Spatial Analysis and GIS

Chapter · January 1993

1 author:

Morton E. O'Kelly
The Ohio State University

4

Spatial analysis and GIS

Morton E. O'Kelly

1. Introduction

Many of us involved in spatial analysis research are excited to see the explosive growth of interest in GIS. It is clear that the rapid growth of GIS has given a big boost to fundamental research in spatial analysis, and in many ways *solidifies* the future of the quantitative focus of the discipline.

There exist potential linkages between many aspects of spatial analysis and the new information processing, data handling, data storage and data display techniques available through GIS. Furthermore, there are strong emerging links to applications in other disciplines with parallel interests in spatial analysis (including ecology, archaeology, natural resources, landscape architecture and geodetic science, etc.) These allied fields have long used spatial analysis techniques (see, for example, Upham, 1979; Bartlett, 1975; and Diggle, 1983) but are now increasingly using GIS as a creative tool.

Against the background of this growth, and from the perspective of a spatial analyst looking out towards the next stages of development in GIS, two major directions which need attention are apparent: First, the traditional methods for displaying data about spatial situations, and for presenting the results of spatial analyses, need to be overhauled in view of the great advances in information processing technology. This line of attack is straightforward: it suggests that new display techniques be added to the output of existing spatial analysis operations. Ideally, the analyst would use the GIS in query mode (Goodchild, 1987) to develop an improved understanding of the properties of a spatial system. While this is easy to state in principle, the actual implementation will stretch current capabilities and will require a rethinking of the conceptual bases for spatial analysis. One such example might arise if improved display techniques increased the accessibility of multiobjective programming, thereby encouraging analysts to take a multiobjective view, and therefore replacing existing uni-dimensional methods. One obvious area for this to play a role is in locational analysis.

Second, spatial analysts need to help GIS end-users and providers to understand and to improve their sets of tools, and to enhance the appropriate levels of theory and modelling capability in real problem-solving situations. This need is particularly acute if *appropriate* analytical methods are to lie at the foundation of developments in GIS. The capability to adapt existing algorithms to the data structures in GIS is a critical component of this research.

Source: Fotheringham & Rogerson, 1994

The overall theme of this chapter is that when GIS integrates spatial analysis at a fundamental level, the full potential of spatial analysis will be unlocked. The case is made by demonstrating that there are potential *new* results and advances that are obtainable by linking spatial analysis and GIS. In the style of a position paper, this effort builds the case for an improved linkage between spatial analysis and GIS by showing that GIS can be improved by adding innovative spatial analysis functions, and in turn by showing how certain spatial analysis operations are enhanced using GIS. The third section of the chapter mentions some the immediate incremental building blocks that are needed to start this transformation of spatial analytical methods, while the concluding section mentions several barriers to the realization of these aims.

2. Innovations in spatial statistics and GIS applications

As discussed in the previous section, the next few years promise an unprecedented opportunity for spatial analysts to promote and contribute to the spread of new tools to the GIS community. This section is designed to make some of the power of these types of tools apparent to the potential user. To attempt to make the discussion more concrete, let us go back to some fundamental building blocks or entities in spatial analysis: i.e. Points, Lines and Areas. Associated with each of these classes of entities are new spatial analytical operations that need to be developed. Couclelis (1991) correctly points out that the focus on space as a container does not provide the platform for answering the complex types of spatial questions posed in planning; however, see more on the site and situation distinction she draws in later paragraphs. In each case the role of new exploratory tools, visualization, and space-time analyses can be seen in slightly different ways.

2.1. Space-time pattern recognition in point data sets

Geographers have been slow to integrate *both* temporal and spatial dimension into GIS. There are formidable technical obstacles, but even at the fundamental research level, analysts have been unable to operationalize the role of space and time in simple models. A concrete example may help. A conceptual model of space and time acting as a constraint on human activity (Hagerstrand) has gone largely untested because of the difficulty of translating activity records from travel diaries into three dimensional diagrams called 'prisms' by Hagerstrand (however, see Miller, 1991). There have been studies of the importance of distance (Gatrell, 1983) and time (Parkes and Thrift, 1980; Goodchild and Janelle, 1984; and Janelle, Goodchild and Klinkenberg, 1988), but despite these impressive research efforts, it is uncommon to consider combined time-space 'reach' as a quantitative tool in measuring transportation plans. Two exceptions include Villoria's (1989) dissertation, which measured the size of activity fields in an urban case study in the Philippines, and Kent's (1984) mapping of activity spaces from archaeological data.

