WHOSE HAND ON THE TILLER? REVISITING "SPATIAL STATISTICAL ANALYSIS AND GIS"

Michael F. Goodchild[1]

ABSTRACT

Anselin and Getis argue in their 1992 paper "Spatial statistical analysis and GIS" that the development of a toolbox of spatial analytic techniques should be directed by the scientists whose work defines the need for such a toolbox. The field of GIS has changed fundamentally since 1992 as a result of new technical developments, including a general move away from the map as the defining metaphor, the influence of the Internet and the World Wide Web, and changes in the practice of software engineering. Science as a whole has also changed, towards a more collaborative model that is more dependent on computational infrastructure. The impacts of space on the methodology of science are also better understood. The Anselin and Getis paper was remarkably prescient in its identification of the major issues that continue to affect the relationship between spatial analysis and GIS. Institutional issues continue to frame the relationship between GIS and spatial analysis, and are best addressed through partnerships.

1. INTRODUCTION

In 1992 Luc Anselin and Art Getis published "Spatial statistical analysis and GIS" in the *Annals of Regional Science* (Anselin and Getis, 1992). The paper was one of a number of explorations of the relationship between spatial analysis and geographic information systems (GIS) that appeared at about that time (Burrough, 1990; Ding and Fotheringham, 1992; Fotheringham and Rogerson, 1994; Goodchild, 1987; Goodchild *et al.*, 1992; Openshaw, 1990), driven perhaps by a perception that the evident growth in GIS as a large and complex software application might eventually benefit science, by providing improved access to many of the tools that researchers had developed over the previous decades. Indeed GIS was being widely hailed in this light (Abler, 1987), and it seemed only a matter of time before virtually all of the known methods of spatial analysis would be available in a single, massive, widely available toolbox.

The Anselin and Getis paper was distinguished from others in the genre by its focus on regional science, and by its strongly stated belief that future developments in GIS should be driven by scientists in the substantive fields of application, including regional science:

> "Some would argue that … using the exisitng GIS tools will lead to different and more interesting questions. Our position is that the technology should be led by

---

[1] National Center for Geographic Information and Analysis, and Department of Geography, University of California, Santa Barbara, CA 93106-4060, USA. Phone +1 805 893 8049, FAX +1 805 893 3146, Email good@geog.ucsb.edu

theoretical and methodological developments in (regional science) itself."
(Anselin and Getis, 1992, p. 30)

In other words, future developments in GIS should be directed by those most familiar with the kinds of questions it was ideally able to answer; the idea that the GIS tail might wag the regional science dog was clearly not as attractive.

Approximately a decade and a half has elapsed since the paper was written, and the landscape of GIS has been changed almost beyond recognition. In this paper I attempt to bring the discussion up to date, to answer some of the questions raised by the authors, and to pose new ones that seem to have arisen recently – and to do so within a somewhat larger context that includes geography and other social and environmental sciences in addition to the regional science of the original paper. The next section provides a brief and I hope accurate summary of the main arguments of the Anselin and Getis paper. This is followed by a review of events and trends within the field of GIS since 1992, and then by a review of major trends affecting science, and particularly the social and environmental sciences. The final section of the paper updates the 1992 comments on the role of regional scientists in directing the development of GIS, by suggesting specific parts that substantive scientists can play in the evolving saga of software development and support.

2. SYNOPSIS OF THE 1992 ARGUMENTS

The model that underpins the Anselin and Getis discussion is particularly elegant in the way it combines widely accepted organizing principles from both GIS and spatial analysis. It is reproduced in Figure 1. It defines GIS operations in four classes -- input, storage, analysis, and output – following many extensive discussions of GIS functionality (see, for example, Maguire, 1991; Maguire and Dangermond, 1991). The analysis function has been seen as the most important by several authors (see, for example, Cowen, 1988), who have argued that it represents the vital transformation of data into useful information, making visible what might otherwise be invisible to the user. The authors then introduced a four-way classification of analysis, into selection, manipulation, exploration, and confirmation. The sharp distinction between analysis and display that was implicit in the classification of GIS was softened somewhat, as display was seen as inherent in each of manipulation, exploration, and confirmation. Selection encompassed sampling and other aspects of what today might be termed *ontology*, or in statistical terms *support*, while manipulation included many of the lower-grade operations of GIS such as buffering, spatial joins, and point-in-polygon operations, that are nonetheless termed *analysis* by many GIS developers (see, for example, Mitchell, 1999).

