2

The State of GIS for Environmental Problem-Solving

MICHAEL F. GOODCHILD

WHAT IS GIS?

The use of the term *Geographic* (or Geographical) *Information System* dates back to the mid-1960s, where it seems to have originated in two quite different contexts. In Canada, it was devised to refer to the use of a mainframe computer and associated peripherals (notably a scanner) to manage the mapped information being collected for the Canada Land Inventory, and to process it to compute estimates of the area of land available for certain types of uses. A rigorous analysis was used to show that a computer was the only cost-effective means of producing the vast numbers of measurements of area required by the project, even with the primitive and expensive nature of digital technology at the time, because manual measurement of area remains an inaccurate and labor-intensive task. Much of the proposed analysis was concerned with measuring areas simultaneously on two maps, to answer questions like "How much area is class 1 agricultural land and not currently used for agriculture?" The ability to overlay two or more maps for analysis (in this case a map of soil capability for agriculture with a map of land use) has always been a strong argument for GIS, because it is so cumbersome by hand.

Almost at the same time, researchers in the U.S. were struggling with the problems of accessing the many different types of data required by the large-scale transportation models then in vogue, and conceived of a GIS as a system capable of extracting appropriate data from large stores, making them available for analysis, and presenting the results in map form (Coppock and Rhind, 1991). Such models combined information on population distributions with other spatially distributed information on places of employment and transportation routes, and required access to data in a variety of formats.

Almost 30 years later, these same arguments are still among the most frequently heard justifications for the use of GIS, particularly in environmental modeling and policy development. GIS is seen as a general-purpose technology for handling geographic data in digital form, and satisfying the following specific needs, among others:

- The ability to preprocess data from large stores into a form suitable for analysis, including such operations as reformatting, change of projection, resampling, and generalization.
- Direct support for analysis and modeling, such that forms of analysis, calibration of models, forecasting, and prediction are all handled through instructions to the GIS.
- Postprocessing of results, including such operations as reformatting, tabulation, report generation, and mapping.

In all of these operations, the typical GIS user now expects to be able to define requirements and interact with the system through a "user-friendly," intuitive interface that makes use of such contemporary concepts as graphic icons and desktop metaphors (Mark and Gould, 1991).

GIS has evolved dramatically since the early days of mainframe computing, particularly in the past 12 to 15 years. Its first commercial successes came in the early 1980s, primarily in resource management, but more recently large markets for GIS software have developed in local government, utility companies, and a host of activities that use geographic data or manage geographically distributed facilities. It is estimated that the global GIS industry grosses over $1 billion annually (for general, popular reviews of the GIS industry see Bylinski, 1989; The Economist, 1992; for a comprehensive industry overview see GIS World, Inc., 1991; for a review of GIS as a whole see Maguire, Goodchild, and Rhind, 1991).

GIS applications now span a wide range, from sophisticated analysis and modeling of spatial data to simple inventory and management. Since the latter account for the lion's share of the commercial market for software, they also dictate the development directions of much of the industry. However, several vendors have chosen to concentrate on the niche market of environmental applications, and to emphasize support for environmental modeling, and GRASS is significant as a public-domain GIS developed by a branch of the military (the U.S. Army Corps of Engineers' Construction Engineering Research Laboratory) and having substantial capabilities for environmental modeling.

Growth has brought confusion, notably to the meaning of the term GIS itself. The 371 software products listed in the GIS World, Inc. (1991) survey include a vast range of capabilities, and run on platforms from the Macintosh to mainframes. Some products are offered on platforms as diverse as the PC and the IBM 3090, while others focus on Unix workstations. All handle geographic data in some form and provide capabilities for input and output. But there is currently no consensus on the minimum set of functional capabilities required to qualify as a GIS. This diversity is illustrated when one compares the relatively focused approach taken by Berry in Chapter 7 with the much broader perspective of Nyerges in Chapter 8.