2.2. A model for space-time data analysis

Assume that the events are located at distinguishable locations, and that a simple date is used to keep track of the time of events. Let $W_{ij} = f(t_i, t_j)$ where t_i and t_j are the times of the events i and j, and further that i and j take place at $p_i = (x_i, y_i)$ and $p_j = (x_j, y_j)$ respectively. Supposing that W_{ij} is a decreasing function of the time between the events

$$W_{ij} = 1 \,/\, \{e + | t_i - t_j |\}.$$

Then, given thresholds D and T, classification of events in space and time can be simply represented in a two-by-two table (Table 4.1).

	SPACE	
TIME	$d(p_i, p_j) < D$	$d(p_i, p_j) > D$
$W_{ij} > T$ highly interactive	CLOSE TOGETHER CONTEMPORANEOUS	FAR APART CONTEMPORANEOUS
$W_{ij} < T$ weakly interactive	CLOSE TOGETHER TIME GAP	FAR APART TIME GAP

The statistical analysis of such 2-way tables is a fundamental one for recognizing whether or not there are significant patterns in space and time (see Knox, 1964). The strength of the temporal interaction is measured as follows: W increases as the event are closer together in time, and D increases as the events are further apart in space. To what extent are observations which are close together in time also spatially clustered? The problem is to recognize clusters or groups in the set. While there are many follow-up papers to Knox's pioneering analysis, the method is problematic because of the non-independence of events (see Glick, 1979). The next section develops a novel mathematical approach to the problem of space-time pattern analysis: one which is of great potential usefulness in identifying the spatial pattern and perhaps helping to uncover the underlying spatial process.

Consider a set of n interacting points in a two-dimensional space. The levels of interactions between the observations are given exogenously, as functions of their temporal separation. Assume that the cluster means must be adjusted to reflect the interaction between the entities. For example, consider a system of $n = n_1 + n_2$ nodes such that the n_1 subset is temporally linked, and the n_2 subset is also highly interactive among themselves, and for the sake of illustration suppose that there are negligible interactions between the subsets. (That is the $(n_1 \times n_2)$ and $(n_2 \times n_1)$ subsystems contain only zero interactions.) If the n_1 and n_2 nodes are plotted graphically, it would be fortuitous if all the n_1 and n_2 nodes could be separated neatly into two easily identifiable spatial groups. Indeed, since there is no requirement of contiguity for the interacting entities, there is no guarantee that a cluster of points n_1 should contain only adjacent nodes. While the conventional geostatistical clustering problem for several groups yields a partition with the property that all the observations which are closer to centroid A than to centroid B are assigned to the same group, this is *not* a property of the interacting cluster problem (see O'Kelly, 1992). A solution to this partitioning problem has recently been proposed by O'Kelly (1992) and this section briefly summarizes the method.

It is required to cluster the n observations into p groups, so that the sum of squared deviations from the cluster means is as small as possible. Assume that the cluster means are adjusted to reflect the interaction between the entities. Further, since it is desirable to place highly interactive observations in the same group, it will be assumed that the penalty for assigning observation i to group g and observation j to group h is an increasing function of the distance between the group centroids and the interaction level W_{ij}. Specifically, the 'cost' of assigning i to g and j to h is:

$$P[i(g), j(h)] = W_{ij}(d_{ig} + d'_{gh} + d_{hj}) \tag{1}$$

where

W_{ij} is the exogenous interaction effect,
d_{ig} is the squared distance from i to cluster center g,
d'_{gh} is the squared distance between the cluster centers,
d_{hj} is the squared distance from j to cluster center h.

Let (X_g, Y_g) be the centroid of group g for $g = 1, \ldots, p$. The distances are defined as: $d_{ig} = (x_i - X_g)^2 + (y_i - Y_g)^2$ for all $i = 1, \ldots, n$ and $g = 1, \ldots, p$; and $d'_{gh} = (X_g - X_h)^2 + (Y_g - Y_h)^2$ for all g and $h = 1, \ldots, p$. In the first part of this chapter the objective is to choose (X_g, Y_g), $g = 1, \ldots, p$, so as to

$$\text{MIN } T = \Sigma_i \Sigma_j W_{ij} \Sigma_g \Sigma_h K_{ijgh} D_{ijgh} \tag{2}$$

where $K_{ijgh} = 1$ if i belongs to group g and j belongs to group h, and $K_{ijgh} = 0$ otherwise; and where $D_{ijgh} = d_{ig} + d'_{gh} + d_{hj}$. Note that the K_{ijgh} integer variables obey the following restrictions:

$$K_{ijgh} = X_{ig} X_{jh} \tag{3}$$

where the X values are allocation variables, that is:

$$\Sigma_j X_{ij} = 1, \text{ for all i and } X_{ij} \text{ is either 0 or 1} \tag{4}$$