[Figure 1 about here]

The authors identified several software strategies that were used *circa* 1992 to deliver the functionality of spatial analysis using GIS. Some techniques were included directly in the GIS itself as core functions, though it was recognized that this strategy was not always

the fastest computationally. Separate modules could be constructed that performed the more elaborate forms of spatial analysis, and coupled to the GIS for purposes of data input, storage, and display, either through the exchange of files or through access to a common database (Nyerges, 1993).

Anselin and Getis adopted the distinction between exploratory and confirmatory analysis, while accepting that the distinction was often somewhat blurred. Exploratory analysis was seen as data driven and inductive, while confirmatory analysis was theory driven and deductive. Exploratory spatial data analysis (ESDA) was a new and exciting field in 1992, building on the improved interactive graphics capabilities that became available first in the Macintosh and in Unix workstations in the 1980s, and later in the PC. Exploratory analysis also included the data-driven approach being advocated by Openshaw (Openshaw *et al.*, 1987; Openshaw, Charlton, and Craft, 1988; Openshaw, Cross, and Charlton, 1990), which saw the search for pattern as an activity that could be essentially independent of any theoretical framework. Today we might find echoes of the same strategy in the use of techniques borrowed from artificial intelligence, including neural nets and self-organizing maps (Fischer, 1998; Fischer and Leung, 1998; Skupin and Hagelmann, 2005).

The fourth section of the paper looked at the relationship between GIS data models and spatial statistics, making the point that discretization of space was an essential step in any representation, and that it impacted the results of analysis in ways that were largely out of the analyst's control. The Modifiable Areal Unit Problem (MAUP) had caught the attention of many researchers in the 1980s, who were dismayed to discover how much their results were affected by the choice of areal units, and therefore by decisions made in statistical agencies that were far from neutral in this respect. The impact of such scale-related effects was discussed at greater length in the fifth section of the paper, titled Implementation Issues, which emphasized the importance of spatial dependence in determining the effects of scale change, and in confounding any attempt to use spatial methods in a confirmatory mode.

In their concluding section, Anselin and Getis argued that to serve the needs of regional science, GIS needed to be "complemented with powerful computational intensive approaches and innovative visualisation. A creative combination of the 'old' spatial analysis with these new technologies to form a 'new' spatial analysis … has not yet been achieved." It was "clearly an overstatement" to suggest that "the recent developments in GIS have already transformed the way spatial analysis is carried out in the field of regional science". "An effective integration (of computationally intensive approaches and innovative visualization) with the new technology may go a long way toward convincing researchers in regional science and other social sciences that the special role of space which underlies the essence of the field merits its own analytical toolbox."

3. GIS SINCE 1992

The term *GIS* had rather different connotations in 1992 than it does today. New technologies have driven a substantial restructuring of the field, as has massive growth in

numerous application areas. In 1992 GIS connoted a single, monolithic software package running on a stand-alone workstation or perhaps a local-area network, and analogous to Microsoft Word or Excel. Its purpose was to relieve its user of tasks that would be too tedious, repetitive, time-consuming, complex, or inaccurate if performed by hand. Analysis of maps and map data (Maling, 1989) has all of these characteristics, and a technology that promised apparently effortless analysis at the speed of light was clearly attractive. In 1992 several vendors offered such packages, including Intergraph, MapInfo, ESRI, Wild, Caliper, and Tydac. Some product differentiation was evident, between large, expensive packages targeted at corporations, government departments, and universities, and small, cheaper packages designed for single users. Several packages were available from the academic community, including Idrisi from Clark University, and in general these offered a more sophisticated range of analytic tools but had more severe limitations in terms of speed and capacity. Finally GRASS offered a large number of useful analytic functions in an open-source package originally developed by the U.S. Army Corps of Engineers.

