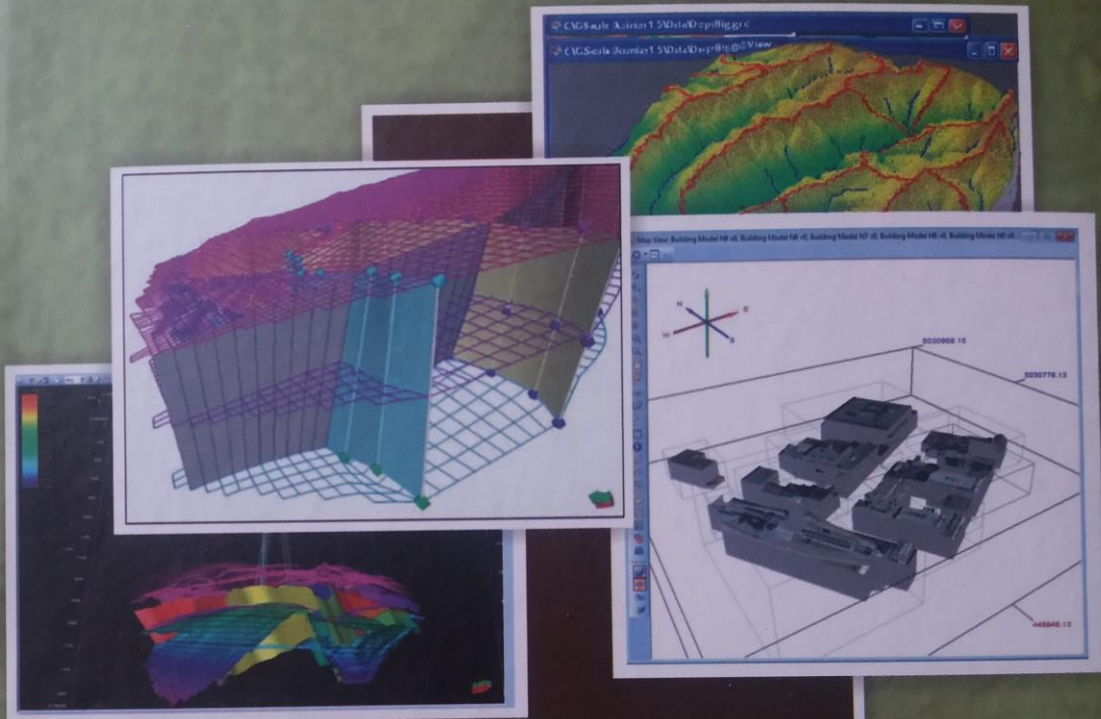




Харківський національний університет
імені В. Н. Каразіна



Кафедра соціально-економічної географії
і регіоналістики



С. В. Костріков, К. Ю. Сегіда

ТЕОРЕТИЧНА І ПРИКЛАДНА ГЕОІНФОРМАТИКА

Навчальний посібник

Sergiy Kostrikov, Kateryna Sehida
Theoretical and Applied Geoinformatics

Textbook Extended Summary

(розгорнуте резюме посібника, *англійською*)

*The new source of power is not
money in the hands of a few,
but information in the hands of many.*

- John Naisbitt
Megatrends

This university textbook is concerned with Geoinformatics theoretical and applied issues within the GIS frameworks. The subject area of Geoinformatics in Post-Soviet Geography overlaps the same disciplinary extent, which the GIS science possesses in contemporary western Geography.

A GIS represents the natural-human environment on a computer screen similar to the way a usual map represent the environment on paper. Nonetheless a GIS provides a power and flexibility, that paper maps could not provide.

