

GIS and Spatial Analysis:
Report on the Specialist Meeting

Edited by

A. Stewart Fotheringham and Peter Rogerson

National Center for Geographic Information and Analysis

Report 92-11

COPYRIGHT

Copyright, © National Center for Geographic Information and Analysis

June 1992

Acknowledgments

We wish to thank Andrew Curtis, Rusty Dodson, Sheri Hudak, and Uwe Deichmann for the notes which they took during the Specialist Meeting, and Andrew Curtis and Connie Holoman for their assistance in preparing this report.

OVERVIEW OF THE INITIATIVE

History of the Initiative

A proposal for a NCGIA-sponsored initiative on GIS and Spatial Analysis was first submitted to the Scientific Policy Committee of the NCGIA on March 17, 1989. It was formally resubmitted on June 2, 1991 after being divided into separate proposals for initiatives on "GIS and Statistical Analysis" and "GIS and Spatial Modeling". The essence of the former of these two proposals was accepted and evolved into the more generic "GIS and Spatial Analysis" initiative that took place with the expectation that an initiative emphasizing spatial modeling would take place at a later date.

The objectives of the initiative are in keeping with the aims of the National Center, as identified in the original guidelines from the National Science Foundation. The solicitation for a National Center for Geographic Information and Analysis circulated by the National Science Foundation in 1987 contained as one of its four goals to

advance the theory, methods, and techniques of geographic analysis based on geographic information systems in the many disciplines involved in GIS research. (National Science Foundation, 1987, p. 2)¹

The solicitation also notes that the research program of the NCGIA should address five general problem areas, including "improved methods of spatial analysis and advances in spatial statistics."

The initial impetus for Initiative 14 is attributable at least in part to this initial call made by the National Science Foundation.

Rationale

Given that there is an increasing trend for GISs to promote their ability to assist in the **analysis** of spatial data, as a complement to their traditional roles as storage and display devices, it is important that the NCGIA be concerned with removing the impediments to the accurate use of spatial analysis within a GIS. Just as the widespread availability of statistical software such as SPSS in the 1960s and '70s led to abuses in spatial statistical analyses, so the availability of spatial analytical modules developed for GISs in the '90s could ultimately discredit the use of GISs for this purpose. We therefore see it as an important mission for the NCGIA to contribute to a better understanding of the role of spatial analysis within a GIS. We feel it is imperative that the NCGIA take a lead in encouraging the **accurate** use of spatial analytical routines within a GIS environment: there are virtually no impediments to their **inaccurate** use, which is really the crux of the problem.

The initiative is based on the premise that the analysis of geocoded data usually invokes special considerations not found in aspatial analysis. Examples include the sensitivity of aggregate data analysis to the definition of data reporting units, edge or boundary effects, and spatial dependency and the definition thereof. A GIS has the potential to highlight the existence of these and other problems but also to perhaps shed some light on potential solutions. The initiative is therefore centered around **two** general questions:

¹National Science Foundation. 1987. National Center for Geographic Information and Analysis. Directorate for Biological, Behavioral, and Social Sciences. Guidelines for Submitting Proposals.

1. How can spatial analysis assist GIS?

2. How can GIS assist spatial analysis?

To give a sense of the direction in which the initiative is headed, listed below are several **representative** (certainly not **exhaustive**) specific questions given to participants prior to the Specialist Meeting:

1. What restrictions are placed on spatial analysis by the modifiable areal unit problem and how can a GIS help in better understanding this problem?
2. How can GIS assist in exploratory data analysis and in computer-intensive analytical methods such as bootstrapping and the visualization of results?
3. How can GIS assist in performing and displaying the results of various types of sensitivity analysis?
4. How can the data structures of a GIS be exploited in spatial analytical routines?
5. What are needs in terms of a user interface and language for spatial analysis performed on a GIS?
6. What are some of the problems in spatial analysis that should be conveyed to a GIS user and how should these problems be conveyed?

Another facet of the initiative conceived at an early stage was its focus on substantive applications in the social sciences. There is indeed a strong interaction between GIS and spatial analysis in the physical sciences as well (see, *e.g.*, the abstracts of the special session organized on GIS and Spatial Analysis in Hydrologic and Climatic Modeling at the most recent annual meeting of the AAG in San Diego), but given the nature of both the division and directorate within NSF receiving the funding, and the proclivities of the participants, a specific focus on the social sciences was deemed appropriate.

Although somewhat blurred, the division between statistical analysis and spatial modeling is nonetheless convenient. The emphasis of this Initiative was on Spatial Data Analysis (SDA), implying statistical analysis, although approximately a quarter of the academic participants were spatial modelers, and one of the four working groups focused on Locational Analysis and Planning in GIS. Out of this group grew a strong vocal request for an initiative devoted solely to the issue of Spatial Modeling and GIS.

Management Issues

The Specialist Meeting of this initiative was held at Humphrey's Half Moon Inn in San Diego between April 15-18, 1992. The location and time were chosen to coincide with the Association of American Geographers Annual Conference. This saved considerably on NCGIA-related travel and it also meant we were able to attract a very good set of participants without offering travel expenses. All local costs were covered by the NCGIA. The superb facilities of the Half Moon Inn contributed to the success of the meeting. Approximately \$2,000 was donated by the Mathematical Models Commission of the International Geographical Union towards the cost of the Specialist Meeting and we are very appreciative of this financial support.

A list of participants at the meeting is provided as an appendix to this report, as is the agenda for the meeting. Of the 40 participants, 26 were academics, the rest being from government agencies and the private sector (GIS and statistical software vendors).

SPECIALIST MEETING SUMMARY

A full paper was required of all academic participants and all complied. Papers will be selected from these, and will be edited into a book in the near future. Based on the contents of these papers, the academic participants were divided into four working groups for days 2 and 3 of the meeting:

Linkages Between Spatial Analysis and GIS

Spatial Data Analysis and GIS

Pattern Recognition, Complexity, and Data Models

Locational Analysis and Planning Applications with GIS

Summaries from each of these four groups are presented later in this document. Day 1 of the meeting consisted of 6 selected papers (Bailey, O'Kelly, Haining, Bracken, Openshaw, Densham) which we felt provided a good basis for further discussion at the meeting and which also provided a common ground for the participants.

Report of the Working Group on Linkages Between Spatial Analysis and GIS

The group began by discussing the relative merits of embedding models and methods of spatial analysis directly within GIS. Such a scheme obviates the need for mastering several software packages, and has the additional advantage of increased speed. However, many techniques of spatial analysis are "cutting-edge", and are not well-known. Such techniques are less likely to be adopted by the private sector market. In these instances, it may be more desirable to link the spatial analysis and GIS systems, and let each do what it does best.

One linkage model would be to use the GIS as a server to store and dispense spatial data to separate, stand-alone spatial analysis functions. Designs such as this would be quite flexible, and would be easier to manage than single, all-encompassing packages. There was some discussion, however, of the need to improve the current GIS input-output "hooks" required for such linkages. Improved hooks would free the academic community from having many tedious programming tasks, and would allow researchers to spend more time on developing the analytical methods themselves. One specific way of improving hooks would be through the addition of an elementary matrix data structure within GIS that would allow additional analysis and display of, *e.g.*, interaction data.

The group talked about the need for a common spatial modeling language. While such a language would provide a common mode of communication, there was no agreement within the group whether such a language should be made into a formal standard, or indeed whether it would ever be able to accommodate the broad range of spatial methods and applications that exist. It was suggested that a three-way taxonomy of spatial operations would be useful: (a) application types, (b) spatial data models/structures, and (c) spatial analysis categories.

The question of whether the Spatial Data Transfer Standard (SDTS) should be adopted was addressed. There were two points of view:

- (1) Yes. This is a national standard; it will be adopted by all federal agencies, and it is likely that vendors will adopt it also.
- (2) No. The standard is too broad and vague. In addition, the purpose of SDTS is to exchange spatial data sets in their entirety. In contrast, the linkage between spatial analysis and GIS often requires the transfer of selected pieces of information from datasets, and the transfer of other data calculated and created from the original data. SDTS may not cover these different circumstances.

Other thoughts and points of discussion:

- (1) Do specialized linkages help to disseminate methods, or do they detract from the higher goal of universal linkage and compatibility?
- (2) Two alternatives for proceeding with the integration of linkages and spatial modeling languages: (a) develop a complete specification of the interface and all functionality, or (b) build small prototypes first.

On the last day of the meeting, this group argued for the development of five prototype models, each of which would be useful in furthering progress in the areas described above. The five prototypical models were chosen to focus upon a broad array of problem types. They are:

- (1) time series
- (2) epidemiology—to include examples making use of point patterns and K-functions
- (3) routing and allocation—to have an application involving networks
- (4) hydrology—to focus upon a problem with 3-D characteristics
- (5) aggregation and disaggregation of areal data

Each of the prototypes would be designed to gain insights on how GIS should evolve for that particular area of analysis. Specifically, through these prototypes one would attempt to discern desirable forms of linkage, desirable data structures, whether the analysis itself is improved via the link to GIS, *etc.* This would be accomplished in part by including for each prototype a list of data structures required, a list of techniques required by the analysis, and by creating a taxonomy of GIS global variables, data structures and functions supporting the analysis for each data model.

