WSG 37/94

Geographic Information Systems, Spatial Data Analysis and Spatial Modelling.
- Problems and Possibilities -

Manfred M. Fischer and Henk J. Scholten
Geographic Information Systems, Spatial Data Analysis and Spatial Modelling. - Problems and Possibilities -

*Manfred M. Fischer and Henk J. Scholten*
Gedruckt mit Unterstützung
des Bundesministerium
für Wissenschaft und Forschung
in Wien

WSG Discussion Papers are interim
reports presenting work in progress
and papers which have been submitted
for publication elsewhere.

ISBN 3 85037 042 9
ABSTRACT

This article is the position paper for the ESF-GISDATA Specialist Meeting on GIS & Spatial Analysis, Amsterdam, 1-5 December 1993. The focus here is on the two major themes of the meeting: Spatial Data Analysis and Spatial Modelling. Special emphasis is laid on specific problems and possibilities for interfacing spatial analysis tools (i.e. spatial data analysis techniques and spatial models) and GIS. Both GIS application fields, the environmental sciences and the social sciences, are taken into consideration.

I INTRODUCTION

Today, geographic information systems incorporate many state-of-the-art principles such as relational database management, powerful graphics algorithms, elementary spatial operations such as buffering in vector maps, polygon overlay with logical operations, interpolation, zoning and simplified network analysis. What is termed spatial analysis and modelling is often no more than map data manipulation such as polygon overlay and buffering. The lack of analytical and modelling functionalities is widely recognized as a major deficiency of current systems.

There is a wide agreement in both the GIS community and the spatial analysis community that the future success of the GIS technology will depend to a large extent on incorporating more powerful analytical and modelling capabilities. In recent years, the theme of GIS and spatial analysis has increasingly been addressed in several conferences, such as the European Congress of the Regional Science Association (RSA) in Lisbon (1991) and Moscow (1993), and in workshops, such as the Workshop in Sheffield (1991) [sponsored by the Economic and Social Research Council (ESRC) as part of its Regional Research Laboratory (RLL) initiative] and the Expert Meeting in San Diego (1992) [sponsored by the US National Center for Geographic Information and Analysis (NCGIA) as part of its initiative 14 and co-sponsored by the IGU-Commission on Mathematical Models]. While the two workshops primarily focused on spatial data analysis techniques, the sessions at the RSA-Conferences paid attention to both major lines of direction in spatial analysis: spatial data analysis and spatial modelling.

Spatial data analytic techniques and spatial models can perform functions which current geographic information systems (GIS) are almost missing but which are important for the sort of questions policy decision makers in private and public organisations are interested in. It is into this context that the need for more basic research in the operation, development and use of spatial data analytical techniques and spatial modelling has been identified as the major focus of the GISDATA Specialist Meeting on GIS & Spatial Analysis, to be held in Amsterdam, December 1-5, 1993.

This position paper is organized as follows. As a starting point the special characteristics of spatial data and some features of the field of spatial analysis are described in section 2. The history and the current achievements of the field of spatial analysis teach us that - despite the large number of diverse contributions of this field in both the environmental and socio-economic application contexts - two main fields can be distinguished: spatial data analysis and spatial modelling. These two fields...
which also determine the structure of the GISDATA Expert Meeting will be briefly discussed in sections 3 and 4. In the final section some implications for this Meeting will be described.

2 ON THE NATURE OF SPATIAL DATA AND SPATIAL ANALYSIS

2.1 Why a Spatial Perspective?

Several arguments might be found why a spatial perspective in data analysis should be adopted (Goodchild et al. 1992). First, space provides a simple, but very useful framework for handling large amounts of data. Second, the spatial perspective permits easy access to information on the relative location of objects and events. Third, a spatial perspective allows objects or events of various types to be linked, in a process formalized in GIS as overlay. Finally, the distance between objects and events reveals to be often an important factor in interactions between them, both in environmental and in socio-economic applications.

2.2 What is Special about Spatial Data?

The ability of geographic information systems (GIS) to handle and analyse spatial data is usually seen as the characteristic that distinguishes GIS from information systems developed to serve the needs for business data processing as well as from computer aided design systems or other systems whose primary objective is map production. Spatial data sets provide two types of information:

- data describing the specific locational position of objects in space (and their topological relationships), so-called positional or topological data, and
- data describing non-spatial attributes of the objects recorded, so-called attribute or thematic data.

