A GIS Methodological Framework Based On Fuzzy Sets Theory For Land Use Management

Dmitry Kurtener
Agrophysical Research Institute,
Grazhdansky prospect 14, St.Petersburg, 195220, Russia
e-mail: Courtene@DK3516.spb.edu

Vladimir Badenko
St. Petersburg State Technical University,
Politechnicheskaia str. 29, St.Petersburg, 195251, Russia
e-mail: badenko@venture.spb.ru

Abstract *This paper considers a GIS methodological framework based on fuzzy sets theory for land use management. Some principles of development of the GIS methodological framework are formulated. Applications of the GIS methodological framework are designed. In particular GIS knowledge management fuzzy models for analysis of soil commutative contamination by heavy metals, for the study of soil acidity, and for evaluation of soil conservation actions are obtained.*

Keywords: *geoinformation systems, fuzzy sets theory, GIS knowledge management, land use prediction models*

1 Introduction

Fuzzy set theory is a useful tool for dealing with knowledge about territory, taking into account uncertainty in the interpretation of quantitative information on land use, particularly when this is automated in Geographic Information Systems (GIS). These and other problems of integration of fuzzy set methods to a GIS environment for land use are discussed in many top-ranking international journals [13-16]. The applications of knowledge management for land use, which may be generated from, or adapted to, fuzzy sets theory and fuzzy logic, are wide-ranging: numerical classification of soil and mapping, land evaluation, modeling and simulation of soil physical processes, soil quality indices and fuzzy measures of imprecisely defined soil phenomena.

This paper shows a methodological framework using GIS for knowledge management for land, based on the fuzzy set theory.

The rest of this paper is organized as follows. Section 2 gives a brief overview of how to consider knowledge management within a GIS framework for the land use application domain. Section 3 introduces the definitions of the fuzzy sets theory which are used in this framework. Section 4 shows examples of the use of the framework for real data, in Russia. Finally, section 5 presents conclusions and directions of ongoing work.

2 Conception of GIS knowledge management for land use

When one considers a methodological framework of GIS knowledge management for land use based on fuzzy sets theory, the most appropriate form of discussion is not detailed consideration, but rather high-level principles. Some principles for creating such a framework are listed below:

1. A methodological framework of GIS knowledge management for land use must be built on a combination of different approaches from fuzzy sets theory.

2. The result of combining these approaches is a hybrid set of models to be applied to the data. These hybrid solutions will be named in this paper GIS knowledge management fuzzy models.

3. A GIS knowledge management fuzzy model must be built considering at least the following (software) modules (see Figure 1):

A module which contains a library of process-oriented (or object-oriented) models and expertise systems that are integrated into the GIS environment.

A module with fuzzy algorithms that is integrated into the GIS environment.

The library of process-oriented (or object-oriented) models and expertise systems is intended to create new data, new information and new knowledge. These new data, information and knowledge are built from applying the fuzzy algorithms to the data in the spatial database, using the process oriented models.

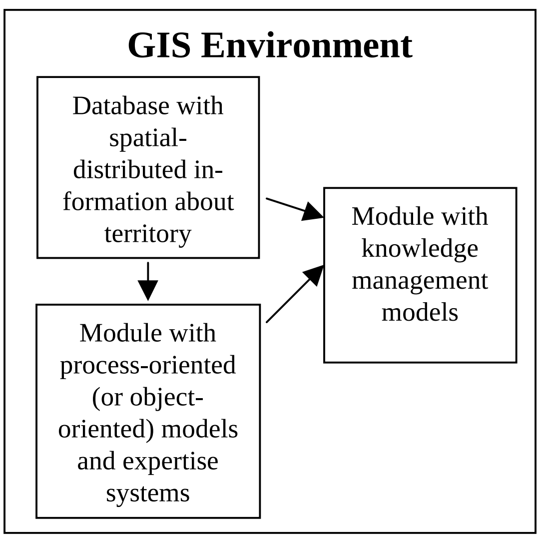

Figure 1: Structure of the GIS knowledge management fuzzy model.

3 Fuzzy sets theory

It is well known that many elements of land plot properties (soil, fertility of soil, microclimate, etc) have uncertainties. Uncertainty is inherent in decision-making processes, which involve data and model uncertainty. These range from measurement errors, to inherent variability, to instability, to conceptual ambiguity, to over-abstraction, or to simple ignorance of important factors.

The fuzzy sets theory originated in the work of Lotfi Zadeh. According to Zadeh [11,12] “The theory of fuzzy sets is, in effect, a step toward a rapprochement between the precision of classical mathematics and the pervasive imprecision of the real world - a rapprochement born of the incessant human quest for a better understanding of mental processes and cognition”.

