

Chapter 1

Introduction: Geospatial Analysis

Yuji Murayama

1.1 What Is Geospatial Analysis?

Although the term “geospatial analysis” is widely used in the academic world, its definition is not clear even in the field of geographical information science (GIS). In this book, we define geospatial analysis as a GIS-based approach to analyze geographically referenced information using methods such as statistics, information theories, computational geometry, and geovisualization techniques. Its goal is to find the driving forces of changes on the earth’s surface, and analyze the geographical phenomena in order to understand their processes and mechanisms. Finally, it allows to suggest appropriate policy planning and decision making for sustainable development of a region to be suggested.

“Geospatial” information is considered a part of “spatial” information, which receives wide recognition in GIS communities. Here, geospatial data are limited to those on the earth’s surface which are geographically referenced by address, place name, latitude/longitude, and so on. Geological phenomena inside the earth, the free atmosphere far above the earth’s surface, and the topography on the Moon and Mars, for example, are not encompassed by the category of geospatial data in this book. Furthermore, architectural concepts like building structure and room arrangement in houses are also outside our research territory, although they are embraced in spatial data. In this regard, field-based empirical studies using geospatial techniques are conducted mainly within the geography-related disciplines.

Y. Murayama (✉)
Division of Spatial Information Science, Graduate School of Life and
Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan
e-mail: mura@geoenv.tsukuba.ac.jp

1.2 Quantitative Revolution

The origin of geospatial analysis can be traced back to the 1950s, when in North America the contribution of academics to society in general was required by the penetration of practicalism (Murayama 2004a). Geography was criticized because its methodological framework was old-fashioned and insufficient to satisfy social needs. Against this backdrop, quantitative methods and theoretical models oriented to law-making were pursued, and these were applied in problem-solving studies (Murayama 2004b). This academic movement was called the *Quantitative Revolution* in geography (Burton 1963). New methodologies, including network approaches using graph theory, central place formulation, point pattern analysis, regionalization and classification techniques employing multivariate analysis, and location-allocation problems were developed (Berry and Marble 1968).

In the 1970s, geospatial analysis with a time dimension, that is, spatiotemporal analysis, emerged and quickly became popular. It embraces time geography, behavioral geography, studies of changes in land use/cover, spatial diffusion and interaction modeling, and so on. A strong emphasis was placed on the conceptual switch from static to dynamic thinking, in other words, from the spatial structure to spatial process studies. In the late 1970s, computer mapping was developed as a type of analytical cartography, paving the way for geovisualization studies linked with GIS (Tobler 1976).

1.3 GIS Revolution

In the 1980s and 1990s, new techniques in geospatial analysis were exploited and implemented in GIS software (Murayama 2001). Advanced geospatial analysis became available without programming via open-source freeware such as GRASS, SPATSTAT, and R Package. Methodologically, great interest was focused on spatial homogeneity and heterogeneity (Anselin 1988). This concept was developed by further studies on local spatial autocorrelation, geographically weighted regression (GWR), spatial weight metrics and prominence, Kernel density, modifiable area unit problems, trend surface, and so on.

Geospatial analysis is interdisciplinary in nature. To enhance its operability, strong cooperation among adjacent disciplines is indispensable. In the 1990s, sophisticated GIS software was invented through teamwork which brought together regional science, information science, computation geometry, and statistics. The analytical functions became powerful due to the integration of raster-based GIS and vector-based GIS. Furthermore, the combination of GIS with remote sensing (RS) enhanced the usability of geospatial analysis, where useful techniques in RS such as spatial metrics, land use/cover classification, fuzzy sets, and so on, were effectively introduced within the GIS framework (Murayama and Thapa 2011a).

Today, tremendous amounts of geospatial data are ceaselessly produced and have become immediately available in the GIS environment. This data-rich condition,

along with increasing computing power, is enabling the development of advanced geospatial data mining. In the 1990s, geocomputation was developed with the use of fine image data from RS, and socioeconomic micro-data such as individual information from censuses and volunteered information (Openshaw and Abrahart 2000; Murayama and Thapa 2011a). The availability of high-resolution images from Quick Bird, IKONOS, and ALOS, for instance, is accelerating the micro-approach within cities, which is called urban remote sensing (Yang 2011). Today, geocomputational techniques are rapidly becoming strong tools for process-based studies where the future is predicted based upon the process from the past to the present. These techniques include neural networks, cellular automata, agent-based modeling, genetic algorithms, self-organizing mapping, and so forth. Since 1996, a geocomputation conference is held every 2 years, where geographers, information scientists, computer programmers, and other experts gather from all over the world to reflect on work within the field of geocomputation from an interdisciplinary perspective (<http://www.geocomputation.org/>).