Besides its collection of tools, GIS now also has a broadly based community of interest, drawn together by a common concern for the computerized handling of geographic data. It includes established disciplines like surveying, remote sensing, geodesy, and cartography, which see GIS as another valuable digital technology with capabilities that augment those of GPS (global positioning systems), image processing, digital cartography etc. In some senses GIS is the common ground between all of these, the broad technology that attempts to integrate data from a number of acquisition systems, and provide it to the user with appropriate analytic tools.

The GIS community also includes specialists in various application fields: local government officials, urban and regional planners, land records administrators, the oil and gas industry, and many others. It includes geographers, planners, resource managers, environmental modelers, geologists, epidemiologists, soil scientists, and representatives of the disciplines that work with geographical data.

The technology that supports this community is complex, but at the same time primitive in the eyes of most of its users. Geographical reality is enormously complex, and it can be represented in digital form in a rich variety of ways. Moreover, the set of GIS functions is long and growing, as uses are found for a greater and greater variety of forms of spatial analysis. Yet ideally, all of this complexity should be presented to the user in a friendly, intuitive manner. The human eye and mind are incredibly powerful processors of two-dimensional data, and compared to them even supercomputers sometimes appear impossibly clumsy. At the same time, the computer is much more efficient at primitive operations like the measurement of length and area and the combination of data from different sources.

This chapter presents an introduction to some of the principles of GIS and the issues surrounding its use. The later chapters on GIS in this book, Chapters 7–9, provide greater depth. Some of the argument in this chapter was published in a previous article (Goodchild, 1991).

PRINCIPLES OF GIS DATA MODELING

Standard models

Many geographical distributions, such as those of soil variables, are inherently complex, revealing more information at higher spatial resolution apparently without limit (Mandelbrot, 1982). Because a computer database is a finite, discrete store, it is necessary to sample, abstract, generalize or otherwise compress information. "Geographical data modeling" is the process of discretization that converts complex geographical reality into a finite number of database records or "objects." Objects have geographical expression as points, lines, or areas, and also possess descriptive attributes. For example, the process of sampling weather-related geographic variables such as atmospheric pressure at weather stations creates point objects and associated measured attributes.

GIS technology recognizes two distinct modeling problems, depending on the nature of the distributions being captured. When the distributions in reality are spatially continuous functions or "fields," such as atmospheric pressure or soil class, the database objects are creations of the data modeling process. The set of objects representing the variation of a single variable is termed a "layer", and the associated models are "layer models" or "field models." However, there are numerous instances where the database objects are defined a priori, rather than as part of the modeling process. The object "Lake Ontario" is meaningful in itself, and has an identity that is independent of any discretization of a binary water/land variable over North America. We refer to these as "object models." In a field model every location by definition has a single value of the relevant variable, whereas in the object model there would be no particular problem in allowing a location to be simultaneously occupied by more than one object. For example the "Bay of Quinte" is also in "Lake Ontario." The term "planar enforcement" is often used to reflect the fact that objects in a field model may not overlap; planar enforcement clearly is not relevant to the object models.

A major difficulty arises in the case of the object models when a well-defined object has no equally well-defined location. For example, the spatial extent of Lake Ontario would most likely be defined by some notion of average elevation, but this is not helpful in deciding when Lake Ontario ends and the St Lawrence River begins. Many geographical objects have inherently fuzzy spatial extents. One common solution to this problem is to allow objects value in a layer. Pipes can cross each other in object models, whereas this would cause problems in a field model. Most manmade facilities are well defined, so the problems of fuzziness noted earlier are likely not important. Another common use of object models is in capturing features from maps.

Object models are also commonly used to capture aspects of human experience. The concept "downtown" may be very important in building a database for vehicle routing or navigation, forcing the database designer to confront the issue of its representation as a geographical object. In an environmental context, McGranaghan (1989a,b) has shown the importance of this for handling the geographical referrents used in herbarium records.