Fuzzy sets theory is a mathematical method used to characterize and propagate uncertainty and imprecision in data and functional relationships. Fuzzy sets are especially useful when insufficient data exist to characterize uncertainty using standard statistical measures (e.g., mean, standard deviation, and distribution type).

An underlying philosophy of the fuzzy sets theory is to provide a strict mathematical framework, where the imprecise conceptual phenomena in decision making may be precisely and rigorously studied, in particular for knowledge management. The fuzzy sets theory includes fuzzy mathematics, fuzzy measures, fuzzy integrals, etc.

Fuzzy logic is a minor aspect of the whole field of fuzzy mathematics. In classical sets theory, the membership of a set is defined as true or false, 1 or 0. Membership of a fuzzy set, however, is expressed on a continuous scale from 1 (full membership) to 0 (full non-membership). We now introduce some formal definitions, which we will use later on [4, 12]:

Definition 1. Let X be a set (universe). D is called a fuzzy subset of X if D is a set of ordered pairs: $D = [(x, \mu_D(x)), x \in X]$, where $\mu_D(x)$ is the grade of membership of X in D. $\mu_D(x)$ takes its values in the closed interval [0,1]. The closer $\mu_D(x)$ is to 1, the more x belongs to D; the closer it is to 0 the less it belongs to D. If [0,1] is replaced by the two element set {0,1}, then D can be regarded as a subset of X.

Definition 2. The α level set of fuzzy subset D is the set of those elements that have at least a membership: $D(\alpha) = [x: \mu_D(x) \geq \alpha]$. A fuzzy subset n^* is called normal if there is at least one z such that $\mu_{n^*}(z) = 1$

Definition 3. A fuzzy subset n^* of the set of real numbers is called convex if

for each real number $(x,y)\ u \in [0,1]$ one has:

$$mn^*(ux + (1 - u)y) \geq min(mn^*(x), mn^*(y)]$$

Definition 4. A fuzzy subset n^* is called a fuzzy number, if n^* is a normal convex fuzzy subset of the set of real numbers.

4 Applications of the GIS methodological framework based on fuzzy sets theory for land use

The GIS methodological framework based on fuzzy sets theory for land use management was considered by Kurtener et al. [5-7]. Some new results are shown in this section.

4.1. Analysis of soil contamination by heavy metals

In the analysis of soil contamination by heavy metals it is very important to obtain thematic maps describing the index of cumulative soil contamination. Application of the GIS methodological framework with this aim gives a possibility for solving of the problem.

A GIS knowledge management fuzzy model includes two sub-models. The first describes an index of soil contamination by *i* - pollutant with the following equation:

$$\mu_i = \begin{cases} 0: & C < th/3 \\ 1{,}5C / th - 0{,}5: & th/3 < C < th \\ 1: & C > th \end{cases} \quad (1)$$

where *th* is threshold of *i* - pollutant, *C* is the current value of this pollutant.

The second sub-models describing index of cumulative soil contamination $\bullet_{com}$ is based on the use of the operation of algebraic sum for fuzzy sets [4]

$$\bullet_{com} = \bullet_1 \oplus \bullet_2 \oplus \ldots \oplus \bullet_k \quad (2)$$

where $\bullet_j$ is index of soil contamination by *i* – pollutant.

On the basis of the model, a software oriented for use with MapInfo Professional GIS software, version 4,0 has been designed.

The approach was used for management of agrarian territories located in suburbs of St. Petersburg. In particular, values of indices of soil contamination by different heavy metals and their cumulative affects are calculated and mapped automatically (Figures 2 – 4).

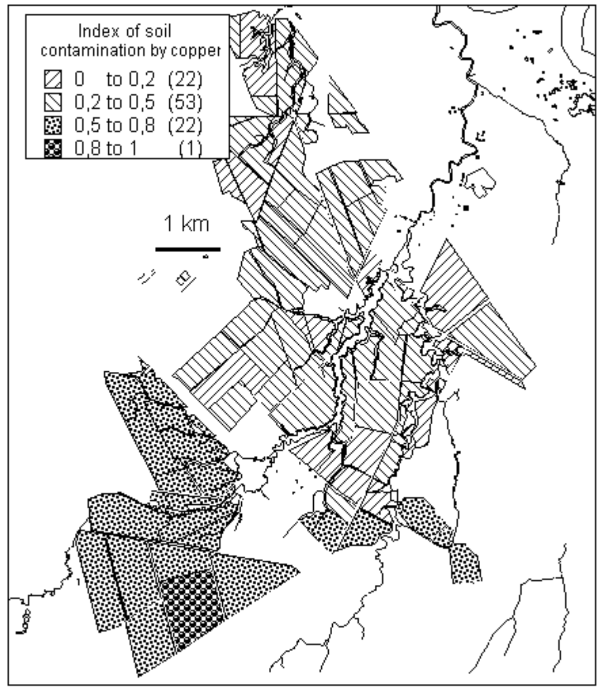

Figure 2: Thematic map describing index of soil contamination by copper.