Report of the Working Group on Spatial Data Analysis and GIS

First Meeting - General Topics

1. What kinds of spatial data analysis do we want to do in a GIS framework, *e.g.*, exploratory versus confirmatory analysis? Much of the subsequent discussion considered SDA as primarily exploratory although the group agreed that the distinction was blurred.
 - (a) Not problematic to produce a shopping list (*c.f.* Bailey's paper) but we need to consider types of data models.
 - (b) Is the link to a GIS much of a problem technically? *c.f.* papers by:
Ding and Fotheringham
Can

Kehris, Rowlinson
Anselin *et al.*

2. Is GIS an appropriate vehicle for SDA? Is it more appropriate for some types of SDA than for others?
3. What kinds of SDA can we do in a GIS framework that cannot be done without a GIS?
 - (a) exploiting multiple layers (coverages) in a database
 - (b) developing spatial equivalents of aspatial methods, *e.g.*, spatial principal components (Cressie)
 - (c) developing multivariate equivalents of univariate methods, *e.g.*, multivariate spatial autocorrelation?
 - (d) modeling relations between different classes of objects
 - (e) using other than Euclidean geometries
 - (f) spatial windowing to exploit power of GIS
 - computing statistics "on the fly" (in real time)
 - correlograms
 - variograms
 - k-functions
 - spatial heteroskedasticity descriptors, *e.g.*, G statistics
 - (g) exploiting GIS power for sensitivity analyses
 - sensitivity of spatial autocorrelation to definition of spatial weight matrix
 - sensitivity of results to zone definition
4. Can we embed econometric and geostatistical models into a GIS framework?
5. How can we best embed space-time models in a GIS?
6. How do we ensure we don't lose sight of the geography? How do we retain our geographical knowledge in the light of masses of information being thrust upon us?

Second Meeting - Specific Topics

1. Produce a taxonomy of ESDA (Exploratory Spatial Data Analysis) techniques with respect to:
 - data models/structures
 - applications
 - visualization
 - exploratory plotting
 - summary statistics
2. How do we detect **spatial** outliers? How do we determine if an observation is sufficiently different from its "neighbors"? When does a cluster of outliers become evidence of non-stationarity?
3. How do we assess the impact of spatial outliers on inference?

4. We need increased breakdown (spatially) of global statistics, *e.g.*,
 - Getis' Gi statistic
 - Pocket plots
 - Variogram clouds

These "place-specific" statistics then lend themselves much more to mapping.

5. Need research on cognitive aspects of Visual Interactive Modeling on a GIS. How do users process information on a screen? What forms of visual display are most meaningful?
6. Need to develop multivariate equivalents to univariate statistics
7. Need new ways of visualizing multivariate data and the results of multivariate analyses.
8. Exploit power of a GIS to perform data-driven sensitivity analyses, *e.g.*, Fotheringham and Wong's work on the modifiable areal unit problem and multivariate statistical analysis. Other areas where sensitivity analysis is needed include spatial interpolation, density estimation, measurement of accessibility, spatial autocorrelation and spatial association.
9. Need research on space-time statistics
 - continuous statistics
 - exploratory/confirmatory
 - monitoring in real time
 - pattern recognition
10. Research into map transformations as a means of displaying information
 - cartograms
 - time-space maps

Report of the Working Group on Pattern Recognition and Complexity

The group on pattern recognition and complexity spent part of the first day discussing the different ways in which complexity had a bearing upon the relationship between spatial analysis and GIS. At least three different definitions of complexity were discussed:

- (1) complexity in modeling—issues associated with large-scale models, and the linkages between their component parts,
- (2) complexity in space-time modeling—issues associated with pattern formation and evolution in nonlinear dynamic spatial models, cellular automata, *etc.*, and
- (3) complexity emerging from the interface between spatial analysis and GIS—issues associated with large databases, and the complex linkages between spatial analysis and GIS that make it difficult to actually use GIS for spatial modeling.

There was also discussion of the importance of improving methods for pattern recognition. This included the potential for matching observed patterns to "reference sets" and the relative merits of using artificial intelligence and neural net approaches to pattern detection.

There was also discussion of the need for multimedia approaches to the representation of patterns and complexity, including interactive and animated representations.

Specific questions and ideas suggested by members of the subgroup included the following:

- (1) How can GIS be configured to support simulation modeling? What data must be transferred in and out of the GIS at each iteration of a model?
- (2) What measures of spatial pattern and complexity can be integrated with GIS?
- (3) There is a need for measures and measurement of complexity found in GIS databases, map layers, *etc.*
- (4) How can "flight simulator" type data modules be developed for theoretical geography?
- (5) How can cartograms and GIS best be developed and integrated?
- (6) Pursue further the integration of neural net and expert systems with conventional GIS.
- (7) Explore the potential and current limitations of artificial neural networks for spatial modeling.
- (8) How can GIS change the way that we think about specific types of modeling (*e.g.*, modeling and visualization of flows in spatial interaction models)?
- (9) Develop good links between GIS data storage and display facilities and such modeling capabilities as estimation, forecasting, and sensitivity analysis.
- (10) How can artificial intelligence (neural networks, genetic algorithms, *etc.*) be integrated with a GIS used as a modeling environment, so that AI pattern recognition can be used interactively with the model? How can patterns like clusters be identified in real time to be then created as objects in a simulation model, as the simulation model runs?
- (11) There is a need for projects detailing what a GIS environment for explanatory (as opposed to prescriptive) spatial models should look like.
- (12) Evaluate the potential for modeling at the level of the individual within GIS. Capture the position and attributes of individuals, and assess the potential impediments (*e.g.*, coping with the data, visualizing results, *etc.*) in constructing such systems.
- (13) Assess different methods of knowledge representation and methods (*e.g.*, fuzzy logic) to evaluate spatial hypotheses.
- (14) Develop approaches to pattern recognition using the methods of mathematical morphology.
- (15) Develop an interactive system for auto/cross correlation and variograms as an aid to the evaluation of local spatial correlation.

On the last day of the meeting, this group produced a list of researchable topics, synthesized in part from the above:

Potential Researchable Topics

- (1) Efforts to make GIS more usable and user friendly for spatial modeling
 - development of "demonstration projects" aimed at determining

- the extent to which existing GIS can be used or adapted for spatial modeling purposes, and
- whether some new form of GIS is needed, and if so, what kind?
- investigation of the differences between model language and GIS language

These demonstration projects would be carried out at different levels of scale and aggregation—including the level of the individual, census tract, and grids of various sizes. Different model types would be used, including dynamic simulation models, cellular automata models, and systems models based on interaction. Particular focus should be given to models where there is a great deal of interaction between the submodels and the units of analysis. Examples here include the interaction between spatial interaction models, economic input-output models, econometric models, *etc.*

(2) Pattern recognition projects

- matching of observed spatial patterns against reference sets, using, *e.g.*, associative memory neural nets in 3 subareas:
 - map pattern recognition
 - space pattern recognition
 - space-time pattern recognition
- development of an interactive pattern recognition module focused upon auto/cross correlation with a movable window
- develop new forms of spatial systems modeling using neural nets, machine learning, *etc.* in the following areas:
 - spatial response modeling
 - urban models of various kinds, and
 - transportation and flow modeling
- a suggestion was made for a workshop on complex dynamics and evolutionary systems in view of the close links between spatial complexity and the facilities of GIS

(3) Topics related to visualization and animation, *e.g.*, the integration of multimedial information, (*e.g.*, visual and auditive information into GIS)

(4) Investigation of the potential for, and the implications of "international analysis servers", where there would be distributed computing in analysis, making only one piece of software at one site necessary.

(5) The further development and assessment of fuzzy logic and Bayesian inference and its extensions, to evaluate spatial hypotheses, *e.g.*, in mineral exploration.

Report of the Working Group on Locational Analysis and Planning Applications with GIS

The discussion of this working group is presented in three parts:

1. a brief summary of the major areas in which further research is needed;
2. a list of topics for further research; and
3. an edited transcript of discussions (presented in Appendix C)

1. Major Areas for Research

1) Exploratory model-building cycles

- : improving the way in which models are formulated. Assisting in the inductive-deductive loop.
- : improving prospects for large-scale (data intensive) urban model-building.
- : manipulation of spatial objects as an aid to modeling and algorithmic development.
- : a drive toward operational urban models that can be manipulated by users.
- : making model-building assumptions explicit. Sensitivity analysis to assist understanding of assumptions and their relaxation.
- : development of dynamic modeling capabilities within GIS.

2). Visualization of Models

- : sensitivity analysis for models.
- : pedagogic developments in the spirit of Sim City.
- : visualization of model results as an aid to model development.
- : spatial aspects of goodness-of-fit.

3) Scale and aggregation issues

- : Is the areal interpolation issue resolvable in a GIS?
- : Diagnostic indication of possible/likely scale/aggregation effects in urban models (*e.g.*, redistricting, distance-decay effects, location-allocation solutions)

2. Specific Topics for Further Research

The ‘A’ List

- 1) The derivation of “applications generators”—objects of code that can be manipulated within different algorithms. An object-oriented programming approach to spatial modeling.
- 2) Derivation of dynamic simulations of urban processes, perhaps utilizing parallel processing architecture. Developing simulation libraries.
- 3) Investigation of optimization routines and data restructuring. Looking at the effects of zone variations on model outputs—sensitivity analysis with respect to scale and aggregation.
- 4) Finding ways to deal with conflicting reporting units for data.

