GIS Based Fuzzy Multi-criteria Analysis for Industrial Site Selection

Svitlana Kuznichenko
dept. of Information Technologies
Odessa State Environmental University
Odessa, Ukraine
skuznichenko@gmail.com

Iryna Buchynska
dept. of Information Technologies
Odessa State Environmental University
Odessa, Ukraine
buchinskayaira@gmail.com

Abstract*—In the paper, the GIS-based multi-criterial model of decision-making support for industrial site selection is proposed. The formalized description of spatial decision-making process is based on the use of multi-criterial decision analysis in a spatial context, where alternatives, criteria, and other elements of solution to the problem have spatial dimensions. The method of decomposition of the set of source objects influencing the decision making on the thematic layers of criteria, is described. The sampling procedure for vector layers for criteria is described in a raster model, which allows a set of cells, attributes of which contain information about the value of function of the effect of layer objects as well as method of determining the set of possible alternatives, taking into account constraints that may be imposed on attribute values. The method of standardization of criteria based on fuzzy logic methods, which allows using expert knowledge in spatial analysis, is proposed. It is shown that phasing of criteria, that is, the transformation of their values of attributes into a fuzzy set on the basis of the expert estimation of a fuzzy membership function allows further combining of criteria with the help of fuzzy rules of output. Fuzzy logic operations such as intersection or union may be used for this purpose. Different methods for determining the standardized weighting criteria and aggregation operators that can be used in the GIS environment, are described. It is noted that it is more reasonable to use the OWA operator, which allows to formalize expert information about the acceptable form of compromise between values according to different individual criteria with the help of a fuzzy quantifier. It is shown that the use of fuzzy logic in the decision making model allows to take into account the uncertainty of the source information and to obtain a more informative combined suitability map by determining the rank of suitability of alternatives, that is, to perform ranking of territories according to the degree of suitability for industrial site selection.

Keywords—geographic information system; decision support system; multiple-criteria decision analysis; fuzzy logic; site selection analysis.

I. INTRODUCTION

Modern geographic information systems (GIS) are an important component of decision support systems (DSS) thanks to advanced functions of preservation, processing and analysis of spatial data, simulation tools and availability of visualization tools. Spatial solutions by their nature are always multi-criteria [1], so DSSs that are designed to support spatial decision-making are often used in cases where a large number of alternatives should be evaluated based on several criteria.

In the last 20 years, GIS actively integrates various methods of multi-criteria decision analysis (MCDA) [2-4]. Separate attempts to fully integrate MCDA and GIS tools in the general interface revealed problems with the lack of flexibility and interactivity of similar systems that can not provide the necessary freedom of action for analysts [5]. Therefore, the choice of procedure and appropriate methods of MCDA, which can provide a better solution to a specific problem, is an urgent task for developers. In addition, preferences of the decision maker (DM), which are often vague, are unimportant, play an unclear role in the MCDA procedure. To take into account subjective fuzzy DM judgments, it is expedient to improve methods of MCDA with the help of the apparatus of "soft" computing, the fuzzy sets theory [6].

II. FORMALIZATION OF THE PROCESS OF MULTI-CRITERIA DECISION ANALYSIS IN GIS

Consider the use of multi-criteria decision analysis to support the adoption of managerial decisions on finding the best location of an industrial object. The general process diagram is shown in Fig. 1. In solving such a task it is important to take into account multiple factors that influence the decision making: the geographical location of the site and its physical characteristics, resource supply of production, transport and social infrastructure, the condition of natural environment and possible negative impact on it, regulatory and legal constraints, etc. There is a complex structure of interaction of various objects and factors of different physical and socio-economic nature. The more precise these factors will be determined at the preliminary stage of study of the problem, the more adequate the model will be. For example, in [7] authors developed a multi-criteria model for making decisions on placement of landfills for solid household waste in the south of the Odessa Region, which took into account physical, environmental and socio-economic factors. In general, 14 criteria were formulated, which were presented in the geo database in the form of vector and raster layers.

