Integrating Dynamic Modeling and Geographic Information Systems

N.W.J. Hazelton, F.J. Leahy and I.P. Williamson

Bill Hazelton has recently completed a Ph.D. in Temporal (4-D) GIS in the Department of Surveying and Land Information, University of Melbourne, Australia. He holds a B.Surv. from the same department and has worked in surveying and mapping. He has wintered in Antarctica as a glaciologist, travelling extensively on over-snow traverses within the Australian Antarctic Territory.

Frank Leahy is a senior lecturer in the Department of Surveying and Land Information, University of Melbourne. He has published and consulted extensively in geodetic surveying, mapping, LIS/GIS, satellite geodesy and related areas.

Professor Ian Williamson holds the chair in the Department of Surveying and Land Information at University of Melbourne. He has published and consulted extensively in the LIS/GIS area.

All three authors are licensed surveyors and members of the Institution of Surveyors, Australia.

Abstract: A major limitation to the development of GIS is that most systems model reality in the same way that a paper map models reality. Until GIS can develop into direct 'reality modelling' systems, rather than 'map modelling' systems, they will not be able to handle 3-D spatial data, nor temporally-referenced data.

A GIS is generally used for some level of decision support. Decision-making relies on the analysis of various options, which are based upon the likely outcomes of different management strategies in the enterprise under consideration. Some form of dynamic modeling is required to simulate future scenarios.

Dynamic modeling cannot be linked with a GIS at present because of incompatibilities in the conceptual frameworks of the two systems. A multi-temporal 4-D GIS would allow the integration of dynamic modeling and a GIS. In this paper, it is argued that the development of a true 4-D GIS is one of the "Grand Challenges" facing GIS in the near future.

In late 1991, David Mark of the State University of New York at Buffalo (an NCGIA Centre) asked whether there were any "Grand Challenges in GIS." Subsequently, the authors of this paper have identified three GIS challenges:

1) That GIS must develop beyond a method for representing maps in a computer, toward a means of representing reality directly as it can be perceived by humans, rather than through an intervening medium (such as a map representation);

2) That as a consequence of (1), GIS must become able to handle 3-D spatially referenced data, temporally referenced data, multiple interpretations of the same reality, and different views of reality dependent upon time;

3) That also as a consequence of (1), GIS must be able to provide decision-support services by such means as predictive and simulation modeling, spatio-temporal analysis, and support for multiple simulated futures.

These three challenges will not be everyone's choice as there are many worthy "Grand Challenges" for GIS. In this paper, it is argued that at least one of them is to develop a four-dimensional GIS. This paper will discuss why a 4-D GIS is a necessary development for current GIS.
The third challenge mentions GIS in providing decision-support services. This application of GIS focuses attention on the roles of modeling and decision-making in GIS use. Cowan's (1988, p. 1554) definition of GIS as a "decision-support system involving the integration of spatially referenced data in a problem-solving environment" emphasizes the use and purpose of GIS, rather than just the technology used to build a system—what GIS should be, rather than the nuts and bolts of its construction.

In this paper, we will consider that a 'whole' GIS consists of five components: hardware, software, data, people and procedures (Dangermond 1988; Turk 1988, 1990, 1992).

Decision Support and GIS

What is involved in making decisions? The essence of decision-making seems to lie in making informed choices among a number of options. 'Hobson's choice' requires no decision-making. Decision-making may be improved by increasing the number of options in the hope of getting better ones, improving the information about the choices, or improving the mental processes involved in making the choice.

Increasing the number of options involves being able to model the consequences of as many decisions as possible; this is a task for some type of predictive model. Improving the information about the various choices involves analysis of the options and comparison between them; this is a task for analytical tools. Improving the mental process is beyond the scope of this paper. Turk (1990, 1992) discusses aspects of this issue.

An example of a simple computer-based decision-support system is a spreadsheet. The user can alter many variables in the models of future options, compare and analyze the results of these models. The user can derive a larger number of models and better analyze them, than by using manual techniques. In analyzing the state of a business enterprise, for example, many different models of interest, tax and depreciation rates may be used to produce a broad selection of options for decision-making.

At the simplest level, a spatial decision-support system (SDSS) is expected to generate and analyze multiple outcomes of possible management options. This requires some form of predictive model, a model in which time is a parameter; such a model is referred to as a dynamic model.