The 'B' List

- 1) Exploratory modeling using operations research methodologies.
- 2) Coupling of location theory (*e.g.*, von Thunen) and GIS.
- 3) Restructuring economic models to include space more realistically. Economic models are often technically sophisticated except in their treatment of space.
- 4) Looking at classification of spatial units with a GIS. For example, small area clusters of similar characteristics for marketing. Sensitivity of clustering results to scale and aggregation.
- 5) Developing fuzzy analytical methods.

The 'C' List

- 1) Comparing potential and actual uses of OR/GIS tools. What methods are/would be used by different institutions?
- 2) Allowing user-defined metrics in models. Moving away from Euclidean metrics.
- 3) Examining the cognitive side of GIS. What type of solutions do users want? What type of visual aids are most effective in conveying model results? How do users process different kinds of visual images? How do users process information in various circumstances, *e.g.*, elimination by aspects versus compensatory?

SYNTHESIS

Despite a spatial database being an essential component of GIS, the explosive development of GIS within the last decade has largely taken place with little reference to the **analysis** of spatial data, apart from simple querying and overlaying routines. The inevitable integration of spatial analysis and GIS prompted the following two questions posed at the beginning of this paper:

1. **How can spatial analysis assist GIS?**
2. **How can GIS assist spatial analysis?**

The first of these two questions is the least interesting academically. There are obvious advantages to a GIS in adding spatial analytical capabilities in terms of increasing functionality and meeting a demand for systems that do something beyond storing, retrieving and displaying large amounts of information. There is also a competitive advantage to those systems that are at the forefront of offering spatial analytical capabilities; an advantage that will probably become increasingly important as the GIS market becomes saturated. The first of the four themes at the specialist meeting, **Linkages between Spatial Analysis and GIS**, explored the technical issues involved in the integration of GIS and spatial analysis.

The second question, we believe, poses a more interesting and challenging task which is to define the advantages to spatial analysis of running spatial analytical routines on a GIS. This question was examined in three areas: **Spatial Data Analysis and GIS; pattern Recognition, Complexity, and Data Models;** and **Locational Analysis and Planning Applications with GIS**. The latter discussion involves spatial modeling and GIS and really only scratched the surface of a very detailed topic that should form the basis for a separate NCGIA

Research Initiative. Discussion on the second topic seemed very diffuse, perhaps because of the complex nature of the problem. By far the most coherent of the discussions came under the first topic concerning the integration of spatial data analysis with GIS and much of that discussion was focussed on the topic of exploratory spatial data analysis (ESDA) and GIS. ESDA seems particularly well-suited to a GIS environment: ESDA can be characterized as "data-hungry" and primarily visual, whereas GIS are "data-rich" and have excellent display capabilities. There are clearly a number of interesting research questions within this topic such as the identification of spatial outliers, the utility of sensitivity analysis and the development of multivariate ESDA routines.

FOLLOW-UP ACTIVITIES

Many of the benefits derived from activities that follow on from the Specialist Meeting are, of course, unmeasurable. The Specialist Meeting brings researchers interested in a topic together, and it is inevitable that both the creative thoughts of individual participants during sessions and the innumerable informal conversations between sessions will lead to both individual and collaborative research. Without soliciting detailed, updated vitae and/or quarterly reports from all participants, the benefits of such fertilization and cross-fertilization will be difficult to evaluate.

Other benefits are more tangible. There are numerous activities in various planning stages that are meant to follow-up on the Specialist Meeting. These include both the individual research activities of the initiative co-leaders and participants, and organizational and publishing activities aimed at a broad group.

Individual research activities include:

- (1) Developing spatial analytical software for use within a GIS. Examples include Ding and Fotheringham's Spatial Analysis Module (SAM) for ARC/INFO; Anselin's Spacestat; and Xia and Fotheringham's ESDA module for ARC/INFO.
- (2) Nonparametric tests for pattern detection and its use in GIS.
- (3) The use of GIS for sensitivity analysis, particularly in examining the modifiable areal unit problem both in statistical analysis and in location/allocation modeling;
- (4) The development of a language for spatial analysis within a GIS.
- (5) The development of ESDA routines for multivariate data.

Organizational and publishing activities and objectives include:

- (1) Publishing an edited collection of selected papers prepared for the Specialist Meeting. Three publishers have expressed an interest in publishing an edited collection of papers from the meeting. We hope to have final details sorted out by the end of June, 1992. There is also the possibility of a special issue of a refereed journal containing papers prepared for the meeting.
- (2) Organization of special sessions on GIS and Spatial Analysis for the annual meeting of the Regional Science Association in November, 1992. Manfred Fischer has already organized one session for these meetings, and we are exploring the possibility of organizing another.

- (3) A summary will be presented at the IGU meetings in August, 1992 to publicize the research agenda to a wider audience. Sessions are also being planned for the RSA meetings in November, 1992, and the ORSA/TIMS meetings in October, 1992.
- (4) Organization of a special session on GIS and Spatial Analysis for the annual meeting of the AAG in April, 1993.
- (5) Other initiatives have had NATO Advanced Study Institutes as part of their follow-up activity, and this possibility will be investigated for I-14. This could form a useful 'closure' for the initiative if held in the summer of 1994.
- (6) A newsletter issued approximately twice per year will be coordinated by Art Getis, at San Diego State University. The newsletter will serve to update Specialist Meeting participants and others interested in the initiative on recent initiative activities and ongoing research, as well as to notify people about upcoming organized sessions on I-14-related topics.
- (7) Visits by participants [or others interested in the initiative] to the Santa Barbara and Buffalo sites. Both sites have facilities for accommodating researchers wishing to focus on research related to the initiative. Researchers typically spend from a week to a month at a site. Visits to the Center afford the opportunity for an intensive, focused period of research effort, and the opportunity for collaboration with Center members. The Center provides office and secretarial services (and, in some cases, may be able to provide a small stipend to offset expenses.).
- (8) A related initiative more focused on spatial modeling is in the planning stages. As noted in the introduction, it was recognized early in the planning process for I-14 that while there was an urgent need to integrate GIS with both spatial modeling and spatial statistics, advances to date, as well as likely short-run advances, would come primarily in the integration of spatial statistics and GIS. Consequently, the attention focus of I-14 was on spatial statistics. A follow-up initiative on both the process of spatial modeling and applications of spatial modeling to the substantive social science areas would complement I-14 very well.

Appendix A: Schedule for the Specialist Meeting

Wednesday, April 15

5 - 6:30 PM: Welcoming Reception: The Gazebo Room

Thursday, April 16

8:45 - 9:15 AM: Opening Remarks: Marina Ballroom
Stewart Fotheringham, Peter Rogerson, Waldo Tobler, Mike Goodchild

9:15 - 10 AM: Linkages Between Spatial Analysis and GIS
Speaker:
Trevor Bailey
Department of Mathematical Statistics and Operational Research
University of Exeter
“A Review of Statistical Spatial Analysis in Geographical Information Systems”

10 - 10:45 AM: Linkages Between Spatial Analysis and GIS
Speaker:
Morton O'Kelly
Department of Geography
The Ohio State University
“Spatial Analysis and GIS: A Position Paper”

11:15 - 12 noon: Spatial Data Analysis in GIS
Speaker:
Robert Haining
Department of Geography
University of Sheffield
“Designing Spatial Data Analysis Models for Geographical Information Systems”

12:30 - 1:30 PM: Lunch and Discussion

1:45 - 2:30 PM: Locational Analysis and Planning Applications with GIS
Speaker:
Ian Bracken
Department of City and Regional Planning
University of Wales, Cardiff
“A Surface Model Approach to the Representation of Population-Related Social Indicators”

- 2:30 - 3:15 PM: Pattern Recognition, Complexity, and Data Models
 Speaker:
Stan Openshaw
Centre for Urban and Regional Development Studies
The University of Newcastle upon Tyne
“Two Exploratory Space-Time-Attribute Pattern Analysers
Relevant to GIS”
- 3:45 - 4:30 PM: Locational Analysis and Planning Applications with GIS
 Speaker:
Paul Densham
Department of Geography
SUNY at Buffalo
“Integrating GIS and Spatial Modelling: The Role of Visual
Interactive Modelling in Location Selection”
- 4:30 - 5:30 PM: Informal panel discussion with private sector and agency participants
- 5:30 - 5:45 PM: Discussion of Friday's activities

Friday, April 17

- 9 - 10:45 AM: Working groups (See accompanying page for assignments):
 •Linkages Between Spatial Analysis and GIS :
 •Spatial Data Analysis in GIS :
 •Pattern Recognition, Complexity, and Data Models :
 •Locational Analysis and Planning Applications with GIS :
- 11:15 AM - 12:15 PM: Continuation of working groups
- 12:30 - 1:30 PM: Lunch and discussion
- 2 - 3:15 PM: Presentation of working groups' findings: Marina Ballroom
- 3:30 - 5:00 PM: Continuation of presentation of working groups' findings and discussion of groups for Saturday

Saturday, April 18

9 - 10:45 AM:	Working groups (same as Friday):
11:15 AM - 12:15 PM:	Continuation of working groups
12:30 - 1:30 PM:	Lunch and discussion
2 - 3:15 PM:	Presentation of working groups' findings: Marina Ballroom
3:45 - 5:30 PM:	Continuation of presentation of working groups' findings and wrap-up

Working Group Assignments for Friday, April 17

Assignments to working groups were made only for academic participants. Agency and private sector participants were free to attend the working group(s) of their choice.