Many factors have contributed to a changing perspective on GIS over the past fifteen years. First, early developments in GIS were built on the map as the primary source of input, and the first applications of commercial GIS were accordingly in areas heavily dependent on maps, such as resource management and forestry. By the early 1990s, however, it had become clear that much could be gained by adding geographic references to the records contained in the otherwise non-spatial but massive databases of utility companies, marketing companies, and other commercial sectors. The leading database vendors, including Oracle and Informix, developed extensions to handle such *spatially enabled* records and to support simple queries, such as "select for me all of the hotels within ten miles of this airport" – and the major GIS vendors responded with products of their own, such as ESRI's ArcSDE.

Second, the Internet became the dominant network, and the World Wide Web emerged as a dramatically effective application. In the second edition of their survey of GIS, Longley *et al.* (1999) admit to having missed completely the impact that this would have on the field when they wrote the conclusion of their first edition in 1991 (Maguire, Goodchild, and Rhind, 1991). Data had always been something of an Achilles heel to GIS, because the conversion of paper maps to digital records was difficult to automate and frustratingly error-prone. But the Internet opened an apparently unlimited potential for sharing of digital geographic data, first with such early applications as ftp and WAIS (Nebert, 1993) and later with Web applications. Today many GIS users have long forgotten the tedium of map digitizing, and instead rely on a vast array of clearinghouses, data archives, digital libraries, and data warehouses to supply the basic data on which GIS depends. The state of the art is perhaps best represented by the Geospatial One-Stop (www.geodata.gov), a U.S. Government effort to provide a *geo-portal*, a single point of entry to a vast, distributed network of geographic data sources. At the same time, much effort has gone into developing the standards and protocols needed to achieve *interoperability* between suppliers and users of data, with their many different data formats and GIS software products. The Open Geospatial Consortium (www.opengeospatial.org) has developed and promulgated many standards, and the Federal Geographic Data Committee

(www.fgdc.gov) continues to have a very effective influence. The term *National Spatial Data Infrastructure* was coined in the early 1990s (Mapping Science Committee, 1993) to describe a vision of an interoperable, networked future, and was authorized by Presidential Order in 1994. Similar efforts are under way in many other countries, and within the European Union (inspire.jrc.it).

This networking of GIS data access has had a fundamental effect on the software, requiring that it contain the data conversion routines needed for interoperability, as well as the tools to support search for data over a distributed network. It has become essential to support *metadata*, the descriptive catalog entries that now allow researchers to specify needs and to search, evaluate, and retrieve suitable data sets. Increasingly, students of GIS find themselves learning as much about the management of geographic data as about the software that runs in their local machine, and there have been calls to drop the S and to refer to the field simply as *GI*.

The Internet has had a further influence in allowing the processing steps that underlie GIS to occur at locations remote from the user. *Client-server* systems divide the processing tasks between the user's own local machine, the client, and a remote and probably more powerful machine known as the server. In the extreme, all processing occurs on the server and the client is reduced to a "dumb terminal" or a simple Web browser such as Microsoft Internet Explorer. This arrangement is often favored by government agencies, which make limited GIS services available in this way and thus avoid having to distribute copies of their data. Many Web sites now offer mapping and simple forms of analysis via servers, allowing users to visualize data, make summary extracts, and even perform simple statistical analyses. The GIS software industry now offers a range of server-based products, such as ESRI's ArcIMS, to provide the necessary services, and MapServer (mapserver.gis.umn.edu) is a popular open-source product. The term *GIServices* is sometimes used to distinguish such client-server configurations from the more traditional GISystem.