Our text introduces Geographical Information Systems as one of the most important and revolutionary tools to become known for geographers in few recent decades. The GIS is an approach with a very broad range of applicability in many areas, and if so, its importance for some particular subject field may not be clearly apparent. Therefore, any applied issue, a GIS is involved in, must be clearly specified. The authors attempt to demonstrate the full power of geographical information systems and technologies and provide basic fundamentals on possible applications in Physical and Human Geography, Geology, Hydrology and Ecology. Authors also try to outline in what way GIS-platforms and its supplementary modeling software have been applied for many years to develop solutions, which contributed significantly to both natural resource and human necessities hot issues – to assess unfavorable earth surface processes; to evaluate water quality and water availability; to prevent floods and other catastrophic event geared by both human and natural forces; to apply the potential analytical capabilities of GIS for urban geosystems.

The textbook is of two parts – theoretical and applied ones. There is a simple explanation in the first part of textbook, that with a GIS you can study not just this map or that map, but every possible map from their long variety. With the proper

data a user can examine whatever wants – natural landscape, city or cities, sea, political boundaries, etc. A GIS map is made up of layers. Each GIS database table, that contains graphic objects, can be displayed as a layer in a Map window. On a paper map, a user cannot remove, for example, cities away from countries, or continents away from the ocean, but on a GIS map he can.

This textbook is a reliable GIS-world guide for all geographers, geologists, and ecologists, from early-year students to experienced scientists, due to obtaining the most out of GIS force.

Part One: Some Issues of Theoretical Geoinformatics

The first chapter of this textbook *Geoinformatics fundamentals* overviews basics of the information science. Two definitions – “Informatics” and “Computer Science” are compared in details with respect of the main information stream considerations by each of these two subject areas. Geoinformatics fundamentals are introduced as several theoretical principles of GIS technology with a broad base of its applicability. The first chapter stresses that a GIS depicts spatially distributed information in a numerical form. GIS history is shortly narrated in the last section of the first chapter. It is explained that numerous scientific and technological advances in the last 25 years of the former century greatly contributed to the development of GIS for geographical, geological and ecological applications. Testing questions for student self-control finalize this chapter as well as all other chapters.

The first chapter includes fourteen visual figures.

Methods of Geoinformatics for spatial analysis in contemporary Geography is the textbook *second chapter*. This chapter starts with explanations that GIS software is understandably worthless for geographers without the appropriate data. The chapter introduces the geographical entities which can be visualized on a map or on a GIS data layer as 0-, 1-, or 2-dimensional spatial objects. There types of spatial objects, which model geographical entities, are points lines, points and polygons, correspondingly.

The second chapter provides an introductory overview of the spatial analysis issues. The latter imply that they describe and research the alteration of one or more geographical entity over the space. Our textbook emphasizes once again the well-known statement: “Spatial analysis is a GIS core”.

Spatial classification is almost the most important GIS-analysis component. The textbook presents classification as a means of combining geographical entities into spatial sets or categories unifying characteristics. The latter are accepted by a GIS as so-called “attributes”. It is taken for granted that a method geographical

entities are classified in a geoinformation system highly impacts GIS-analysis output. Therefore it is very crucial to define the basics of classification in the right way.

There are twenty three figures and three tables in this chapter.

The **third chapter** of the textbook is entitled *Geographical data – GIS data structures*. The chapter is about the geographical data. The latter are stored in a GIS with the digital representation form. GIS data bring the information about the localized placement and attributes of real geographical entities.

There are raster data structures and vector data structures in a GIS database. All other structures are, speaking generally, only derivative ones. A raster layer depicts GIS-features as a matrix of cells, and these cells are normally square. Raster cells are always two-dimensional, although geographical entities, which they represent, may be 0- or 1-dimensional. Vector data structure portrays GIS-features as points, lines, or areas (polygons). Vector structures associate topology – it is, for instance, contiguity or connectivity of geographical entities as GIS features. Highly usable structures within vectors are triangular irregular networks (TINs). Vector data versus raster ones can be discussed according to the research goals defined and to three following factors: if there is clear discreteness among a set of geographical entities; for what exact subject this GIS application is intended; peculiarities of the source data.

Grid data, while they are similar to raster structures, can be distinguished by the mandatory attribute presence for each cell. There are both numerical and classified layers among grid structures.