We describe the method of decomposing a plurality of objects belonging to the investigated territory and influencing decision-making in the thematic layer of criteria.

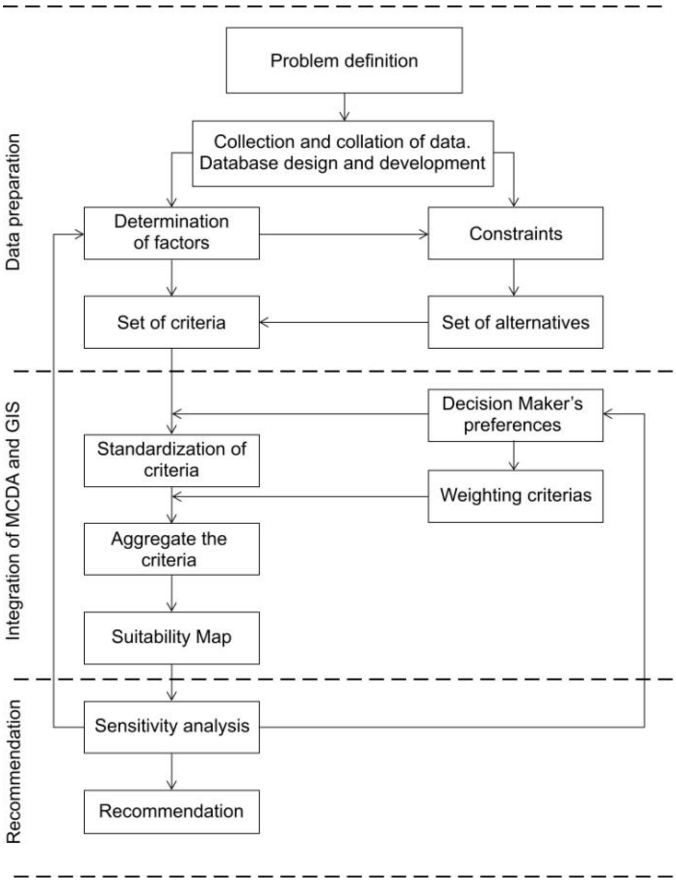

Fig. 1. General diagram of the process of multicriterial decision analysis in GIS

Let us imagine some finite set of objects that influence the solution be given:

$$O=\{o_i\}=\left\{\left(G,\{I_j\}\right)_i\right\},\ i=\overline{1,n},\ \ j=\overline{1,m}, \qquad (1)$$

where G is information on the spatial position of the object; I is attributive information about the object; n is the total number of objects belonging to the investigated area and affect the decision; m is the number of attributes of the object.

It is necessary to select a set of O_j subsets that influence the decision on any factor (availability of transport infrastructure, type of soil, ecological safety, etc.) from a set of objects O and combine them into separate vector layers of criteria.

$$O=\bigcup_{j=1}^{t} O_j,\ O_j\in O. \qquad (2)$$

The method of decomposing objects involves performing an analysis of their spatial and attributive information. Decomposition is usually performed according to following features:

- the set of geometric properties $G'=\{g_1, g_2, g_3\}$, where g_1 is point objects; g_2 is linear objects; g_3 is polygon objects;
- the set of attributive properties $I'=\{Q, N\}$, where Q is a set of qualitative properties that determines belonging of an object to a certain thematic group (transport infrastructure, water objects, settlements, etc.); N is the set of quantitative characteristics of the properties of the object (for example, for entities belonging to the "Settlements" thematic group, one can make a decomposition according to the population size).

Thus, belonging of objects to a certain layer of criteria can be determined by following set of properties:

$$S=\langle G', I'\rangle. \qquad (3)$$

Schematically, the process of decomposition of the set of objects O on thematic layers of criteria is shown in Fig. 2.