In principle any GIS function, and any form of spatial analysis, could be offered as a GIService. In practice a limited number of simple GIS functions are now available as commercial or public services, including geocoding (the task of converting street addresses to coordinates, *e.g.*, www.travelgis.com/geocode/), wayfinding (the task of generating driving directions from an origin to a destination, *e.g.*, www.mapquest.com), gazetteer lookup (the task of converting a placename to coordinates, *e.g.*, www.alexandria.ucsb.edu), location analysis (*e.g.*, a demographic analysis of the neighborhood of a potential retail site, offered by many market research companies), and spatial search (*e.g.*, search for hotels within a given distance of an airport, www.expedia.com). Some of these are based on a viable business model, and it seems that the functions most likely to be offered as GIServices are those that (1) depend on access to a large and rapidly changing database that individuals would not be able to keep up to date, and (2) require a level of complexity of analysis that would be beyond the average user. There is clearly potential here for providing the kinds of exploratory and confirmatory spatial analysis discussed by Anselin and Getis as GIServices, but to date there appear to be no obvious examples in practice.

Third, the past fifteen years have seen a radical rethinking of practice in software engineering. The single, monolithic package of 1992 has largely been replaced by re-usable software components that can be mixed and matched for specific applications. Microsoft's COM/OLE and .Net environments allow software components to be mixed across boundaries that once seemed impenetrable, such as between Excel and ESRI's ArcGIS (Ungerer and Goodchild, 2002), allowing applications to combine functions from both under the direction of scripts written in standard languages such as Visual Basic or Python. Instead of a single package, vendors now offer a variety of extensions for specific purposes, leading to a growing sense of segmentation in the GIS market. This sense has been reinforced by the advent of object-oriented data modeling (Zeiler, 1999; Arctur and Zeiler, 2004), which allows the basic data objects of a GIS (points, lines, and areas) to be specialized for particular application domains.

Today the old sense of GIS as a well-defined type of software has been significantly eroded. Some vendors have chosen to adopt descriptions that better reflect their target application domains, while others emphasize the ability to customize a range of products to specific needs. An adjective seems more appropriate than a noun in this new more complex world, and in recent years the term *geospatial* appears to have gained some traction, as evidenced by the recent renaming of what was the Open GIS Consortium and by extensive restructuring within the U.S. Geological Survey.

The geospatial world of today is clearly a much broader domain of data, tools, services, and concepts than the limited GIS world of 1992. Many of the statistical packages now include limited support for spatial analysis: an extensive set of geospatial tools can be found in Matlab, and the R community has developed a particularly rich set (http://cran.r-project.org/src/contrib/Views/Spatial.html). Links have been constructed between GIS packages and simulation environments such as Stella and Repast, and what were previously considered functions exclusive to GIS, such as simulated fly-by, are now readily available in Google Earth and other *geo-browsers*. A social or environmental scientist needing tools to support spatial analysis now has a vast array of options, many of which would no longer involve anything recognized as a GIS. The idea examined by Anselin and Getis in 1992, of whether the particular needs of spatial analysis justify the development of a special toolbox, no longer seems as relevant – the importance and special nature of spatial analysis is clearly demonstrated by the vast array of data, tools, and services that are now available, whether or not they are labeled as GIS.

Moreover, the days of monolithic software environments are now over, and it is no longer appropriate to envision a day when all forms of spatial analysis will be available in a single toolbox. Just as geo-portals provide a single point of entry to a complex, distributed array of data, it seems appropriate to envision a day when a single point of entry will provide access to a set of distributed, interoperable tools and services. Already some sites provide searchable directories of tools (see for example www.csiss.org), and some geo-portals provide searchable directories of GIServices in addition to data (see for example www.geographynetwork.com). But as yet there are no universal standards for

catalogs of tools (Crosier *et al.*, 2003), or standards for interoperability. Thus the necessary supporting infrastructure for such a vision still needs substantial work.

4. THE BROADER CONTEXT

4.1 Cyberinfrastructure

The topics discussed in the previous section – the need for interoperability, access to distributed data, and GIServices – are representative of much broader trends within computation and within the infrastructure it provides for science. Various names have been given to this new, distributed form of computing, including *cyberinfrastructure*, and much has been written about its potential. In the U.K. the Economic and Social Research Council has made substantial investments in building a cyberinfrastructure for social science through the e-Social Science initiative (www.ncess.ac.uk), and in the U.S. the National Science Foundation (NSF) has recently established an Office of Cyberinfrastructure (www.nsf.gov).

