*Geographic Information Technology Training Alliance (GITTA) presents:*

# Introduction to spatial analysis

**Responsible persons: Helmut Flitter, Robert Weibel, Stephanie Rogers**

# Table Of Content

# 1. Introduction to spatial analysis

Operations and data analysis tools, especially spatial analysis, are important components for the functionalities of a GIS. In the module "Basic Spatial Analysis" background information is provided about the performance and the application of operations in Spatial Analysis (SA). The module "Intermediate Spatial Analysis" offers advanced information about this topic in terms of special operations.

## Learning Objectives

- You know the origin of SA as well as some relevant literature about this.
- You are able to describe the term "Spatial Analysis" concisely and know the position of SA in the GIS context.
- You are able to give an overview of essential SA functions in GIS and are able to assign these functions to various criteria.

# 1.1. GIS: Background and overview

This lesson will introduce the subject of spatial analysis (SA). In the first part, the historical background of SA in GIS will be discussed, followed by the concept of SA. Typical functions of SA are given in the last section.

## 1.1.1. Roots of Spatial Analysis in GIS

Despite their central role in geographic information systems, most spatial analysis methods emerged before and independently from GIS technology. Most of these methods have been integrated later into the GIS technology. GIS platforms provide several functions such as data acquisition, data management and visualization. The combination of these functions with analytical operations makes them even more efficient.

The origin of spatial analysis, known in the context of GIS of today, goes back many years. A selection of early articles about spatial statistics and quantitative spatial analysis is provided in the reader written by Berry et al. (1998). Some of these articles were written in the 1930s, but most of them were written in the 1950s and 1960s. The use of statistical and other quantitative methods for the analysis of spatial patterns and processes were especially prominent and developed in spatial sciences, which focus on the analysis of those spatial patterns and processes. Examples of such disciplines are regional science, quantitative geography or landscape ecology. In geography, the use of quantitative methods was coupled to a change of paradigm, called the "quantitative revolution". This trend was particularly pronounced in the 1960s and 1970s. It is well described in the books written by Abler (1971) as well as by Haggett (1965) and Haggett et al. (1977). At that time, geodetic sciences did not focus on the development of SA methods. However, since the 1980s, as GIS became a more integral platform, there was a growing convergence of methods used in SA involved in the development of GIS technology sciences. This development is well illustrated in the book written by Bill (1999).

With his background in geomatics, he devoted a large part of the content to data analysis, focusing specifically on SA . But which methods were developed before the advent of GIS technology? As stated in Fisher (1999), these were mainly quantitative methods for the characterization and analysis of:

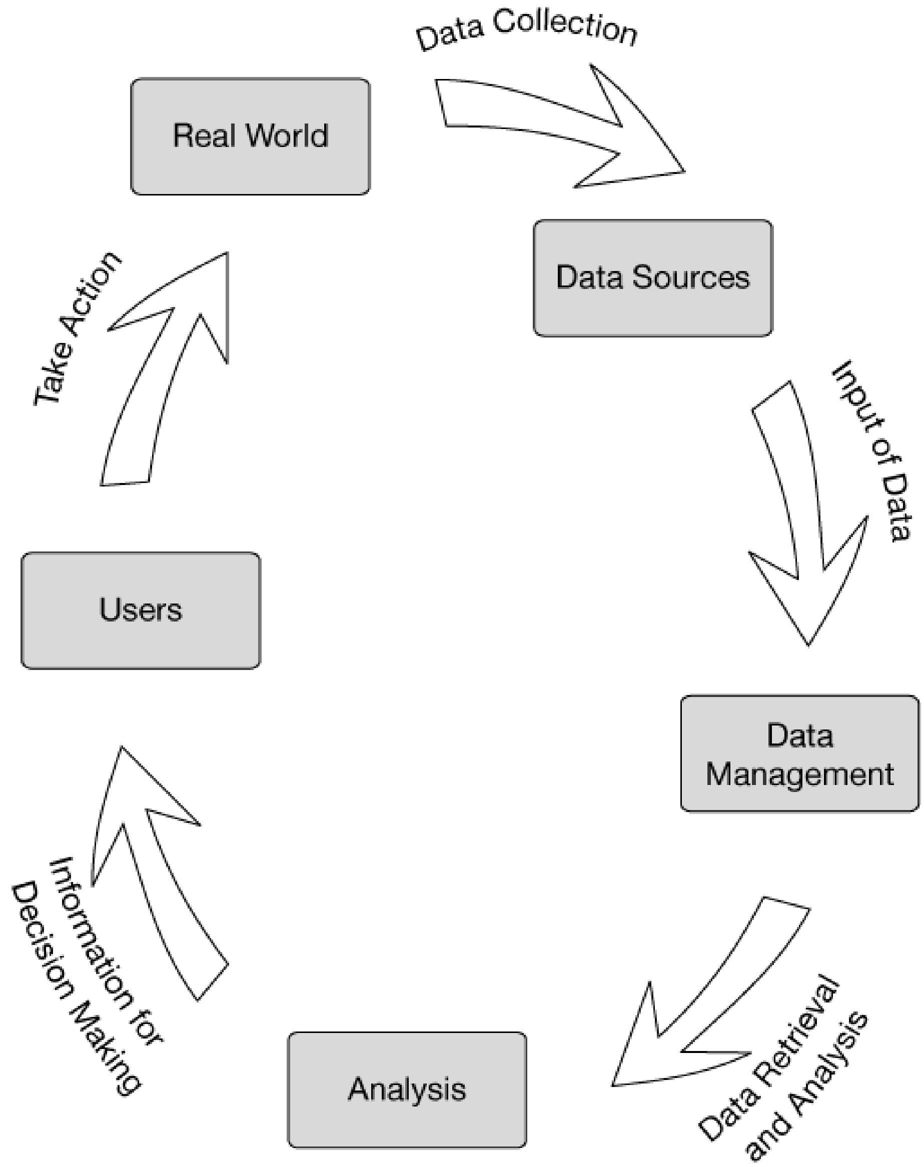
- Patterns (e.g. distribution and arrangement)
- Shapes of geographic features (points, lines, areas and surfaces).