A SDSS requires not only a GIS to manage the data, but the ability to operate with time as a parameter, handle dynamic models and their output, and manage spatio-temporal analysis of data. This requires integrating dynamic models with GIS.

Modeling and GIS

In the broadest sense, modeling is an integral part of GIS, but often seems to be treated as though it were quite a separate process. A GIS attempts to create a model of the 'real world' or, at least, certain aspects of it. In this section, a brief overview of current GIS modeling concepts is presented in order to provide an understanding of how GIS and modelling fit together.

Modeling is a very general term. In its most general sense, modeling is the creation of an abstract process that attempts to emulate another, often more complex, process. The mathematical model "1+1=2" is an abstraction process that describes how people can count two of any type of object. The mathematician Hamming has stated that mathematics is a remarkably good abstraction of reality. He notes (1980, p. 81) that:

I have tried, with little success, to get some of my friends to understand my amazement that the abstraction of integers for counting is both possible and useful. Is it not remarkable that 6 sheep plus 7 sheep make 13 sheep; that 6 stones plus 7 stones make 13 stones? Is it not a miracle that the universe is so constructed that such a simple abstraction as a number is possible? To me this is one of the strongest arguments of the unreasonable effectiveness of mathematics. Indeed, I find it both strange and unexplainable.

Models have several constraints, among them: they should be operationally simpler than the process that they emulate; they should be quicker than the original process in most cases; and you should be able to use them without affecting the state of the original process.

Modeling is a very broad concept and people have differ-
ent ideas of what is meant by the term, so one can begin by categorizing modeling. These categories are by no means exhaustive, nor are they mutually exclusive. However, they include most GIS modeling concerns and allow a look at some general characteristics that are associated with specific types of models.

Within a ‘whole’ GIS, modeling may be generalized into four categories:

1) Reality Modeling: The means by which a GIS stores a model of ‘reality’. Different GIS use different representations of ‘reality’, e.g., grid-cell structures, discrete irregular polygons, and 2-D and ‘2.5-D’ representations of the surface of the Earth. This is primarily a function of GIS software.

2) System Modeling: How a GIS as a whole will work, its relationship to the rest of an organization, data requirements, outputs, efficiency.


4) Dynamic Modeling: These models utilize time, as well as some form of deterministic, stochastic or chaotic model, to generate future scenarios.

Reality Modeling

Reality modeling is the model that the GIS ‘has’ (or was designed with) of the ‘real world’. Unfortunately, current GIS tend to be map-representation systems, rather than reality-representation systems. This can severely limit their view of the world and the types of reality models that are possible within GIS. Many of the reality models in current GIS are very close to paper maps.

System Modeling

This type is concerned with modeling the GIS, rather than working with data within the GIS. The types of things modeled include: the operations that the system will be required to perform; the data required for various operations; the usage patterns of the hardware and software; the communications within the ‘whole’ GIS (e.g., how the GIS meshes in with the rest of the organization); and the types of GIS specialists that are required to operate the system. A discussion of some aspects of LIS/GIS system modeling can be found in Williamson and Hunter (1991).

As will be realized, most of these models will tend to be conceptual models of the system and will, or should, be undertaken before the purchase of a GIS is considered. As this is not the main area of interest in this paper, we will only note its critical importance to the overall efficiency and effectiveness of a ‘whole’ GIS.

Empirical Modeling

‘Empirical’ means resting upon trial or experiment, or known only by experience. In the GIS case, empirical modeling is that derived from experience of the real world, rather than from purely theoretical considerations or a priori reasoning. Burrough et al. (1988) divide modeling into the two general categories of ‘empirical’ and ‘process’ (called ‘dynamic’ in this paper).

Burrough et al. (1988) further divide empirical modeling into two general sub-categories. These are threshold models and regression models. There are other possible classification systems, but this is a more useful taxonomy for GIS.

Threshold models use the boundary values of specific attributes in the GIS to determine output values. The actual values of these boundary points are determined by experience. For example, it may be decided from experience that land suitable for residential development has the following attributes: slope < 7 percent; soil fertility less than 6 (on some previously decided scale); is not currently designated as prone to flooding; is within 500 meters of a major road; is within 800 meters of a public transport connection; and is within 1,000 meters of water, gas, electricity and sewerage connection points. These threshold values are then used to develop a model to determine suitable land from the data in the GIS. One may go further and decide to provide ‘gradations’ of suitability for each attribute. One may employ ‘fuzzy’ set theory to reduce the loss of information that tends to occur with the strict application of purely Boolean algebra.