Examination of ESRI coverages and geodatabases concludes this chapter. *Arcview 3.x* GIS coverage was once followed by *ArcGIS* geodatabase (GDB). This one is shortly discussed as an object-oriented data model, that introduces geographic features and attributes as objects and the relationships between GIS-objects. GDB stays within a relational database management system. It can keep objects, such as feature classes, feature datasets, attribute tables, and classes of relationships.

The third chapter contains thirty six visual figures and nine tables.

The **fourth chapter** is *Basic operations and procedures in GIS*. This chapter describes GIS analytical abilities (which are divided for operations and procedures with detailed explanations) within there broad categories: basic GIS database operations and procedures on spatial and non-spatial data; executed on a single spatial data layer; performed on several (multiple) data layers. Some particular steps of these operations are described and discussed due to the geoeological subject area.

The operations over raster layers are grouped into local, focal and zonal operations. General GIS-operations predestination is that one, according to which a geoinformation system is employed to choose data that correspond to a certain criterion. Data selections and queries can hardly be overvalued in this aspect. Data selections with GIS can be done: interactively; by a look-up table; with determining numerical threshold; and by Boolean logic. Operation of recalculation take math expressions to generate new attribute data from existing attributes.

Fist operations related to ones performed on a single spatial data layer are format conversions and measurements. The feature according to which a GIS can read and transform the format variety is a characteristic of a powerful GIS. Many measurements can be executed on spatial data with a GIS. A whole measuring range starts from simple distance and area measurements to complicated fractal measurements and complex landscape ones.

Spatial data aggregation, as one more single data layer operation, diminishes excess of spatial details in a grid or raster data layer. Examining of moving window operations, juxtaposition analysis, and buffer zones finalize this operation type.

Such operations as geometric transformations, clipping and overlay are normally performed on multiple data layers. With geometric transformation a GIS provides spatial adjustment between tow layers so, that they would be overlaid correctly one on another of the same data. Spatial subsetting of one layer with another layer boundaries is titled “clipping”. The broad range of overlay operations is made up of graphical overlay, logical overlay, and arithmetic one. Few examples of the regional environmental tasks are provided with their solutions through GIS-operations.

This chapter includes forty six figures and two tables.

The *fifth chapter* of the textbook – *Digital elevation model and environmental geoinformation model* – overviews broad range of topography and other landscape component presentation in GIS. Digital representation of topography has become widely available with GIS-tools for few recent decades. Earth surface topography is a crucial factor of many driving forces in geographical phenomena, because of its impact on solar radiation input, runoff and channel flow, and mass movement. Under the gravity force action the land surface is sculptured by water, wind, ice, and other impacts. This sculpturing produces the landforms with which a geographer, or a geologist are concerned.

A GIS can produce a complete picture of their concern with a Digital Elevation Model (or a Digital Terrain Model) – a raster, or a grid DEM, and with TIN-model. The latter is considered to be a more efficient tool of topographic presentation, than a DEM. Conventionally, a GIS uses some particular analysis as

topographic decomposition of landforms. For instance, the fluvial topography (erosion relief) may be represented as hierarchical system, which has been produced by number of the landforms. Having combined the topographic analysis with the GIS-approach, a scientist gets the opportunity to fix the fluvial relief as known in form of the complicated system of subbasins - parts of the higher order basin – if this scientist builds a watershed geoinformation model. This model belongs to a broad class of the environmental GIS-models.

Chapter Five outlines the variety of GIS-models of watersheds, which have been developed and elaborated concerning different components of basin environment. The landscape component modeling normally simulates other than topography and hydrology constituents of natural systems in a watershed. Different approaches to the environmental GIS-modeling have been examined and grounded - theoretical, empirical, and "rational" one.