The spatial (or geo-) referencing can take several forms. The reference may locate a single point to an exact location in space (e.g. the position of a firm). It can be a set of references locating a more complex entity in space, e.g. the route of a road, or it can be a reference to an area. Often positional references used for compiling spatial data include administrative units such as census tracts, land parcels, municipalities, counties and road networks. The primitive elements of spatial data are referred to as objects. Point, line and area are primitives. The elements of the object classes are characterized by attributes (non-spatial information).

Major problems encountered in dealing with spatial data include the issue of errors in spatial databases, data integration, confidentiality, and appropriate spatial analysis. In a spatial database errors arise in measuring both location and attribute values (source errors) and are related to the computerized process responsible for storing, retrieving and manipulating spatial data. This issue received increasing attention in recent times (Goodchild and Gopal 1989).

A frequent problem for GIS users with different types of data is the issue of data integration. By integration we mean the process for making different data sets compatible with each other. Data incompatibilities may be caused by the use of
different spatial referencing systems, different statistical or temporal coverage, different degrees of generalization, and locational errors. This problem is fundamental to socio-economic GIS applications and has received increasing attention recently (Flowerdew and Green 1991).

As the size of a geo-reference unit decreases, the number of objects located within its boundaries also tends to decrease. At some small unit size, then information needed would violate confidentiality laws. It is confidential only when a sufficient number of individual facts are aggregated together so that one cannot ascertain information about individuals. In some socio-economic contexts, such as, for example, in industrial geography and housing research, events are sufficiently rare that even when data are tabulated at the county level confidentiality issues might arise.

Finally, because objects or events distributed over space tend not to be independent, statistical analysis problems are encountered with geo-referenced data in contrast to more traditional forms of data. Complications are caused by two special features of spatial data: spatial dependence and spatial heterogeneity. Spatial dependence refers to the relationship between geo-referenced data due to the nature of the variable under consideration, and the size, shape and configuration of the spatial units used as geo-referencing framework. The smaller the size of spatial units is, the greater is the probability that nearby spatial units will be spatially dependent. Spatial heterogeneity arises when spatial uniformity of the effects of spatial dependence and/or of the relationships between the variables at hand are lacking.

These spatial effects may invalidate many of the properties of standard statistical procedures and lead to errors in the statistical inference, misleading indications of model validity, and other problems. Consequently, spatial data analysis must go beyond standard statistical analysis (Anselin and Getis 1992). But unfortunately, there seems to be no applicable technology available which is suited to deal with very large data sets.

2.3 Spatial Analysis is more than Statistical Spatial Data Analysis

The origins of spatial analysis lie in the development of quantitative geography and regional science in the early 1960s. The use of quantitative (mainly statistical) procedures and techniques to analyse patterns of points, lines areas and surfaces depicted on analogue maps or defined by coordinates in 2- or 3-dimensional space characterize the initial stage. Later on more emphasis was placed on the indigenous features of geographical space, on spatial choices and processes and their implications for the spatio-temporal evolution of complex spatial systems.

Figure 1: Two Major Fields of Research in Spatial Analysis
Spatial analysis as it has become over the past three decades is more than spatial statistics (geo-statistics or spatial econometrics). A closer look at the development and the current achievements of spatial analysis shows us that - despite the very large number of rather diverse contributions in this area - two main fields can be identified for which spatial analysis may claim to have provided an important contribution (see figure 1):

- **Statistical spatial data analysis** providing more adequate and specialized frameworks and methodologies to deal with a wide range of spatial effects and spatial process models,

- **Spatial modelling** including a wide range of different models such as, for example, deterministic and stochastic process models as well as environmental policy models in the environmental sciences, and location-allocation models, spatial interaction models, spatial choice models and regional economic models in the social sciences.

These two fields which have the potential to enrich GIS technology will be considered in some more detail in the following sections.

### 3 SPATIAL DATA ANALYSIS AND GIS

#### 3.1 Basic Functions of a GIS

It is widely recognized that geographic information systems may be considered to perform four basic functions on spatial data (see, for example, Goodchild 1987, Anselin and Getis 1992, Fischer and Nijkamp 1992c):

- **data input** (data model, data measurement),
- **data storage, retrieval and database management**,
- **data analysis** (data manipulation, exploration and confirmation),
- **output** (display and product generation).

