INTEGRATING TRADITIONAL SPATIAL MODELS OF THE ENVIRONMENT WITH GIS

Karen K. Kemp
Assistant Director
National Center for Geographic Information and Analysis
University of California, Santa Barbara, USA
email: kemp@ncgia.ucsb.edu

Beginning with the premise that environmental science disciplines have traditionally used conceptual spatial models which are different both from those used in other disciplines and from those provided by the digital data models of current GIS, a consideration of the fundamental role that phenomena which vary continuously across space play in environmental models suggests that continuity may provide an essential unifying theme. This in turn may provide an important basis for the design of interoperable GIS for environmental modeling purposes. Issues arising from a fundament assumption of continuity and themes for continued research are raised.

INTRODUCTION

The increasingly sophisticated tools for spatial modeling and analysis provided by today's GIS are now leading to a revolution in environmental modeling, one which encourages scientists to incorporate spatial processes and relationships in their models. However, the driving force for the design of most widely used GIS packages has not been environmental science. As a result, translation of the unique spatial concepts and models which have evolved independent of GIS in the various environmental sciences is not always obvious or without misapplication. Some effort has been directed recently to the development of software interfaces which will permit translation of generic spatial models such as "field" into the standard data models provided by today's GISs (cf Laurini and Pariente 1996; Vckovski and Bucher 1996). Likewise, the relationship between the real world and how it is represented in GIS has also been the subject of discussion (Burrough and Frank 1995; Couclelis 1992; Csillag 1996; Goodchild 1992; Goodchild 1993; Kemp 1996a; Kemp 1996b; Nyerges 1991; Peuquet 1990). However, these advances have yet to fully address the specific needs of individual environmental scientists as they attempt to make use of spatial information and the new spatial technologies.

This premise behind the research outlined here is that traditional (pre-GIS) approaches used by various environmental modeling disciplines to represent the spatial extents of their phenomena of interest and to implement the interactions between them differ from each other and from those provided by GIS. By examining the underlying bases which led to the development of these discipline specific representations, it should be possible to determine which are the critical aspects needing particular attention in GIS/modeling interfaces. The somewhat surprising conclusion that there may, in fact, be more similarities than differences in how environmental modelers conceptualize space points at a potential basis for true integration of environmental models and GIS.

The following paper documents a preliminary study carried out through in-depth interviews with a range of scientists building or implementing environmental models. The opportunity for this study was provided during a brief sabbatical visit to the Australian National University and CSIRO, both in Canberra, Australia, during October 1996.

PRELIMINARY PREMISES

The following four premises formed the initial basis for the study. A later section explains how these ideas have since been modified as a result of the interviews. The term "conceptual spatial model" refers to the analog models used to constrain or inform data collection activities and/or used during the conceptualization of process models.

1. The conceptual spatial models used by environmental modelers differ in significant ways from the spatial data models provided in current GI systems. Simple mappings between these models are not currently possible. This implies that environmental modelers generally must modify their models in order to use GIS. This is sometimes difficult and may result in incorrect use of available data and misinterpretation of model results.
2. The objects of study, traditional sampling designs and modeling techniques used by individual environmental science disciplines lead to discipline specific conceptual spatial models. These conceptual spatial models vary significantly between disciplines. Thus, there are significant differences in how different sciences discretize space, sample spatially distributed phenomena and extrapolate from their discrete samples to the phenomena being studied.
3. However, it is possible to deconstruct these differences such that the fundamental common characteristics of conceptual spatial models can be identified and measured.

4. These characteristics can be used to develop interoperable interfaces, data models or other elements of GI systems which will enable environmental modelers to use them more efficiently.

In order to find support for these premises and to set the basis for activities related to the fourth item above, the following steps were planned.

1. Describe and characterize conceptual spatial models used by environmental modelers for data collection and for model development
2. Using this information, devise direct mappings between these conceptual spatial models and digital spatial models.
3. Using these mappings, outline some improvements to the design of new interfaces, data models and/or other GIS components in order to improve the efficiency with which environmental modelers can use GIS and to assist in the development of interoperable components for GIS .

CONCLUSIONS ABOUT THE PREMISES

Premise 1: The conceptual spatial models used by environmental modelers differ in significant ways from the spatial data models provided in current GI systems.