Regression models use the results of experimentally observed regression analyses to decide the control parameter values. In this case, the resulting values are some function of the weighted values of the input values. For example, a regres-
sion model for residential suitability of land may be:
\[ S = \text{Flooding}/(2 \times \text{slope} + 3 \times \text{fertility} + 3 \times D_{\text{road}} + 5 \times D_{\text{public tran}} + 4 \times D_{\text{service}}) \]

This is in very approximate correspondence to the example given for threshold models. GIS are well suited to the use of this type of model, as well as being able to generate the data for the initial regression analysis.

Empirical models comprise practically all models used with current GIS and, effectively, queries proceed based on some form of empirical model. Further, current GIS are designed to support this type of modeling at a very general level.

**Dynamic Modeling**

Dynamic models include various types of models referred to as process models, simulations, predictive models, etc. These models generate some form of prediction of the future behaviour of a system. The prediction may be a series of future states, the expected probability of future states of the system, or some other descriptive forms. In most cases, dynamic modeling uses time as one of its arguments.

Dynamic models for GIS would tend to generate future scenarios from existing ones, and a set of rules for prediction, which determine the type of dynamic model used. Burrough et al. (1988) describe two general types of dynamic models: 'deterministic' and 'stochastic'. Hazleton (1991b) adds 'chaotic models' to these.

A deterministic dynamic model represents some observed relationship in terms of various processes in the system under consideration. The resulting values are generally determined by solving differential equations describing the process. The basic concept of a deterministic model is that it attempts to model the real world from fundamental physical or chemical laws, i.e., from some theoretical basis, rather than from an empirical basis. Deterministic models often build up a model of a system from many smaller models, rather like stones in a dry stone wall.

A stochastic dynamic model is used to model the average behaviour of a large number of events, based on probability theory. If a stochastic process, such as a Markov chain, is used to model the system's behaviour, then the model predicts not an exact result, but one probable realization of many possible outcomes. The model also should provide some statistical estimate of the actual likelihood of this specific outcome.

The traditional examples of the different behaviours of a physical system are the clock and the coin-flip. "Utter regularity and utter randomness are the dynamical legacy of two millennia of physical thought." (Crutchfield and Young 1990). However, time-dependent behaviour necessarily incorporates elements of both these processes. In ergodic theory, the formal model of a random process is the Bernoulli flow, \( B \). A completely determined Laplacian system, such as a perfect clock, can be denoted as \( P \), where the completely predictable pattern is repeated every \( t \) seconds.

\( B \) and \( P \) can be considered as the basic processes with which to model the complexity of non-linear systems. If some system is found to exhibit some repeating characteristics in the observed data, the observations can be described as having been produced by some variant of \( P \). If they are completely unpredictable, their generating process is essentially the same as \( B \). Any real system \( S \) will contain elements of both processes. We can ask, "Is it always true that the case of some observed behaviour can be decomposed into these two separate components?" Is \( S = B \otimes P \)?

Ergodic and probability theories state that in general this cannot be done so simply. It has been shown that "there are ergodic systems that cannot be separated into completely random and completely predictable processes (Ornstein 1989). The Wold-Kolmogorov spectral decomposition states that although the frequency spectrum of a stationary process consists of a singular spectral component associated with periodic and almost periodic behaviour and a broadband continuous component associated with an absolutely continuous measure, there remains other statistical elements beyond these (Kolmogorov 1977a, 1977b; Wold 1954; Crutchfield and Young 1990, p. 224) These statistical elements are generally described as 'chaotic'.

As with building a dry stone wall, finding a suitable model for part of a larger modeling system is a matter of searching the pile of available stones until
a stone of the appropriate size and shape is found. Using the wrong stone distorts and weakens the wall and may cause it to collapse. Bulldozing stones together does not form an effective wall, nor does careless placement of the stones. Modeling building, like building stone walls, is as much art as science; an acquired skill that requires patience and knowledge.

Dynamic modeling is a rather more uncertain area than empirical modeling, and should not be approached casually. The need for much more data than are required for empirical modeling places pressure on the most expensive area of GIS operation—data collection—and introduces the need to consider surrogate data sets.