A digital elevation model and flow algorithms represent the preliminary modeling conditions. A GIS-model of a watershed is proved to be a research-based simulation model. Explanation of topographic visualization as GIS typical depiction of spatially distributed data concludes this chapter. New trends of topographic visualization is the recent development of a bridge from efficient three-dimensional visualization to the virtual reality with further environmental visualizations on the base of routine mapped data.

The fifth chapter possesses twenty three visual figures.

The *six chapter*, which is entitled *Spatial analysis and some components of thematic GIS-modeling* introduces more in details some issues from those ones, which the textbook authors have already touched initially in the second chapter. This chapter summarizes a GIS as some kind of “super-map”, specialized software that joins geographical data through spatial analysis (where things are?) with descriptive (attributive) information. The latter provides a GIS-user with object / process / phenomenon properties (what thing are like).

The chapter characterizes in what way the spatial resolution (the geographical scale) defines the net of connections and relationships between the natural-human environment and a model constructed in a GIS.

Actually there is a complete introductory overview of models and modeling in the GIS aspect. It also overviews both early computer applications in Physical and Human Geography and existing main groups of theoretical and applied models for the object-process-phenomenon simulations in some environmental domains. Authors make a stress in this chapter on dynamic simulation models as on the conceptual computer-based models that simulate how a complex system will change over time. It is clearly explained in this chapter, in what way GIS thematic models, in particular – geomorphic and hydrological ones - provide geographers

with means for extrapolating measurements made in a field and integrating geographical information over a space and time. GIS – environmental models integration is the most advanced topic of geographical research for the time being.

As far as geovisualization of the modeled results is concerned, the chapter explains why unlike a routine paper map, where “what a user sees is what he / she can obtain”, with geovisualization a GIS can have many various layers of information underneath its surface.

The chapter examines thoroughly math modeling of both fluvial topography processes, and initial runoff initiation. The surface topography and hydrological event simulation with both deterministic and random approaches has been considered. Chapter Six of the textbook proves, that a GIS should contain and properly handle both the topographic, and channel network information. The delineation algorithm of a complete erosion network should be based on the determined geomorphologic model, accounting for information related to channel network and watershed morphology. The watershed GIS-model has been further outlined (after introduced initially in the fifth chapter) as a set of its constituents: mathematic model of the formalized fluvial topography representation, the algorithm of hydrological runoff routing through a digital elevation model, the general and particular algorithms of the channel network definition.

The six chapter also presents our original concept of the watershed distributed models of the channel peak flow concerning both spring floods and summer rainfalls. Some basics of the GIS-models for watershed hydrology have been discussed as well as key issues relevant to the distributed hydrological modeling provided. Several various definitions of a distributed model are emphasized. The algorithmic calculation of the initial basin area is examined in details upon hydrological flow simulation on the digital elevation model matrix. In addition the geospatial data concept has been outlined in details, since this data accepted as the initial data sources for the GIS-modeling in the relevant domain. This chapter sections gives some relevant examples of the original modeling software.

GIS-modeling of geological environment is overviewed in one of the concluding chapter sections. Textbook authors both refer to well-known solutions in the area as, for example, Petrel-Eclipse software, and introduce the original concept of the modeling entity “Random Resource Cube + Pillar Grid”. It is explained in what way a mineral deposit has been modeled with it. There have been several samples of the geological data integration, analysis and visualization. These examples have been related to the deposit producing layers simulation and visualization in the software with further localization of promising boreholes.

Thematic GIS modeling is concluded in this chapter by examples of GIS-tools involvement for data processing, modeling and visualizing on the base of the

LiDAR-technology remote sensing for the urban geosystem analysis. The GIS-approach is employed for some feature analysis of the urban system dynamics. Some of this approach constituents, GIS-interface and specialized software functionality are considered in the last section of the six chapter. Several evident advantages of such GIS applications are proved in this section due to municipal management challenges.

There are forty one visual figure in the six textbook chapter.