By definition, environmental models are environmentally determined. Thus since many environmental phenomena are fields (phenomena for which a value exists at all locations and which may vary continuously across space), environmental models are fundamentally continuous. Hence, environmental *modelers* generally have a continuous view of the world. There are several data models for representing fields in GIS (including cellgrids, planar enforced polygons, TINs, contour lines, pointgrids and irregular points). Methods for manipulating data in these data models are widespread and robust. For example, watershed models which model the flow of water across surfaces are often implemented as finite element solutions (where finite elements are expressed in GIS as polygons). Ground water models likewise often use gridded finite element structures (where finite elements are stored as cellgrids). Therefore, it may be concluded that the discrete digital data models provided by GIS do not present *conceptual* problems to the modelers.

Premise 2: The objects of study, traditional sampling designs and modeling techniques used by individual environmental science disciplines lead to discipline specific conceptual spatial models.

If environmental modelers generally do perceive their phenomena as continuous or see their phenomena as being environmentally determined, then their conceptual spatial models do not vary significantly between disciplines when analysis depends on a context of the continuous environment. However, the objects of study do vary from superimposed continuously varying phenomena (such as pressure and temperature surfaces in climatology), to objects embedded in continuous matrices (such as faults and intrusions in structural geology), to independent objects (such as individual mammals in wildlife biology). On the other hand, environmental determinism is a fundamental principle in the prediction of the occurrences of many of the phenomena and so they can all be seen to exist within a continuous matrix or at least on a continuous probability surface. Thus, again, continuity provides a common context.

In some sciences, traditional data collection and representation techniques have relied on the discretization of both space and the phenomena being studied. This is particularly true in soil science, geology and vegetation ecology. In these cases, data collection requires experts who interpret the environmental clues, some of them unspecified and unmeasureable, and make conclusions about the distribution of classes of the phenomenon being mapped. The data which is ultimately recorded (i.e. mapped) is not the fundamental observed phenomena, but an inferred classification. An assumption of continuous change across space in the class of the phenomenon does not exist in these data collections.

However, it has long been recognized that this assumption of discontinuity, of homogeneous regions with distinct boundaries, in disciplines such as soils or vegetation science is invalid (cf. Burrough et al. 1977; MacIntosh 1967). These phenomena which are strongly influenced by environmental gradients do vary significantly over space. For many environmental modeling purposes, classified data collection techniques do not result in satisfactory digital records of the phenomena. They do not match the scientists' conceptual models of their phenomena.

Fortunately, the ability to store and manipulate large spatial data bases and the powerful new spatial technologies have begun to allow environmental modelers to move the digital representations closer to these continuous conceptual models. At several different locations, researchers are now working to develop models of soil formation and vegetation growth which are based on continuous environmental determinants such as elevation and rainfall (see for example Burrough et al. 1992; Gessler et al. 1996; Kavouras 1996; Lees 1996; Mackey 1996). These environmental models allow soils or vegetation to be described by a number of different parameters, and, only when necessary, classified accordingly. Classes can be extracted for any set of criteria using various statistical techniques.

If these newest efforts to model soils and vegetation on continuous bases are successful, as seems likely, the contention that significant conceptual differences do not exist between the different sciences themselves becomes even more well founded. However, significant differences do still exist, but these come between the conceptual models of the environmental scientists and those of the environmental managers for whom the models are often developed (Burrough and Frank 1995; Couclelis 1996). At the management end of modeling applications, continuous results are often too difficult to integrate conceptually, particularly when there are several environmental gradients involved. Classification allows many different factors to be summarized and understood conceptually, though not necessarily analytically.

Premise 3: It is possible to deconstruct these differences such that the fundamental common characteristics of conceptual spatial models can be identified and measured.

As the above discussion has asserted, there are no fundamental differences in conceptual models between environmental science disciplines. Thus, environmental phenomena as continuous fields, in some cases with embedded objects, may provide the unifying theme, the fundamental common characteristic.

However, this does not mean everything will need to be represented as continuous fields. It is possible to conceive of and model a continuous environment composed of homogeneous discrete units such as watersheds. It is generally accepted that if the processes being studied operate at a regional scale, subregional size areas below this scale, such as small watersheds, provide sufficient variation for the modeling effort. This, in fact, is the basic premise of the finite element models so widely used in watershed modeling (Vieux et al. 1990). This permits an assumption of homogeneity even within a continuous context and suggests that further consideration should be given to the issue of scale and its relationship to classified continuous phenomena.