•Linkages Between Spatial Analysis and GIS

Trevor Bailey
Mike Goodchild
Morton O'Kelly
Gerard Rushton

•Spatial Data Analysis in GIS

Ayse Can
Robin Dubin
Noel Cressie
Tony Gatrell
Art Getis
Robert Haining

•Pattern Recognition, Complexity, and Data Models

Mike Batty
Manfred Fischer
Stan Openshaw
Waldo Tobler
Roger White

•Locational Analysis and Planning Applications with GIS

Ian Bracken
Paul Densham
Robin Flowerdew
Paul Longley
Bill Macmillan
Bruce Ralston

Appendix B: Abstracts of Papers Prepared for the Specialist Meeting

A Review of Statistical Spatial Analysis in Geographical Information Systems[†]

Trevor C. Bailey

*Department of Mathematical Statistics and Operational Research
University of Exeter*

Two constantly recurring themes in the GIS literature are on the one hand, the potential that GIS technology should afford for more easily accessible and more sophisticated analysis of spatially referenced data; and on the other, the limited functionality of the standard analysis options provided in many commercial GIS packages and their inadequacy for research in many disciplines. This gap between what one feels the technology should be able to provide in the way of analysis and what is actually on offer has become something of a cliché. This paper is a review of the potential that exists and the progress that has been made, in exploiting the capability of GIS in the statistical analysis of spatial data. The paper identifies eight groups of statistical spatial analysis methods considered to be those that are most generally useful in relation to GIS. The potential benefits of close linkage between each of these groups and GIS are then analyzed in detail, together with the progress that has been made towards realizing such benefits. Notwithstanding several recent research initiatives, it is concluded that many of the potential benefits remain to be exploited. Various reasons which may account for this are analyzed and possible directions for developments in this area are suggested.

Urban Models in Computer-Graphic and Geographic Information System Environments[†]

Michael Batty

*National Center for Geographic Information and Analysis, and Department of Geography
State University of New York at Buffalo*

Geographic information systems (GIS) are beginning to have an impact not only upon the ways in which decision and policy problems are resolved but also on the ways analysts and model-builders can articulate and communicate their intellectual understanding of spatial systems. Because the wave of interest in GIS is so great, serious attempts are now being made to couple such systems to a diverse set of spatial-analytic models and methods which hitherto have been developed without such systems in mind. GISs are in fact just one from a wider set of techniques which enable the development of tools for visualization and in this paper, we will describe one such development within which traditional urban models can be extended and made more relevant to policy makers.

We begin by sketching the context to such work, outlining the important issues involved in visualization in this domain. We first define the elements of such visualization in terms of the model-based processes which characterize applications, emphasizing data exploration, model calibration, prediction and prescription; and then we concentrate these processes through model inputs, outputs and causal structures illustrating the operation of various model sectors in terms of different modes of spatial statistical and graphic analysis. These ideas are implemented through windowing systems which mix text, numerics and graphics.

[†] Indicates abstract written by the original author of the paper. Other abstracts were prepared by the editors of the Technical Report.

We then illustrate some of these ideas with an elementary model of residential location and show a typical run of the model in terms of its operation in a UNIX environment which we access through proprietary windows-based software. This enables us to intelligently discuss the problems of linking conventional model-based techniques to graphic software in general, GIS in particular. This approach is currently being extended to a model of the Buffalo-Niagara region and the paper concludes with a sketch for future work.

A Surface Model Approach to the Representation of Population-Related Social Indicators

*Ian Bracken
Department of City and Regional Planning
University of Wales, Cardiff*

In his paper, Bracken addresses the problems associated with the analysis of data collected for geographic regions by developing a surface model to represent the data. The intent of the model is to represent population and other related variables “using a structure which is independent of the peculiar and unique characteristics of the actual spatial enumeration”. Bracken's specific aim is to transform both zone-based and point-based data into “something approaching a continuous surface”. Ultimately, he ends up with data that are mapped onto a fine and variable resolution grid. He achieves this by estimating the population of cell as a weighted average of the recorded populations at points j , where the weights are determined by the strength of the connection between cell i and point j . Weights are determined by a distance decay function, with weights outside a given distance set equal to zero. Bracken's principal objectives are to “provide a form of data representation that is inherently more suitable to the display, manipulation, and portrayal of socioeconomic information”, and to facilitate the integration and analysis of multiple sources and types of data by using a common geographical structure.

Computation of Spatial Autocorrelation Statistics Using a Topological Vector Data Model

*Ayse Can
Department of Geography
Syracuse University*

There have been recent efforts at interfacing GIS with external software as a means for increasing the analytical capabilities for the specific needs of the user. This development has been facilitated by increasing computational and storage capacities, as well as an increasing availability of software on a variety of hardware platforms. This paper considers how the ARC/INFO data model can be used to construct two types of commonly used weight matrices and then uses these in the computation of two spatial autocorrelation statistics, namely Moran's I and Geary's C . This is accomplished by interfacing ARC/INFO's data structure with an algorithm written in C programming language. This algorithm can be used as either a stand-alone program or can be executed from within ARC/INFO using the macro developed.

Multivariate Spatial Statistics in a GIS†

*Noel Cressie & Jeffrey D. Helderbrand
Department of Statistics
Iowa State University*

One of the strengths of a geographic information system (GIS) is its capability to handle multiple layers of information through common georeferencing. Thus, statistical methods that are both multivariate and spatial should, along with the GIS, be an integral part of the “analysis engine” that interfaces between multivariate spatial data and spatial models. This paper presents methods for multivariate spatial prediction, multivariate spatial modeling, and multivariate spatial data analysis.

Integrating GIS and Spatial Modeling: the Role of Visual Interactive Modeling in Location Selection

*Paul Densham
National Center for Geographic Information and Analysis, and Department of Geography
State University of New York at Buffalo*

Densham’s paper focuses explicitly upon the integration of GIS and spatial models. He begins by defining integration from the programmer’s perspective, the user’s perspective, and a conceptual perspective. He then describes and evaluates five location analysis software systems from these perspectives. Densham notes that in all of the systems, “data is transferred via files rather than data structures in RAM” and that “[u]ltimately, this characteristic is the Achilles heel of these systems because it constrains the form of human-computer interaction which can be supported.” In particular, it limits the potential for visual interactive modeling. Visual interactive modeling affords the capability of linked, complementary model representations and manipulation (e.g., simultaneous portrayal and updating of map, graphical, and tabular representations).

The Integration of Spatial Analysis and GIS†

*Yuemin Ding & A. Stewart Fotheringham
National Center for Geographic Information and Analysis, and Department of Geography
State University of New York at Buffalo*

It is widely expected that future GISs will have increased analytical capabilities that will take them beyond being efficient display and database management devices. Several attempts have already been made to link existing analytical software to various GISs. However, a problem with all of these attempts is that the user is forced to switch back and forth between the GIS operating environment and the analytical software. In this paper we present a Statistical Analysis Module, SAM, that runs totally within the operating environment of a GIS and utilizes a command structure that makes running the package transparent to the user of the GIS. We believe that the development of this module is an important step in linking spatial analysis with GIS technology.

Estimating Correlograms : A Monte Carlo Study

*Robin A. Dubin
Department of Economics
Case Western Reserve University*

In analyzing data that is spatially arrayed, it is often necessary to measure the correlations that occur between the observations. These correlations may be summarized by a correlogram, which expresses the correlation between two observations as a function of the distance separating them. The parameters of the correlogram are generally unknown and must be estimated. In this paper the effectiveness of three methods of correlogram estimation, *traditional*, *Maximum Likelihood*, and *Generalized Least Squares* are examined through the use of Monte Carlo experiments. Spatially autocorrelated data was simulated and used with each of the three techniques to estimate the parameters of the correlogram. Because the parameters of the correlogram are known, the effectiveness of each of the estimation techniques can be assessed, with the estimations gained from the maximum likelihood method providing the most reliable results.

Expert Systems and Artificial Neural Networks for Spatial Analysis and Modeling: Essential Components for Knowledge-Based Geographical Information Systems

*Manfred M. Fischer
Department of Social and Economic Geography
Vienna University of Economics and Business Administration*

This paper addresses three major deficiencies of current geographic information systems: the logical foundation based on the classical concept of Boolean logic and classical set theory; the limited analytical functionality; and the low level of intelligence in terms of knowledge representation. The removal of these three areas of deficiency is crucial to the further development of GIS into knowledge-based systems. Two major components of these improved systems are rule-based expert systems and artificial neural networks for spatial analysis and modeling. The paper then summarizes the expert system approach, and emphasizes the characteristics and principles of neurocomputing. The paper concludes with some potential applications of artificial neural networks and their importance to knowledge based geographical information systems.