The way in which geographic space (i.e. the 'real' geographical system) is measured and structured in a spatial database depends on the data model. There are two fundamental approaches to represent the spatial component of geographic information: the vector and the raster data model. The first approach views geographic space as an empty space populated by objects that are represented as points, lines or areas, after adequate generalization of the form of objects. This view seems to be more important for spatial analysis in the social sciences where discrete entities are considered as interacting over space. The second approach subdivides the space into a set of fields which may be modelled in different ways such as, for example, a raster of cells (e.g. a remote sensing scene), a raster of regularly spaced point samples (e.g., a digital elevation model), a set of non-overlapping space-exhausting polygons (e.g., a map of soil cover), and a map of digitized isolines (e.g., a contour map). This view is more relevant in the environmental sciences. It is important to note that the choice of the data model also determines the set of analysis that can be carried out (Goodchild et al. 1992).
Based on the view by Anselin and Getis (1992) the interaction between the four basic functions of a GIS may be displayed in a simplified form as shown in figure 2. At the one end of the figure there is the 'perceived real' geographic space (simply termed 'perceived reality'), at the other the GIS user concerned with research and/or policy issues to be solved. In between there are the four functions data input, data storage, data analysis and output. All these functions are available in existing GIS. The analysis functions, however, are generally limited to data manipulation procedures.
only, including partitioning, aggregation, overlay and interpolation procedures. This is what GIS vendors generally understand under 'spatial analysis'.

Two issues will be briefly addressed in the sequel, first, the question which statistical spatial data analysis techniques should be considered for inclusion in a GIS to increase the analytic functionalities of a GIS, and, second, how the integration of GIS and statistical spatial data analysis might be achieved.

3.2 Which Statistical Spatial Data Analysis Techniques?

Despite recognition that spatial analysis is central to the purpose of a GIS, the lack of integration of GIS and spatial data analysis, and the relative simplicity of many GIS is seen as a major impediment to the full utilization of the potential of geographical information systems. The need to develop GIS-relevant statistical spatial analysis tools is currently the subject of research projects at both the NCGIA and the RRLs (see Openshaw 1991a,b, Anselin and Getis 1992, Goodchild et al. 1992, Fotheringham and Rogerson 1993). There are four major areas where statistical spatial data analysis techniques might be strengthen current GIS-practice:

- the sampling of (observational) objects from the database and the choice of the adequate spatial scale of analysis,

- data rectification to compare variables which are defined for the same study area, but for a different and incompatible set of zones,

- exploratory spatial data analysis (ESDA) aiming to explore and exploit the GIS database to arrive at new insights, including the search for data characteristics such as trends, spatial outliers, spatial patterns and associations,

- confirmatory or explanatory spatial data analysis (CSDA) concerned with systematic analysis of data and hypothesis testing based upon specified assumptions.

Methods of data sampling, such as, e.g., the DIST (Dependent Areal Units Sequential Technique) procedure aiming to identify areas to be included in a sample so as to maximize the information content of the sample (Arbia 1991), may have an important role to play for extracting information from the databases.

A frequent problem for GIS users in a socio-economic context is the need to transfer data between incompatible sets of areas, more generally between incompatible spatial referencing systems. Some aspects have already been analysed, but more attention needs to be paid to developing practical and operational procedures. Recently, Flowerdew and Green (1991) developed a promising method based on the expectations maximum likelihood (EM) algorithm which offers a flexible and effective way to improving the various methods of areal interpolation at least for the large class of cases where the variable to be interpolated is a count. It seems likely to be possible to integrate the method with GIS software in a user friendly way to circumvent the problem of incompatible areal units.

There is a clear need for a quantitative exploratory style of spatial analysis which can complement the map-oriented nature of GIS, especially in data rich and theory poor
GIS environments. Generally, there is a broad agreement in the GIS and spatial analysis communities that general data exploratory procedures developed in mainstream exploratory data analysis would be of great value within GIS. The techniques considered as primary candidates include techniques such as, e.g., box plots, Tukey stars, nearest neighbour, K-function plots, scatter plot matrices, autocorrelation tests, classification and spatial classification procedures (regional typification, regionalisation) for (spatial) data simplification, outlier and spatial outlier detection methods. But there are many practical difficulties which have to be faced in this context. Especially, spatial data suffer from errors, which non-spatial data not do, and these errors propagate and further reduce the quality of the data (Openshaw 1991a,b). Some techniques appear to be limited and might be not easy to adapt for very large GIS data sets.