One of the most definitive documents in this general trend is the Report of the Blue-Ribbon Advisory Panel on Cyberinfrastructure, generally known as the Atkins report after the study committee's chair (www.nsf.gov/od/oci/reports/atkins.pdf). It argues persuasively that cyberinfrastructure can not only improve the ability of scientists to do what they already do, but also underpin a new kind of science that is more computationally intensive, more collaborative, and more visual than before. The report distinguishes between two traditional kinds of science, one inductive and data driven and the other deductive and theory driven, and argues that cyberinfrastructure can enable a third kind that has elements of both, and that relies heavily on simulation. This vision has clear echoes of the comments made by Anselin and Getis in 1992, regarding the need for computationally intensive tools and better methods of visualization, and it reflects trends that have been evident within the social and environmental sciences for at least the past decade. It could be argued that this is a case of the computational tail wagging the substantive dog. But the Atkins report clearly leaves the role of defining needs to the substantive sciences and to the fundamental questions they need to answer.

In principle one might expect GIS to provide an ideal environment for this third kind of science. It is digital, and it allows rules and functions representing process to be simulated on the data contained in its databases. But in practice, as Anselin and Getis noted in 1992, the macro or scripting languages used by GIS to program complex tasks have had a reputation for being awkward to use and slow to execute. Traditional GIS software has been designed for the comparatively leisurely pace of analysis, rather than the computationally intensive pace of simulation.

However this situation has changed significantly in the past decade. Certain GIS packages, notably GRASS and the PCRaster package produced by the University of Utrecht (pcraster.geo.uu.nl) have been designed specifically to support simulation, and reference has already been made to recent efforts to link GIS to simulation environments such as Repast. Moreover, the comments in the previous section suggest that it is no

longer useful to ask whether some somewhat arbitrarily delimited type of software known as GIS can or cannot perform simulation – rather, the broader set of geospatial tools able to support simulation is now clearly rich and powerful (Maguire, Batty, and Goodchild, 2005).

Nevertheless, the community identified with GIS remains somewhat limited in its perspective, and has not yet built the kinds of bridges to larger communities focused on simulation, in domains such as atmospheric science, geophysics, or oceanography, all of which are clearly embedded in geographic space. Moreover, simulation technology is very advanced in domains that deal with other spaces, including aeronautical and structural engineering. As yet, no GIS vendor supports the representation of partial differential equations, either as finite difference or finite element approximations, and there is little support for simulation of processes in three spatial dimensions.

4.2 New methodologies

Cyberinfrastructure is only one manifestation of other, more fundamental trends in science. It argues for the use of information and communication technologies (ICT) to support collaboration between scientists, reflecting a general trend away from the single-investigator style of science to a more cooperative mode in which teams of specialists combine their expertise to solve complex problems. The software systems needed to support massive simulation of complex systems, such as those that underlie global climate models, are far too elaborate for any one person to know completely. To most scientists, computational tools will have to be black boxes, defined by their inputs and outputs rather than by their contents, and no one individual will be able to meet the traditional standard of scientific reporting: to provide sufficient detail to allow another scientist to replicate the experiment.

Unfortunately this problem is all too well known to scientists who use computational tools developed and marketed by the commercial sector, and particularly geospatial tools. The GIS industry has been driven largely by commercial applications, and science has always been only a small fraction of its market. Standards of documentation that are adequate for commercial applications may well fail the scientific test of providing sufficient detail for replicability. For example, it is frequently impossible to determine exactly how certain GIS functions operate, or to determine the degree of noise introduced in certain operations such as raster-vector conversion or projection change.

The widespread use of GIS is raising other issues of a methodological nature. Anselin and Getis refer to spatial dependence and spatial heterogeneity as the two defining characteristics of spatial data, noting the problems that the former causes for many methods of statistical analysis. Tobler's First Law of Geography (Tobler, 1970) asserts that positive spatial dependence is endemic in geographic data, in obvious violation of the independence assumption of many statistical tests, and ensuring that any significance test that results in the acceptance of a null hypothesis of no spatial dependence is almost certainly making a Type II statistical error.

