*i*Journals: International Journal of Social Relevance & Concern (IJSRC)
ISSN-2347-9698
Volume 11 Issue 1 January 2023

Spatial Analysis of Epidemiological Situation Using the New GIS Technology to Identify Its Sources

Authors: V.P. Belogurov; Yuliya Yanush

DOI: 10.26821/IJSRC.11.1.2023.110107

ABSTRACT

Methods of spatial analysis, suitable for solving the problem of identifying the source of an epidemiological situation in conditions of incomplete information are being viewed in the article. The proposed technology is based on a probabilistic approach to spatial data analysis with statistical calculations apply.

INTRODUCTION

Geographic Information System (GIS) is an organized set of hardware and software funds, geographic data and personnel, designated for efficient receipt, storage, update, processing, analyzing and obtaining images of all types of geographically referenced information [1]. With the help of GIS it is possible to perform advanced complex spatial operations that otherwise would be very complex, time-consuming or impractical.

The most important advantage of geographic information systems over the rest of the information systems and databases is the ability of spatial analysis and creation of fundamentally new environmental models based on this analysis.

GIS is a unique and powerful tool meant for spatial analysis of territories. The use of GIS raises traditional geographic analysis to a new level, enriching it with many new functions and capabilities, precise accuracy and reliability of measurements. Many operations of spatial analysis which required a lot of time and resources when using paper maps can be calculated within seconds using GIS.

Spatial analysis uses statistical methods of data analysis widely. Using statistical functions in spatial analysis makes it possible to enlarge the range of methods that can be used to solve a variety of tasks related to georeferenced objects.

1. GIS TECHNOLOGIES USED FOR IDENTIFYING EPIDEMIC SOURCES

This article discusses the situation of the intestinal disease epidemic, which is caused by poor quality food, i.e. the sources of such a phenomenon potentially are food items (in the conditions of the task given). The selected topic is an example of tasks that can be solved using standard GIS tools and the proposed methods. Functionality of the GIS products (for example, ESRI's ArcGIS) are quite wide and provide opportunities for performing various types of data operations [2], however, under the conditions of this problem, functionality is not sufficient for achieving the result. Therefore, a new analysis technique is being developed, which together with standard (for a GIS application) functionalities makes it possible to find a solution to the problem.

The term standard functionalities of the geographic information systems mean implementation of spatial analysis, as well as obtaining statistical information about data and change of statistical functions.

2. SPATIAL ANALYSIS AND STATISTICAL FUNCTIONS AS GIS METHODS IN SOLVING THE PROBLEM

According to the "Features of the presentation data in GIS "[3] all operations of spatial analysis in GIS can be divided into two groups:

- simple spatial analysis;
- sophisticated spatial analysis.

Simple spatial analysis includes the following operations:
1. Determination of the objects coordinates on the map. It is a standard feature of most GIS packages. The coordinates are displayed in the selected system.

2. Measurement of distances, areas and perimeters.

It is also a standard GIS cartometric function, present not only in fully functional GIS packages, but also in many electronic atlases and reference books.

3. Analysis of relations is the choice of objects on different map layers by their relative position (search spatial queries).

Spatial queries are applied to two layers, in the conditions of the problem being solved these are layers containing affected and potential sources; the objects of one layer in relation to the objects of another layer can be:

inside the selected objects of the specified layer; intersect selected objects of the layer; contain the objects of the layer; touch selected objects of the layer; stay at a given distance from selected objects of a given layer; be completely inside the selected layer objects.

Complex spatial analysis includes the following operations:

1. Build buffer zones around objects.

The result of such an operation is usually a new polygon data layer that includes the terrain within a certain distance from these objects. The calculation of buffer zones makes it possible to perform complex search spatial tasks queries such as finding objects of the same layer within a given distance relative to the selected objects of another layer. There is a possibility of building buffer zones of different radii to one object, i.e. buffer zones are built for several values of the selected parameter, for example zones at a 10m, 50m and 100 meter distance from the objects. So, this operation can be used to solve the problem at the stage of analysis of finding one layer objects in the buffer zones of different radii of another layer objects.

2. Operations with digital elevation models, including complex morphometric analysis and analysis of visibility / invisibility zones.

3. Network analysis - network analysis tools (such as ESRI's ArcGIS Network Analyst) allow build optimal routes on a real street-road network with its capabilities, limitations of movement directions, turns, access ability of the streets. A wide range of tasks is being solved, for example, finding a route in a transport network with minimal time costs or traffic flow calculation.

4. Overlay operations - overlay operations of several layer objects on top of each other with a goal of getting a new layer (vector map algebra). Overlay operations are widely used in territory analysis from the point of view of relative position of object zones of influence, using such operations you can analyze the source objects relative location.

Statistics is a branch of knowledge in which the collection, measurement and analysis of mass statistical (quantitative or qualitative) data takes place.

Also, statistics can be defined as knowledge of organizing, collecting, presenting, analyzing and interpreting quantitative data for making the most effective decision [4].

In GIS software geostatistical analysis is presented as one of the functional opportunities (in the ArcGIS software package. The ESRI Geostatistical Analyst allows you to make the geostatistical analysis). Geostatistical Analysis allows you to predict and calculate the static probability of certain properties, phenomena and objects within a given volume, display the results obtained in the form of graphs and maps, which are closely related to each other.