The lack of a priori hypotheses and theory seems to make ESDA more important in a GIS world than explanatory spatial data analysis. Nevertheless, confirmatory data analysis is also of relevance. But very little has been achieved in terms of confirmatory spatial data analysis. Spatial autocorrelation problems have been well analysed. Despite most applications of explanatory data analysis are non-spatial applications of regression analysis and fail to exploit the information on the topology of the observations. This is primarily caused by the lack of adequate software (see Haining 1990). Recently, some advances have been made to develop a number of macros and specialised software to carry out spatial data analysis within standard statistical/econometric packages (see Griffith 1989, Anselin 1989, Bivand 1991). So far the linkage between this software and a GIS is very limited (Anselin and Getis 1992).

There seems to be also a need to develop new innovative spatial data analysis tools which are different in methodology and more suited to data rich and theory poor GIS environments. Neurocomputing might offer the basis for such new approaches which may establish a new style and era and lead to an automated modus operandi in spatial data analysis in general, and in ESDA and CSDA in particular. Neural based approaches seem to be especially well suited to cope with problems such as very large volumes of spatial data, missing data, noisy data and fuzzy information for which conventional statistical techniques may be inappropriate or cumbersome to use (see Openshaw 1992, Fischer 1993a, Fischer and Gopal 1993).

3.3 Linking Spatial Data Analysis and GIS

The problem of linking new spatial analytic functions to standard GIS is beginning to emerge as an important research area (see, e.g., Goodchild et al. 1992, Anselin and Getis 1992). Various logical ways of coupling spatial data analysis and GIS can be identified and there is work underway to explore these various possibilities. In principle, Goodchild (1992) distinguishes three major approaches:

- **full integration** of spatial analytic procedures within the GIS software,
- **close coupling** between statistical spatial data analysis software and GIS, and
- **loose coupling** where an independent spatial data analysis module relies on a GIS for its input data, and for such functions as graphic display, via import and export data in a common format.
The first approach aiming to integrate spatial analysis techniques fully within GIS software is essentially not yet realized, due to the lack of analytical capabilities in most commercial GIS. Possibly, Openshaw's idea to develop a spatial analysis tool kit comes close to this approach (Openshaw 1990). This approach would have several important advantages. Especially, software would be well documented and supported by the vendors. The spatial analysis techniques would be available to all users of the GIS software. But it might be difficult to persuade GIS vendors to adopt such an approach without corresponding pressure from the market for spatial data analysis facilities (Goodchild et al. 1992).

The second approach seems to be much more realistic in nature. The unresolved problem refers to the nature of the user-friendly interface for the linkage. Possibilities range from source code through subroutine libraries, high-level languages, sequences of commands, menus to virtual reality. These offer the developer access to the standard user-interface facilities of the GIS package in question and to (locational and attribute) data structures held by the GIS. Even though this approach links a statistical package to a GIS, it is limited to simple descriptive measures up to now (see, e.g., Kehris 1990). An effective form of close coupling which data could be passed between a GIS and a spatial analysis module without loss of higher structures, such as topology, object identity, metadata, or various kinds of relationships, is still missing (Goodchild 1992). A challenge for the future is to develop a standard spatial query language for spatial data which permits users to access spatial data without knowing the particular data structures being used in the database (Goodchild et al. 1992).

The third approach seems to be the strategy most adopted in practice. It uses a GIS package for the things it is best suited and, when spatial analysis facilities are required, taking the output from GIS as input into statistical software package. The coupling is achieved via the interchange of ASCII files or the transfer data in common data formats between the packages. The major drawback of this strategy is the failure to take the distinctive characteristics of the GIS database into account for use in spatial data analysis.

4 SPATIAL MODELLING, GIS AND DECISION SUPPORT

4.1 Spatial Models

Spatial analysis goes far beyond data sampling, data manipulation, ESDA and CSDA into areas of spatial modelling encompassing a large and diverse set of models in both the environmental and the social sciences. In the environmental sciences the models range from physical dispersion models (e.g. for suspended particulates), chemical reaction models (e.g. photochemical smog), and biological systems models (e.g. for ecosystems in water) through deterministic process and stochastic process models (e.g. for erosion and groundwater movement) to comprehensive models of regional environmental quality management and integrated models for environment-energy-economic assessment.

