

CHAPTER 3

Geographic Information Science

*Paul A. Longley, Michael F. Goodchild,
David J. Maguire and David W. Rhind*

3.1 Geographic Information Systems and Science

Science is concerned with the discovery and organization of knowledge. By employing scientific methods, scientists create empirically testable theories and models of observable events. It is common to subdivide science into natural sciences—including biological, environmental and physical—and social sciences—economics, psychology, politics, etc. Science can be practiced in both a pure and applied way. Geographic information science (henceforth GIScience) deals with the science of information that can be used to describe, explain and hence predict the forms and processes that characterize the Earth surface and near surface. In practice, it is often impossible to study events and occurrences without considering both natural and human-induced processes (e.g., the impacts of deforestation on subsequent environmental disasters), and thus GIScience spans natural-social and pure-applied scientific divides. GIScience unifies the ways in which we study the surface of the Earth and the impact of those living upon it, using geography as an organizing framework to structure the information that is available to us.

GIScientists use models, or simplifications, to represent aspects of the seemingly infinite complexity of the Earth in digital information systems (see also Longley, Chapter 30, this volume). Some of these models are predominantly descriptive—data or information models—and are primarily used in studies concerned with surface morphology or form. These are used to store vast amounts of information about the Earth. Other models are mainly analytical and are used to understand the processes that operate upon it. These processes may be static or dynamic, and process models may be used both inductively and deductively to search for, and test, theories and hypotheses about geographic events. Because of the complexity of environmental and social systems there are only a few geographic laws of nature. This can frustrate our attempts to model the complexities of the real world, to explain how things work, and to make predictions about future outcomes.

Core to the remit of Geographic Information Systems (GIS) is our ability to create and share representations of the world that often lie beyond our direct experience. But how do we know that a shared representation of the world is fit for the purpose for which it has been designed or applied? This question may in turn invite questioning of the provenance of the data that have been used to create the representation, the ways in which diverse data sets have been structured and assembled together, or the decision-making context in which an application is developed. Other contributions to this manual have addressed the rapid development of geographic information (GI) technologies, developments in software and organizational work flow, changes to the data economy of GIS, and broadening of the range of settings in which basic GIS skills may be acquired (see McDonnell and Burrough, Chapter 2, this volume; see DeMers, Chapter 4, this volume; Longley et al. 2005b). Implicit in these various discussions is the notion that all but the most basic concepts in GIS—such as distance measurement, topology, proximity and connectivity—are fast changing, and hence transitory. Yet this is not the full story, for the apparent rapid evolution of GI software conceals rather a lot that is enduring about the application of GIS, in terms of core organizing principles and techniques.

This is, of course, widely recognized, and many of the enduring characteristics of GIS receive attention in this manual, alongside discussion of what is new. Our own objective

here is to introduce the remit of GIScience as weaving many of these different threads into a distinctive area of research and applications activity. The context for this is disarmingly simple: if the real world is of seemingly infinite complexity and the process of representation requires us selectively to discard nonessentials, then the field of GIScience is necessary to guide us in this task, and to suggest ways in which the quality of the resulting partial representations may be evaluated. This requires us to move beyond basic principles and routine applications, to “the science behind GIS” (Goodchild 1992). The developing field of GIScience provides much of the direction, growth and impact of GIS and related technologies. If information science studies the fundamental issues arising from the creation, handling, storage, and use of information, then the remit of GIScience may be defined as study of the fundamental issues arising from geographic information, as a well-defined class of information in general. Other terms have much the same meaning: *geomatics* and *geoinformatics*, *spatial information science*, or *geoinformation engineering*. Each of these terms suggests a scientific approach to the fundamental issues raised by the use of GIS and related technologies, though they all have different disciplinary roots and emphasize different ways of thinking about problems.

Invocation of the term “GIScience” implicitly differentiates between straightforward spatial query operations (such as those involving distance measurement, overlay analysis, buffering, optimal routing and neighborhood analysis) and the much broader and generalized range of transformations, manipulations and techniques that form the bedrock to the field that is often described as *spatial analysis* (see Longley, Chapter 30, this volume). The term is also used to describe the use of GIS in support of scientific research, in the social or environmental sciences, in accordance with the norms and practices of science. Furthermore, discussion of GIScience also requires that we consider the scientific activities associated with the field, its developing organizational infrastructure, and the types of uses to which it may be put.

Insofar as we are concerned with the successful application of GIScience, the social, organizational/management and even political contexts in which decisions are made also need to be considered. This is often apparent when considering the big challenges facing humankind—such as climate change, natural hazards like tsunamis, major migrations, or global terrorism, to name but a few. For example, the precise location of the December 2005 tsunami off Indonesia was predicted in a published map in 1979, yet no warning systems were installed or other preparations made.