Finally, object models can be conceptualized as the outcome of simple scientific categorization. The piecewise approximation inherent in field model (4) assigns locations to a set of discrete regions, in the geographical equivalent of the process of classification. In geomorphology, the first step in building an understanding of the processes that formed a given landscape is often the identification of "landforms" or "features," such as "cirque" or "drumlin." Band (1986), among others, has devised algorithms for detecting such objects from other data. Frank and Mark (1991) have discussed the importance of categories in the GIS context, and there is growing interest in understanding the process of object definition and its effects. For most purposes, environmental data modeling is dominated by the field view, and its concept of spatially continuous variables. But the object view is clearly important, particularly in interpreting and reasoning about geographical distributions.

Network models

Both field and object models have been presented here as models of two-dimensional variation. An important class of geographic information describes continuous variation over the one-dimensional space of a network embedded in two-dimensional space. For example, elevation, flow, width, and other parameters vary continuously over a river network, and are not well represented as homogeneous attributes of reaches. Models (1), (2), (4), (5), and (6) can all be implemented in one-dimensional versions, but none is supported in this form in any current GIS.

Choosing data models

In principle, the choice of data model should be driven by an understanding of the phenomenon itself. For example, a TIN model will be an appropriate choice for representing topography if the Earth's surface is accurately modeled by planar facets. Unfortunately other priorities also affect the choice of model. The process of data collection often imposes a discretization, the photographic image being a notable exception. The limitations of the database technology may impose a data model, as, for example, when a "raster" GIS is used and the choices are reduced to field models (2) and (4), or when a "vector" GIS is used and a cell grid must be represented as polygons. Finally convention can also be important, particularly in the use of certain data models to show geographic variation on maps. For example, digitized contours are used in spatial databases not because of any particular efficiency—in fact accuracy in a field sense is particularly poor—but because of convention in topographic map-making.

Relationships

A digital store populated by spatial objects—points, lines, and areas—would allow the user to display, edit, or move objects, much as a computer-aided design (CAD) system. However, spatial analysis relies heavily on interactions between objects, of three main forms:

- Relationships between simple objects, used to define more complex objects (e.g., the relationships between the points forming a line);
- Relationships between objects defined by their geometry (e.g., containment, adjacency, connectedness, proximity); and
- Other relationships used in modeling and analysis.

Examples of the third category of relationships not determined by geometry alone include "is upstream of" (connectedness would not be sufficient to establish direction of flow, and a sink and a spring may not be connected by any database object). In general, a variety of forms of interaction may exist between the objects in the database. In order to model these, it is important that the database implement the concept of an "object pair," a virtual object that may have no geographical expression but may nevertheless have attributes such as distance or volume of flow.

Recent trends in data modeling

Recently there has been much discussion in the GIS community over the value of "object orientation," a generic term for a set of concepts that have emerged from theoretical computer science (see, for example, Egenhofer and Frank, 1988a,b). Unfortunately the debate has been confused by the established usages of "object" in GIS, both in the sense of "spatial object" as a point, line, or area entity in a database, and also "object model" as defined here.

Three concepts seem particularly relevant. "Identity" refers to the notion that an object can possess identity that is largely independent of its instantaneous expression, with obvious relevance to the independence of object to have "multiple representations"—spatial extents that vary with scale. A river, for example, might be a single line at scales smaller than 1:50,000, but a double line at larger scales. Both geometric and topological expressions vary in this case as the object changes from line to area.

The field models

The purpose of field models is to represent the spatial variation of a single variable using a collection of discrete objects. A spatial database may contain many such fields or layers, each able in principle to return the value of one variable at any location (x, y) in response to a query, and fields may be associated with variables measured on either continuous or discrete scales. Because information is lost in modeling, the value returned may not agree with observation or with the result of a ground check; so accuracy is an important criterion in choosing between alternative data models. We define the accuracy of a field measured on a continuous scale as $E(z-z')^2$ where z is the true value of the variable, as determined by ground check, and z' is its estimated value returned from the database; for a field measured on a discrete scale, accuracy will likely be defined by the probability that the class recorded at a randomly chosen point is indeed the class at that point on the ground (further discussion of the measurement of accuracy can be found in Chapter 9). Note that the true value may be inherently uncertain because of definition or repeated measurement problems.