For many models, especially the physical dispersion models, chemical reaction models and biological systems models the geographical location of the site is not considered to be of overriding importance. The variation in model results over space can be achieved by obtaining the inputs from different spatial entities (e.g. polygons) and relating the model inputs to the area in question.
Process models include deterministic versions which attempt to describe a particular process in terms of known physical laws, and stochastic model versions which aim to describe a particular process such as erosion, groundwater movement and absorption of pollutants in terms of stochastic theory. For such processes there is an interaction between the spatial process and a substrate. This substrate provides a 1-, 2- or 3-dimensional framework within which the process model can operate. The substrate itself may be unaffected by the process as in the case of simple runoff or may be modified as in the case of an absorption of pollutants, erosion or transport. The basic substrate (e.g. catchment or slope facets) is called a finite element. The space in which the process is being modelled is disaggregated into a set of finite elements which are usually assumed to be internally homogenous. 2-dimensional models generally use small pixels. In contrast to 2-dimensional models, 3-dimensional models used in groundwater modelling cause much greater storage problems. For example, already a 60x60x60 pixels array requires 216,000 memory elements (Burrough 1991).

Models of regional environmental quality management are comprehensive environmental policy models which are useful in assessing a number of alternative options of managing environmental quality according to some economic criteria. Typically, they include five major components: first, an economic model of production and consumption; second, a model of residuals and emissions (and occasionally assimilation of wastes into media and biota); third, an environmental model which translates the space-time pattern of environmental discharges into spatio-temporal states of natural environment (e.g. Gaussian plume type model of atmospheric dispersions of $SO_2$ and particulates aquatic ecosystem models which estimate residual concentrations); fourth, a model of effects of such concentrations or receptors in terms of damage functions; and finally, an analytical component on management strategies which compares the benefits and costs associated with alternative environmental management strategies (Lakshmanan and Bolton 1986).

Integrated models for environment-energy-economic assessment link energy system models, models of the national-regional economy and models of environmental emissions in the form of integrated impact models to assess interrelated energy supply, environmental quality and economic policies. They usually share some common characteristics. First, they contain some version of a macroeconomic model; second they represent production and consumption; third, there are clear linkages between the energy sector, the rest of the economy and the environment; fourth, they have a reference energy supply system and depict energy technologies, demands and pieces, and finally the represent space in a discrete manner (multiregional framework) (Lakshmanan and Bolton 1986). These models seem to have taken the form of drawing the model components not only from the environmental and ecosystems modelling traditions, but also from different analytical traditions in the social sciences.

In the social sciences in general and in quantitative geography as well as in regional economics in particular a wide range of spatial models has been developed in the past decades for the purposes of describing, analysing, forecasting and the policy appraisal of economic developments within a set of localities or regions. Such models deal not only with internal structures of regions and relationships within one region, but also with interregional interrelationships. The tradition of models in quantitative geography and regional science has shown a strong orientation towards location-allocation problems in space (e.g., siting of retail centres, employment location) (Baumann et al. 1983, 1988), witness the wide variety of spatial interaction models.
(see, e.g. Haynes and Fotheringham 1988) dealing with flows of people, commodities and resources between regions (journeys-to-work, journeys-to-shop, recreational trips, migration, spatial interaction on a transportation network, interregional commodity trade, capital and information flows). Most of these models were static in nature and regarded space as a set of discrete points (areas, grids) rather than as a continuum. The problems of ecological fallacy and spatial auto-correlation are essentially caused by the sometimes arbitrary spatial demarcation in spatial modelling and have led to various misspecified models (see, e.g., Baxter 1987), while the early studies in spatial choices and processes during the sixties and early seventies were dominated by spatial interaction (gravity) models justified by probability and entropy maximizing formulations. This lack of behavioural context begun to be criticized in the 1970s and gave rise to the study of individual choice behaviour in various contexts (journey-to-work, journey-to-shop, migration etc.) This, in conjunction with the parallel development of discrete choice models made it possible to propose new disaggregate choice based alternatives. Unlike spatial interaction models, these alternatives could link explicitly individual decisions at the micro level with population flows and other observables at the macro-level (see Fischer et al. 1990).