In summary, thus far:

- Modeling is the process of producing an abstraction of the ‘real world’ so that some part of it can be more easily handled. Most models are designed to represent a simplified version of some natural process or situation. Modeling and GIS are inseparable, GIS itself being a means of modeling reality. Unfortunately, most current GIS model paper maps, rather than directly modeling reality.

- Modeling is an intrinsic part of GIS, both within the ‘software’ component, and as a part of the ‘procedures’ component of a ‘whole’ GIS. System modeling is an important part of the conceptual development of a ‘whole’ GIS, and is used to help design the GIS.

- Empirical modeling is the imposition of certain empirical models from the ‘procedures’ component of a ‘whole’ GIS upon the ‘data’ component; this is used to derive new information from existing data. All current GIS allow this type of modeling, and it is the basis of most queries of current GIS.

- Dynamic modeling is the extension of empirical modeling to include the time domain. In this type of modeling, time is an essential parameter of the model. Few current GIS can handle temporally referenced data in any quantity.

### Dynamic Modeling

Dynamic models are a subset of mathematical models, which are in turn a subset of symbolic and procedural models. Dynamic models can be used as a generic term to describe all models that use time as a parameter. Table 1 shows the major classifications of mathematical models.

The most important classifications of models can be considered as contrasts between what are almost diametrically opposed design factors: distributed vs. lumped, dynamic vs. static, continuous vs. discrete, and deterministic vs. stochastic vs. chaotic. These factors affect the manner in which the model is designed and built, the types of problems that it can handle and its implementation on a computer. It also influences how the model can be tied into a GIS.

Of particular interest to GIS are distributed models, which use spatial location as a parameter in the model, although the discrete nature of data representation in a GIS means that truly distributed models cannot directly use GIS data. GIS can be used to prepare data for use by lumped parameter models, although this, in some senses, degrades the use of the GIS. Current GISs are probably best suited to ‘distributed-lumped’ models that allow lumped models to be used for small regions over the entire GIS data structure. Table 2 shows the internal requirement for lumped and distributed models.

The ‘distributed-lumped’ modeling approach is sound, provided that the models are supplied with data that describe the spatial variation of the GIS attribute data as accurately as the model can describe the process it is modeling. If this is not so, then the model’s results will be unreliable. The costs and benefits of quantified resource assessment and improved decision support need to be considered carefully, to ascertain

<table>
<thead>
<tr>
<th>TABLE 1 Major Classification Divisions of Mathematical Models (from Kecman, 1988, p. 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>According to</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>A Mutual dependence of variables</td>
</tr>
<tr>
<td>B Memory length changes with time</td>
</tr>
<tr>
<td>C Dependence on spatial co-ordinate</td>
</tr>
<tr>
<td>D Processing of material or energy</td>
</tr>
<tr>
<td>E Randomness of variables</td>
</tr>
<tr>
<td>F Time as an independent variable</td>
</tr>
</tbody>
</table>
TABLE 2.
Internal Requirements for Lumped and Distributed Models. The horizontal 'axis' of the table can be considered to indicate the degree of spatial representation in the model, while the vertical 'axis' of the table represents the degree of temporal representation in the model. (ODE = Ordinary Differential Equations; PDE = Partial Differential Equations). (From Kecman 1988, p. 12).

<table>
<thead>
<tr>
<th>Lumped Parameters ≠ f(x, y, z)</th>
<th>Distributed Parameters = f(x, y, z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>algebraic equation</td>
<td>one-dimensional ODEs ≠ f(t)</td>
</tr>
<tr>
<td>representations:</td>
<td>multidimensional elliptic PDEs</td>
</tr>
<tr>
<td>• tables</td>
<td></td>
</tr>
<tr>
<td>• nomograms</td>
<td></td>
</tr>
<tr>
<td>• characteristics</td>
<td></td>
</tr>
<tr>
<td>Static</td>
<td></td>
</tr>
<tr>
<td>• system of 1st order ODEs</td>
<td>hyperbolic PDEs = f(t)</td>
</tr>
<tr>
<td>Dynamic</td>
<td></td>
</tr>
<tr>
<td>• ODE of n-th order</td>
<td>parabolic PDEs = f(t)</td>
</tr>
<tr>
<td>• several ODEs of varying order</td>
<td></td>
</tr>
</tbody>
</table>

where and when the additional investment in data and computing resources will be justified.