The *GIS revolution* has brought methodological and conceptual shifts: from model-driven to data-driven, from deductive to inductive, from top down to bottom up, from aggregate to disaggregate, from discrete to continuous, from lagged time to real time, from static to dynamic, from quantitative to qualitative, and from linear to non-linear (Murayama and Thapa 2011b). From this revolution, it could be argued that geospatial analysis with GIS is more a useful tool in formulating hypotheses than it is in verifying them.

The book consists of 17 chapters. Chapter 1, the introduction, traces the evolution/origin of geospatial analysis using GIS-based techniques, and discusses the usefulness of theories, techniques, and methods for analyzing geographical processes occurring on the earth's surface. Then an overview of each chapter is given.

1.4 Geospatial Data Acquisition and Processing

In Part I, geospatial data acquisition and processing are discussed (Chaps. 2 and 3).

Chapter 2 focuses on the preprocessing of remotely sensed data and classification methods. Information extraction from satellite image data requires appropriate image-processing methods and techniques. The overall aim is to extract the information and explore the spatial patterns of land-surface objects, that is, land cover patterns from the classified satellite imagery known as a thematic map. Furthermore, this chapter presents theoretical procedures and techniques to produce useful information from such thematic maps. Significant technological advancements in data acquisition and analysis in recent years have made it easier to analyze the spatiotemporal dynamics of landscape changes using multispectral classification techniques with various statistical rules. Finally, this chapter discusses the various methods of assessing the classified image ranging from field-based validation to visual interpretation with high-resolution satellite images.

Chapter 3 focuses on geospatial data collection methods, database design and construction, and modeling with GIS. Geospatial data collection, including remote sensing, field surveying, and other in-house GIS data conversion processes (i.e., scanning, georeferencing, digitizing, etc.), is an important task for many geospatial information users. Traditional field data collection (i.e., pen-and-paper-based) is a bulky and time-consuming task. However, recent developments in mobile communication, global navigation systems, the Internet, and portable computational devices such as smartphone, netbooks or ultra-mobile personal computers (UMPCs) allow us to carry out field data collection in a timely manner. Moreover, under the client-server setting for field data collection, a field user may take advantage of digital repositories prepared for data collection (i.e., base maps, satellite images, and other ancillary data) as well as information resources more generally available via the Web. Proper geospatial data collection and conversion are required to support spatial analysis with GIS, which is vital for accurate decision making. Geospatial data processing is at the heart of the task in many GIS analyzes.

1.5 Geospatial Theories and Methods

GIS is designed to store, retrieve, manipulate, analyze, and visualize geospatial data. On the other hand, the uncertainty which affects the accuracy of maps and geospatial analysis results always exists in the data because of the limitations of human cognition of geographical phenomena or the resolving power of surveying instruments. Part II is composed of nine chapters (Chaps. 4 through 12), and aims to review the existing theories and methods in geospatial analysis, particularly in land use/cover modeling and GIS network data models.

Chapter 4 analyzes the fuzziness of geographical phenomena and classifies them into three aspects: the fuzziness of the distribution of the geographical phenomena or concept of a geographical entity, the fuzziness derived from spatial relationships, and the fuzziness derived from geospatial analysis and spatial reasoning operations. Fundamental fuzzy set theory, including the definition of a fuzzy set, fuzzy set operations, and fuzzy relationships and membership functions, is summarized in the representation of such fuzzy phenomena. Applications of fuzzy set theory in GIS are reviewed according to the sequence of fuzzy representations of geographical entities and their distribution, fuzzy spatial relationships, and fuzzy operations in spatial reasoning. Then, an example of a field-based integrated spatial reasoning model in the case of a constraint satisfaction problem (CSP) is given to synthetically illustrate the above-mentioned three fuzzy aspects. Finally, the author considers the future research directions of fuzzy set theory in applications using GIS.