Six field models are in common use in GIS:

1. Irregular point sampling: the database contains a set of tuples $<x, y, z>$ representing sampled values of the variable at a finite set of irregularly spaced locations (e.g., weather station data).
2. Regular point sampling: as (1) but with points regularly arrayed, normally on a square or rectangular grid (e.g., a Digital Elevation Model).
3. Contours: the database contains a set of lines, each consisting of an ordered set of $<x, y>$ pairs, each line having an associated z value; the points in each set are assumed connected by straight lines (e.g., digitized contour data).
4. Polygons: the area is partitioned into a set of polygons, such that every location falls into exactly one polygon; each polygon has a value that is assumed to be that of the variable for all locations within the polygon; boundaries of polygons are described as ordered sets of x,y pairs (e.g., the soil map).
5. Cell grid: the area is partitioned into regular grid cells; the value attached to every cell is assumed to be the value of the variable for all locations within the cell (e.g., remotely sensed imagery).
6. Triangular net: the area is partitioned into irregular triangles; the value of the variable is specified at each triangle vertex, and assumed to vary linearly over the triangle (e.g., the Triangulated Irregular Network, or TIN, model of elevation) (Weibel and Heller, 1991).

Other possibilities, such as the triangular net (6) with nonlinear variation within triangles (Akima, 1978), have not received much attention in GIS to date.

Each of the six models can be visualized as generating a set of points, lines, or areas in the database. Models (2) and (5) are commonly called "raster" models, and (1), (3), (4), and (6) are "vector" models (Peuquet, 1984); storage structures for vector models must include coordinates, but in raster models locations can be implied by the sequence of objects. Models (3) and (6) are valid only for variables measured on continuous scales.

Models (4), (5), and (6) explicitly define the value of the variable at any location within the area covered. However, this is not true of models (1), (2), and (3), which must be supplemented by some method of spatial interpolation before they can be used to respond to a general query about the value of z at some arbitrary location. For example, this is commonly done in the case of continuous-scaled variables in model (2) by fitting a plane to a small 2x2 or 3x3 neighborhood. However, this need for a spatial interpolation procedure tends to confound attempts to generalize about the value of models (1), (2), and (3).

In practice, model (6) is reserved for elevation data, where its linear facets and breaks of slope along triangle edges fit well with many naturally occurring topographies (Mark, 1979). It would make little sense as a means of representing other variables, such as atmospheric pressure, since curvature is either zero (within triangles) or undefined (on triangle edges) in the model. Models (2) and (4) are frequently confused in practice, since the distinction between point samples and area averages is often unimportant. Models (1) and (3) are commonly encountered because of the use of point sampling in data collection and the abundance of maps showing contours, respectively, but are most often converted to models (2), (4), (5), or (6) for storage and analysis. The ability to convert between data models, using various algorithms, is a key requirement of GIS functionality.

The object models

Objects are modeled as points, lines, or areas, and many implementations make no distinction in the database between object and field models. Thus a set of lines may represent contours (field model) or roads (object model), both consisting of ordered sets of x,y pairs and associated attributes, although the implications of intersection, for example, are very different in the two cases.

Object models are commonly used to represent man-made facilities. An underground pipe, for example, is more naturally represented as a linear object than as a identity and geographic expression in GIS. "Encapsulation" refers to the notion that the operations that are possible on an object should be packaged with the object itself in the database, rather than stored or implemented independently. Finally, "inheritance" refers to the notion that an object can inherit properties of its parents, or perhaps its component parts. As a geographical example, the object "airport" should have access to its component objects—runway, hangar, terminal—each of which is a spatial object in its own right.

Of the three concepts, inheritance seems the most clearly relevant, particularly in the context of complex objects, and in tracking the lineage of empirical data. It seems increasingly important in the litigious environment that surrounds many GIS applications to track the origins and quality of every data item.