3.2 Core Organizing Principles and Remit of GIS and GIScience

GIS helps us to manage what we know about the world—to hold it in forms that allow us to organize and store, access and retrieve, manipulate and synthesize, visualize and communicate spatial data. Many of the routine operations that are core to GIS implicitly invoke hypotheses with respect to the data and information that are available—in the spirit of deductive reasoning. Induction also plays an (increasingly) important part in GIS-based analysis, whereby data “mining” is used to identify what to leave in (and, hence, what to take out of) a representation, and what weight to assign that which is left in. Yet these complementary procedures of induction and deduction can often raise questions that are at the same time frustrating and profound. For example, how does a GIS user know that the results obtained are accurate? How might the quality of the input data be ascertained, with respect to other validating sources that might be available to us? How can we be sure that the visual medium of a GIS does not obscure the underlying messages of a representation? What principles might help a GIS user to design better maps? How can GIS be fine-tuned to assimilate the

limits of human perception, cognition and decision making? Some of these are questions of GIS design, and others are about GIS data, methods and system design. GIScience takes this a stage further, towards developing models that improve our understanding of the effects of underlying processes upon unique places. They all arise from practical use of GIS, but also relate to its core underlying principle and techniques.

A good starting point to develop an empirical understanding of the remit of GIScience is the “First Law of Geography,” often attributed to geographer Waldo Tobler. This can be succinctly stated as “everything is related to everything else, but near things are more related than distant things.” This statement of geographical regularity is key to understanding how events and occurrences are structured over space, and presents an empirical agenda for GIScience. It can be formally measured as the property of *spatial autocorrelation* (Longley et al. 2005a) which, along with the property of *temporal autocorrelation* (“the past is the key to the present”) established as the uniformitarianism principle by Hutton over 200 years ago, suggests a fundamental geographic premise: the geographical context of past events and occurrences can be used to predict the future. However, the observable regularity in spatial and temporal structure that allows us to generalize and predict—on the basis of observed distributions of events and occurrences—falls far short of the degree of regularity worthy of being described as laws of GIScience. To take a building metaphor, in attempting to explain evolving geographic phenomena, no representation is ever founded upon rock; but rather, the structures of our representations are erected upon piles that have been sunk sufficiently deeply into the mud to sustain them. Thus a representation of the local market for a grocery store may reduce the complexity of human decision making to distance minimization in ways that appear crass, simplistic or primitive, but the point is that such simplification may be both necessary and sufficient to devise usable forecasts of store turnover. From this perspective, “good GIS” is about retaining the most significant spatial and temporal events, and discarding detail that may be considered to be irrelevant for practical purposes. What we know already about the problem (our substantive expertise, such as how far people are prepared to travel to buy groceries, or how grocery shopping fits into daily activity patterns) may be used deductively to frame the problem, and principles from GIScience help guide us towards identifying the precise attenuating effect of distance (upon store patronage, for example).

3.3 GIScience Agendas

GIScience has evolved significantly in recent years—some idea of the range of current interests in the field can be gained from the program of the GIScience biannual conference series (Longley 2005a). Efforts to enumerate the constituent issues of GIScience began with the US National Center for Geographic Information and Analysis in the early 1990s and the twenty systematic research initiatives that it spawned. Since then the University Consortium for Geographic Information Science has assumed responsibility for developing an extensive research agenda (McMaster and Usery 2005), which has been modified over time to keep up with a changing and expanding set of issues. There are currently (October 2006) thirteen long-term issues:

- Spatial Data Acquisition and Integration
- Cognition of Geographic Information
- Scale
- Extensions to Geographic Representations
- Spatial Analysis and Modeling in a GIS Environment
- Uncertainty in Geographic Data and GIS-Based Analysis
- Future of the Spatial Information Infrastructure
- Distributed and Mobile Computing

- GIS and Society: Interrelation, Integration, and Transformation
- Geographic Visualization
- Ontological Foundations for Geographic Information Science
- Remotely Acquired Data and Information in GIScience
- Geospatial Data Mining and Knowledge Discovery.

A different, and disarmingly simple, way of viewing the remit of GIScience is provided by the Varenus project; for a summary see Longley et al. 2005a, Chapter 1. Here, GIScience is viewed as anchored by three concepts—the individual, the computer, and society. These form the vertices of a triangle, and GIScience lies at its core. The various terms that are used to describe GIScience activity (such as those enumerated in the previous paragraph) can be thought of as populating this triangle. Thus research about the individual is dominated by cognitive science, with its concern for understanding of spatial concepts, learning and reasoning about geographic data, and interaction with the computer. Research about the computer is dominated by issues of representation, the adaptation of new technologies, computation, and visualization. And finally, research about society addresses issues of impacts and societal context.