Dynamic models can be deterministic, stochastic or chaotic in nature. Deterministic models are designed to produce 'exact' results, while stochastic models have their parameters determined by probability distributions. Chaotic models arise when situations of considerable complexity are modeled. The chaotic factor is not designed into the model, but appears by itself as a function of model complexity.

The complexity of a model is a function of the complexity of the system that it is trying to model. Because highly complex systems are not isomorphic to a purely deterministic model, a purely stochastic model, or any combination of the two (Crutchfield and Young 1990), it is necessary to consider the additional effects on the model caused by the complexity of the system. Complexity theory can be used to model some of these effects, and much of this modeling is complementary to the modeling of these effects using information theory.

Why Dynamic Modeling is Needed in GIS

"The amount of time spent by generalists in making technically based decisions is in inverse proportion to the complexity of the subject matter." Jones’ Seventh Law, (Jones 1990, p. 176).

The quotation illustrates the problems of decision-making in an increasingly complex society. Specialists gather data in their respective fields, but because of the depth of knowledge required in many specialist areas, they often do not have a broad view beyond their own discipline. Generalists, on the other hand, do not have the depth of knowledge of the specialists, but strive to obtain an understanding of the whole situation. This is becoming increasingly difficult. As the quotation suggests, if the matter is highly technical, the generalists spend little time on the technical side, effectively leaving the decision to the specialists. This situation is caused by poor, or non-existent, decision support and communication between specialists and generalists.

Rhind (1988) has pointed out, "virtually all GIS developments thus far have resulted in 'data retrieval and sifting' engines; modeling work has not yet been brought together with this technically accomplished sub-culture". Densham and Goodchild (1989) have discussed the use of GIS as a decision-support tool. They note that current GIS have limited analytical capabilities and do not support decision-making adequately. Densham and Goodchild (1989) suggest that important areas of consideration for a GIS-based decision-support system include analytical modeling and expert knowledge.

Goodchild (1987) has suggested that the ability of GIS to adopt an extensive set of spatial analysis tools is limited by the inadequacies of the data structure, and the potential for development in output from GIS is limited by the query modes available to GIS, which are in turn limited by the cartographic model underlying the design of GIS in general. Goodchild (1991) also discussed the means of linking GIS and spatial data analysis concluding that the diversity of GIS data structures and the "unbounded nature of spatial data analysis" meant that development of a common
interface was unlikely. However, he felt that a close coupling of GIS and spatial data analysis was the best method of improving the current low level of GIS functionality in spatial analysis, together with making the discrete nature of GIS data structures explicit in the development of spatial analysis tools.

It has been recognized that a major intention of GIS is to act as a decision-support tool (Cowan 1988). The degree of support provided is often a matter of debate. Decision support can range from the provision of a topographic map, to a screen display of a modeling and analysis outcome that says “Do This.” ‘Advanced’ decision support requires that the analysis of information in a GIS be a component part of a larger information system (the ‘procedures’ component of a ‘whole’ GIS). Hopper (1991) has pointed out that there is a dearth of study into the total flow of information through organizations, and of research into the value of information per se. A significant part of decision support is attempting to evaluate the consequences of management decisions or policies; this requires some form of predictive modeling.

Decision support means that all sides of an argument, all reasonable outcomes of a management decision, must be considered. The ‘technological determinism’ of Veblen (1924), where basic decisions are shaped by available technological capacity rather than the traditional political process based on ideology and value systems, has led us to a polluted, unhappy world, where people feel alienated from the decision-making process. Long-term planning requires that the consequences of various courses of action are considered, and that decisions are made based on what is in accord with human values, rather than what is technologically possible.

Perhaps the most important aspect of decision support is that it attempts to link the specialist, who has mastery of a narrow area of knowledge, with the generalist, who has a grasp of the larger situation. In modern democracies, the specialists (often in the public service, which is not directly accountable to the public) take power from the generalists (such as government ministers, who are directly accountable), “because the specialists understand their own particular area of expertise—but not the whole—while the political generalists do not understand the parts and are gradually losing their grip on the whole” (Jones, 1990, p. 175). Computerized decision-support systems may be a way to bring specialist knowledge together, so that the whole may be more easily viewed.