Initially, a geometric approach dominated the field, which was particularly concentrated on point pattern analysis as well as the characterization of networks. Later, the focus changed to the development of methods. The analysis of the intrinsic properties of the geographical area (e.g. the relative distance between spatial objects), processes of spatial preferences (e.g. positioning of shopping centers) and the analysis of spatial interactions, became more important in the development of methods. Within this time of development, methods of multivariate statistics were increasingly used and adjusted to the needs of spatial science; standard statistical packages have been linked to GIS for exploratory data analysis, statistical analysis and hypothesis testing. Examples for such statistical methods are stated below:

- Regression analysis
- Principal component analysis
- Linear discriminant analysis
- Trend surface analysis

As these methods were developed in other disciplines rather than spatial science, their use in spatial related analysis caused several problems as these methods did not take spatial heterogeneity or spatial dependence of samples distributed in space into account. As a consequence, methods for geostatistical analysis were developed from the late 1960s that perform better with spatially related data.

By the late 1970s, geometry#oriented operations of spatial analysis (e.g. spatial query, point pattern analysis, polygon overlay, buffering etc.) and methods of geostatistical analysis were being made available in GIS. These days, typical commercial GIS packages provide a wide range of functions for geometric and geostatistical analysis. A big advantage of the incorporation of analysis functions in GIS is the connection between the tools of data acquisition, data modeling, data management and visualization. This enables the entire cycle of spatial data processing to be completed using one software to answer complex questions.




## 1.1.2. Description of the Term "Spatial Analysis"

The definition of the term "spatial analysis" raises some problems, as stated by Bailey:
"*One difficulty experienced in any discussion of links between GIS and spatial analysis is clarification of exactly what is to be considered as spatial analysis. The problem arises because, by its nature, GIS is a multi-disciplinary field and each discipline has developed a terminology and methodology for spatial analysis which*

*reflects the particular interests of that field. In the face of such a diversity of analytical perspectives, it is difficult to define spatial analysis any more specifically than as: a general ability to manipulate spatial data into different forms and extract additional meaning as a result*" (Bailey 1994, p. 15)

This means that the questions and methods of spatial analysis grew "naturally". They have been developed in various sciences related to GIScience, which focus on different interests and topics of investigation. For example, a distance query to reveal all ski resorts within a certain distance could be a simple data retrieval for some, but for others a query could represent a complete spatial analysis.