It is possible to identify a number of themes that link the different taxonomies and characterizations of the field. First, it follows that if we select what to leave in and what to take out of a representation, then any representation is necessarily incomplete and hence presents an inherently *uncertain* view of the world. Ongoing research in GIScience is attempting to understand the outcomes of choice, convention and chance upon what a representation does *not* reveal about the world. Formal theories have been developed based in the frameworks of geostatistics and spatial statistics, implementing many ideas of geometric probability. Techniques have been devised for simulating uncertainty in data, and for propagating uncertainty through GIS operations in order to provide confidence limits on results.

Second, research in GIScience is having profound implications for “low order” as well as “high order” geographic concepts. GIScientists have attempted to write a formal theory of geographic information, replacing the somewhat intuitive and informal world of raster and vector data and topological relationships that existed prior to the 1990s. Formal theories of topological relationships between geographic objects have been developed and GIScientists have formalized the fundamental distinction between object-based and field-based conceptualizations of geographic reality. The narrow and pragmatic view of GIS has hitherto been of software dominated by either raster or vector data structures, in turn reflecting the analytic roots to particular suites of software (McDonnell and Burrough, Chapter 2, this volume). Yet these new ideas are now becoming embedded in the standards and specifications promulgated by the International Organization for Standardization (ISO) and the Open Geospatial Consortium (OGC).

Third, GIScientists have investigated the social contexts in which GIS are applied, and the ways in which the technology both empowers and marginalizes. This work was stimulated in the early 1990s by a series of critiques of GIS from social theorists, following which it became clear that the broader social impacts of the technology were an important subject of investigation. Critics have worked alongside GIScientists to develop insider views of the ways in which GIS may be acquired and manipulated by the powerful, sometimes at the expense of the powerless. Active research communities in GIS and Society and Public-Participation GIS attest to the compelling nature of these arguments. Such research, especially of both closely and loosely coupled interactions between human activities and natural events, is of course also being carried out by many others. In the extreme case, terrorist actions designed to undermine financial systems would have a downstream effect on the capacity of some nations to sustain and protect their citizenry—the first duty of the state. Perhaps surprisingly, GIScience can be embedded successfully in modeling to assess probabilities and amelioration options.

From a broader perspective, developments in GIScience can be defined through their relationship to other, larger disciplines. Information science studies the nature and use of

information, and in this context GIScience represents the study of a particular type of information. In principle all geographic information links location on the Earth's surface to one or more properties, and as such the field of GIScience is particularly well-defined. For this reason, many have argued that geographic information provides a particularly suitable testbed for many broader issues in information science. For example, the development of spatial data infrastructure in many countries has advanced to the point where its arrangements can serve as a model for other types of data infrastructure. Metadata standards, geoportal technology, and other mechanisms for facilitating the sharing of geographic data are sophisticated, when compared to similar arrangements in other domains.

3.4 Conclusion

At its core, GIS is concerned with the development and transparent application of the explicitly spatial organizing principles and techniques of GIScience, in the context of appropriate management practices. GIS is a practical problem-solving tool that is judged by the "success" of its applications. The spatial dimension to problem solving is special because it poses a number of unique, complex and difficult challenges that are investigated and researched through GIScience. Together, these provide a conduit for committed applications specialists to pursue their interests through vocation, in academic, industrial, and public service settings alike. All sciences have their tools and systems: astronomers use telescopes to view stars and information systems to record their characteristics; biologists use electron microscopes to visualize the structure of cells and supercomputers to simulate ecological systems; and computer scientists develop new computer architectures using design software. GIScientists also have their tools—geographic information systems—which are a fundamental and integral part of pursuing GIScience. These tools are applied across the whole arena of the natural and social sciences to enhance our understanding and explore better ways of making the world a more sustainable place.

References

- DeMers, M. N. 2008. A hitchhiker's guide to geographic information systems: A multi-disciplinary introduction. In *Manual of Geographic Information Systems*, edited by M. Madden, Chapter 4 in this volume. Bethesda, Md.: ASPRS.
- Goodchild, M. F. 1992. Geographical information science. *International Journal of Geographical Information Systems* 6:31–45.
- Longley, P. A. 2008. Spatial analysis and modeling. In *Manual of Geographic Information Systems*, edited by M. Madden, Chapter 30 in this volume. Bethesda, Md.: ASPRS.
- Longley, P. A., M. F. Goodchild, D. J. Maguire, and D. W. Rhind. 2005a. *Geographic Information Systems and Science*. 2d ed. Chichester: John Wiley & Sons.
- . 2005b. *Geographical Information Systems: Principles, Techniques, Management and Applications*. 2d ed, abridged. Hoboken, New Jersey: John Wiley & Sons.
- McDonnell, R. and P. Burrough. 2008. Principles of geographic information systems. In *Manual of Geographic Information Systems*, edited by M. Madden, Chapter 2 in this volume. Bethesda, Md.: ASPRS.
- McMaster, R. B. and E. L. Usery, eds. 2005. *A Research Agenda for Geographic Information Science*. New York: Taylor & Francis.