GIS is often touted as “an integrating technology”, able to link together many data strands into a common structure. GIS, holding the data, must link together users and organizations, by providing support for the decisions that the organization must make collectively. Dynamic modeling, as a method both of simulating the future and of perceiving reality, will be one of the GIS tools involved in this process.

In summary, dynamic modeling in GIS is needed:

- To take GIS beyond being a “data retrieval and sifting engine”;
- For GIS to become a decision support tool;
- For providing GIS with other models of reality; and
- To bring generalists and specialists together and remove the alienation between the decision-makers and those affected.

Four-Dimensional GIS

All current GIS can handle 2-D spatial data; this is the minimum expectation of GIS. The limitations of GIS at this level are more the limitations of the particular application’s data structure, or the limitations of that particular implementation. For example, a raster GIS provides very little topological structure to assist in network analysis; however, a vector-based GIS may not have a network analysis module, or it may be poorly implemented.

The main limitation in 2-D GIS is that it is trying to model a physical world that is, essentially, 3-D. The effects of this limitation are that elevations are difficult to represent, or are treated as another attribute, rather than as a necessary part of space itself. Being attributes, only one elevation can exist at any one point. Overhangs cannot be represented, and it is not possible to represent things that are buried. For example, two buried objects that are in the same horizontal locations, but at different depths, cannot be represented. This is a major limitation for GIS use in geological
applications, for modeling buildings, and for considering many natural processes that occur at different levels.

**Dynamic Modeling and GIS**

A major problem in linking other applications to GIS is that encountered when trying to link GIS and dynamic modeling. Discussion of some GIS limitations to the integration of GIS and dynamic modeling may be found in Hazleton et al. (1990a) and Hazelton (1991a, 1991b).

Dynamic models that assist environmental management can benefit by being linked to GIS. The GIS provides data collection, storage, manipulation, and presentation facilities for a higher-level management system, of which the dynamic modeling system is another part. A management system that combines dynamic models and GIS would seem to have considerable utility, as it would also allow the large amounts of data associated with the dynamic modeling to be managed more easily. It also allows a GIS to be used as a better tool for decision support.

Investigation of ways of linking dynamic models and GIS has shown that the two systems are not compatible (Hazelton, 1991a, 1991b). There are two areas of incompatibility: 1) Current 2-D GIS cannot handle large quantities of temporally referenced data, such as that generated by dynamic models; and 2) GIS are restricted to working in what is essentially a 2-D environment. Dynamic models for general-purpose use would have to handle 3-D spatial data.

There are two ways of dealing with this problem of incompatibility.

First, one can try to live with it. This has been tried with varying degrees of success. It has required complex linkages and data-exchange methods, and most of the solutions have been ad hoc and not in any way generic solutions. For example, Osborne and Stoorvogel (1989) describe a system that allows temporally-referenced data to be used in a GIS, specifically using a model to 'grow' a forest, then using a GIS to analyse the results. That this operation requires two separate GIS (IDRISI and ARC/INFO), the growth modelling system (FIBER, which creates its own temporal database) and a suite of Pascal programs to link them together shows the complexity involved. This approach also requires the use of time purely as an attribute, something that is undesirable when dealing with spatio-temporal data, as it denies the possibility of time having a topological structure.

Secondly, one can alter the components to achieve compatibility. One could try to change the dynamic models to 2-D, but this, naturally, defeats the purpose of the model, and sensibly, no-one has tried it. Far better would be to develop a 4-D GIS and allow it to interact directly with the dynamic modeling system being used.

Two-dimensional GIS have major limitations, if it is wished to incorporate many additional features with GIS, or to extend GIS into new applications. Some of these limitations are being overcome by the development of 3-D GIS (Raper 1989; Pigot 1991).

**Space, Time and Space-time**

Newton-Smith (1986) expressed the view that no satisfactory philosophical understanding of the nature of space, time and space-time can be achieved by remaining at the purely semantic level. Questions about meaning take one to physics, and certain results in physics take one back to meaning. While space and time are apparently well-understood concepts, it is very difficult to explain what they actually are, owing to them being so intimately connected with such a wide range of other fundamental ideas. It is only by considering the use of space and time in a given situation that it is possible to understand what they mean in that context.