The development of regional and interregional economic models occurred almost contemporaneously with the growth of interest in national-level input-output modelling. The regional input-output model may be clearly seen as a model of production side of a regional economic system in which demand and changes in demand create the signals for production to take place. One of the most promising developments in (inter)regional input-output modelling in recent years has been the extension of the modelling framework in two major ways, first, the extension of the input-output model itself, and, second, the linkage of input-output analysis to take account of the wider socio-economic context in which individual activity takes place (Hewings and Jensen 1986). Various extensions have been suggested, e.g. to take into account labour markets and migration (see Batey 1985).

Other models are concerned with the regional labour and housing markets, land use and transportation, and multi-objective decision analysis (see Nijkamp 1986). In general, spatial models can be used for three purposes, viz. forecasting and scenario generation, policy impact analysis and policy generation and/or design. A major hurdle in the way of practical application and interpretation of spatial models is the specificity of each practical situation: the spatial arrangement, size, shape and organization of the basic spatial units (cells or lattice) for which data is gathered affects the spatial process which can be identified in the parameters. The parameters are not independent of the spatial framework used. This problem which has been extensively recognized in spatial statistics has been only relatively recently approached in spatial modelling. But there are now extensive simulation results available which illustrate the effects of different spatial frameworks in spatial modelling (see Openshaw 1977 for spatial interaction models, Baumann et al. 1983, 1988 for multiregional labour market modelling).

All these models can perform functions which proprietary geographic information systems are almost consistently missing, but which are important for the sort of questions policy decision makers are interested in. On the other hand, large sized models tend to become unmanageable, caused by the increase of computer speed, availability of library programs for econometric routines and computerised data banks. GIS seems to offer a great potential to facilitate the task of the researcher.
4.2 Interfacing Spatial Modelling and GIS

GIS and spatial modelling can be considered as complementary approaches especially in two respects (Fischer and Nijkamp 1992b):

- in the area of spatial pattern and flow recognition, where spatial differences - in multiple dimensions - can be shown either by statistical representation or by GIS computer graphics and maps; the choice for one of the two representations depends on the complexity of the pattern to be represented;

- in the area of explanatory and predictive analysis, spatial models are usually much more powerful than GIS in carrying out precisely numerical experiments, but the final results can again be used by a GIS as an input for a user-friendly computer presentation.

The main problem is that current GIS are less suitable for exploratory or predictive modelling compared to conventional spatial modelling. Nevertheless, it is noteworthy that some progress has recently been made, e.g., in the field of spatial location-allocation modelling, where traditional spatial interaction tools have been linked to a GIS representation of the resulting patterns and flows. To some extent one may claim that GIS seems to be a more proper tool for perception/visualization methods for impact analysis rather than a direct impact tool itself.

Also in dynamic modelling the same remarks essentially hold. GIS is so able to produce dynamic maps in connection with the result of a dynamic descriptive explanatory or predictive dynamic model. In this context GIS is a meaningful vehicle for dynamic scenario analysis. Good examples can be found in Grossmann and Eberhardt (1992) who suggest to link complex aggregated dynamic feedback models, simple generic dynamic models (in particular object oriented models) and physical dispersion models (e.g., for heat conduction, diffusion of noise or transport of gaseous pollutants) with GIS to produce a time series of maps (so-called dynamic maps).

In recent years, much attention has been given to the coherent and joint use of GIS and spatial models in spatial planning. This raises the question how the interface between GIS and spatial modelling might be achieved. In principle, one can think of several strategies. At one extreme, there is the strategy to entirely incorporate the GIS into the modelling framework (see figure 3a). This is, however, unlikely in that most geographic information systems contain many more functions that are relevant to any single spatial model or set of spatial models. At the other extreme, there is the strategy favoured by the designers of GIS to entirely embed spatial models into the GIS and to consider spatial models as simply additional functions to a GIS (see figure 3b). There are several examples such as SPANS, TRANSCAD and ARC-INFO which embed spatial models into GIS. But any geographical systems which is extensive enough to embrace a set of relevant more complex planning models is unlikely to be in an integrated package and must be a set of relatively independent modules which may be interlinked in diverse ways (see figure 3c) (Batty 1992a).