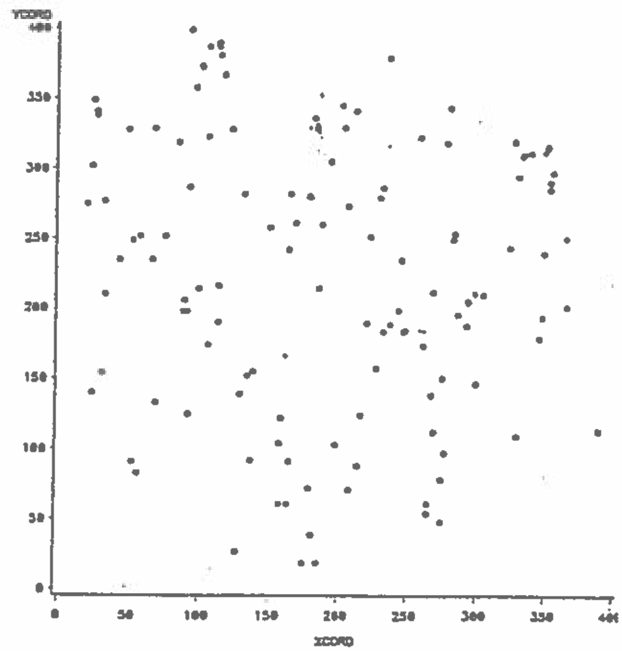

Figure 4.1. Point data set n = 125 from M. R. Klauber and P. Mustacchi (1970).

A useful property of the problem is that the use of a squared distance term yields a linear system of equations for the coordinates of the cluster centroids. These equations are derived and solved repeatedly, for a given set of cluster allocations. A sequential reallocation of the observations between the clusters is then performed: that is K_{ijgh} is initially assumed to be fixed and this is equivalent to starting the problem with a known partition of the observations between p groups. The model solved in O'Kelly (1992) discusses the iterative reassignment of observations to clusters.

As an example of the procedures explained in the previous paragraphs, consider the 125 points shown in Figure 4.1. The observations are used because they present a convenient source for a set of x, y locations and a time stamp for each event. No substantive contribution to the original data context is attempted here, rather suppose that these are the locations of fires in a city, and we are interested to see if there are clusters of events in space, in time, or in space and time. For the sake of illustration, a simple set of interactions between entities is modeled as $W_{ij} = 10./(EPS + | t_i - t_j |)$ where EPS is a small constant (set to 0.1) to prevent division by zero if the two events occur at exactly the same time. The result of clustering 125 observations into 4 groups in shown in Figure 4.2 which shows conventional group centroids, using a 'X' symbol, and the cluster membership of the data points. No attempt has been made here to find the optimal number of groups, and it is

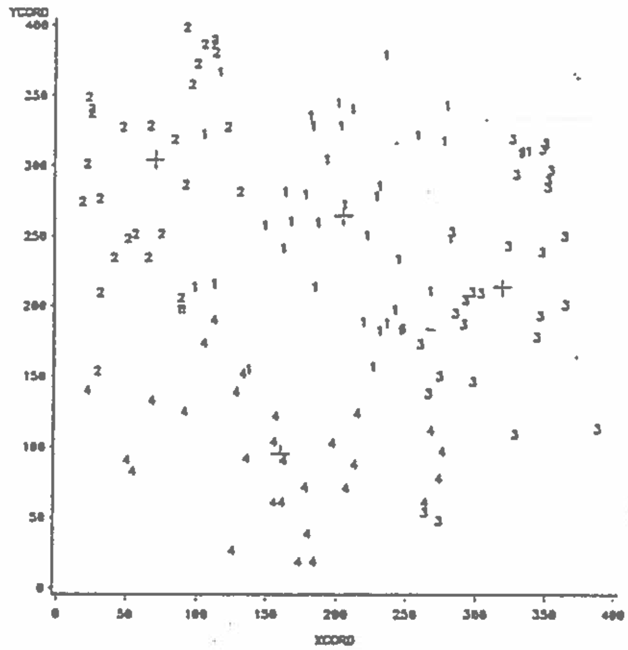

Figure 4.2. Preliminary example: clustering into 4 groups.

recognized that many different partitions of the data could be produced by altering the initial partition, or by changing convergence criteria.