The *seventh chapter – Spatial interpolation and geostatistical modeling* – concludes our consideration of Geoinformatics theoretical issues in Part One. This chapter presents that spatial statistical sampling methods supply a data subset about “a whole world”, providing data for some places, but not for other geographical locations. Spatial interpolation and geostatistical modeling are elaborated to fill the empty space between data samples around chosen points. This empty space is normally the most part of the mapped area, therefore both spatial interpolation and geostatistics are the only effective solutions, which can provide a geographer / a geologist / an ecologist with qualitative data.

The seven chapter introduces the interpolation classification for several major groups. Each group is considered more in details. Thus a group of methods called “Interpolation through boundary determination – deterministic methods” unites a number of interpolating tools, which effectiveness mainly depends on the discreteness of geographical entities and their clear boundaries. There are quite a few quantitative methods for boundary determination examined within this group besides in addition to the human interpolation method. The moving split-window, textural analysis, Thiessen polygons, and fuzzy sets are considered in detail in the second section of this chapter.

The second group of interpolating methods to be examined is the “interpolation of continuous data” set. This set contrary to interpolation through boundary determination interpolates data across a surface mathematically definable. In their turn all methods of this group can be divided for two subsets: whole area interpolation methods and local interpolators. The former methods are Fourier series and the trend surface analysis, where interpolation is grounded on all sampling points in the area of interests. The latter methods are moving average windows, splines, spatial autocorrelation and Geostatistics.

The authors discuss Geostatistics as a complete subject area with great potential to contribute to GIS-modeling. This method, which provides tools for calculating the locational effect on sample variability, is considered in the last section of the seventh chapter. Geostatistics includes a regression method – models of semivariogram. It also employs an interpolation method - different types of so-called “kriging”. The latter uses a semivariogram model to calculate a trend

surface map. This map is based on empirical points and includes the spatial variability of a geographical entity within the area of interest.

Authors introduce in the last subsection of this chapter their original solution for a locational task with mathematical means of the simple kriging method. It theoretically proves the main geostatistical premise, that statistical set variants located closely, usually will be more close in their attributes, than sample points located farther apart. This theoretical task solution finalizes the last chapter of Part One of the textbook.

The seventh chapter contains ten visual figures.

Generally, Part One of the textbook explains many possible benefits of GIS-tools for geographers, geologists, and ecologists though GIS operability and utilities. The greatest GIS significance is the quality of its spatial analysis functions in the following domains: geometric measurements of distance, circumference, center of gravity coordinates, area and volume; table operations that can make logical withdrawal of attribute information from map data; nearest-neighbor analysis that searches for spatial relationships between objects; spatial searches that are made; overlay analyses that create new thematic maps by juxtaposing layers with different attributes; network analysis that enables continuous measurements, searches for shortest routes, and the specific channel network analysis for watershed environment.

Part Two: Applied Geoinformatics: working in the interfaces of some GIS-platforms and modeling applications

The first chapter of Part Two, the *eighth chapter* of the textbook introduces a reader to basic concepts of *ArcView 3.x Desktop GIS* through practical tasks. A reader takes a closer look and examines this GIS by opening a project and looking through each of its component in the first exercise of this chapter. Thus you can learn that *ArcView* unites set of GIS-features to their attributes in the themes and manages these themes in a view. The second exercise of the chapter teaches how you can create an *ArcView* theme from a variety of geographical data sources. In the third chapter exercise you get to know how to use *ArcView* tools for classifying and presenting information in the themes, while in the fourth one – how to apply marker, pen and fill symbols to each classified unit and how to change symbol color.

The fifth exercise, in which you have to complete a GIS-project in municipal management, summarizes applied GIS skills you have to possess from the eighth chapter topics.

There are thirty eight figures in this chapter.

The *ninth chapter* of the textbook is *ArcGIS samples of data geovisualization, attribute and spatial queries*. This chapter briefly overviews in its first section key abilities of the premier desktop GIS, which is *ArcGIS* for the time being. This GIS compounds geographic data into a series of thematic layers and attribute tables. All geographic datasets in a GIS must be georeferenced, thus they have real environmental locations and overlay one another. Geovisualization as working with maps, geoprocessing as producing the entity “data + tool=new data”, and a geodatabase as a core geographical information model, all these three determine three main source of *ArcGIS* power.