Chapter 5 shows that in geospatial analysis of geographical phenomena, a region or a city under study might be divided into several small areal units such as a regular square tessellation or administrative units emerging in irregular shapes. The spatial interrelation between areal units can be expressed as different definitions of the

weight coefficient. For example, the spatial structure of areal units might be defined as the spatial contiguity, which is treated as a spatial weights matrix \mathbf{W} with a binary variable. This chapter also explains how to define and create the spatial weight, which is an expression of the spatial dependence between areal units. Four types of weight functions are introduced here. An area which has special geometric attributes and keeps significant spatial correlation close to adjacent areas is called a *prominent area*. As an application of the spatial weights matrix, the prominence of irregular areas can be measured by the *prominence index*, which is a stationary distribution of a Markov chain transition matrix that is identical to a spatial weight matrix. An empirical study in this chapter shows that generalized weight matrices are more appropriate for measuring the prominence rather than the distance decay and k -order.

Chapter 6 explains GWR, which is a technique for spatial statistical modeling used to analyze spatially varying relationships between geographical variables. Unlike the traditional regression framework, GWR allows local rather than global parameters to be estimated. In this chapter, the authors discuss the theoretical basis of the GWR method and modeling. Some of the current best practices for understanding urban and regional problems, for instance, regional analysis of wealth and land use/cover, driving forces behind deforestation and afforestation, and so on, are highlighted. GWR is one of the recent developments of local spatial analytical techniques, and it has been part of the growing trend in GIS toward local analysis.

Chapter 7, which deals with Bayesian theory, focuses on the weight of evidence (WofE) method and its applications in GIS. WofE, which is entirely based on the Bayesian approach of conditional probability, is traditionally used by geologists to point out areas which are favorable for geological phenomena such as seismicity and mineralization. Recently, the WofE method has been used to combine spatial data from a variety of sources to describe and analyze interactions, provide evidence for decision making, and construct predictive models. This chapter discusses the theoretical basis of WofE and presents best modeling practices. Basically, this method concerns the likelihood of detecting a certain event, which could be a given category of land use/cover change such as a change from an agricultural area to a built surface, in relation to potential evidence (proximity to urban centers, roads, water, etc.), often called the driving factors of change.

Spatial simulation models such as the Markov cellular automata (MCA) are critical for land use/cover change modeling because models are needed to gain insights into land use/cover change processes at many spatial and temporal scales. The MCA model combines cellular automata with Markov chains and GIS-based techniques such as multicriteria evaluation (MCE) and WofE in order to simulate land use/cover changes. The Markov chain process controls temporal dynamics among the land use/cover classes, while spatial dynamics are controlled by local rules determined either by the cellular automata mechanism or by its association with transition potential maps computed by WofE and MCE techniques. Chapter 8 reviews the methodological developments of the MCA model, as well its current status and future prospects based on criteria such as modeling techniques, data requirements,

calibration, and validation. Thus, issues raised in this chapter could contribute to the improvement of future MCA land use/cover change models.

Chapter 9 outlines the progress of artificial neural networks, architectures, algorithms, and future developments in geospatial analysis with GIS. Their applications are reviewed with land use/cover change analysis and modeling. Many artificial neural network architectures have been developed over the past years. One of the most popular is the multilayer perceptron (MLP) neural network. From a geospatial analysis point of view, MLPs have been shown to be a universal and highly flexible function approximation tool for any data. This chapter gives a comprehensive review of the history and basic architecture of MLPs. The use of MLPs for land use/cover classification is presented as a representative type of data-pattern recognition. Finally, future trends in the development of MLPs are briefly summarized.

Chapter 10 reviews multiagent simulation models to understand land-use change management and sustainability in the area of forest loss (deforestation). An attempt is made to assess the sustainability of deforestation management from the perspective of individual decisions made at the farm or household level. Agent-based models (ABMs) continue to receive wide attention as a method of modeling complex real-world applications. Founded on the pretext of understanding the non-linearity of natural systems, the multidisciplinary ABM originates primarily from artificial intelligence. This chapter delves into the history of complex systems and connects it to present-day fundamental principles of defining and designing agents. It then highlights the basic tenets of agent modeling implementation, which include a discussion on available ABM development platforms. This chapter also contextualizes the agent, its environment, and its interaction when applications of ABM in land use/cover modeling are presented and addressed. All this adds to the available literature on ABMs in the hope that by presenting both the past and present, modelers will be able to distinguish between and develop new ABM approaches for the future.