The attempt at a definition fails. There is no possibility to define the term of spatial analysis but only to describe it. In order to give an idea, the following descriptions can be mentioned:

"*A general ability to manipulate spatial data into different forms and extract additional meaning as a result.*" (Bailey 1994, p. 15)

"*In broad terms one might define spatial analysis as the quantitative study of phenomena that are located in space.*" (Bailey et al. 1995, p. 7)

The next section offers a wider and more complex view on the term "spatial analysis", which can be given various meanings. But first the position of SA in the GIS context should be discussed. There are the two major streams within quantitative SA: geometrically oriented and geostatistically oriented. A third stream, visual data analysis, recently arose. It is well known that maps can provide information about spatial patterns and processes (more details are provided in the module "Basic Presentation"). These three streams mentioned can be characterized as follows, using different approaches:

**Geometric SA:**

The geometric approach is focused on geometric criteria (location of objects and attributes), and has mainly a descriptive effect. It cannot be used for hypothesis testing. Some examples using geometric SA include: the analysis of point distributions, network analysis (route calculation, shortest path), polygon overlay, analysis of distance relations, shape analysis, or the calculation of slope and exposure in elevation models.

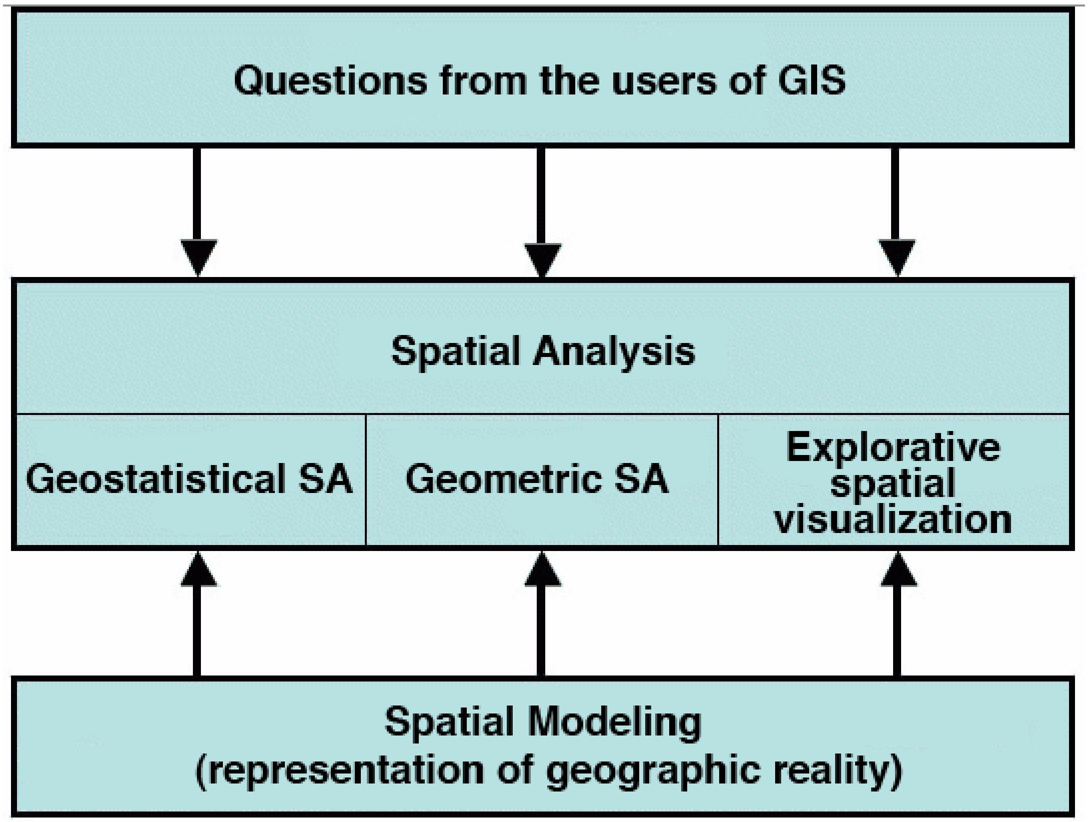
**Geostatistical SA:**

The geostatistical approach refers to spatially distributed (random) variables. This approach uses statistical methods for description but also for hypothesis testing. Some examples using geostatistical SA include: multivariate statistics, spatial correlation analysis (spatial autocorrelation), and geostatistical interpretation.

**Explorative spatial visualization:**

This is a purely visual approach, in visualization as well as in interpretation. Visualization is a qualitative and explorative method to explore new data, to identify outliers and to formulate hypothesis.