Spatial heterogeneity raises even greater methodological issues, because it suggests that any attempt to find universal principles that apply everywhere on the Earth's surface is fundamentally problematic. Instead, analysis should focus on estimating and interpreting the inevitable variation in parameters, adopting a methodological position that is somewhere between the traditional nomothetic and idiographic extremes. So-called *local* or *place-based* analysis is more consistent with this position, and is now represented by a range of techniques that includes Anselin's LISA (Anselin, 1995) and the geographically weighted regression of Fotheringham, Brunsdon, and Charlton (2002).

Substantial progress has been made in building geospatial tools to support spatial analysis in the presence of both spatial dependence and spatial heterogeneity over the past 15 years. ESRI's ArcGIS supports a range of geostatistical techniques through its Geostatistical Analyst, and extensions are also available for the analysis of point patterns and other spatial statistical tests. Anselin's own GeoDa (geoda.uiuc.edu) has been developed as a stand-alone package but using standard GIS data structures, and has proven very popular (by late 2007 over 20,000 copies had been downloaded). The issues raised by Anselin and Getis in 1992 are now rapidly becoming the basis of standard practice.

5. WHOSE HAND ON THE TILLER?

Anselin and Getis concluded with comments about the need for regional scientists to play a central role in directing the future of GIS. As noted in the previous section, GIS has been a largely commercial product, and its development has been driven by its market, where the emphasis has been on such applications as forestry and utility management rather than on science. More effort has gone into providing rapid responses to simple queries from massive databases than into the kinds of sophisticated spatial analysis demanded by the social and environmental sciences.

The commercial nature of GIS is both a blessing and a curse, of course. On the one hand it is doubtful if GIS could have survived commercially based on the science market alone, and the existence of large commercial applications has provided access by scientists to a well-engineered, well-supported range of products. The distinction between commercial and science applications is also quite blurred, with many companies and agencies making use of the more sophisticated analytic functions of GIS and many universities using GIS to maintain inventories of their own physical plant. On the other hand, academics are traditionally leery of commercial motives, and as noted in the previous section there are clear differences between modes of operation in the commercial world and the norms of the scientific method.

Moreover, many significant contributions have been made by the non-commercial GIS sector, by such products as Idrisi, PCRaster, and GRASS. But these always run the risk of being attacked as unfair competition, and the early support of GRASS by the U.S. Army Corps of Engineers had to be withdrawn, at least in part for this reason.

Nevertheless, there are several reasons why the academic sector continues to exert a greater amount of influence on the future of GIS than its significance as a market would suggest. First, academics are also educators, and have a strong influence on the knowledge and predilections of future generations. A commercial GIS vendor looking to long-term success will clearly want to curry favor with the academic sector, by discounting software and providing other forms of support. If academics feel the need to educate students in particular techniques of spatial analysis then commercial GIS vendors may well choose to support those techniques, whether or not the market for them is viable. Sophisticated features also add to the perceived attractiveness of a product, whether or not they are actually used, just as many customers demand features in other products that they will never learn to use.

Second, the academic sector has a well-recognized duty to reflect and comment on all aspects of human society, and commentary on GIS and its impacts has become a significant subject of scholarship in geography, planning, and related disciplines. Academics have pointed to the importance of uncertainty, and the inherent vulnerability of GIS when uncertain results are used to regulate the use of land (Goodchild and Gopal, 1989); and they have questioned many of the assumptions underlying GIS analysis (Pickles, 1995). While the impacts of these critiques may not be apparent in GIS software products, they have undoubtedly altered the context in which GIS is used.

Finally, the academic sector is the primary source of fundamental innovation in GIS. The research domain variously known as geographic information science, geomatics, geovisualization, or geoinformatics has grown rapidly in the past decade, and has spawned journals, conferences, and organizations (www.ucgis.org). GIScience research has led to new insights into such fundamental issues as scale, new methods of representation that go far beyond the traditional GIS data models, and new techniques for addressing uncertainty. Such research increasingly provides the framework for new geospatial standards, and for improvements in user interface design.