Chapter 11 discusses multicriteria decision analysis (MCDA). Since its development in the 1970s, the analytic hierarchy process (AHP) has been an important tool for decision makers and researchers. It presents a flexible, step-by-step, and transparent way of analyzing complex problems in a MCDA environment based on experts' preferences, knowledge, and judgments. There has been a growing interest in using this method in the last two decades or so, and its scope of application has been expanding, especially in the field of geospatial analysis. This chapter also reviews the basic principles of AHP, its historical development, and its applications as a decision support tool for GIS-based MCDA. Major findings show that AHP has been implemented in various fields of geospatial analysis in various countries around the world. This indicates how versatile and useful AHP is as a decision support tool. However, the review also shows that researchers have still not achieved consensus on certain issues concerning the implementation of AHP as a weighting method for GIS-based MCDA. Some of these issues include considerations about the method of capturing expert opinion using the pair-wise comparison method, the method of aggregation of individual expert's ratings (in cases where consensus ratings are not used), and the method of standardizing the individual factors involved in a GIS-based MCDA. These issues are crucial; thus, careful attention is needed

when using AHP as a decision support tool for GIS-based MCDA. Nevertheless, because of AHP's effectiveness in evaluating problems involving multiple and diverse criteria and the measurement of trade-offs, its simplicity and robustness, and its precision and ease of use, its applications will undoubtedly continue to expand in fields of both non-spatially and spatially based decision-making environments.

Chapter 12 discusses the GIS network model and its applications. A network is an interconnected set of points and lines that represent possible routes from one location to another. Road network models play a critical role in urban planning, emergency preparedness, retail market and market competition analysis, public facility management, and other planning and decision-making processes. Understanding the road network patterns in urban areas is important for human mobility studies, because people live and move along the road networks. Network data models allow us to solve daily activities such as finding the shortest path between two locations, looking for the closest facilities within a specific distance, and estimating driving time. A network model can include a multilayer model representing, for example, a railway system, a subway system, and a bus system to solve problems using multiple modes of transportation in an urban area. Many commercial GIS data models composed of layers such as points (nodes) and lines (links) comprise separate layers. This is called a layer-based approach. These nodes and links can also be represented as object classes; this type of model is known as an object-oriented network data model, and is still in the design phase. The development of 3D network models and concepts in GIS can solve these complex multilayer network solutions. Moreover, the combination of Internet technology and user-friendly Web-GIS provides an opportunity to perform interactive network analysis to make spatial decisions in a timely manner for local residents and city planners through Web-based GIS systems.

1.6 Applications in Geospatial Analysis

Part III provides five applications (Chaps. 13 through 17) in geospatial analysis employing GIS, remote sensing, and global positioning systems.

The complexity of urban systems requires integrated tools and techniques to understand the spatial process of urban development and project future scenarios. In this connection, Chap. 13 aims to simulate urban growth patterns using a Bayesian probability function in the Kathmandu metropolitan region in Nepal. Like many cities in the developing world, it has been facing rapid population growth and daunting environmental problems. Three time-series land-use maps at a fine scale (30 m resolution) derived from satellite remote sensing covering the last three decades of the twentieth century were used to clarify the spatial process of urbanization. Based on historical experiences of land-use transitions, the authors adopted the WofE method integrated in a cellular automata framework to predict the future spatial patterns of urban growth. The authors extrapolated urban development patterns to 2010 and 2020 under the current scenario across the metropolitan region. Depending on local characteristics and land-use transition rates, this model produced a noticeable

spatial pattern of changes in the region. Based on the extrapolated spatial patterns, urban development in the Kathmandu valley is projected to continue through both infilling in existing urban areas and outward rapid expansion toward the east and south. Overall development will be greatly affected by the existing urban space, transportation networks, and topographic complexity.

Chapter 14 discusses land suitability assessment using a fuzzy MCE model, which is an important step for sustainable land-use planning in the rural landscapes of developing countries. Taking the Tam Dao National Park Region, Vietnam, as a case study, this chapter demonstrates how the results of the assessment are used for sustainable land-use planning decisions. Land degradation is recognized as one of the major threats to the buffer zones of protected areas (PAs) in Vietnam. In particular, the expansion of land degradation into the PAs is exerting pressure on biodiversity conservation efforts. This degradation is partially the result of mismanagement: the utilization of the land is often mismatched with the inherent nature of the land. Identification of the spatial distribution of suitable areas for cropland is essential to allow sustainable land-use recommendations to be made. The authors delineate the areas in the region which are suitable for cropland using a GIS-based MCE of biophysical factors and Landsat ETM+imagery. GIS is used to generate the factors, while MCE is used to aggregate them into a land suitability index. The results which indicate the location and extent of crop farming areas at different suitability levels can be used to identify priority areas for crop farming and sustainable land-use management. The GIS–MCE approach provides an effective assessment tool for land-use managers working in the PAs of Vietnam.