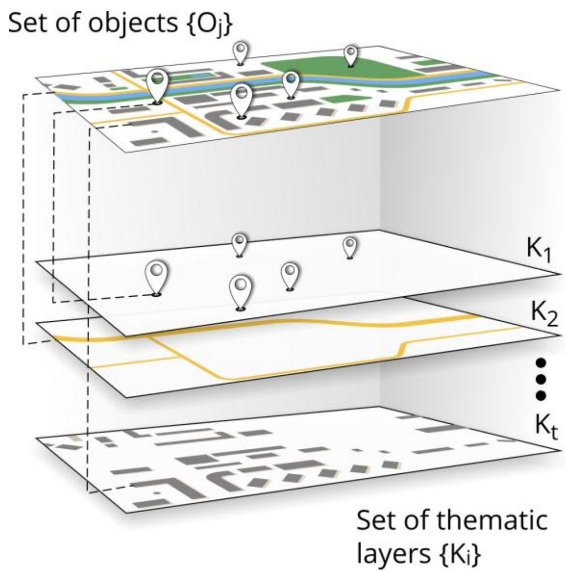

Fig. 2. Scheme of decomposition of objects in thematic layers

After the process of decomposing objects and structuring the problem, we obtain a vector map K representing a set of thematic vector criteria layers K_i (Fig. 2):

$$K=\{K_i\},\ i=\overline{1,t}, \qquad (4)$$

$$K_i=\{O_j^{\,i}\},\ j=\overline{1,l}, \qquad (5)$$

where i is a number of map layer K, j is an object number in the i-th layer.

For conducting a spatial modeling, it is convenient to use a raster data model. Therefore, it is advisable to represent the received vector layers of objects as a set of cells (pixels) in a GIS raster model, which has the form of a two-dimensional discrete rectangular grid of nxm cells, where $\Delta x=\Delta y=\Delta r$ is a cell size:

$$A=\{a_i \mid a_i=n\Delta r, m\Delta r\} \qquad (6)$$

The set A is a set of alternatives. To reduce the equation (5), it can be written as follows:

$$A=\left\{a_i \mid i=\overline{1,n\cdot m}\right\} \quad (7)$$

It is important to choose such a sampling procedure for vector layers of the criteria in the raster, which will receive a set of cells whose attributes contain content information about the value of impact function of objects of the layer. For example, attributes can be derived from vector maps that contain point objects of observation points by the value of some factor using different methods of interpolation.

Often, the distance measurement is used to study the relationship between objects and their interaction, for example, using the Euclidean metric, the value of which between two point objects O_1 (x_1, y_1) and O_2 (x_2, y_2) is calculated by equation:

$$ED(O_1,O_2)=\sqrt{\left(x_1-x_2\right)^2+\left(y_1-y_2\right)^2}. \quad (8)$$

In the case of a raster data model, the distance from any cell of the raster to the object O_i will be equal to the minimum distance from this cell to each cell that covers the object being investigated.

After completing the sampling procedure, since attributes are variable solutions, you can represent the result of the solution as x_{ij}, that is, the value of the j-th attribute according to the i-th alternative:

$$X=\left\{x_{ij} \mid i=\overline{1,n\cdot m},\ \ j=\overline{1,t}\right\} \quad (9)$$

There may be restrictions imposed on the set of alternatives A: onto attribute values (non-spatial constraints) or onto placements (spatial constraints). For example, in [7], taking into account the State Building Regulations of Ukraine, vector layers of restrictive zones around reserves, airfields, forests and forest plantations, agricultural lands were created with the help of the buffering procedure.

The general restrictive vector layer K^{constr} was constructed using an overlay union operation:

$$K^{constr}=\bigcup_{r=1}^{R}K_r,\ R\subset T, \quad (10)$$

where K_r is a thematic vector boundary layer; R is the set of thematic vector layers on which the general boundary layer is constructed; T is a total number of thematic vector layers.