Encapsulation seems to present the greatest problems for modeling using GIS. In a modeling context, the operations that are permissible on an object are defined by the model, and are therefore not necessarily treatable as independent attributes of the object. This issue seems particularly important in the context of the discussion of object orientation in location/allocation modeling by Armstrong, Densham, and Bennett (1989). For example, one can rewrite the shortest path problem by treating each node in the network as a local processor, making it possible to encapsulate the operations of a node with the object itself. It is possible that this process of rewriting may lead to useful insights in other models as well.

A related debate is that over procedural and declarative languages: A user should be able to declare "what" is required (declarative), and not have to specify "how" it should be done (procedural). But are these largely distinguishable in a modeling context, and do they imply that the modeler should somehow surrender control of the modeling process to the programmer?

The role of data models in environmental analysis and modeling is clearly complex. Models written in continuous space, using differential equations, are independent of discretization. But for all practical purposes modeling requires the use of one or more of the data models described here. Perhaps the greatest advantage of GIS is its ability to handle multiple models, and to convert data between them.

GIS AND ENVIRONMENTAL PROBLEM-SOLVING

One of the strongest and most successful application areas for GIS has been in addressing problems of the environment. GIS were acquired by many forest resource companies and regulatory agencies in the early 1980s, and subsequently by other environmental agencies such as EPA (U.S. Environmental Protection Agency), the National Parks Service, and the Bureau of Land Management. Virtually all North American resource management agencies now have some form of GIS program and associated policies. However, this pattern gives no impression of the diversity of GIS applications within agencies. The following subsections provide a categorization of different types of applications within the broad area of environmental problem-solving.

Mapping

In principle, it is possible to make a clear distinction between GIS and digital cartography. The latter deals with map features and with associated attributes of color, symbology, name or annotation; provides capabilities for input, editing, and output; deals with such cartographic features as legends, neatlines, and north arrows; and includes algorithms for projection and scale change. GIS, on the other hand, frequently allows geographic entities to have multiple attributes, frequently includes capabilities for storing and handling relationships between entities, and includes the capabilities of digital cartography in its input and output subsystems.

Of particular importance to this discussion is the manner in which digital cartography and GIS handle notions of continuous variation. Consider the elevation of the Earth's surface as an example of a variable conceived as varying continuously across geographic space and having a unique value everywhere in that space. Cartographically, such continuous variation must be represented in the form of map features, typically contour lines and irregularly spaced spot heights, and digital cartography provides capabilities for processing these as line and point objects. From a GIS perspective, however, numerous other methods exist for creating a digital representation of continuous variation. The objects used in that representation should be viewed as artifacts, and hidden from the user so as not to obscure the latter's concept of a continuum.

In practice, however, many applications of GIS technology turn out to be little more than digital cartography. The map is a very persuasive form of data display, and a computer-drawn map carries the authority of a powerful technology. Many GIS contain user-friendly map editing and formatting subsystems, and allow the user with little cartographic training to make a convincing product quickly and efficiently. But cartographers have accumulated centuries of knowledge and experience about the effective visual communication of geographic information, and GIS may be doing its users a disservice if it encourages them to ignore this (Buttenfield and Mackaness, 1991).

Data preprocessing

Because there are so many ways of representing geographic variation, and such diversity in data structures between GIS products and databases, most available GIS include substantial facilities for input and output of data in different formats. Some of these are sanctioned by official standards: the major U.S. federal effort at standardization has produced a Spatial Data Transfer Standard (SDTS) that has been accepted (Federal Information Processing Standard (FIPS) 173 (DCDSTF, 1988)).

Other useful preprocessing functions include the ability to extract information in a user-defined window; scale and projection change, including knowledge of such coordinate systems as UTM (Universal Transverse Mercator) and SPC (State Plane Coordinates); and capabilities for resampling. All of these help to give GIS an important role as a manager of data, particularly if data must come from diverse sources in mutually inconsistent formats.

The chapters of this book provide strong support for this argument for GIS as a data integrator. Models for specific environmental processes, such as non-point-source pollution, have often been developed without reference to GIS, and without the ability to interface with GIS software. But models that integrate more than one process, or attempt to provide links to policy development, such as those described in Chapters 26–34 of this book, are much more likely to rely on GIS for preprocessing of data. Moreover, it is clear that in areas like global climate modeling, the need to integrate data on a variety of themes, such as soil or land cover, is leading to increasing concern with GIS (see Chapters 10–13).