Areal Interpolation and Types of Data†

*Robin Flowerdew & Mick Green
North West Regional Research Laboratory
Lancaster University*

Incompatibility between the areal units for which data are available frequently obstructs the analysis of spatial data. It can be overcome by interpolating the data from one set of zones to another. It should be remembered that the interpolated figures have some degree of error attached to them, which will affect the reliability of any subsequent analysis. Obviously, other things being equal, the better method used for interpolation, the more reliable the results of the analysis. A number of approaches can be taken to the interpolation problem, and this paper is based mainly on methods developed to take advantage of other information that may be available and relevant to the problem. The type of method that is appropriate depends on the nature of the data and the probability model assumed to have generated it. This paper attempts to outline the methods that can be used for some different types of data.

Problems in Spatial Analysis from a GIS Perspective†

*A. Stewart Fotheringham & Peter A. Rogerson
National Center for Geographic Information and Analysis, and Department of Geography
State University of New York at Buffalo*

Increasingly, methods of spatial analysis are being integrated within geographic information systems. As this integration occurs, it is important to ensure that (i) users of GIS recognize the limitations of spatial analysis, (ii) researchers continue to work on removing the existing impediments to accurate spatial analysis, and (iii) developers of GIS consider these limitations. In this paper, we discuss nine general impediments that arise in spatial analyses that span a diverse range of substantive applications. Geographic information systems offer not only the opportunity to integrate various methods of spatial analysis, but also the chance to learn more about the underlying impediments.

Spatial Point Process Modeling in a Geographical Information Systems Environment†

*Anthony Gatrell
North West Regional Research Laboratory
Lancaster University*

This paper describes two possible approaches to the problem of embedding statistical spatial analysis functionality into a GIS environment. The first is to link a proprietary GIS to user supplied FORTRAN programs and (if required) to other proprietary software products. The second approach is to develop new spatial analysis routines within a statistical programming environment. Both approaches have been motivated by applications in geographical epidemiology, with the main emphasis in the current research program on spatial point process modeling. For instance, we might seek to describe the spatial structure of a point pattern and then to fit a statistical model to a distribution of points, as a means of testing and explicit hypothesis. Empirical work on these problems is possible in Britain because much disease data is routinely available as a set of point events.

Some Thoughts on Developing “Proximal” Data Bases

*Arthur Getis
Department of Geography
San Diego State University*

The solution to spatial problems in a geographic information systems environment inevitably must be based on the special character of spatial data. In recent years there have been several works concerning the attention that must be given to the spatial data if they are to be used in the analysis. This paper outlines a procedure that would enhance the capability of a GIS to display and identify spatial heteroskedasticity, and in particular spatial dependency. This is achieved by extending data bases to include statistical information about spatial units in the vicinity of the spatial unit, and in so doing extending the view of space to include both “site” and “situational” characteristics and their interactions.

Spatial Analysis with GIS: Problems and Prospects†

Michael F. Goodchild
National Center for Geographical Information and Analysis
University of California, Santa Barbara

In principle, GIS provides an ideal platform for supporting a wide range of analyses using geographic data. In practice, the linkage of GIS with analysis is impeded by numerous factors, and has been more successful in some fields than in others. This paper looks at the problems of integrating GIS with analysis in general, and at the prospects for greater in the future.

The paper presents a broadly based classification of methods of spatial analysis. Efficient support of any class requires that the appropriate data model be recognized by the GIS. In some cases, methods of analysis are written for continuous space, without explicit discretization, and thus cannot be implemented intact. Many methods of spatial analysis require a data model that abstracts space to a simple matrix of interactions between objects; the paper discusses the implications of simple 'hooks'. In other cases, the data model required to support analysis includes features such as time, or the vertical dimension, that are not commonly available in current GIS.

Designing Spatial Data Analysis Modules For Geographical Information Systems

Robert Haining
Department of Geography
University of Sheffield

The paper begins by subdividing spatial analysis into three components: statistical spatial data analysis (SDA); map-based analysis; and mathematical modeling. The paper then concentrates on the integration of SDA and GIS which it is argued is necessary if GIS is to develop into a universal geographic tool. With the aid of an empirical example based on intra-urban mortality rates, the paper examines the following six questions that arise if a linkage between SDA and GIS is to take place successfully:

1. What types of data can be held in a GIS?
2. What classes of questions can be asked of such data?
3. What forms of SDA are available for tackling these questions?
4. What minimum set of SDA tools are needed to support a coherent program of SDA?
5. What are the fundamental operations needed to support these tools?
6. Can the existing functionality within GIS support these fundamental operations?

An Analysis of Changing Local Taxation Regimes Using a Street-Level Database†

Gary Higgs, Paul Longley & David Martin
Wales and South West Regional Research Laboratory
University of Wales, Cardiff

The UK has experienced different local taxation regimes in each of the last three financial years, namely the property-based household rates; the personal community charge or 'poll tax', and the hybrid personal community charge with neighborhood 'transitional relief'. In this paper we examine the impact of these changing policies in the Inner Areas of the City of Cardiff, highlighting the importance of historical rateable values and household sizes. Using a purpose-built street-level database, the implications of the different taxation systems are examined at increasingly

detailed geographic scales, and the complexity of their impact is illustrated. The results are compared to previous studies of intra-urban transference of the burden of taxation in the British Isles. The inappropriate nature of the neighborhood areas used for transitional relief allowances are highlighted in this analysis.

Optimization Modeling in a GIS Framework: The Problem of Political Redistricting

*W. Macmillan & T. Pierce
School of Geography
University of Oxford*

Advances in computing, as had earlier advances in mathematical modeling from a geographic perspective, promised much but delivered relatively little. This is in part due to the limited mathematical modeling in contemporary geographical information systems. This paper addresses the need for the expanding or rebuilding of these models in a GIS framework. It explores one of a number of related attempts to introduce optimization models, used prescriptively, into the GIS framework for the issue of congressional redistricting in the USA.

This paper describes part of the work involved in the development of a set of active redistricting algorithms based on mathematical optimization models and their operationalization as TransCAD procedures. Some of these algorithms utilize NAG routines but it is argued that the mathematically interesting features of the redistricting problem make it desirable to use non-standard algorithms. In particular, a simulated annealing approach is described which appears to be successful in circumventing the difficulties inherent in the structure of the redistricting problem, with reference to one application using the ANNEAL algorithm. A prototype redistricting decision support system using TransCAD-ANNEAL and present limitations of the algorithm is then considered.

Spatial Analysis and GIS: A Position Paper

*Morton O'Kelly
Department of Geography
Ohio State University*

This paper addresses two major issues in the integration of GIS and spatial analytic methods. One is the need to overhaul the traditional methods of displaying data and presenting results in the light of new technology. The other is the recognition that users of a GIS with expanded capabilities in terms of spatial analysis will probably be relatively poorly trained in the use of such methods. Some sort of guidance in their use will therefore be advantageous if we are to fully exploit the power of the integration and not simply allow increased abuse of spatial analysis.

By addressing both of these issues we can ensure that the integration of spatial analysis and GIS produces benefits to both components and not simply to the GIS. These two advantages are demonstrated in a variety of spatial analytic methods including space-time pattern recognition in point data sets; spatial interaction; spatial autocorrelation; and spatial situation variable measurement.

Two Exploratory Space-Time-Attribute Pattern Analysers Relevant to GIS

Stan Openshaw
CURDS
University of Newcastle

In his paper Openshaw contends that exploratory methods of spatial analysis need to be given greater emphasis. He focuses specifically upon arguing the need for pattern detectors that are (1) not scale specific, (2) highly automated, and (3) have the flexibility of including human knowledge. Openshaw categorizes the types of information contained in a GIS database into (1) geographic information, (2) temporal information, and (3) attribute information. He argues that previous attempts to find spatial pattern, temporal pattern, space-time interaction, and clusters of highly correlated attributes are too limited, and that all three should be viewed simultaneously in “tri-space”. Openshaw offers two types of pattern analyzers —the first is an algorithm based upon grid search and Monte Carlo significance tests, and the second involves the genetic evolution of pattern detectors through the survival of good detectors and the death or weeding out of poor ones. Openshaw demonstrates the approaches through an application to crime pattern detection. He concludes by reiterating the need for methods that explore GIS databases in a way that allows such activity to be “fluid, creative, and artistic”, and in a way that stimulates “the user's imagination and intuitive powers”.

Object Oriented Spatial Analysis†

Bruce A. Ralston
Department of Geography
University of Tennessee

Geographic information systems are powerful tools, but their power is in many ways limited. There are many tools of spatial analysis which, at present, are difficult to use in a GIS. We present an approach which facilitates the integration of spatial analysis tools with GIS packages, and which can be adapted to a large number of spatial analysis procedures. The basic paradigm used is that of object oriented analysis. We use an object oriented programming language, C++, to develop a set of classes and procedures for handling models of spatial processes which require the formulation of some type of matrix. We then apply these tools to a set of increasingly complex optimization problems. The OOP approach allows us to concentrate more on the relationships between spatial entities than on how the computer code should process those entities. We end by outlining how an applications generating program could be used to construct the necessary child classes and relationships so that building spatial models in a GIS would be much easier than it is at present.