Theoretical physics has shown that space and time can be considered to be parts of a single entity: space-time. Most implementations of temporality in databases treat the two as quite separate. Langran (1989b) has reviewed the literature relevant to temporal databases in GIS and discussed the methods available for storing and retrieving temporally referenced data in relational DBMS. These methods considered relatively slow rates of change in the database and had major problems handling a join, or GIS overlay, operation.

As stated in their paper, the methodology proposed by Lan-
gran and Chrisman (1988) relies heavily on a temporal attribute database. All the spatial and topological data in the database’s domain are held in one layer, while the topology of the line segments is used to reconstruct the polygonal topology after a time slice is selected. This also may involve a topological ‘clean up’ for the selection of any time slice. As an alternative, the hybrid GIS modeling system described by Osborne and Stoggenke (1989) shows the degree of complexity involved.

Data from different time instants, or epochs, can of course be stored in a conventional GIS, as described in Hunter and Williamson (1990). There are many examples of using GIS to compare the state of things at different epochs. However, a 4-D GIS should be able to store data that represent the complete history of ‘real world’ objects, rather than just occasional ‘snapshots’ of the state of things. It also should allow spatial, temporal and spatio-temporal analyses in the same manner as current GIS perform spatial analysis. A dynamic modeling capability is also a necessary component of the entire system. A 4-D GIS can naturally store and process ‘non-current’ data, i.e., data that represent the history of an object or region.

Hunter (1988) argues the case for retention of non-current data in a manner that allows easy retrieval, and the importance of having these data available. Vrana (1989) discusses the need for recognition of the temporal component of data in LIS, pointing out that the interdependence of the spatial, temporal and thematic attributes of features is not a by-product of such systems, but their very heart. Hunter and Williamson (1990) describe a method of storing data for a digital cadastral database, where the changes in the graphical component were quite slow, compared with potential changes in the attribute database. Here the non-current graphical data are stored in a reference file and ‘switched on’ when necessary by the attribute database. As the data system was a CADD system linked to a database (Intergraph’s IGDS and DMRS, running on an Intergraph VAX), topological structure was not built on the data until a specific instance (time-based) was selected.

Hazelton et al. (1990b) proposed a classification of temporal GIS, based on the expected rates of change over time of the graphical and attribute components of the system. Where the attribute database changed more frequently than the graphical database, the GIS was termed ‘Slow’, while the reverse situation was termed ‘Fast’. A digital cadastral database would therefore be ‘Slow’, while an environmental monitoring GIS may well be ‘Fast’.

Hazelton et al. (1990a) discussed a simplified version of a spatio-temporal GIS data structure. The essence of this data structure was to treat the history of ‘real world’ objects as 4-D objects, and to develop a 4-D GIS to handle these types of objects. Many aspects of this conceptual GIS data structure are developed by analogy from 2-D GIS. This GIS would be a vector-based system, rather than a cell-based system.

Bell et al. (1990) describe the basic approaches they intend to follow to create a 4-D cell-based GIS. The aims are similar to Hazelton et al. (1990a), but the data structures are simpler. Both these approaches have a place in the range of 4-D GIS, as cell-based and vector-based systems have their roles in 2-D GIS.

Langran (1989a, 1989b) and Langran and Chrisman (1988) can be said to sum up the situation of spatio-temporal data structures as follows:

- Spatio-temporal indexing of data is a major problem;
- Structures for storing spatio-temporal data will not be simple;
- There is unlikely to be a single spatio-temporal data structure that will be widely acceptable or useful.

However, despite some use of GIS to store temporally referenced data, there does not seem to have been an implementation of a true 4-D GIS to date.

**Integrating Dynamic Modeling and GIS**

In the preceding discussion, we have argued that the most general solution to integrating dynamic modeling and GIS is the development of a 4-D GIS. This leads to questions about the role of dynamic modeling in a 4-D GIS, and how the two systems should work together.

If we consider using a straight-forward cell-based GIS such as GRASS (USA-CERL, 1988), and we wish to use a purely Boolean empirical model, we can use the module ‘com-
bine'. If we wish to use a 'graduated' Boolean empirical model, we can use the module 'weight'. If we wish to use a regression model, we can use the module 'Gmanpcalc'. Such operations can be done in various other ways in other GIS, but the facilities generally exist as part of the basic GIS package. The input required to run the model is 2-D data, a model and its parameters; and the output is further 2-D data.