GIS in modeling

Many applications of GIS in environmental modeling have followed the scheme outlined in the previous sections, where a GIS is used to preprocess data, or to make maps of input data or model results. We refer to this mode as "loose coupling," implying that the GIS and modeling software are coupled sufficiently to allow the transfer of data, and perhaps also of results in the reverse direction. In many instance loose coupling requires the development of a linking module of specialized code, if the GIS is not capable of providing data in the format(s) required by the modeling software.

Closer forms of coupling are possible if the GIS and modeling module share the same data structures, obviating the need for a linking module and allowing both systems to interact with the same database. Nyerges discusses issues of coupling in more detail in Chapter 8. At the highest level of sophistication, models are calibrated and run directly in the GIS, using the GIS command language. Berry describes the concept of cartographic modeling in Chapter 7, and illustrates the use of a GIS to perform complex analysis of layered data. However, the current generation of GIS command languages falls well short of satisfying the requirements of this style of environmental modeling. To support modeling directly in the GIS's command language, it would be necessary for the GIS's data model to match that of the environmental model. For example, if an atmospheric model requires that space be represented as a set of square finite elements, it would be necessary for the GIS database also to represent spatial variation in terms of these same square finite elements. In other words, the GIS data model must match the needs of environmental modeling.

In practice, GIS data models tend to have more in common with maps than with the finite elements of environmental models. GIS databases have often been constructed from mapped information, and maps are a common form of output. Thus elevation is often represented by digitized contours, since this is the preferred method of map representation, as noted earlier. Soil or land cover maps are often represented as collections of digitized polygons. The space of environmental models is often conceived as continuous, and discretized as finite elements only for the purpose of numerical analysis. In GIS, however, the finite elements tend to be the basis of conceptualization, and frequently it is the cartographic model that dominates the choice of finite elements. Based on these arguments, the following are suggested as the minimum requirements for an environmental modeling language interface to GIS:

- In the first instance, the user should be able to work with symbolic representations of continuous geographic variation (e.g., T and h as representations of the continuous variation of temperature and elevation, respectively).
- The language should allow such continuous variations (fields) to be combined symbolically (e.g., $T = 20 - h/100$ to compute ground surface temperature by approximating the adiabatic lapse rate).
- As far as possible, the language should hide issues of discrete digital representation from the user. This is straightforward when T and h are both evaluated for the same set of sample points, as in the examples and language presented by Berry in Chapter 7. Where operations change the discrete representation, the user should provide the necessary information. For example, $T(100) = T$ might resample T to a 100 m sampling interval, using a default method of resampling, and assuming that the sampling interval of T was much smaller than 100 m.
- The language should include all of the common primitives of environmental modeling. For example, although both discrete and continuous scalar fields are handled, in current GIS products there is no support for the concept of a vector field. Related operators, such as grad or the dot product, are not supported, and neither are simple methods of display.

GIS in policy

Some of the greatest interest in the use of GIS for environmental problem-solving has come from those who would apply the technology to translate the results of environmental modeling into policy. Postprocessing is essential if the results of a spatially distributed model are to be used for policy development. Results must often be aggregated by administrative unit, or brought into consistency with social and economic data for comparison and correlation. Displays must be developed to present the results of modeling in convincing form. Finally, increasing use is being made of the paradigm of spatial decision support, in which the technology is made available directly to decision-makers for scenario development, rather than being confined to use by analysts.

As the chapters of this book progress from models to policy, the need for a technology that can cross the gap between rigorous science and responsible policy formulation will become increasingly clear. We spend vast sums on collecting raw geographic information with technologies such as remote sensing, and on modeling environmental processes, and yet it often seems that the biggest problem of all is the translation of this knowledge into useful and effective policy.