Premise 4: These characteristics can be used to develop interoperable interfaces, data models or other elements of GI systems which will enable environmental modelers to use them more efficiently.

Interoperability works best when based on a common conceptual reality. Objects and phenomena should be conceptualized within their physical environment and their attributes and relationships expressed in ways which allow interfaces to translate these generic qualities into system specific values. This means that if reality forms the central interface between different environmental models and spatial databases, all data can be passed through the interface, conceptually returning it to its expression in the physical environment before it is redefined as required for specific software.

Some effort has been directed at itemizing these generic qualities of the physical environment which we seek to model (Burrough and Frank 1995; Couclelis 1996). Methods for quantifying these characteristics are the subject of further research by this author. It is possible to conceive of a software product which would assist environmental scientists and managers to identify and measure the critical characteristics of the environment which determine how it will be modeled and to understand and express the spatial and aspatial components relative to their problems. Such a product might, for example, construct objects (in an OO sense) ready for computation.

ISSUES

These conclusions suggest a number of critical issues which need to be addressed if a functional link between conceptual models and GIS data models can be found.

Is continuity, possibly with embedded objects, "the" conceptual model for environmental modelers?

Do all environmental sciences work in the continuous model? Is geology fundamentally different given that there are discrete geologic objects within continuous matrices as well as continuously varying rock bodies which are discontinuous at boundaries? Can a conceptual temporal model be used to combine and explain intersecting lithologies?

The need for classification remains.

What is the role of classification in sciences with continuous views of their phenomena? Is the need to classify during data collection now unnecessary given current computing power and massive digital storage? What do boundaries mean in the environmental sciences? How do boundaries based on varying criteria affect ecological theory? Does the identification and study of pattern require classification? Must we eventually classify in order to understand?

Classification is a scale issue.

How does scale affect our ability to conceptualize continuous phenomena using discrete representations? Can varying process scales be integrated by using a continuous conceptual model of the phenomena?

Do managers need different spatial models?

Is there a difference between modeling for prediction versus modeling for description and/or management? Do managers need a more discrete (i.e. classified) view of space or do we simply need to educate managers to work with data in forms other than classified maps?

Expert knowledge plays a major role in the understanding and modeling of environmental systems.

How is expert knowledge incorporated into models of processes? What role does it play in classification and data collection? How can we be explicit about the incorporation of expert knowledge in data collection and modeling activities? Can modelers replace or simulate the expert knowledge of the field scientists?

Conceptual temporal models also need to be addressed.

Which sciences assume change and which are static? Historical and episodic events affect the environment but these cannot be represented or modeled well. This is also a scale issue. What about continuity in time? Can space be substituted for time or vice versa (e.g. succession demonstrated by going up an elevation gradient or astrophysical location equating with time)?

Can models be usefully classified as either spatial or aspatial?

Is there a significant difference? Are aspatial models just spatial models at regional scales in which spatial heterogeneity is not relevant at large process scales? How do aspatial models and aspatial data incorporate space? How is space despatialized for aspatial modeling and data collection? How do aspatial models represent changes which have an impact over space?

CONCLUSION

Powerful new tools and paradigm changes are leading to a revolution in environmental modeling. The opportunity to build models which represent continuous variation across the landscape are changing the ways in which we gather data and describe the environment. These in turn provide greater opportunity to provide case-specific information for environmental managers. Much work remains to be done. The results of this preliminary study will inform further detailed research into conceptual spatial models and continuity as an integrating medium. Further work on this theme will be incorporated into efforts related to NCGIA's new research initiative on Interoperating GISs.

ACKNOWLEDGMENTS

These conclusions arise from discussions with the scientists and modelers in Canberra Australia at the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Divisions of Water Resources, Soils, Forestry, Wildlife and Ecology and Information Technology, also at the Australian National University (ANU) at the Centre for Resource and Environmental Studies (CRES) and the Department of Geography. My thanks to everyone who gave me some of their time and many of their ideas. I hope I will have a chance to continue these discussions in the future. Also, I would like to thank CSIRO Division of Information Technology, CSIRO Division of Water Resources, the ANU Department of Geography and the NCGIA who provided support for this research. Research at the NCGIA is supported by grants from the National Science Foundation (SBR 88-10917).

