some were joined into a zonal region along the important transportation axes.

**Conclusion**

Region building is one of the commonest applied problems encountered in location analysis. This chapter has shown advances that have been made in regionalization methods. Regional divisions represent a compromise between spatial contiguity on the one hand and grouping counties with like characteristics on the other. The number of possible regional divisions or combinations in any study area is usually very large indeed. Thus, any study area is usually very large. Thus, any proposed scheme is less likely to be a single sharply peaked optimum than one of a set of rather similar near-optimal solutions. This chapter also reveals the practicability of fractal dimension measures for homogeneity and compactness, just as some literature shows; the differences in this research is that it is applied in analysis not of inner urban areas but of urban clusters. In the course of the analysis, applying a scaling action may help detect the change threshold range. So the fractal dimension and its incidental scaling behaviour may complement each other when performing an analysis of homogeneity and compactness. The urban sprawl intensity is a good normalized index for analysis and comparison of different urban sprawls and also is a base to detect the urban cluster pattern by spatial autocorrelation measures, which is very practical and applicable by means of a Moran scatter map and hot/cold spot detection.

**Acknowledgement**

The authors would like to express our gratitude to Gilles Vüidel for having developed the pertinent computer program Fractalyse, used for fractal pattern analysis, and also to Luc Anselin for the program GeoDa, used for spatial autocorrelation analysis.
Urban sprawl and region building

Note

1 See Ronghua et al. (2008).

References

Gu Chaolin