After performing the rasterization and reclassification operations of the K^{constr} layer C, a raster, cells of which are invalid alternatives, have a value of 0, cells that are possible alternatives – value 1, can be obtained.

$$C=\left\{c_i \mid c_i\in\{0,1\}\right\},\ i=\overline{1,n\cdot m}. \quad (11)$$

To determine the set of possible alternatives A' from the set of alternatives A, we must remove the set of bounding cells by conjunction operation.

III. Use of fuzzy logic apparatus to standardize criteria

Layer criteria typically have different ranges or scale values of attributes. The normalization procedure allows you to transfer output values of attributes from the unprocessed scale to the [0, 1] scale.

The description of spatial information based on methods of fuzzy logic is based on transformation of values of attributes of the i-th layer in the sense of degree of belonging to the fuzzy set a_i:

$$a_i=\left\{(x,\mu_a{}^i(x)) \mid x\in U\right\},\ \ \mu_a{}^i(x):x\to[0,1], \quad (12)$$

where x is the value of the attribute, and U is a continuous set of attribute values.

The membership function $\mu_a(x)$ indicates the degree of membership of the attribute x to to the fuzzy set a_i. Typically, the membership function is built under participation of an expert (expert group), so that the degree of membership is approximately equal to the intensity of manifestation of some factor. In practice, following types of membership functions are applied: linear, triangular and trapezoidal (linear-lump); nonlinear (Gaussian function, sigmoid function, spline).

Fuzzification of criteria, that is, conversion of their attribute values to a fuzzy set, based on expert assessment of the fuzzy membership function, allows further combining the criteria with the help of fuzzy rules of output. Fuzzy logic operations such as intersection or merge may be used for this purpose.

The standard fuzzy intersection of sets $a_1, a_2, \dots, a_t$ for all $x\in U$ is defined as follows:

$$\bigcap_{i=1}^{t}\mu_a{}^i(x)=\min\left[\mu_a{}^1(x),\mu_a{}^2(x),\dots,\mu_a{}^t(x)\right]. \quad (13)$$

The standard fuzzy union of sets $a_1, a_2, \dots, a_t$ for all $x\in U$ is defined as follows:

$$\bigcup_{i=1}^{t}\mu_a{}^i(x)=\max\left[\mu_a{}^1(x),\mu_a{}^2(x),\dots,\mu_a{}^t(x)\right]. \quad (14)$$

The use of a fuzzy intersection operation (13) leads to alternative ranking based on only the lowest rank, that is, it is a pessimistic approach to decision making. Fuzzy union operation (14) takes into account only best evaluations of all criteria.

IV. Methods for determining the normalized weight of criteria

Using multi-criteria decision analysis involves assigning weight criteria to specify their relative importance. In the case of t criteria, the set of weights is defined as follows:

$$W=\left\{w_i \mid \sum w=1,\ i=\overline{1,t}\right\}. \quad (15)$$

The easiest way to evaluate the importance of criteria is to rank, that is, to streamline criteria by an expert in order of importance. Once the rating is set, we can calculate weights according to the equation:

$$w_i = \frac{t - r_i + 1}{\sum \left(t - r_j + 1\right)}, \tag{16}$$

where w_i is the normalized weight for the i-th criterion, t is the number of criteria considered $(j = 1, 2, ..., t)$, and r_i is the rank position of a criterion.

Weights of criteria can be found directly by experts on the basis of a given scale, for example, from 0 to 100. In this case, the normalized weight of a criterion is calculated as follows:

$$w_i = \frac{w_i'}{\sum w'}, \tag{17}$$

where w_i is the normalized weight for the i-th criterion, and w'_i is the score for the i-th criterion.

The normalized weights of criteria can be calculated by the Analytical Hierarchy Process (AHP) [8], which is based on a pair comparison of criteria using the 9-point fundamental Saaty scale of absolute numbers. According to the results of the pair comparison of t criteria, we can construct a matrix $(t \times t)$ in which each element a_{ij}, $i, j = 1, 2, ..., t$ is the estimation of a pair comparison of the i-th criterion with the j-th criterion. For the matrix, own numbers and their own vectors are calculated, and a vector of local priorities is formed.