Chapter 15 presents the application of the neighborhood interaction method in land-use analysis. Local spatial interactions between neighborhood land-use categories (i.e., neighborhood interactions) are important factors which affect urban land-use change patterns. Therefore, they are key components in cellular automata-based urban geosimulation models that aim at forecasting urban land-use changes. In this chapter, the authors interpret the similarities and differences of neighborhood interactions in three metropolitan areas in Japan, namely Tokyo, Osaka, and Nagoya, to provide empirical material to understand the mechanism of urban land-use changes. Neighborhood interaction reveals the effect of spatial autocorrelation in the spatial process of urban land-use changes in the three metropolitan areas, which correspond with the agglomeration of urban land-use allocation in Japan. Neighborhood interactions amidst urban land-use changes among the three metropolitan areas generally showed similar characteristics. The regressed neighborhood interaction coefficients in the models may represent the general characteristics of the neighborhood effect on urban land-use changes in cities. The results provide very significant material for exploring the mechanism of urban land-use changes, and the construction of universal urban geosimulation models which may be applied to any city.

Walkability is a well-known measure of how conducive an area is to walking to and from chosen destinations. The calculation of a walk score is widely used in accessibility studies to determine the ease or difficulty of travel by foot between one point and another. Based on this situation, Chap. 16 proposes an integrated methodology (remote sensing, GIS, and spatial Web technology) to model urban green space walkability which will enable local residents to make informed decisions

that will improve their living conditions and physical health in relation to the neighborhood environmental quality. The authors discuss the modeling of urban green space walkability by utilizing a Web-based GIS to calculate eco-friendly walk scores based on the presence of green spaces by integrating advanced land observing satellite data and other GIS datasets. We use this spatial Web technology to help local residents make decisions related to neighborhood environmental quality, such as how to choose an eco-friendly living space when buying a home, or how to find the shortest or greenest route to walk to improve their health.

Chapter 17 deals with watershed evaluation using geospatial techniques. Severe watershed degradation continues to occur in the tropical regions of southern Africa. This has raised interest in harnessing and manipulating the potential of watershed resources for human benefit as populations grow. The Songwe river is one such degrading watershed that causes biennial flooding, among other problems. In this study, climatic, land use, topographic, and physiographic properties were assembled for this watershed, and were used in a process-based GIS with the aim of determining the hydrological sediment potential of the Songwe river watershed and quantifying the possibilities of reservoir sedimentation. The study further aimed at determining the critical sediment-generating areas for prioritized conservation management, and the relationship between the increasing flood events in the floodplains and the rainfall trends. Based on the evaluation of hydrological runoff processes using the Pan-European Soil Erosion Risk Assessment (PESERA) model, the estimated amount of sediment transported downstream is potentially huge. It was established that most of the sediment generation was occurring in the upper sub-basin, and specifically from built-up villages and degraded natural land. These trends have not only caused an increase in flooding events in the lower sub-basin, but also pose a great sustainability risk of sedimentation to the proposed reservoir.

References

- Anselin L (1988) *Spatial econometrics: methods and models*. Kluwer Academic, Dordrecht
- Berry BJL, Marble DF (1968) *Spatial analysis: a reader in statistical geography*. Prentice-Hall, Englewood Cliffs
- Burton I (1963) The quantitative revolution and theoretical geography. *Canadian Geographer* 7 (4): 151–162
- Murayama Y (2001) Geography with GIS. *GeoJournal* 52:165–171
- Murayama Y (2004a) How has GIS contributed to urban geography? *Annu Rep* 30:5–8, Inst. Geosci., Univ. Tsukuba
- Murayama Y (2004b) American influence on Japanese human geography: a focus on the quantitative and GIS revolutions. *GeoJournal* 59:73–76
- Murayama Y, Thapa RB (2011a) Spatial analysis and modeling in geographical transformation process: GIS-based applications, vol 100, *GeoJournal library series*. Springer, Dordrecht
- Murayama Y, Thapa RB (2011b) Spatial analysis: evolution, methods, and applications. In: *Spatial analysis and modeling in geographical transformation process: GIS-based applications*, vol 100, *GeoJournal library series*. Springer, Dordrecht, pp 1–26
- Openshaw S, Abraham RJ (2000) *GeoComputation*. CRC Press, London
- Tobler W (1976) Analytical cartography. *American Cartographer* 3:21–31
- Yang X (2011) *Urban remote sensing: Monitoring, synthesis and modeling in the urban environment*. Wiley, Chichester