FUNCTIONALITY FOR ENVIRONMENTAL ANALYSIS

The statistical packages such as S, SAS, or SPSS are integrated software systems for performing a wide variety of forms of analysis on data. By analogy, we might expect GIS to integrate all reasonable forms of spatial analysis. However, this has not yet happened, for several reasons. First, while the analogy between the two systems may be valid, there are important differences. The statistical packages support only one basic data model—the table—with one class of records, whereas GIS must support a variety of models with many classes of objects and relationships between them. Much of the functionality of GIS must therefore be devoted to supporting basic housekeeping and transformation functions that would be trivial in the statistical packages.

Second, spatial databases tend to be large, and difficult and expensive to create. This is particularly true of imagery, where a single analysis might require gigabytes (10^9 bytes) or even terabytes (10^{12} bytes) of data, and of representations of geographical variation in three dimensions. While many users of statistical packages input data directly from the keyboard, it is virtually impossible to do anything useful with a GIS without devoting major effort to database construction. Recently there have been significant improvements in this situation, with the development of improved scanner and editing technology.

Third, while there is a strong consensus on the basic elements of statistical analysis, the same is not true of spatial analysis. The literature contains an enormous range of techniques (for examples, see Serra, 1982; Unwin, 1981; Upton and Fingleton, 1985), few of which could be regarded as standard.

Because of these issues and the diversity of data models, GIS has developed as a loose consortium, with little standardization. While ESRI's ARC/INFO and TYDAC's SPANS are among the most developed of the analytically oriented packages, they represent very different approaches and architectures. Among the most essential features to support environmental modeling are:

- Support for efficient methods of data input, including importing from other digital systems;
- Support for alternative data models, particularly models of continuous spatial variation, and conversions between them using effective methods of spatial interpolation;
- Ability to compute relationships between objects based on geometry (e.g., intersection, inclusion, adjacency), and to handle attributes of pairs of objects;
- Ability to carry out a range of standard geometric operations (e.g., calculate area, perimeter length);
- Ability to generate new objects on request, including objects created by simple geometric rules from existing objects (e.g., Voronoi polygons from points, buffer zones from lines);
- Ability to assign new attributes to objects based on existing attributes and complex arithmetic and logical rules;
- Support for transfer of data to and from analytic and modeling packages (e.g., statistical packages, simulation packages).

Because of the enormous range of possible forms of spatial analysis, it is clearly absurd to conceive of a GIS as a system to integrate all techniques, in contrast to the statistical packages. The last requirement listed proposes that GIS should handle only the basic data input, transformation, management, and manipulation functions, leaving more specific and complex modeling to loosely coupled packages. Whereas the statistical packages are viable because they present all statistical techniques in one consistent, readily accessible format, GIS is viable for environmental modeling because it provides the underlying support for handling geographical data, and the "hooks" needed to move data to and from modeling packages, at least until command languages can be developed that are more suitable for environmental modeling.

CONCLUSIONS

GIS is a rapidly developing technology for handling, analyzing, and modeling geographic information. To those sciences that deal with geographic information it offers an integrated approach to data handling problems, which are often severe. As with all GIS applications, the needs of environmental modeling are best handled not by integrating all forms of geographic analysis into one GIS package, but by providing appropriate linkages and hooks to allow software components to act in a federation.

Models that lack a spatial component clearly have no use for GIS, but increasing concern for geographically distributed models of the environment, particularly as inputs to environmental policy, has led to increasing interest in GIS in the past decade in the environmental modeling community, and this book describes much of that interest. The ability of GIS to integrate spatial data from different sources, with different formats, structures, projections, or levels of resolution is a powerful aid to spatially distributed models, particularly those models that integrate more than one process. Finally, the need to integrate environmental information with administrative, political, social, and economic data in developing environmental policy and regulating the use of the environment provides another powerful incentive to remove some of the impediments to greater integration of GIS and environmental modeling.

This chapter has provided a brief overview of GIS and some of the issues involved in its use in environmental problem-solving. Many of its themes are explored in greater detail later in this book in Chapters 7–9.