6. CONCLUSION

In hindsight, Anselin and Getis provided a remarkably prescient analysis of the issues surrounding GIS and spatial analysis in the early 1990s. They accurately anticipated the growing need for tools that were more computationally intensive, with better support for advanced visualization, in a clear call for what today has become a much more broadly based interest in cyberinfrastructure. They also recognized the importance of spatial dependence and spatial heterogeneity as the defining characteristics of geographic data, and their implications in the need for new forms of analysis and new supporting tools.

The software landscape has changed dramatically since 1992. GIS is no longer a monolithic, stand-alone application, but instead encompasses a range of product types and a range of new functions, and new, broader terms such as geospatial have been adopted to try to capture this new complexity. The idea of a single toolbox has been replaced by the concept of interoperable software modules, operating in a mix of hardware architectures that ranges from the hand-held PDA (personal digital assistant)

through the desktop workstation to the remote server. New techniques of search allow researchers to discover and access these modules quickly, though much work remains to be done on appropriate catalogs and methods of description.

Like almost all of us, Anselin and Getis missed the massive transformations caused by the Internet and the Web that began in 1993 and still continue today, and the impacts that these transformations are having on the practice of science. They were correct, however, in arguing that it should be the basic questions addressed by science, and the needs of scientists for techniques to answer those questions, that should drive future development of tools. The GIS steam-engine has always been driven by a range of applications, only some of which are concerned with fundamental science questions, and the needs of researchers have always had to battle with the needs of other users. Anselin's own GeoDa is an excellent example of a feasible strategy in this environment – a package for sophisticated spatial analysis that is designed to interoperate with standard GIS products, but is designed and programmed by a team of scientists. But it is difficult for the academic marketplace to provide an income stream, and such projects must therefore be funded by grants, an erratic source at best. Software produced by academic teams is rarely engineered to the same standards as commercial software, and it is difficult to provide the same kinds of support.

In short, while much has happened in the decade and a half since the publication of the Anselin and Getis paper, its institutional context remains much as it was in 1992. The commercial GIS sector is large and growing, and able to produce and support complex "industrial-strength" software. The academic sector has only a limited ability to influence the commercial sector's directions, in integrating the kinds of tools needed to support research, and in ensuring adherence to the norms of the scientific method. As in many other fields, it is clear that innovative partnerships represent the best way forward, integrating the work of the two sectors and allowing their different objectives to be harmonized.

REFERENCES

Abler, R.F., 1987. The National Science Foundation National Center for Geographic Information and Analysis. *International Journal of Geographical Information Systems* 1: 303-326.

Anselin, L., 1995. Local indicators of spatial association -- LISA. *Geographical Analysis* 27: 93-115.

Anselin, L. and A. Getis, 1992. Spatial statistical analysis and geographic information systems. *Annals of Regional Science* 26: 19-33.

Arctur, D. and M. Zeiler, 2004. *Designing Geodatabases: Case Studies in GIS Data Modeling*. Redlands, CA: ESRI Press.

Burrough, P.A., 1990. Methods of spatial analysis and GIS. *International Journal of Geographical Information Systems* 4: 221-223.

Cowen, D.J., 1988. GIS versus CAD versus DBMS: what are the differences? *Photogrammetric Engineering and Remote Sensing* 54: 1551-1555.

Crosier, S.J., M.F. Goodchild, L.L. Hill, and T.R. Smith, 2003. Developing an infrastructure for sharing environmental models. *Environment and Planning B: Planning and Design* 30: 487-501.

Ding, Y. and A.S. Fotheringham, 1992. The integration of spatial analysis and GIS. *Computers in Environmental and Urban Systems* 16: 3-19.

Fischer, M.M., 1998. Computational neural networks -- a new paradigm for spatial analysis. *Environment and Planning A* 30(10): 1873-1892

Fotheringham, A.S., C. Brunsdon, and M. Charlton, 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Hoboken, NJ: Wiley.

Fischer, M.M. and Y. Leung, 1998. A genetic-algorithms based evolutionary computational neural network for modelling spatial interaction data. *Annals of Regional Science* 32: 437-458.