If we use a dynamic model to generate future scenarios of a system, the result is a series of 'time-slices', each of which shows the state of the system at a particular instant of time. Each time-slice may include several attribute 'layers' or 'geographical extents'. The input required for this type of model is several 'time-slices' of 2-D data (to allow calibration and verification), a model, its rules and parameters; the output is several 'time-slices' of 2-D data. This provides a simple means of telling the difference between the two types of modeling in a conventional, 2-D GIS.

However, in a 4-D GIS, the input and output of any type of modeling is rather different. An empirical model in general will take some 4-D data and a model, and produce 4-D data. A dynamic model will also take 4-D data and a model, and produce 4-D data. The dividing line starts to blur.

Calkins (1988) produced the following taxonomy of GIS uses, which contains six groups. These are, from highest to lowest usage:

1. Inventory
2. Map-based Presentations
3. Query and Map-based Presentations
4. Simple Map Comparison
5. Complex Spatial Modeling
6. Spatial Decision Support

At the lowest level, Inventory, the GIS requires no capabilities other than storage and retrieval. At higher levels, more modeling in general, and more complex modeling in particular, is introduced. Level 3 requires the use of empirical models at a simple level, and probably indicates the lowest level of what could be clearly stated to be a 'proper' GIS, rather than a digital mapping system. Level 5 requires the use of dynamic models, and as such is going beyond the capabilities of most GIS on the market. Level 6 goes beyond what we currently refer to as 'GIS', as in a Spatial Decision Support System we could expect GIS to be just another tool available within a much broader application system.

If we consider the use of GIS at Levels 3 and 4 (where most 'true' GIS operations are performed) we notice that empirical modeling is such a necessary part of GIS that it has been subsumed into the GIS. Average users think about using the GIS to find the most suitable housing sites, for example, rather than thinking about using a model to find them. However, the same users think about a dynamic model as being quite separate from the GIS, as this is their usual experience. That this discussion has presumed various levels of GIS usage are 'true' or 'proper' GIS operations, is an indication of this conceptual separation.

In a 4-D GIS, we may expect that dynamic modeling will become part of the essential fabric of the GIS, as empirical modeling has become in 2-D GIS. The definitions of modeling in 4-D will have to be revised, both because of this incorporation and the differences between empirical models and models based upon 'fundamental laws.'

Conclusion

In this paper we have argued that GIS is severely constrained by its dependence upon the cartographic metaphor as its model of reality. Development of GIS beyond a 'map representation' system requires GIS to become a 'reality modeling' system. At the most basic level, GIS must be able to handle truly 3-D spatial data, and temporal data.

With the ability of GIS to incorporate time on the same basis as space, dynamic models that use both time and space as parameters can be used to model the consequences of various management options in many application areas, such as agricultural and environmental management. Dynamic models in general will require the management of spatial data that is truly 3-D, over time. This will necessitate the use of a 4-D GIS.

The extension of GIS beyond the current 2-D design will require re-thinking of the relationship between modeling and GIS. In the same way that 2-D GIS has subsumed empirical
modeling, 4-D GIS will subsume dynamic modeling.

At its most fundamental, GIS has been considered to be a decision-support system that handles spatial data. At present, the decision-support role is limited by the inability to produce and analyse future scenarios in any depth or quantity. Four-dimensional GIS provides the basis for development of better systems for decision support.

Acknowledgements.

This work has been primarily sponsored by Intergraph Corporation, under the supervision of Prof. Ian Williamson. Andrew Turk, Walter Hesse and other researchers in the Department of Surveying and Land Information have been involved in many discussions with the authors about these issues, and have contributed to the development of the ideas expressed in this paper. Marcus Lester, Simon Pigot, John Herring, Tom Watson, Gail Langran and Nick Chrisman also discussed a number of the issues involved in this paper and the background to it.

Note

1. This is akin to 'Gell-Mann's totalitarian principle' in quantum mechanics: "whatever is not forbidden is compulsory." Fitzgerald (1970).

References

Bell, S.B.M., Chadwick, A., Cooper, A.P.R., Mason, D.C., O'Connell, M., and Young, P.A.V. 1990. "Handling Four Dimensional Geo-Coded Data." Proceedings of the 4th International Symposium on Spatial Data Handling, held in Zürich, Switzerland, August, 1990, pp. 918-927.