These combined approaches offer a range of possible methods in spatial analysis as shown in the following image. Spatial analysis is not used as an end in itself but to get answers. Therefore, methods to represent phenomena and processes are needed but also appropriate methods to analyze these phenomena and processes. That means that a direct relationship exists between the methods of modeling and representation of geographic reality and spatial analysis techniques. The methods of spatial modeling and their impact on spatial analysis are discussed in the module "Basic Spatial Modeling".



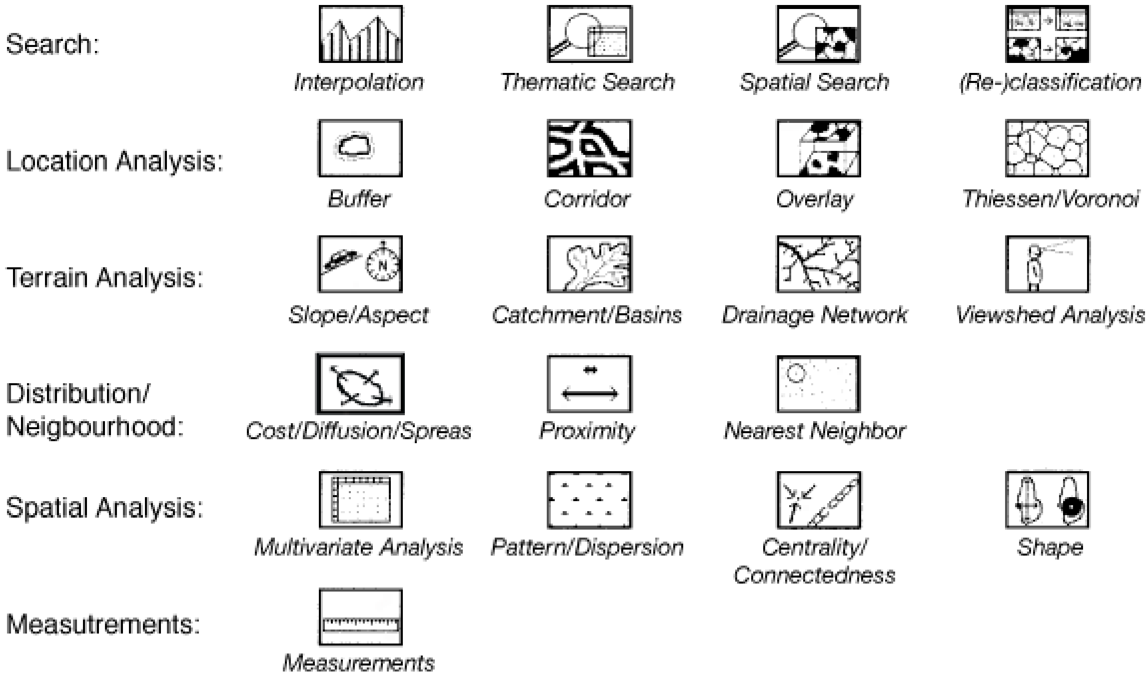

## 1.1.3. Functions of Spatial Analysis

What operations are typically used as spatial analysis functions in GIS? Chou (1997, p. 15) provides the most efficient description. Three different types are illustrated:

- Attribute query
- Spatial query
- Derive new data from existing data

Note that only some of these operations generate new data. The first two functions mentioned are simple queries, and the result consists of a selection of objects from the databases.
Spatial analysis functions can also be classified in regards to the data type involved in the spatial analysis (point, line, network, polygons/areas, surface), the data structure (vector vs. raster), or the conceptual model of space (discrete entity vs. continuous field (Burrough et al. 1998)).
Bailey et al. (1995) and Abler et al. (1971) propose another differentiation of the functionality of spatial analysis. They distinguish the functions by their level of dynamics: E.g. static data (point distribution, surface area etc.), interactions between objects in space (for example interactions between two economic centers), or analysis of spatio-temporal changes). Depending on the author's point of view and his knowledge, organization and classification of SA functions are defined differently.
The lesson of the module "Basic Spatial Analysis" is organized using a mixed approach. The module is composed of lessons dedicated to the different analysis and application functions such as terrain analysis, accessibility analysis, and suitability analysis, etc. Albrecht (1996) provides a nice classification of SA functions used in GIS (see also the illustration below). This classification was developed to deliver a universal interface for GIS. That is why this classification has two advantages: It gets the user's aspect (not the technical

aspect), it is the briefest description, but it is a complete description of SA functions available in GIS (at least in commercial GIS). The term "spatial analysis" could be irritating. For this reason, a better description of SA would be "pattern analysis".