While conventional clustering has been in use for many years as a means of spatial pattern analysis (see Baxter, 1971, for early examples) the extension of clustering techniques to allow for interdependence between the clusters adds significant benefits in GIS applications such as pattern recognition – e.g. hypothesis testing for patterns of arson and crime. There exists some strong potential applications of these tools, e.g. K-means program/ISODATA package; similar problems arise in the case of organizing large operations like the postal service, mail order deliveries and retail market delimitation.

Both the OR questions (concerning the optimal partitioning) and the GIS question (concerning the display/visualization of large data base) require further work.

2.3. Lines and flows: spatial interaction

Consider the location of facilities, using an interaction model to drive the choice of sites. The probabilistic allocation of demand to these facilities is a classic problem in spatial interaction theory. Techniques to locate these facilities in a manner which optimizes the spatial interaction properties are well known (see Hodgson, 1978; and O'Kelly, 1987.) However, a new feature which is essential to these models, was omitted until the graphic capabilities of GIS revealed its importance.

Spatial interaction concepts are linked to covering and optimal location models in the following way: it is apparent from the visualization tools that the role of spatial constraint in interaction models has not been fully integrated into locational analysis and indeed the interplay with the space-time prism model is a very interesting possibility. If a facility has a known range, and therefore can only cover a demand point in the fixed radius, then this fact ought to be built in to the spatial location model. The hybrid concept of probabilistic interaction towards facilities that have a covering radius introduces some subtle new spatial situations. Observe that there is no guarantee that a specific demand point has any facility within in its range. The demand from that origin is therefore uncovered, but in spatial interaction terms some demand must be sent to an artificial destination to represent the uncovered demand.

To pull out the general point: spatial analysts are concerned to link methods from OR etc., and to develop novel visualization tools to help with analysis of the model. The algorithm does the work, and then the solutions are posted back to the GIS to validate and query the quality of the solution. Questions are clear: are the correct assumptions built into the optimization tool? Are the key parameters and mechanisms represented in the model? What data/empirical estimates need to be adjusted as a result of the analysis? The entire model-building exercise is viewed as an iterative that cycles back to the beginning to check the validity of the model – a step which is too often replaced by hand waving.

The problem is a generic one – whenever a complex model produces output in the domain of X, the sensitivities of outputs to changes in the key parameters, or to changes in the assumptions about the role of critical variables, must be investigated. The answers in this kind of research hinge on both a fundamental substantive understanding of the problem, and a mathematical insight to the model.

2.4. Spatial autocorrelation

As another example of the marriage of spatial models and GIS, consider the example of spatial autocorrelation. The correct measurement of spatial autocorrelation is a necessity, but one which is open to a wide variety of subtle variation. Clearly GIS allows these measures to be gathered rather easily, since adjacency is one of the properties of spatial entities that is maintained in GIS databases. This is possible because of the increased availability of adjacency 'facts' from topologically integrated databases. Even with spatial autocorrelation at hand, the analyst must decide how to make appropriate statistical use of the indices, and be cognizant of the role of units and scale of analysis on the results (Chou, 1991; Newsome, 1992). Also, the theory underlying the appropriate statistical operators for various measures of spatial autocorrelation needs to be fully understood by the end user (see, for example, Haining, 1978). There is little to be gained by making spatial autocorrelation one of the many descriptive statistics collected from a spatial database, unless the sophistication of the user is sufficient to make correct use of this information (see Getis, 1991; Chou, 1991).

2.5. Spatial situation vs. spatial site

A final example grows from the theme developed by Couclelis (1991) about the importance of site and situation and is discussed in terms of Fotheringham's competing destination (CD) models, and the possibility of enhanced measurement of CD effects in GIS.

An example of the kind of situation effect which is widely known is the Fotheringham (1983) measure of competition between destinations. The measure is important to spatial interaction models and can be expected to be an easily calculated variable in GIS implementations of spatial interaction models. Moreover, the GIS focus allows us to see that the conventional CD measure can be generalized once the ideas of neighborhood variables and adjacency are fully integrated.