Geostatistical Analyst provides a complete set of spatial analysis tools - from methods allowing to investigate the initial data, up to post-processing (processing of results) with an assessment of non-certainties (reliability of the results obtained by analysis). Geostatistical Analyst provides solution of two important tasks: Research analysis Spatial Data Lysis (ESDA) and Construction surfaces (real and predicted). There are a number of spatial analytic tasks that use different combinations of each tool types. When creating the resulting (output) map reliable default parameters for each of the available methods can be used [5]. Geostatistical analysis provides a basis for an intuitive look based on a statistical approach and a

more complete understanding of the sometimes unobvious essence and relationships of data layers. However, even such a powerful toolbox does not give a solution to this problem without the use of additional measures to eliminate this deficiency. So the following technique has been developed.

3. TECHNOLOGY FOR IDENTIFYING THE SOURCES OF THE EPIDEMIC

The proposed technology for solving the problem of identifying the epidemic sources when looking for solutions is based on uneven distribution of the objects affected by the epidemic. Every object has several attributes that characterize its location in a given territory (for example measures, home and work addresses). Certain parameter that indicates the presence or absence of relationship between objects is calculated. The relationship between the affected objects and the potential source objects of the epidemic is investigated.

To solve this problem, it is necessary to analyze 2 main classes - a class of objects suffering from an epidemic (let's call it Class 1). This class is characterized by home and working addresses of each of the class elements. Class 2 is a class of potential sources of the epidemic, for it the main attributes are those characterizing the food services area: the owner of the area, product suppliers, etc.

To apply this development, you need to prepare data in a certain way:

1. Generate layers of point objects (a separate layer to display each of the significant attributes of each class)

2. Create logical connections between objects in the classes and interclass relationships. The sequence of actions when solving the problem is as follows:

1. Display objects of both classes on the map;

2. Select the area that includes all Class 1 objects (one attribute at a time) using one of the methods standard for a GIS program;

3. Divide the resulting area into 2^n equal parts so that the number of parts is at least the number of objects in Class 1;

4. Select the reach zones of Class 2 objects (using buffers or spatial query) in relation to objects of Class 1;

5. Calculate the uniformity of Class 1 objects distribution relative to created parts (2^n).

In order to test the hypothesis of a uniform distribution of value *X* general population (Class 1 objects), i.e. according to law

$$f(x) = \begin{cases} \frac{1}{b-a}, x \in (a,b) \\ 0, x \notin (a,b) \end{cases} \quad (1)$$

It is necessary to estimate the parameters *a* and *b* - the end of the interval in which possible values *X* were observed, according to the formulas (*a** and *b** denote the estimates of the dimensions):

$$a^* = \bar{x}_в - \sqrt{3}\sigma_в \qquad b^* = \bar{x}_в + \sqrt{3}\sigma_в.$$

Then you should find the differential function of the estimated distribution:

$$f(x) = \frac{1}{b^* - a^*} \quad (2)$$

Compare empirical (for function 2) and theoretical (for function 1) frequencies using Pearson criterion, taking the number of degrees of freedom $k = s - 3$, where s is the number of intervals into which the sample is divided [6].

If the objects are evenly distributed, then you can conclude that a given attribute of Class 1 objects does not carry a solution to the problem, therefore, it is necessary to explore other significant attributes of Class 1objects by repeating this sequence of actions for them.

If the distribution of Class 1 objects is different from uniform (in accordance with the same criterion), then it is necessary to analyze the mutual location of Class 1 objects according to this attribute and location of Class 2 objects. Highlighting the reach zones of Class 2 objects, you can estimate the number of Class 1 objects that fall into these zones. Then the most likely of the potential sources of the epidemic will be Class 2 object, in the smallest buffer zone from which the maximum number of Class 1 objects is located.

The program is developed in the Visual Studio environment 6.0, in the Visual Basic programming language, and Avenue is used for scripting. The user interface is not developed, the program is an add-on module for ArcGIS programs 9.x and ArcView 3.x from ESRI.

4. CONCLUSION

GIS - technologies provide powerful functional analysis capabilities of spatial data that raise geographical analysis to a new, higher level, which allows improve the quality and speed of work when solving various kinds of problems, making decisions.

The functionality of the spatial analysis is very wide, rich toolkit makes it possible to conduct almost any analytic operation with data.

The proposed technology is based on a combination of standard spatial analysis functions of ESRI's ArcGIS software package with a probabilistic approach which is new for GIS technologies and helps to provide assessment hypotheses about the presence of spatial relationships.

The application of this approach to solving problems associated with mass phenomena, the sources of which need to be identified, allows you to find a solution even in conditions of significant incompleteness of the initial data.

REFERENCES

[1]. Roger Tomlinson. Thinking about GIS, Russian translation. – M. : Data+, 2004.

[2]. Svitlychnyj O.O., Plotnysckij S.V. Basics of geoinformatics: Study guide / In general Ed. Svitlychnyj O.O. - Sumi: VTD «Universytetscka knyga», 2006.- 295 p. (in ukrainian)

[3]. Features of data presentation in GIS, electronic tutorial (http://mylearn.ru). (in russian)

[4]. Electronic encyclopedia (http://ru.wikipedia.org/wiki). (in russian)

[5]. Materials of articles on the Internet site Date+ (http://dataplus.ru). (in russian)

[6]. Gmurman V. E. Guide to solving problems in probability theory and mat. statistics. Uch. allowance for VTUs. Edition 2 add. – M.: «Vuschaja schkola», 1975. – 333 p. (in russian)