*Classification of spatial analysis functions* (Albrecht 1996)

There is a very pragmatic approach in the following list provided by Goodchild (1990) based on Goodchild (1987). It provides the typical function range of commercial GIS and operation components in data analysis. Other books provide overviews and discussions on the topic of the typical SA functions used in GIS, such as Aronoff (1989), Bill (1999), Burrough et al. (1998), or Jones (1997).

**Data analysis functions according to Goodchild (1990)**

**Counting and measuring:**

- Measure number of items: The ability to count the number of objects in a class.
- Measure distances along straight and convoluted lines: The ability to measure distances along a prescribed line.
- Calculate bearings between points: The ability to calculate the bearing (with respect to True North) from a given point to another point.
- Measure length of perimeter of areas: The ability to measure the length of the perimeter of a polygon.
- Measure size of areas: The ability to measure the area of a polygon.
- Measure volume: The ability to compute the volume under a digital representation of a surface.

**Function of spatial analysis:**

- Point in polygon: The ability to superimpose a set of points on a set of polygons and determine which polygon (if any) contains each point.

- Line on polygon overlay: The ability to superimpose a set of lines on a set of polygons, breaking the lines at intersections with polygon boundaries.
- Polygon overlay: The ability to overlay digitally one set of polygons on another and form a topological intersection of the two, concatenating the attributes.
- Sliver polygon elimination: The ability to delete automatically the small sliver polygons which result from a polygon overlay operation when certain polygon lines on the two maps represent different versions of the same physical line.
- Nearest neighbor search: The ability to identify points, lines or polygons that are nearest to points, lines or polygons specified by location or attribute.
- Shortest route: The ability to determine the shortest or minimum cost route between two points or specified sets of points.
- Contiguity analysis: The ability to identify areas that have a common boundary or node.
- Connectivity analysis: The ability to identify areas or points that are (or are not) connected to other areas or points by linear features.
- Network analysis: Simple forms of network analysis are covered in shortest route and connectivity. More complex analyses are frequently carried out on network data by electrical and gas utilities, communications companies etc. These include the simulation of flows in complex networks, load balancing in electrical distribution, traffic analysis, and computation of pressure loss in gas pipes. In many cases these capabilities can be found in existing packages which can be interfaced to the GIS database.

**Statistical analysis functions:**

- Create lists and reports: This is the ability to create lists and reports on objects and their attributes in user-defined formats, and to include totals and subtotals.
- Calculate – arithmetic: The ability to perform arithmetic, algebraic and Boolean calculations separately and in combination.
- Complex correlation: The ability to compare maps representing different time periods, extracting differences or computing indices of change.

**Terrain modeling:**

- Spot heights: Given a digital elevation model, interpolate the height at any point.
- Heights along streams: Given a digital elevation model and a hydrology net, interpolate points along streams at fixed increments of height.
- Contours (isolines): Given a set of regularly or irregularly spaced point values, interpolate contours at user-specified intervals.
- Elevation polygons: Given a digital elevation model, interpolate contours of height at user-specified intervals.
- Watershed boundaries: Given a digital elevation model and a hydrology net, interpolate the position of the watershed between basins.
- View shed: Identification of the cells in an input raster that can be seen from one or more observation points.
- Generate view shed maps: Given a digital elevation model and the locations of one or more viewpoints, generate polygons enclosing the area visible from at least one viewpoint.

- Calculate slopes along lines (gradients): The ability to measure the slope between two points of known height and location or to calculate the gradient between any two points along a convoluted line which contains two or more points of known elevation.
- Calculate slopes of areas: Given a digital elevation model and the boundary of a specified region (e.g., a part of a watershed), calculate the average slope of the region.
- Calculate aspect of areas: Given a digital elevation model and the boundary of a specified region, calculate the average aspect of the region.
- Locations from traverses: Given a direction (one of eight radial directions) and distance from a given point, calculate the end point of the traverse.

**Complex analysis:**

- Combination of the analysis functions mentioned above.

Try to assign the functions defined by Goodchild (1990) to the spatial analysis functions determined by Albrecht (1996). Specify Goodchild's with Albrecht's categories. Where do problems occur? Are there functions defined by Goodchild which cannot be assigned to Albrecht's categories (that could possibly indicate that Albrecht's classification is not universal)?