The starting point is $A_{ij} = \sum_{k \neq j,\, k \neq i} W_{ik}\, f_i(d_{jk})$, which is the Fotheringham measure of CD for facility j, from i's perspective. There are three reasons that the accessibility of j is measured with an 'i' subscript: (1) the calculation avoids the inclusion of i; (2) the attraction variable W is possibly calculated differently from each zonal perspective; and (3) the distance decay function is specific to zone i. Notice that this calculation includes all potential competitors of j, including some that might reasonably be expected to avoid interaction with i. The following set of generalizations emerge, and these are measures which would be easy to calculate in a database that keeps track of the 'neighbourhood' variables. Some care would have to be taken to ensure that the data structure allowed the efficient collection of the facts needed for the measures below. Two examples of each of three types of generalization are given here:

1 Restrictions on the list of competitors of j, from i's perspective

(a) $A_{ij} = \sum_{k \neq j,\, k \in S(i)} W_{ik}\, f_i(d_{jk})$, where $S(i) = \{k |\, d_{ik} \geq L\}$

This measures the accessibility of j to those facilities 'k' which are further than L from i (e.g. consistent use of the masking radius used in Fotheringham, would set L = 160).

(b) $A_{ij} = \sum_{k \neq j,\, k \in T(i)} W_{ik}\, f_i(d_{jk})$, where $T(i) = \{k |\, d_{ik} \leq d_{ij}\}$

This measures the accessibility of j to those facilities 'k' which are closer than d_{ij} to i (e.g. it could also be defined with inequality reversed).

2 Restrictions on the list of competitors of j, from j's perspective

(a) $A_{ij} = \Sigma_{k \neq j,\ k \in S(j)} W_{ik} f_i(d_{jk})$, where $S(j) = \{k | d_{jk} \leq Z\}$

This measures the accessibility of j to those facilities 'k' which are closer than Z to j (e.g. this is called a traffic shadow effect, in Taaffe, 1956).

(b) $A_{ij} = \Sigma_{k \neq j,\ k \in T(j)} W_{ik} f_i(d_{jk})$, where $T(j) = \{k | D_{jk} = 1\}$

This measures the accessibility of j to those facilities 'k' which are in the set of pairs for which $D_{jk} = 1$. These could be pairs of twin cities, or other close substitutes.

3 Restrictions on the list of competitors of j, from (i, j)'s perspective

(a) $A_{ij} = \Sigma_{k \neq j,\ k \in S(i,j)} W_{ik} f_i(d_{jk})$, where $S(i, j) = \{k | d_{ij} + d_{kj} \leq R\}$

This measures the accessibility of j to those facilities 'k' which are in the set of pairs (j, k) for which the sum of distances from i to j and from k to j is less than some budget. This could be used to define an interaction space such as a corridor.

(b) $A_{ij} = \Sigma_{k \neq j,\ k \in T(i,j)} W_{ik} f_i(d_{jk})$, where $T(i, j) = \{k | d_{ik} + d_{kj} \leq R_1,$
$d_{ik} + d_{kf} \leq R_2,\ d_{jk} + d_{kf} \leq R_3\}$

This measures the accessibility of j to those facilities 'k' which are in the action space bounded by distances of R_1, R_2, and R_3 from the fixed points i, j, k and f. These could be pairs of interactive points, or corridors pointing towards some particularly attractive alternative such as the city center 'f'.

Important research questions could then be answered in spatial analysis: such as further empirical evidence of the role of the competition effect, especially in the light of the generalized measurement of spatial competition. With these tools realism is added to the theoretically justified measures of competition between alternatives proposed by Fotheringham (1983).

3. Progress via incremental improvements

There are many spatial analytical tools which do not currently exist in GIS (and some which may be inherently impossible to implement because of data structures – see Couclelis, 1991; Goodchild, 1987). In addition GIS prompts us to see new ideas in spatial analysis, which otherwise might not be clear. But there are barriers to the realization of this potential.

To get from the current level to some of the tools discussed here, major fundamental research will have to be done on some basic steps. These intermediate steps include:

1. Exploratory data analysis, to search for previously unseen processes within the data and test new analytical techniques for finding new instances of a sample pattern.

2. Systematic integration of existing models of spatial processes such as interaction and gravitational concepts, spatial autocorrelation measures etc. This includes investigation of the problem of 'modifiable areal units'.
3. Spatial process and spatial patterns through time also need to be investigated.