Spatial Models and GIS in Designing and Managing Emergency Response Systems

*Gerard Rushton
Department of Geography
San Diego State University*

This paper compares present literature concerning the design and management of an emergency response system from both a GIS and a spatial analytical viewpoint. GIS tend to improve decision making from an organizational systems perspective, with the view that the better the information, the better the decision. The spatial analyst's approach uses mathematical analysis to make better decisions, with collaboration with the appropriate party to collect data at the appropriate level of aggregation, then employing the correct methodology to obtain a solution. This paper suggests that the future lies in a hybrid of the two approaches, combining the information richness of the GIS with the analytical depth of spatial analysis, as long as a way can be found around the restricting entrenched cultures of the two camps.

Global Analysis

*Waldo Tobler
National Center for Geographical and Information Analysis
University of California, Santa Barbara*

Both the spatial analysis and GIS literature concentrates on local and not global problem solving. This *flat earth* attitude is perpetuated in the methods of geometry and statistics that still pervade the educational system. This paper suggests that what is needed, especially with the advent of GIS, is a rethinking of the techniques and attitudes of statistics and spatial modeling, incorporating a more spherical approach. There is also a need for GIS to deal with and display spherical data, combined with an Exploratory Global Analysis and Display system for a more evolved 'world-wide' outlook.

Cellular Dynamics and GIS : Modeling Spatial Complexity†

*Roger White and Guy Engelen
Department of Geography
Memorial University of Newfoundland*

Both observation and theory support the view that spatial complexity is a necessary feature of many systems, and that far from representing mere randomness, it constitutes the real, information-rich order inherent in these systems. cellular automata are developed to model the dynamics of spatially complex systems. The approach permits an increase of several orders of magnitude in the spatial resolution of dynamic models when compared to traditional techniques for modeling regional dynamics, and thus permits modeling to address the issues of spatial complexity directly. The cellular approach is illustrated in two-models - an urban land use model which is purely cellular, and a model of a Caribbean island which combines cellular and traditional techniques. Since the models require large quantities of data with high spatial resolution, the full power of the cellular technique can only be realized when the models are integrated with a GIS. More generally, cellular automata have the potential for converting a GIS into a dynamic modeling tool.

Appendix C: Transcript of Discussions, Working Group on Locational Analysis and Planning Applications with GIS

Paul Longley - Looking at redistributive consequences of local revenue raising regimes. Looking at the geography of property tax systems, with considerations to the socio-economic characteristics of each of the areas. He is interested in the emergent pattern of gainers versus losers under the different regimes. At the moment there is a backlash against the community charge, with the government attempting to bail out those who lost the most. However this has been done using arbitrary community designations. The geography of built form has been replaced by a geography of population (which forms the basis of revenue raising). However this system has not produced any patterns as such because of the arbitrary use of boundaries. This means that any future tax regime problematic as everyone has a favorite regime based on arbitrarily chosen areas.

*: This shows a problem with integrating different data sources.

This was the first time someone attempted to use the register of electors with the property tax register. No-one has discussed the problems caused by the different socio-economic effects of both regimes.

*: A GIS would be useful if it could integrate different sources of data.

In the UK, the private sector uses Pinpoint Data Analysis. This is digitized to the scale of every address. The Cardiff City Council bought it, but elected not to use it because of the amount of errors.

*: What methods are available for putting different geographically scaled data together?

(Flowerdew) Goodchild did some work on merging one scale with another in a GIS. He had set areal units, used polygon overlay by proportioning (if one overlaps another polygon by 15% then presume 15% of the population is captured within that area).

(Bracken) In UK, it's a messy job; there's no consistency with the Census because of the attempts to even out the work loads of the enumerators. The population also changes. The Census areas rest reasonably well with the local government areas. All the different services have different zonal areas, *e.g.*, the Police). The post codes (which provide a lot of useful information) do not have defined boundaries as such.

The post code is attached to a national file grid reference which is accurate to approximately 100m. The grid reference is taken from the most extreme NW or SW address on the delivery strip. The Pinpoint Data set has 1m accuracy of every address. But, as said before, the data set is full of errors. The census files also contain a lot of errors.

*: How are errors dealt with?

An important tool in a GIS is linking data. Error identification is important. It is a complex and serious issue. But how do you approach a checklist of spatial errors?

(Ralston) Should we decide on an agenda for GIS development, whether it should be aimed at vendors, developing systems that will sit on the top of present systems, or should we aim at the end users who may or may not have a masters degree?

(Bracken) What is our position as academics? Should we concern ourselves with policy decisions, or should we concentrate on the research and let others use it? There are a lot of people who can't use present systems; does this mean the systems should be simpler—easier to use—but is that our role? Should we be more frontier like? Should we be researching more efficient algorithms?

(Longley) We should attempt to understand each others' experiences and attempt to understand whether GIS is adjunct, and integral part or just a trendy repackaging of what we have always been using.

(Macmillan) In reference to Openshaw's paper, we should sit down and put the spatial interaction model into the GIS. GMAP started to do this but moved outside the academic community.

(Densham) The problem with the presentations was that they were still thinking about GIS and models as being separate. Instead we need to bring them together. Need to start to think about 'pipes' of data.

(Ralston) Think of objects of data - what kinds of interactions are relevant and what information is needed to be known about them, and how one changes the other.

(Flowerdew) What is the difference between objects and data?

(Ralston) Objects have operations associated with them. The need is to try and merge the data and the operation together.

(Bracken) What can we learn from other fields which have gone through this process whereby information is pushed around? Are these problems unique to just spatial situations?

(Densham) In the business area, they just look at spreadsheets. They just consider the relationship between data, but not in a spatial context. When we think about spatial representation, tabular information doesn't help. Management science is not into object oriented programming because of other problems (not spatial).

(Ralston) They are also just interested in efficiency, they want to make things faster.

(Densham) They have no concept about errors of data aggregation (they tend to use census tract data). They are very good modelers who have no idea of the other problems involved.

(Bracken) In one rural project, a lot of information was digitized using SPANS. This produced maps and tables which added 8 decimal places. People looked at this, saw the amount of digital places, and commented on how accurate it must be although they were probably digitized from crumpled maps. After all you can't argue with figures.

Does this mean there is more responsibility on us to identify and solve these kinds of problems?

(Macmillan) The standards issue—the academic community needs to present standards.

(Flowerdew) Is there any systematic error in the system. Are biases and characteristics present?

(Macmillan) Should we be looking at the systems, like Transcad, with a more open architecture, to see how we can use and modify them? Should we work more closely with the vendors.

(Ralston) Should we build a research agenda of when not to buy a system? We don't need all these extra details, they get in the way of people learning the system (Information overload, confusion *etc.*)

(Macmillan) ARC/INFO was so complex that it had a very long learning curve. Now it's packaging only small elements of the overall system.

*: But when should some of these tools be used? This technique should be used when.....

(Longley) In what applications does a GIS provide a better tool than exploratory data analysis?

(Densham) The big role for GIS is in visualization.

(Longley) Then are we in a transitory phase? We look at it then get in with our models as before.

(Densham) We are at the start and we need to get through today's GIS to get to a better stage.

(Macmillan) The problem is we don't have common elements.

(Longley) MAUP is one—the range of scale and aggregation effects.

(Bracken) Are there other common problems?

(Macmillan) But we don't get to see these problems until we're well into a specialized problem.

(Densham) What is common is model management. How can people get into and use models easily.

(Ralston) An object oriented approach would write a set of codes that would incorporate a lot of models.

(Densham) About half a dozen data structures can support most models. Then need to define the data structures so as they can incorporate visualization.

(Ralston) There should be two different emphases. Should be spatial statisticians and spatial modelers.

(Flowerdew) Are they deterministic rather than stochastic?

(Macmillan) Yes. the difference between inductive work and inference. I have a very theory down approach, avoiding data whenever possible. GIS gives me a way of having to address data. There may be errors, but still be better off than just theory. Other people have just data and try and squeeze some theory out of it.

(Densham) Many kinds of tools are available. You can give them a choice of 1 out of 4 options, giving a purely intuitive approach. Now can give them the ability to do 'what-if' scenarios but within a theory driven framework.

(Macmillan) There must be a core of techniques that can stimulate ideas without worrying too much about errors. Should not worry too much about individual parameters or the error attached to them, just see them as a guide.

Personal Agendas

(Bracken) Data linking—transforming social data into raster. Know something about the errors involved, then integrate into RS data for urban areas to model social data. Trying to refine knowledge about how data sheds light on processes and change. It doesn't need a GIS. It has GIS implications but not GIS based research.

(Macmillan) Why do I need GIS? For respectability. I'm interested in theories and models. Such spatial data may be useful. There is a need to disguise models from the user. It's a presentation issue, to be user friendly but have the correct guts. It will breathe some life into evaluation methods of SIM and SEM. At the moment I'm interested in integrated dynamic modeling with GIS. For grand urban transportation models—it would provide an all seeing all dancing version. Large scale systems approach urban models—only when you see the whole thing can you see what's going on. GIS provides the tool where we can address issues in ways marginally better than now. Don't have to get optimum routines. It is a means of explanatory modeling. Like a 'grown-up' version of SIM CITY. There are data problems, but we still get a better feel for what's going on. It's a way of clarifying your assumptions. (Not an end in itself but a means to an end). With STELLA, when you operationalize it you realize there are gaps, but it makes you have to think in more sensible ways. We now have good theory giving good models but without the technical means to implement them. We are still playing, but there is some constructive purpose—needs more than high tech veneer, it can make for more sensible ways of thinking.