Two applications of *ArcGIS* – *ArcMap* and *ArcCatalog* can provide a sophisticated desktop mapping, and you learn in the second section of this chapter how this GIS applications give the power of geographical analysis to an average GIS user. Actually you will know how to reference spatial data to real environment.

Due to the limited textbook volume the authors were able to choose only few topics for practical exercises from the “*ArcGIS* universe”. Some of these are attribute and spatial queries. You can select GIS-features spatially with Select Feature tool by clicking on them one after another or drawing a rectangle over them. You can select a feature by its attributes by entering a value, or by writing a statement in Query Builder. You can also use Find tool.

The ninth chapter ends with two tasks to be completed for self-control in the forth section, and these tasks require basic *ArcGIS* skills to demonstrate.

There are twenty eight figures and two tables in the ninth chapter.

The introductory section of the *tenth chapter* – *Spatial modelling in MapInfo application – Vertical Mapper software* focuses on the basic *MapInfo Professional* GIS functionality, since this GIS environment is a modelling domain for *Vertical Mapper*. The basics of this modelling and display software you examine in the second section by a relevant exercise completion. There are several frequently used *Vertical Mapper* functions available through its GUI focused in this exercise.

In the exercise of the third section of this chapter you will have to analyse the transmitter signal strength data, which is stored in an Excel spreadsheet. It is necessary then to import this spreadsheet into *Vertical Mapper* and make data mappable. Then, you will have to change the cartographic projection for the file and aggregate sample points.

Creating numeric grid files using interpolation methods are considered in two exercised of the fourth section in chapter ten. Firstly, you will create a numerical grid from elevation data. This data is in a point file, and you have to use the Triangular Irregular Network interpolation method to build a grid. You take into account, that TIN-method is usually applied to data, which requires no regional

averaging, for instance – to elevation data. Secondly, you will employ IDW-interpolation as a moving average interpolation technique. The latter is usually applied to highly variable data. When it is possible to come back to the field and record a new value that is statistically different from the original one, but it is still within the main trend for this area. Examples of this type of data include environmental monitoring data. For instance, with this data taken from around a nuclear plant you can interpolate a surface polluted by radioactivity. Such example considered in details in the second exercise of this chapter last section. c

The tenth chapter of the textbook includes forty nine figures.

The last *eleventh chapter* of the textbook focuses on the *GIS-Module Ukrainian* package, which is unique development of one from textbook authors. This software mirrors definite progress in complex usage of digital elevation data, geomorphic-hydrological and human impact simulation modelling, and GIS technique to evaluate the time-and-space distribution of geomorphic and hydrological processes, and assess the environment of watersheds.

Nine detailed exercises cover all main menu options of *GIS-Module Ukrainian: File, Modelling, Selections, Research Maps, View, Morphology – Morphometry, Hydrology, and Windows*.

The eleventh chapter proves, that significance of the river basin research proceeds from a watershed role as a basic mapping unit in solutions of many environmental issues. Strong spatial aspect of this research approach must be naturally combined with conventional techniques of GIS modelling. The software introduced consider a river basin as an “environmental system”. Thus *GIS-Module Ukrainian* necessarily implies quantitative research of basin areas and produces valid criteria for this environmental system component definitions. “What-If” interactive modelling of *Ukr_GIS* helps in managing water resources by achieving the following goals: quantification of the water resources for designed reservoirs and identification of types of probable erosion processes (as consequences of water flow regulation by a reservoir). Watershed 3D-visualization with *Ukr_GIS* can assist in Human Geography issues: for instance, how human settlements might share space in watersheds.

There are ninety four visual figures and two table in this last chapter of the textbook.

The textbook ends with there-language glossary in Geoinformatics, the acronym dictionary, the list of references, and the literature in GIS and Geoinformatics recommended for students.