The advanced tools discussed in this chapter must not be viewed as modules to be added 'piecemeal' to the functionality of a GIS: rather these are themes which underlie the development of a complex set of new procedures. These themes will surface in various ways and in differing proportions depending on the specific domain of research. What are the implications for spatial analysis? While there are many (see Openshaw, 1991; Nyerges, 1991; and others) just two issues are mentioned here. The first is that exploratory data analysis (EDA) will be the tool of choice; the second requires analysts to rethink the role of geometric spatial analysis. These issues are dealt with in turn.

Exploratory data analysis

Techniques for sifting through large data streams and innovative visualization techniques are needed to allow the analyst to assimilate the quantity of data presented (e.g. imaging systems; panel surveys; census products). The major problem becomes that of deciding what is important in the vast quantities of data which are generated in large models and GIS. The key technical advance will be in pattern recognition, which intelligently allows the user to sift through the data, reduce dimensionality, find patterns of interest, and then order the GIS to find other instances or similar occurrences. This sounds simple, but is difficult to implement when the size of the underlying data base is of the order of multiple gigabytes.

Also in the arena of *exploratory data analysis* and *categorical data analysis* is the important point that large-scale data surveys in the *social sciences* are producing very large amounts of data. The quantity of these data may discourage analysts from embarking on useful research. In some cases the quality of the data are also in question, especially in the marketing and retail analysis arena where the results of cluster analysis are used to estimate the micro demographics of spatial zones. Other high-quality examples which are currently available include the *Annual Housing Survey*, data collected from continuous work history, the Italian census as discussed by Openshaw, the Cardiff panel data survey, etc. In the presence of such large data sets the need for novel visualization techniques in the social sciences is just as pressing as in the physical realm.

Geometric pattern analysis

At the earliest stage of quantitative analysis, simple small-scale analyses were performed. Typically, these data analyses considered small sample case studies, and involved a proof of concept kind of approach. The full 'industrial strength' scale up of these methods was constrained by the power of data processing technology, and by the lack of large digital data bases. Now, with the advent of digital databases, conventional spatial statistical studies using nearest neighbour analysis, quadrat analysis, or other spacing statistics (e.g. Getis and Boots, 1978; and Rogers, 1974) can be re-invigorated when matched to very large data sets, and to efficient computational geometry routines. A broader issue remains: are the types of questions and analyses carried out earlier still interesting? Are these techniques tractable for data sets of arbitrary size? Are heuristics needed? Do these breakthroughs allow us to give trustworthy answers (Openshaw, 1987; Openshaw *et al.*, 1987) to reasonable questions about spatial patterns? Researchers need to re-examine the tool kit of spacing, quadrat and nearest neighbour techniques, and assess the impact of larger data volumes and improved computational technique on their applicability.

While short-run questions challenge us, spatial scientists must also aim for the integration of spatial models into GIS and anticipate the successful marriage of theory and practice, yielding solutions to problems that were previously thought to be unmanageable.

4. Conclusion: barriers to improving GIS via spatial analysis tools

As a conclusion to this chapter, let us examine, briefly, the capability of GIS is for spatial analysis tasks, and highlight some of the barriers to realizing the potentials mentioned in this paper.

One of the observations that is made repeatedly is that there is a mismatch between the spatial analytical capabilities of the research community and the applied tools available in GIS and in use by practitioners. It would seem to be obvious that GIS would be improved immediately if only all the research capabilities of the best spatial analysts were somehow 'built in' to GIS. This has not happened and the result is that exciting research tools are not quickly adopted by GIS and are inaccessible to the practitioner. Three reasons for the breakdown in transfer of technique to technology appear: first, because the non-geographers using GIS do not think in spatial terms, and therefore do not intuitively ask questions about spatial pattern and spatial association. For this group, space as a container is a perfectly acceptable medium, and the GIS as a spatial data handling tool provides all the functionality (see Couclelis, 1991). Second, because the research community has not been able to justify the models using real world 'value added' terminology. Therefore, research tools remain in the laboratory, without being implemented in commercial packages. And third, because the real world is driven by market concerns and perceives spatial research tools as limited by the size of the market (see similar point in Goodchild, 1987, p. 333).