In order to control the consistency of expert assessments, two related characteristics, the Consistency Index, C.I. and the Consistency Ratio, C.R., are introduced. The reasonable level of consistency in paired comparisons is $C.R.<0.10$, while $C.R.\geq 0.10$ indicates conflicting expert judgments.

V. Methods of aggregation

Aggregation of attributes according to different criteria can be accomplished using various methods of MCDA, which are implemented in GIS. The easiest method is the weighted linear combination (WLC) method, which is based on finding of the average value. The alternative membership function is calculated as follows:

$$\mu_a^{WLC}(x_i) = \sum_{j=1}^{T} w_j \mu_a^{\ j}(x_i). \tag{18}$$

where $\mu_a^j\ (x_i)$ is the function of the membership of the alternative to the i-th criterion, and w_j is the normalized weight of the i-th criterion and $\sum_{j=1}^{T} w_j = 1.$

The weighted product method (WPM) uses the multiplication operation:

$$\mu_a^{WPM}(x_i) = \prod_{j=1}^{T} \left(\mu_a^{\ j}(x_i)\right)^{w_j}. \tag{19}$$

An alternative to the GIS aggregation operators considered is the OWA operator, which was developed in the context of the theory of fuzzy sets [9]. It includes a weighted averaging for specific cases, and maximum and minimum operators – as extremums. The method has two sets of scales: the importance of the criterion and order one. By changing weighting rates of order, you can create maps for different decision making strategies. The OWA operator is flexible and allows to formalize expert information on the permissible form of compromise between values according to different individual criteria with the help of a fuzzy quantifier.

Finally, we note the importance of analyzing the sensitivity of evaluation results to the change in parameters of the model of multicriterial task before formation of final recommendations of DMs, which usually involves analyzing sensitivity of results of ranking alternatives to change in weight rates of criteria.

VI. Conclusion

Based on the proposed GIS-based multi-criteria decision support model, a composite map of suitability can be constructed and ranking of territories according to the degree of suitability for placement of industrial objects can be completed. Application of fuzzy logic in the model of apparatus allows to take into account expert knowledge and judgment, which partially compensates for the lack of information through the use of experts' experience.

References

[1] S. Chakhar and V. Mousseau, "Spatial multicriteria decision making" // Shehkar S. and H. Xiong (Eds.), Encyclopedia of GIS, Springer-Verlag, New York, 2008, pp.747–753.

[2] S. Chakhar and J.M. Martel, "Enhancing geographical information systems capabilities with multicriteria evaluation functions", Journal of Geographic Information and Decision Analysis, Vol. 7, No. 2, 2003, pp. 65–71.

[3] J. Malczewski, "GIS–based multicriteria decision analysis: a survey of the literature", International Journal of Geographical Information Science, 20 (7), 2006, pp. 703–726.

[4] J. Malczewski, "Review Article on the Use of Weighted Linear Combination Method in GIS: Common and Best Practice Approaches", Transaction in GIS 4(1), 2000, pp. 5–22.

[5] K. Lidouh, "On themotivation behind MCDA and GIS integration", Int. J. Multicriteria Decision Making, Vol. 3, Nos. 2/3, 2013, pp. 101–113.

[6] L. A. Zadeh, "Fuzzy sets", Information and Control, 8 (3), 1965, pp. 338–353.

[7] S. Kuznichenko, L. Kovalenko, I. Buchynska and Y. Gunchenko, "Development of a multi-criteria model for making decisions on the location of solid waste landfills", Eastern-European Journal of Eenterprise Technologies, №2/3(92), 2018, pp. 21–31.

[8] T. Saaty, "Decision making with the analytic hierarchy process", Scientia Iranica, 9(3), 2002, pp. 215–229.

[9] R. R. Yager, "On ordered weighted averaging aggregation operators in multicriteria decision making", IEEE Transactions on System, Man, and Cybernetics, 18, 1988, pp.183–190.