REFERENCES

Armstrong, M.P., Densham, P.J., and Bennett, D.A. (1989) Object oriented locational analysis. *Proceedings, GIS/LIS '89*, Bethesda, MD: ASPRS/ACSM, pp. 717–726.

Band, L.E. (1986) Topographic partition of watersheds with digital elevation data. *Water Resources Research* 22(1): 15–24.

Buttenfield, B.P., and Mackaness, W.A. (1991) Visualization. In Maguire, D.J., Goodchild, M.F., and Rhind, D.W. (eds.) *Geographical Information Systems: Principles and Applications* (2 volumes), London: Longman, and New York: Wiley, pp. 427–443.

Bylinski, G. (1989) Managing with electronic maps. *Fortune* (April 24): 237.

Coppock, J.T., and Rhind, D.W. (1991) The history of GIS. In Maguire, D.J., Goodchild, M.F., and Rhind, D.W. (eds.) *Geographical Information Systems: Principles and Applications* (2 volumes), London: Longman, and New York: Wiley, pp. 21–43.

DCDSTF (Digital Cartographic Data Standards Task Force) (1988) The proposed standard for digital cartographic data. *The American Cartographer* 15(1): 9–140.

Economist, The (1992) The delight of digital maps. *The Economist* (March 21): 69–70.

Egenhofer, M.J., and Frank, A.U. (1988a) Object-oriented databases: database requirements for GIS. *Proceedings, IGIS: The Research Agenda*, Washington DC: NASA, Vol. 2, pp. 189–211.

Egenhofer, M.J., and Frank, A.U. (1988b) Designing object-oriented query languages for GIS: human interface aspects. *Proceedings, Third International Symposium on Spatial Data Handling*, Columbus, OH: International Geographical Union, 79–96.

Frank, A.U., and Mark, D.M. (1991) Language issues for GIS. In Maguire, D.J., Goodchild, M.F., and Rhind, D.W. (eds.) *Geographical Information Systems: Principles and Applications* (2 volumes), London: Longman, and New York: Wiley, pp. 147–163.

GIS World, Inc. (1991) *1991–92 International GIS Sourcebook*, Fort Collins, CO: GIS World, Inc.

Goodchild, M.F. (1991) Integrating GIS and environmental modeling at global scales. *Proceedings, GIS/LIS '91*, Bethesda, MD: ASPRS/ACSM, Vol. 1, pp. 117–127.

Maguire, D.J., Goodchild, M.F., and Rhind, D.W. (eds.) (1991) *Geographical Information Systems: Principles and Applications* (2 volumes), London: Longman, and New York: Wiley.

Mandelbrot, B.B. (1982) *The Fractal Geometry of Nature*, San Francisco: Freeman.

Mark, D.M. (1979) Phenomenon-based data structuring and digital terrain modeling. *GeoProcessing* 1: 27–36.

Mark, D.M., and Gould, M.D. (1991) Interacting with geographic information: a commentary. *Photogrammetric Engineering and Remote Sensing* 57(11): 1427–1430.

McGranaghan, M. (1989a) Incorporating bio-localities in a GIS. *Proceedings, GIS/LIS '89*, Bethesda, MD: ASPRS/ACSM, pp. 814–823.

McGranaghan, M. (1989b) Context-free recursive-descent parsing of location-descriptive text. *Proceedings, AutoCarto 9*, Bethesda, MD: ASPRS/ACSM, pp. 580–587.

Peuquet, D.J. (1984) A conceptual framework and comparison of spatial data models. *Cartographica* 21(4): 66–113.

Serra, J.P. (1982) *Image Analysis and Mathematical Morphology*, New York: Academic Press.

Unwin, D.J. (1981) *Introductory Spatial Analysis*, New York: Methuen.

Upton, G.J.G., and Fingleton, B. (1985) *Spatial Data Analysis by Example*, New York: Wiley.

Weibel, R., and Heller, M. (1991) Digital terrain modelling. In Maguire, D.J., Goodchild, M.F., and Rhind, D.W. (eds.) *Geographical Information Systems: Principles and Applications* (2 volumes), London: Longman, and New York: Wiley, pp. 269–297.