For theoreticians the first priority is *not* that ideas be commercially applied, as the worth of a method is not judged solely in terms of its marketing in a commercial package. However, if geographers want to participate fully in this environment, *they have to be more concerned with explaining research results in a manner which is accessible and where the marginal value of a enhancement to a research tool is clear.* This may mean being careful not to leave the discussion in the hands of some less technical marketing person. Changes *have* taken place, and the realities point to a much more commercially oriented research environment than before. It is important that the people who were instrumental in nurturing this branch of the discipline learn the additional skills necessary to ensure the timely and accurate usage of sophisticated spatial tools. These commercial prospects also introduce heightened competition, ethical issues, and legal issues, and these are part of the costs associated with increased real world interest in the realm of spatial analysis. As a side-benefit, the GIS explosion has ended the debate about 'relevance' of quantitative methods, and the growth in GIS enrollment in graduate programs has reinvigorated the quantitative courses that go along with this track; it is instructive to reassess some of the essays in Gaile and Willmott (1984), and in Macmillan (1989) in the light of the fast pace of GIS growth. Openshaw (1991) puts this point of view in especially clear terms.

Taking stock, there is a lot to be happy about – the reinvigorated role of spatial analysis, and the increased visibility of the research workers at the cutting-edge of this discipline. However, specialists in this area must consider the challenge of taking research tools out of the lab and into the marketplace for consumption by end-users.

Appendix

Throughout this chapter I have kept the discussion 'neutral' with respect to software and commercial packages. This set of notes adds specific names and products in parenthetical form.

1. Examples on the cutting edge include McDonald's use of GIS platforms to maintain customer spotting data, and the portrayal of multi-media reports on retail sites. Nevertheless, it is not surprising that the majority of commercial real estate research departments are using simple demographic desk top mapping packages, and are a long distance away from adopting optimal centralized site selection. Even in redistricting software, there is an emphasis on mapping impacts of boundary changes, driven by the analysts judgement, rather than using algorithms to define optimal partitions. In any business there is a spectrum of innovativeness all the way from visionary advanced guards (e.g. McDonald's and Arby's) to relatively staid followers, who are not interested in innovation. There is a market for spatial analysis, but at the moment it is directed towards simple

rather than complex spatial data handling tasks. The market will grow more quickly as the adoption of more sophisticated ideas penetrates the user community, and trained spatial analysts will be called on to explain and implement technical advances.

2. The SPANS package allows the demand from a set of areas to be allocated among a fixed set of facility locations.
3. Many associate GIS with the commercial software packages that are on the market today. These packages often provide elementary analytical tools, but quite often they stop short of the full flexibility of the individual's expertise. That is, individuals writing their own personalized code would add features that would not have broad appeal, or indeed broad acceptance and understanding. Thus there is a problem – it is not possible to tackle cutting-edge research problems solely within the confines of commercial GIS, because by definition they only include the more widely accepted and known tools. The exception would be that the GIS may provided a 'macro', 'procedure' or 'application' writing capability, through which the skilled analyst can build a complex model from fundamental building blocks. Some packages have such diverse tools that it is possible to create macros for novel spatial analytical techniques.

Acknowledgements

The support of the National Science Foundation, SES 88-21227 and SES 89-46881 for 'Models of the Location of Hub Facilities' is gratefully acknowledged. NSF Grant SES 88-10917 supported participation in the NCGIA I–14 initiative. Support from the Ohio Supercomputer Center for 'A clustering approach to the planar hub location problem' is also appreciated. Research assistance has been provided by Harvey Miller, and his comments on the manuscript are appreciated. An earlier version of this chapter was presented at San Diego State University, 14 February 1991, where useful comments were received from Stuart Aitken, Serge Rey, Gerry Rushton and Art Getis. Discussions at the I–14 meetings with Howard Slavin and Noel Cressie were valuable. Comments have also been received from the Geodetic Science seminar at The Ohio State University.

References

Bartlett, M. S., 1975, *The Statistical Analysis of Spatial Pattern*, London: Chapman and Hall; New York: Wiley.

Baxter, R. S., 1971, *An urban atlas – Reading*, prepared within the Urban Systems Study under the sponsorship of the Centre for Environmental Studies. Cambridge, University of Cambridge Department of Architecture.

Chou, Y. H., 1991, Map resolution and spatial autocorrelation. *Geographical Analysis*, 21(3), 228–246.

Couclelis, H., 1991, Requirements for planning-relevant GIS: a spatial perspective, *Papers in Regional Science*, 70(1), 9–19.

Diggle, P., 1983, Statistical Analysis of Spatial Point Patterns (Mathematics in biology), London: Academic Press.