Up to now, GIS and spatial models have become closest in addressing problems of decision-making in form of spatial decision support systems. The term decision support system was coined in the late 1970s to embody a framework for integrating database management systems with analytical and operation research models,
graphical display and tabular reporting capabilities, and the expert knowledge of decision makers to address business problems (strategic planning, scheduling of operations and investment appraisal). It is this usage which has been adopted by those researchers working with operation research models in a spatial setting (Batty 1992). Models based on the location of facilities ranging from those based on formal mathematical programming and optimization to those based on more informal processes of heuristic research have been quite widely developed as part of spatial decision support systems (SDSS) (see Densham and Rushton 1988, Densham 1991).

4.3 Some Examples

From a broad spectrum of interesting and promising approaches we have selected a few examples in the interface between environmental and social sciences which, however, by far do not claim to be representative nor exhaustive.

In a number of projects, IIASA has developed and (partly) implemented several computer based information and decision support systems on a regional or national scale. These systems designed for regional development planning generally include (Fedra et al. 1991):

- *simulation and optimization models* representing procedural, quantitative and numerical assessment tools,
- *databases* containing non-spatial data,
- *geographical information systems* providing access to spatial data such as environmental properties, land use, infrastructure, resources, site-specific information (e.g. industrial plants),
- *expert systems* representing qualitative, symbolic and logical assessment tools and knowledge about the problem domain, and
- *graphical user interface* integrating the above mentioned components into one coherent system and providing the user with interactive access.
Three examples may be chosen to illustrate the activities taking place at the International Institute for Advanced System Analysis IIASA in Laxenburg (Austria) (Fedra et al. 1991). The first example refers to the Regional Air Pollution Management System. Sponsored by the former CSFR Federal Commission for the Environment, IIASA started to develop a regional environmental information system for air quality management problems such as emission control options to meet specific environmental standards in Northern Bohemia. The system provides interactive tools to display and analyse environmental control measures including investment requirements based on cost functions for various pollution control technologies, focuses on air pollution from industrial sources, and uses a system of nested multi-layer air quality models covering both the entire region as well as selected urban and industrial centres.

The second example refers to Environmental and Technological Risk Assessment on a National Scale. Main emphasis of the ongoing research for the Dutch Ministry of Housing, Physical Planning and the Environment is the integration of a series of technological and environmental risk assessment models and tools into a coherent environmental information and decision support system. Key components encompass an object-oriented GIS, databases on major risk installations, sources of pollution, and a hazardous substances database. Safety reports as required by the Seveso Directive in Dutch legislation can be compiled and analysed for both fixed installations as well as along transportation routes based on technological data and spatially variable population density, land use, and meteorological data.

The third example is a Model-Based Information and Decision Support System for Integrated Regional Development of China's Shanxi Province. The system was designed to assist the regional government in questions of integrated regional development planning and comprises in its current stage of development the following major modules:

- the Knowledge-Based Integration Manager (an experimental version of a LISP-based model integration manager based on an object-oriented approach),
- the Macroeconomic Symbolic Simulator providing the user with the possibility to conduct a dynamic simulation of the macroeconomic behaviour of the Shanxi province,
- a number of conventional econometric models including multi-criteria optimization and scenario comparison modules,
- a global investment analysis module for the integrated economic development of the province,
- an investment distribution model describing conflict resolution between urban and rural development in terms of investment distribution based on the theory of cooperative games,
- a module for interregional comparison at a macroeconomic level based on macroeconomic indicators to facilitate comparison of different regions or different development stages of the same region,
- a transportation system analysis model allowing to analyse investment requirements for capacity extensions of the current transportation system,
- a prolog-based model of spatial choice and siting for analyzing the feasibility of industrial or socio-economic activities,
• a spatially disaggregated optimization model describing a broad set of industries (including mining, the building materials sector, the energy production sector, chemical and metallurgical industries) to analyse and optimize the distribution of production capacities, and using mathematical programming techniques to balance the material flows,
• an energy demand simulation model describing energy intensive industries in terms of their energy demand, water demand and basic economic behaviour (including investment),
• a coal mine development model for the coal economy in the province, based upon dynamic simulation concepts,
• a hydro-economic simulation model providing a dynamic analysis of water demand-supply budgets for river basins,
• an air pollution simulation model based on an extended Gaussian plume equation and describing the concentration and deposition of substances in time and space,
• a geographical database including topics such as mines, mineral resources, industrial locations, road networks etc.,
• a specific interactive and graphical stand-alone implementation of a discrete multicriteria decision support system from IIASA's Dynamic Interactive Decision Analysis and Support Systems program family.