(Ralston) How to get the tools operational to make it easier for decision makers. I believe an object oriented approach gives a way of putting together the software which would allow such 'model-based management'. I would also put data in. Need sophisticated users who would think more about spatial problems and less about writing code. Then academics can adapt the private sector stuff.

(Densham) GIS allows to move from paper to a completely digital environment in order to solve models. Is there any benefit to be gained from going digital? It's quicker and more efficient, but it has other potentials as well. We now have these abilities; data management and display, and how to integrate them.

(Ralston) How can we use the users' expertise? We need integrated systems. We need to know why they don't like something.

(Macmillan) The user needs to know about some of the potential problems, in case they become hooked up in a local optimum.

(Flowerdew) There is an areal interpolation problem. Originally hoped that a set of software would be developed whereby routines could be tagged onto something like ARC/INFO so the user can have some idea about the problems of areal interpolations. Using GIS for emergency planning. Interests in geo-demographics—especially clustering. These aspects are linked to GIS—especially visualization. Interested in the MAUP (and the development of social indexes and scales). Also, the local denominator problem—to pick places at the worse social scale—if you use the finest resolution. MAUP and SIM get different solutions according to scale. Look at models at different scales. Functional distance measure at different scales. Not really in GIS. Statistical analysis—incorporating K functions into GIS as separate add ons. Extending this to space time clustering. Present GIS are woefully inadequate in straightforward statistical analysis. Even straightforward simple things. Need vendors to have more simple techniques—don't need to concentrate so much on high level techniques (K function *etc.*) when can't do the simple things.

Saturday

Operational Research Facilities

Three Broad Areas:

a: Optimization—any problem that can be put into optimization form (*e.g.*, LA, economic equilibrium —these are descriptive rather than predictive)

b: Stochastic process

c: Simulation

Q: Is applied OR our focus?

(Macmillan) Should be looking at overall context — incremental advances.
What sort of partnerships should we forge—OR vendors or GIS vendors?

(Flowerdew) We could talk about the different types of OR models and how they can fit into GIS.

(Macmillan) Should we have generic optimization routines or should there be separate tools for separate problems.

Generic Optimization - take LA problems, they have generic routines for all—but less efficient than having 1 routine for each data set.

But may want to have a suite of routines.
May want to use code for SIM or other model.

Q: How do you nest data in a hierarchical way?

All sorts of OR and GIS available, therefore problem could be in bringing these softwares together—you can get somewhere but then get frustrated. Then have to break into the code to do what you want to do because of spatial problems like contiguity checks.

One problem is that most OR are written in Fortran and these don't allow for dynamic arrays.

(Ralston) Should talk about solution techniques not formulations. Users have real problems with formulation techniques.

There are a lot of cases where we know what the problems are but we have no idea at how to take the database and generate the correct files. Need to ask questions (or partial questions) to generate code stubs which makes the code writing easier.

Need to focus on the 'art' rather than the 'science' of OR.

(Macmillan) If we had an OR library underneath the GIS, we could make use of the library by an interface.

STELLA—has pipes and tools—you put all the links together in the way you want—it tells you about the difference equations, but won't run if not logically approached.

(Flowerdew) So we need code that will generate a set of prescriptions.

(Ralston) If we treat as object oriented programming. There is a matrix, the coefficients are functions of how the columns and rows interact. Each has its own property. *e.g.*, May have a row with a demand node. Look at the column (virtual functions) and then row—see if there is an interaction between the two—then, if not, return 0 as the coefficient. If get a link, with a point that has demand—you get a value. This generates the matrix. The user has to worry about how the links correspond to the demand nodes. Need to have some knowledge about the constraints, and what the variables are. To generate new demand, need new column, don't have to re-write the code.

Every row has a right hand side (float) and a direction (character). Then have a function called a 'constructor'. Take as an argument either node pointer, link pointer and polygon (all in rows).

(Macmillan) Is this a way for people who don't have an algebraic understanding to solve the problem?

(Ralston) No. For complex problems you use the same structure; just redefine the spatial primitives.

(Ralston) Should we give examples of how powerful object oriented programs are?

(Bracken) Need to get two types of data into the same system—more than an overlay—need to sample one with reference to the other. At the moment no GIS does that. Need to formulate some rules. Use the properties of one data structure to structure another. Need a set of rules for data manipulation.

(Macmillan) Could look at what has been done already. Might be useful to distinguish between discrete and continuous optimizing problems.

(Bracken) Modeling—put census data onto a regular grid. Could try to vary the size of the cells—to get a homogenous number of people per cell. At the moment don't know how to do this.

Cells in forms of cartograms in response to the size of data. The objective—in a visual way—want to represent data in a more meaningful way—don't see it as an analytical problem. How will the application generator help me? (like the American districting problem)

(Macmillan) The districting problem is messy because it is highly convex. May find that the algorithms are increasingly not suited to the solutions we need.

(Flowerdew) Some look for what are the best algorithms of what they have, the most efficient.

(Macmillan) Very problem specific—don't know of the global optimum. Can judge algorithms by how quickly they converge. If non-convex don't know how close you are to the solution. Often don't have bounds on an objective function. Therefore don't have any method to compare different algorithms.

(Ralston) Application generators. Flexible way of dealing with the density estimation—controlling of geography in optimization. Spatial data optimization. There is a whole class of optimization that is needed to be solved.

- spatial
- statistical
- then consider how to incorporate into a GIS.

Can use optimization in whole host of areas. Therefore need optimization routines (either 1 for each problem, a library of routines, or a 'jumbo all-encompassing' routine.)

Application generator part of the package.

Connection - data restructuring.

How to move towards optimization in GIS. Lots of problems, can't conceive of objective function (don't know how to interrelate). Mechanism needs to be set in more exploratory framework.

(Macmillan) Problems that revolve around 'what-if' situations in the real world. GIS offers way of exploring ideas rather than just data.

(Flowerdew) Are there any characterizations of simulations? Is there a distinction between spatial and non-spatial simulations?

(Macmillan) At the moment basically everything is aspatial simulation. Do for every point then hook them all up. STELLA—when the problem gets complex spatially as well as temporally then it loses its ability to give insights.

(Bracken) Maybe give more emphasis to GIS with just an interpretive role and not 'all singing all dancing'. GIS may help people think more spatially, *e.g.*, banks just use OR cost-benefit analysis, no spatial impact. Economists are starting to become more spatially oriented—they have 3 regions sometimes.

(Ralston) Are we using tools in GIS correctly? Misuse of OR. Data quality may not be as important. Should be exploring rather than coming up with answers. Can always warn people to problems that may accompany the answer. May give a range of values and say try again. Could try and inhibit people from stupid things 'Gentle Policing'.

(Bracken) Would policing stop exploration? There are a whole set of issues about sensitivity analysis.

(Macmillan) Optimization and statistics. If want to do least squares, would it be beneficial to access a quadratic algorithm that could be used in other areas— or do you want one specific algorithm? Scale and algorithm effects—ability to insert certain criteria in the data manipulation you're dealing with.

Connection between dynamic simulation and GIS.

Dynamic simulation doesn't work well over space.

(Macmillan) We may be able to overcome this now we are starting to parallel program, which will eventually change we model, with a movement away from linear algorithms. Have a parallel problem with simulation rather than optimization.

Appendix D: List of Initiative 14 Specialist Meeting Participants and Declines

List of Participants

T. R. Bailey
Department of Mathematical Statistics
and Operational Research
University of Exeter
Laver Building, North Park Road
Exeter EX4 4QH
UNITED KINGDOM
Phone: 392-264474
Fax: 392-263108
Email: bailey.tc@exeter.ac.uk or
tcb@msor.exeter.ac.uk

Michael Batty
NCGIA/Department of Geography
SUNY at Buffalo
Buffalo, NY 14261
Phone: 716-636-2545
Fax: 716-636-5957
Email: geobatty @ubvms.bitnet

Graeme Bonham-Carter
Geological Survey of Canada
Energy, Mines, and Resources
601 Booth Street
Ottawa, Ontario K1A 0E8
CANADA
Phone: 613-996-3387
Fax: 613-996-3726
Email: graeme=bonham-carter%geochem
%gsc=ottawa@gsc.emr.ca

Ian Bracken
Department of City and Regional Planning
University of Wales, Cardiff
P. O. Box 906, Colum Drive
Cardiff CF1 3YN
UNITED KINGDOM
Phone:
Fax: 222-874845
Email: bracken@towp.abcy.cardiff.ac.uk

Ayse Can
Department of Geography
Syracuse University
343 H. B. Crouse Hall
Syracuse, NY 13244-1160
Phone: 315-443-5639
Fax: 315-443-4227
Email: acan@suvvm.bitnet

Noel Cressie
Department of Statistics
Iowa State University
102 Snedecor Hall
Ames, IA 50011
Phone: 515-294-3440
Fax: 515-294-4040
Email: sl.nac@isumvs.bitnet

Andrew Curtis
NCGIA/Department of Geography
SUNY at Buffalo
Buffalo, NY 14261
Phone: 716-636-2545
Fax: 716-636-5957
Email: v070sk9f@ubvms.bitnet