Fotheringham, A. S., 1983, A new set of spatial interaction models: the theory of competing destinations, *Environment and Planning A*, **15**, 15–36.

Gaile, Gary, L. and Cort, J. Willmott (Eds.), 1984, *Spatial Statistics and Models*, Dordrecht; Boston: D. Reidel Pub. Co.; Hingham, MA: distributed in the USA by Kluwer Boston Academic Publishers.

Gatrell, Anthony C., 1983, *Distance and Space: a Geographical Perspective*, Oxford: Clarendon Press; New York: Oxford University Press.

Getis, A. and B. Boots, 1978, Models of Spatial Processes: an Approach to the Study of Point, Line, and Area Patterns. Cambridge (Eng.); New York: Cambridge University Press.

Getis, A., 1991, Spatial interaction and spatial autocorrelation: a cross-product approach, *Environment and Planning A*, **23**, 1269–1277.

Glick, B., 1979, Tests for space-time clustering used in cancer research, *Geographical Analysis*, 11(2), 202–207.

Goodchild, M. F., 1987, A spatial analytical perspective on geographical information systems, *Int. J. Geographical Information Systems*, 1(4), 327–334.

Goodchild, M. F. and Janelle, D., 1984, The city around the clock: space-time patterns of urban ecological structure, *Environment and Planning A*, **16**, 807–820.

Haining, Robert P., 1978, Specification and estimation problems in models of spatial dependence. Evanston, Ill.: Dept. of Geography, Northwestern University.

Hodgson, M. J., 1978, Towards more realistic allocation in location-allocation models: an interaction approach, *Environment and Planning A*, **10**, 1273–1285.

Janelle, D., M. F., Goodchild and Klinkenberg, B., 1988, Space-time diaries and travel characteristics for different levels of respondent aggregation, *Environment and Planning A*, **20**, 891–906.

Kent, S., 1984, Analyzing activity areas: an ethnoarchaeological study of the use of space, Albuquerque: University of New Mexico Press.

Klauber, M. R. and Mustacchi, P., 1970, Space-time clustering of childhood Leukemia in San Francisco. *Cancer Research*, **30**, 1969–1973.

Knox, E. G., 1964, The detection of space-time interactions. *Applied Statistics*, **13**, 25–30.

Macmillan, Bill, 1989, *Remodelling Geography*, Blackwell: Oxford.

Miller, H., 1991, Modeling accessibility using space-time prism concepts within geographical information systems, *Int. J. of Geographical Information Systems*, 5(3), 287–301.

Newsome, T. H., 1992, *Measuring Spatial Pattern in Census Units: Residential Segregation in Franklin County, Ohio*, Unpublished Ph.D. Dissertation, Department of Geography, The Ohio State University, Columbus, Ohio.

Nyerges, T., 1991, Geographic information abstractions: conceptual clarity for geographic modeling, *Environment and Planning A*, **23**, 1483–1499.

O'Kelly, M. E., 1987, Spatial interaction based location-allocation models, *Spatial Analysis and Location-Allocation Models*, A. Ghosh and G. Rushton (Eds.), New York: van Nostrand Reinhold.

O'Kelly, M. E., 1992, A clustering approach to the planar hub location problem, *Annals of Operations Research*, 40, 339–53.

Openshaw, S., 1987, An automated geographical analysis system. *Environment and Planning A*, **19**, 431.

Openshaw, S., Charlton, M., Wymer, C. and Craft, A., 1987, A Mark I Geographical Analysis Machine for the automated analysis of point data sets, *Int. J. Geographical Information Systems*, 1(4), 335–358.

Openshaw, S., 1991, Commentary: A view on the GIS crisis in geography, *Environment and Planning*, 23, 621–628.

Parkes, D. and Thrift, N., 1980, Times, spaces, and places: a chronogeographic perspective. Chichester (Eng.); New York: John Wiley.

Rogers, A., 1974, Statistical analysis of spatial dispersion: the quadrat method. London: Pion; New York: Distributed by Academic Press.

Taaffe, E. J., 1956, Air transportation and United States urban distribution, *The Geographical Review*, 46(2), 219–238.

Upham, S., 1979 (Ed.), Computer graphics in archaeology: statistical cartographic applications to spatial analysis in archaeological contexts. Tempe: Arizona State University.

Villoria, O. G., 1989, An Operational Measure of Individual Accessibility for use in the Study of Travel-Activity Patterns, Unpublished Ph.D. Dissertation Department of Civil Engineering, Ohio State University.