Comparatively less comprehensive, but more GIS based is Despotakis' (1991, 1992) hybrid GIS-DSS model for sustainable development planning. In a large scale regional development study Despotakis (1991) focused attention on principles and actual policies for co-evolutionary planning and development of the Greek Sporades islands. A geographical information system (GIS) which may assist Sustainable Development (SD) for an area was developed in this study. This GIS-SD system was linked to a Decision Support System (DSS) for an evaluation of its spatio-temporal results in the context of selected SD criteria. The whole system was then applied to the test area of the Greek Sporades islands in the Aegean sea, for which relevant data at various spatial levels and with different degrees of reliability were selected. The results showed that GIS may be successfully employed to assist, monitor and control ecologically sustainable economic development for a region. This is achieved by applying certain strategic policies which are based on optimum - from the sustainability point of view - development scenarios generated by the system. The underlying model system was based on a combination of meso-economic models for the area and micro land use developments (including the marine environment). GIS appeared to be able to create the intermediate link between these layers of the model system, while the results - originating from scenario and simulation experiments can be visualized in an attractive computer graphic way. Various development alternatives for the islands group could next be evaluated by means of multi-criteria analysis.

Several other examples may be found in Fischer and Nijkamp (1992a), such as the ISP computer system for interactive spatial planning for generating and evaluating land use plan alternatives (Roy and Snickars 1992) and a prototype GIS for urban program impact appraisal in St. Helen, North West England (Hirschfield et al. 1992).

5 SUMMARY AND OUTLOOK

There is a wide agreement in both the GIS and the modelling communities that the future success of GIS technology will depend, among other factors, on the extent to
which it can incorporate more powerful analytical and modelling functionalities.
Spatial analysis encompassing spatial data analysis and spatial modelling, which deals
with the description, explanation and prediction of spatial phenomena occurring in
spatial and/or space-time systems, offers a wide range of methodologies and
procedures which are highly relevant to GIS research.

Geographic information systems are often either explicitly or at least implicitly
designed to assist decision makers. The capabilities of current geographic information
systems for specialist forms of spatial search and decision making are limited. A
successful GIS has to attempt to support spatial data analysis and spatial modelling,
and produce results required by decision-makers in environmental, urban and regional
planning. Scholars around the world are currently looking at fields of artificial
intelligence (especially expert systems) to make geographic information systems and
spatial analysis more efficient and more friendly to users with limited computer
analytical literacy. From the viewpoint of the next generation of more intelligent
graphic information systems spatial analysis tools and current GIS technology
may be considered as intermediate tools for developing knowledge-based information
systems supporting activities in urban, regional and environmental planning and
decision making.

The idea of the GISDATA Specialist Meeting on GIS & Spatial Analysis to be held in
Amsterdam, December 1-5, 1993 is to stimulate fundamental research in this area in
bringing together scholars from across various European countries, and to establish a
research network in which the participants of the meeting will form the initial nucleus.
This network should encourage the sharing of data and computer programs, evaluate
and promote alternative innovative spatial information processing approaches, and
generally work towards developing a critical mass of scholars on model-based GIS in
Europe, which is considered to be critical to foster the diffusion of GIS-technology
and its further development.

The meeting itself will focus on two major lines of research: Exploratory Spatial Data
Analysis (ESDA), and Explanatory Spatial Data Analysis and Spatial Modelling. In
both areas emphasis will be laid on methodological issues such as to explore and
evaluate the potential of new innovative methodological approaches, as well as
environmental applications and socio-economic applications.

REFERENCES

GAUSS, Discussion Paper No. 2, Institute of Mathematical Geography, Ann Arbor.

Anselin, L. (1991a) Quantitative methods in regional science: Perspectives on
research directions, in Boyce, D., Nijkamp, P. and Shefer, D. (eds) Regional Science:

Department of Geography, University of California, Santa Barbara, CA.

information systems, in Fischer, M.M. and Nijkamp, P. (eds) Geographic Information
Systems, Spatial Modelling, and Policy Evaluation, pp. 35-49. Springer-Verlag,
Berlin et al.