## 1.2. Summary

SA is a method used to manipulate spatial data and to produce results. It is used for example to classify objects, to calculate the distance between them or to compute their area. The range of methods consists of the geometric and geostatistical analysis as well as the interactive methods of spatial visualization. In GIS, these methods are used in combination.

In geography, SA techniques have been used as quantitative methods since the 1960s and 1970s. Since the 1980s there has been an increasing convergence between the methods used in SA. Since the embedding of data acquisition, data modeling, data management and visualization methods into GIS, the data processing became more efficient. Initially, the development of analysis methods in GIS was focused on geometry. Thus the analysis was heavily focused on characterization of point distributions and networks. Only later the methods became more focused on intrinsic properties of the geographic space. But at that time, the specific characteristics of spatial variables such as heterogeneity of the measured values as well as the spatial dependence had been neglected. Nowadays, GIS incorporates a big range of geometric and geostatistical analysis functions. The functions of a spatial analysis include classification, distance calculation, area analysis, terrain modeling, visualization and others.

## 1.3. Glossary

**Explorative spatial visualization:**
This is a purely visual approach, in visualization as well as in interpretation. Visualization is a qualitative and explorative method to explore new data, to identify outliers and to formulate hypothesis.

**Geometric SA:**
The geometric approach is focused on geometric criteria (location of objects and attributes), and has mainly a descriptive effect. It cannot be used for hypothesis testing. Some examples using geometric SA include: the analysis of point distributions, network analysis (route calculation, shortest path), polygon overlay, analysis of distance relations, shape analysis, or the calculation of slope and exposure in elevation models.

**Geostatistical SA:**
The geostatistical approach refers to spatially distributed (random) variables. This approach uses statistical methods for description but also for hypothesis testing. Some examples using geostatistical SA include: multivariate statistics, spatial correlation analysis (spatial autocorrelation), and geostatistical interpretation.

## 1.4. Bibliography


- **Abler, R.; Adams, J.S.; Gould, P.**, 1971. *Spatial Organization: The Geographer's View of the World*. Englewood Cliffs, New Jersey: Prentice Hall.
- **Albrecht, J.**, 1996. *Universal Analytical GIS Operations: a task-oriented systematisation of data structure-independent GIS functionality leading towards a geographic modelling language*. (PhD Thesis). Institut für Strukturforschung und Planung in agrarischen Intensivgebieten (ISPA), Universität Vechta.
- **Aronoff, S.**, 1989. *Geographic Information System: a management perspective*. Ottawa: WDL Publications.
- **Bailey, T.C.**, 1994. A Review of Statistical Spatial Analysis in Geographical Information Systems. *Spatial Analysis and GIS (Editor: Fotheringham, S. and Rogerson, P. Publisher: Taylor and Francis, London.)*, 13-44.
- **Bailey, T.C.; Gatrell, A.C.**, 1995. *Interactive Spatial Data Analysis*. Harlow-New York: Longman.
- **Berry, B.J.L.; Marble, D.F.**, 1998. *Spatial Analysis: A Reader in Statistical Geography*. Englewood Cliffs, New Jersey: Prentice Hall.
- **Bill, R.**, 1999. *Grundlagen der Geo-Informationssysteme*. Karlsruhe: Wichmann Verlag.
- **Burrough, P. A.; McDonnell, R. A.**, 1998. *Principles of Geographical Information Systems*. New York: Oxford University Press.
- **Chou, Y.-H.**, 1997. *Exploring Spatial Analysis in GIS*. Santa Fe: OnWord Press.
- **Fisher, M.M.**, 1999. Spatial Analysis: Retrospect and Prospect. *In:* **Longley, P. A.; Goodchild, M. F.; Maguire, D. J.; Rhind, D. W.**, ed. *Geographical Information Systems*. New York: John Wiley.
- **Goodchild, M.F.**, 1987. A Spatial Analytical Perspective on Geographical Information Systems. *International Journal of Geographical Information Systems*, 1(2), 327-334.
- **Goodchild, M.F.**, 1990. Spatial Analysis Using GIS. *In: Fourth International Symposium on Spatial Data Handling*. Zurich: University of Zurich.
- **Haggett, P.**, 1965. *Locational Analysis in Human Geography*. London: Edward Arnold.
- **Haggett, P.; Cliff, Andrew D.; Frey, A.**, 1977. *Locational Models. Locational Analysis in Human Geography. Volume I and II*. London: Arnold.
- **Jones, Ch.**, 1997. *Geographical Information Systems and Computer Cartography*. Singapore: Longman Singapore Publishers.