Fotheringham, A.S. and P. Rogerson, editors, 1994. *Spatial Analysis and GIS*. London: Taylor and Francis.

Goodchild, M.F., 1987. A spatial analytical perspective on geographical information systems. *International Journal of Geographical Information Systems* 1: 327-34.

Goodchild, M.F. and S. Gopal, 1989. *Accuracy of Spatial Databases*. Basingstoke, UK: Taylor and Francis.

Goodchild, M.F., R.P. Haining, and S. Wise, 1992. Integrating GIS and spatial analysis: problems and possibilities. *International Journal of Geographical Information Systems* 6: 407-23.

Longley, P.A., M.F. Goodchild, D.J. Maguire, and D.W. Rhind, 1999. *Geographical Information Systems: Principles, Techniques, Management, and Applications*. New York: Wiley.

Maguire, D.J., 1991. An overview and definition of GIS. In D.J. Maguire, M.F. Goodchild, and D.W. Rhind, editors, *Geographical Information Systems: Principles and Applications*. Harlow, UK: Longman Scientific and Technical, pp. 9-20.

Maguire, D.J., M. Batty, and M.F. Goodchild, 2005. *GIS, Spatial Analysis, and Modeling*. Redlands, CA: ESRI Press.

Maguire, D.J. and J. Dangermond, 1991. The functionality of GIS. In D.J. Maguire, M.F. Goodchild, and D.W. Rhind, editors, *Geographical Information Systems: Principles and Applications*. Harlow, UK: Longman Scientific and Technical, pp. 319-335.

Maling, D.H., 1989. *Measurements from Maps: Principles and Methods of Cartometry*. New York: Pergamon.

Mapping Science Committee, National Research Council, 1993. *Toward a Coordinated Spatial Data Infrastructure for the Nation*. Washington, DC: National Academy Press.

Mitchell, A., 1999. *The ESRI Guide to GIS Analysis*. Redlands, CA: ESRI Press.

Nebert, D., 1993. Implementation of Wide-Area Information Server (WAIS) software to disseminate spatial data on the Internet. *Proceedings, International ESRI User Conference, Palm Springs, CA*.

Nyerges, T.L., 1993. Understanding the scope of GIS: its relationship to environmental modeling. In M.F. Goodchild, B.O. Parks, and L.T. Steyaert, editors, *Environmental Modeling with GIS*. New York: Oxford University Press, pp. 75-93.

Openshaw, S., 1990. Spatial analysis and geographical information systems: a review of progress and possibilities. In H.J. Scholten and J.C.H. Stillwell, editors, Geographical Information Systems for Urban and Regional Planning. Dordrecht: Kluwer, pp 153-163.

Openshaw, S., M. Charlton, C. Wymer, and A. Craft, 1987. A Mark I geographical analysis machine for the automated analysis of point data sets. *International Journal of Geographical Information Systems* 1: 335-358

Openshaw, S., M. Charlton, and A. Craft, 1988. Searching for leukaemia clusters using a geographical analysis machine. *Papers of the Regional Science Association* 64: 95-106

Openshaw, S., A. Cross, and M. Charlton, 1990. Building a prototype geographical correlates exploration machine. *International Journal of Geographical Information Systems* 4: 297-311.

Pickles, J., editor, 1995. *Ground Truth: The Social Implications of Geographic Information Systems*. New York: Guilford.

Skupin, A. and R. Hagelman, 2005. Visualizing demographic trajectories with self-organizing maps. *GeoInformatica* 9(2): 159-179.

Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46: 234-240.

Ungerer, M.J. and M.F. Goodchild, 2002. Integrating spatial data analysis and GIS: a new implementation using the Component Object Model (COM). *International Journal of Geographical Information Science* 16(1): 41-54.

Zeiler, M., 1999. *Modeling Our World: The ESRI Guide to Geodatabase Design*. Redlands, CA: ESRI Press.

FIGURE CAPTIONS

1. Functions of a GIS (Figure 1 from Anselin and Getis, 1992).