Lee De Cola
United States Geological Survey
National Mapping Division
521 National Center, Room 2D330
12201 Sunrise Valley Drive
Reston, VA 22092
Phone: 703-648-4178
Fax: 703-648-5585
Email: ldecola@resdgv11.er.usgs.gov

Paul Densham
NCGIA/Department of Geography
SUNY at Buffalo
Buffalo, NY 14261
Phone: 716-636-2545
Fax: 716-636-5957
Email: geopjd@ubvms.bitnet

Uwe Diechmann
NCGIA
3510 Phelps Hall
University of California, Santa Barbara
Santa Barbara, CA 93106-4060
Phone: 805-893-8224
Fax: 805-893-8617

Rusty Dodson
NCGIA
3510 Phelps Hall
University of California, Santa Barbara
Santa Barbara, CA 93106-4060
Phone: 805-893-8224
Fax: 805-893-8617
Email: dodson@toast.ncgia.ucsb.edu

Randall Downer
Applied Biomathematics
100 North Country Road
Setauket, NY 11733
Phone: 516-751-4350
Fax: 516-751-3435
Email: in%dowranh@sbbioum.com

Robin Dubin
Department of Economics
Case Western Reserve University
Cleveland, OH 44106
Phone: 216-368-3981
Fax: 216-368-5039
Email: rad4@po.cwru.edu

Chuck Ehlschlaeger
US Army Construction Engineering
Research Lab
2902 New Mark Drive
Champaign, IL 61821-1076
Phone: 217-398-5464
Fax: 217-398-5470
Email: chuck@cerl.cecer.army.mil

Manfred M. Fischer
Department of Economic and
Social Geography
Vienna University of Economics
and Business Administration
Augasse 2-6, A-1090
Vienna
AUSTRIA
Phone: 222-31336-4836
Fax: 222-31336-703
Email: mmf@awiwuw11.bitnet

Robin Flowerdew
Department of Geography
University of Lancaster
Lancaster LA1 4YB
UNITED KINGDOM
Phone: 524-65201
Fax: 524-844037
Email: gya@014@central1.lancaster.ac.uk

Stewart Fotheringham
NCGIA/Department of Geography
SUNY at Buffalo
Buffalo, NY 14261
Phone: 716-636-2545
Fax: 716-636-5957
Email: geoasf@ubvms.bitnet

Tony Gatrell
Department of Geography
University of Lancaster
Lancaster LA1 4YB
UNITED KINGDOM
Phone: 524-65201, x 3754
Fax: 524-844037
Email: gya007@central1.lancaster.ac.uk

Art Getis
Department of Geography
San Diego State University
San Diego, CA 92182
Phone: 619-594-6639
Fax: 619-594-4938
Email: agetis@sciences.sdsu.edu

Michael Goodchild
NCGIA/Department of Geography
University of California, Santa Barbara
Santa Barbara, CA 93106
Phone: 805-893-8224
Fax: 805-893-8617
Email: good@goodrs.geog.ucsb.edu

Bob Haining
Department of Geography
University of Sheffield
Sheffield S10 2TN
UNITED KINGDOM
Phone: 0742-768555
Fax: 0742-722199
Email: r.haining@primea.sheffield.ac.uk

Dale Honeycutt
Environmental Systems Research Institute
380 New York Street
Redlands, CA 92373
Phone: 714-793-2853, x 1383
Fax: 714-793-5953
Email: dhoneycutt@esri.com

Sheri Hudak
NCGIA
3510 Phelps Hall
University of California, Santa Barbara
Santa Barbara, CA 93106-4060
Phone: 805-893-8224
Fax: 805-893-8617
Email: sheri@pollux.geog.ucsb.edu

Clifford Kottman
Intergraph Corporation
2051 Mercator Drive
Reston, VA 22091-3413
Phone: 703-264-7150
Fax: 703-264-7131

Paul Longley
Department of Geography
University of Bristol
University Road
Bristol BS8 1SS
UNITED KINGDOM
Phone: 272-303030
Fax: 222-874845
Email: LONGLEY@bristol.gma.ac.uk

Bill MacMillan
School of Geography
Oxford University
Mansfield Road
Oxford OX1 3TB
UNITED KINGDOM
Phone: 865-271919
Fax: 865-271-929
Email: geog1@vax.oxford.ac.uk

R. Douglas Martin
STATSCI
1700 Westlake Avenue, N., Suite 500
Seattle, WA 98109
Phone: 206-283-8802, x229
Fax: 206-283-8691
Email: doug@statsci.com or
doug@stat.washington.edu

Morton O'Kelly
Department of Geography
The Ohio State University
103 Bricker Hall
190 North Oval Mall
Columbus, OH 43210-1361
Phone: 614-292-8744
Fax: 614-292-6213
Email: okelly+@osu.edu

Stan Openshaw
CURDS
Newcastle University
Newcastle upon Tyne NE1 7RU
England
UNITED KINGDOM
Phone: 91-222-6000, x7577
Fax: 91-232-9259
Email: stan.openshaw@uk.ac.newcastle

Bruce Ralston
Department of Geography
University of Tennessee
408 G and G Building
Knoxville, TN 37996-1420
Phone: 615-974-2418
Fax: 615-974-6025
Email: ralston@utkvxl.bitnet

Peter Rogerson
NCGIA/Department of Geography
SUNY at Buffalo
Buffalo, NY 14261
Phone: 716-636-2545
Fax: 716-636-5957
Email: geopar@ubvms.bitnet

Peter Rosenson
Geographic Files Branch
U. S. Bureau of the Census
Washington, DC 20233
Phone: 301-763-7024
Fax: 301-763-4749

Gerard Rushton
Department of Geography
San Diego State University
San Diego, CA 92182
Phone: 619-594-2599
Fax: 619-594-4938
Email: grushton@sciences.sdsu.edu

Howard Slavin
Caliper Corporation
1172 Beacon Street
Newton, MA 02161
Phone: 617-527-4700
Fax: 617-527-5113

Waldo Tobler
NCGIA/Department of Geography
University of California, Santa Barbara
Santa Barbara, CA 93106
Phone: 805-893-3831
Fax: 805-893-3146
Email: tobler@pollux.geog.ucsb.edu

Roger White
Department of Geography
Memorial University of Newfoundland
Elizabeth Avenue
St. John's, Newfoundland A1B 3X9
CANADA
Phone: 709-737-8193, 709-737-7417
Fax: 709-737-4000
Email: rwhite@kean.ucs.mun.ca

List of Declines

Carl Amrhein
Department of Geography
University of Toronto
Toronto, Ontario M5S 1A1
CANADA

Luc Anselin
NCGIA/Department of Geography
University of California, Santa Barbara
Santa Barbara, CA 93106
805-893-8617

Richard Appelbaum
Department of Sociology
University of California, Santa Barbara
Santa Barbara, CA 93106

Peter Burrough
Department of Geography
Universiteit Utrecht
Postbus 80.115,
Utrecht
THE NETHERLANDS

Anne Case
Department of Economics
Princeton University, Box 430
Princeton, NJ 08544
609-258-4000

C. K. Chow
Program Manager, IBM Corporation
NCGIA, 3510 Phelps Hall
University of California, Santa Barbara
Santa Barbara, CA 93106-4060
805-893-8641

Richard Dunn
Department of Geography
University of Reading
Whiteknights, Reading RG6 2AB
UNITED KINGDOM
734-318673
in%"dunn@uk.ac.bristol"

Clive Granger
Department of Economics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508

Dan A. Griffith
Department of Geography
Syracuse University
Syracuse, NY 13244-1160

Steve Guptill
United States Geological Survey
National Mapping Division
519 National Center
Reston, VA 22092

Gary Hamilton
Department of Sociology
111 Young Hall
University of California, Davis
Davis, CA 95616
916-752-0782

Peter Haslett
Department of Statistics
University College Dublin
Dublin 2
REPUBLIC OF IRELAND

Leonard Hochberg
Department of Sociology
McClatchy Hall
Stanford University
Stanford, CA 94305
415-855-9314

Giulio Maffini
INTERA TYDAC Technologies Inc.
2 Gurdwara Road, Suite 510
Nepean, Ontario K2E 1A2
CANADA
Phone: 613-226-5525
Fax: 613-226-3819

Ross MacKinnon
Faculty of the Social Sciences
SUNY at Buffalo
Buffalo, NY 14261
716-636-3101
proross@ubvms.bitnet

John O'Loughlin
Institute of Behavioral Science
University of Colorado
Campus Box 487
Boulder, CO 80309-0487
303-492-1619
oloughlin_j@cubldr.colorado.edu

Burkhard Schaffrin
Department of Geodetic Science and Surveying
The Ohio State University
1958 Neil Avenue
Columbus, OH 43210-1247
614-292-6753

William Skinner
Department of Anthropology
University of California, Davis
Davis, CA 95616
916-752-8685

John Sprague
Department of Political Science
Washington University
St. Louis, MO 63160
314-889-5000
in% "c38991js@wuvmd.bitnet"

Harvey Starr
Department of Political Science
University of South Carolina
Columbia, SC 29208
803-777-7292

Nancy Tosta
Teale Data Center
2151 Markham Way
Sacramento, CA 95818

Michael Ward
Department of Political Science
University of Colorado
Campus Box 333
Boulder, CO 80309
303-492-7871