REFERENCES

Burrough, P. A., Brown, L., and Morris, E. C. (1977). Variations in vegetation and soil pattern across the Hawkesbury Sandstone plateau from Barren Grounds to Fitzroy Falls, New South Wales. *Australian Journal of Ecology*, 2:137-59.

Burrough, P. A., and Frank, A. U. (1995). Concepts and paradigms in spatial information: Are current geographical information systems truly generic? *International Journal of Geographical Information Systems*, 9(2):101-116.

Burrough, P. A., MacMillan, R. A., and vanDeursen, W. (1992). Fuzzy classification methods for determining land suitability from soil profile observations and topography. *Journal of Soil Science*, 43(2):193-210.

Couclelis, H. (1992). People manipulate objects (but cultivate fields): beyond the raster-vector debate in GIS. In *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, A. U. Frank, I. Campari, and U. Formentini, eds., Springer-Verlag, pp. 65-77.

Couclelis, H. (1996). Towards an Operational Typology of Geographic Entities with Ill-defined Boundaries. In *Geographic Objects with Indeterminate Boundaries*, P. A. Burrough and A. U. Frank, eds., Taylor &Francis, pp. 45-55.

Csillag, F. (1996). Variations on hierarchies: Toward linking and integrating structures. In *GIS and Environmental Modeling: Progress and Research Issues*, M. F. Goodchild, L. T. Stayaert, B. O. Parks, C.

Johnston, D. Maidment, M. Crane, and S. Glendinning, eds., GIS World Books, Fort Collins, CO, pp. 433-437.

Gessler, P., McKenzie, N., and Hutchinson, M. (1996). Progress in Soil-landscape Modelling and Spatial Prediction of Soil Attributes for Environmental Models. In *Proceedings of Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, Santa Fe, NM. National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA. CD-ROM and WWW.

Goodchild, M. F. (1992). Geographical data modeling. *Computers and Geosciences*, 18(4):401-408.

Goodchild, M. F. (1993). Data models and data quality: Problems and prospects. In *Environmental Modeling with GIS*, M. F. Goodchild, B. O. Parks, and L. T. Steyaert, eds., Oxford University Press, New York, pp. 94-103.

Kavouras, M. (1996). Geoscience Modelling: From Continuous Fields to Entities. In *Geographic Objects with Indeterminate Boundaries*, P. A. Burrough and A. U. Frank, eds., Taylor &Francis, pp. 313-323.

Kemp, K. K. (1996a). Fields as a framework for integrating GIS and environmental process models. Part one: Representing spatial continuity. *Transactions in GIS*, 1(3):in press.

Kemp, K. K. (1996b). Fields as a framework for integrating GIS and environmental process models. Part two: Specifying field variables. *Transactions in GIS*, 1(3):in press.

Laurini, R., and Pariente, D. (1996). Towards a Field-oriented Language: First Specifications. In *Geographic Objects with Indeterminate Boundaries*, P. A. Burrough and A. U. Frank, eds., Taylor & Francis, pp. 225-235.

Lees, B. (1996). Improving the spatial extension of point data by changing the data model. In *Proceedings of Third International Conference/ Workshop on Integrating GIS and Environmental Modeling*, Santa Fe, NM. National Center for Geographic Information and Analysis, University of California, Santa Barbara. CD-ROM and WWW.

MacIntosh, R. P. (1967). The continuum concept of vegetation. *Botanical Review*, 33:130-187.

Mackey, B. (1996). The role of GIS and environmental modelling in the conservation of biodiversity. In *Proceedings of Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, Santa Fe, NM. National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA. CD-ROM and WWW.

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

Peuquet, D. J. (1990). A conceptual framework and comparison of spatial data models. In *Introductory Readings in Geographic Information Systems*, D. J. Peuquet and D. F. Marble, eds., Taylor & Francis, London and Bristol, PA, pp. 250-285.

Vckovski, A., and Bucher, F. (1996). Virtual Data Sets - Smart Data for Environmental Applications. In *Proceedings of Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, Santa Fe, NM. National Center for Geographic Information and Analysis, University of California, Santa Barbara. CD-ROM and WWW.

Vieux, B. E., Bralts, V. F., Segerlind, L. J., and Wallace, R. B. (1990). Finite Element Watershed Modeling - one-dimensional elements. *Journal of Water Resources Planning and Management - ASCE*, 116(6):803-819.