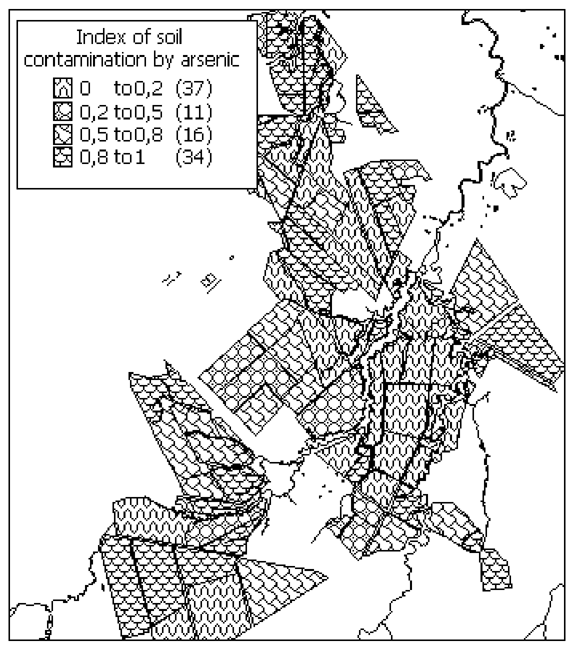

Figure 3: The thematic map describing index of soil contamination by arsenic

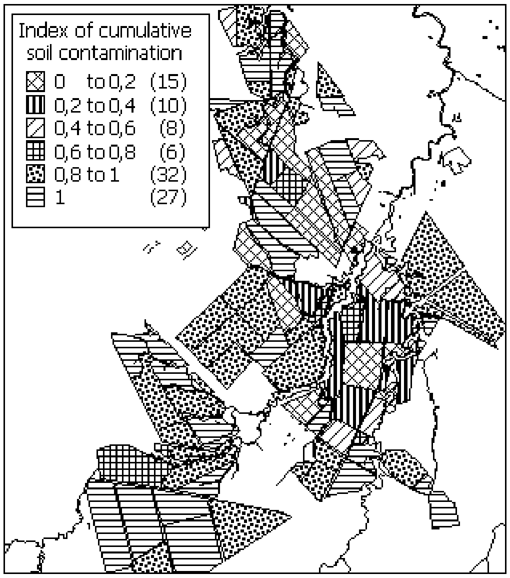

Figure 4: Thematic map describing index of cumulative soil contamination

4.2 Multiple assessment of land parcels for soil conservation

The second example of the use of the methodological framework concerns assessment of land parcels, having soil conservation as a goal. In the process of monitoring natural resources it is a need to estimate an effect of soil conservation actions taking into account ecological, agricultural, and socio-economic factors.

In this case, the GIS knowledge management fuzzy model is based on a multicriteria theory of multiple assessment of landscape parcels and choice of alternatives [3]. In particular, consider there is a set of *m* landscape plots as described in equation (3):

$$A=\{a_1, a_2, \dots a_m\} \qquad (3)$$

and a fuzzy set of criteria described in (4):

$$C=\{\mu_c(a_1)/a_1, \mu_c(a_2)/a_2, \dots, \mu_c(a_m)/a_m\}, \qquad (4)$$

where the membership function $\mu_c(a_i)$ expresses the experts knowledge about grade of landscape plots satisfaction to criteria *C*.

If there are several criteria $C_1, C_2, \dots, C_n$, and the coefficients of relative significance of these criteria are α_1, $\alpha_2, \dots \alpha_n$, the rule for selection of the best land plot taking into account α_i will be written as intersection of C_i.

$$D=C_1^* \cap C_2^* \cap \dots \cap C_n^*;$$

where is

$$C_i^* = C_i^{\alpha_i}, \quad \alpha_i \geq 0, \quad i=\overline{1,n}; \quad \frac{1}{n}\sum_{i=1}^{n}\alpha_i = 1.$$

The coefficients of relative significance α_i are found by comparison of pairs of criteria. To start this assessment, these criteria are initially entered into matrix *B***.** Elements b_{ij} of matrix *B* are defined in Table 1 and must satisfy the conditions: $b_{ij} = 1$, $b_{ij} = 1/b_{ji}$. For example, if user is estimated relative importance of criteria C_i and C_j *as* equilibrium so element $b_{ij} = 1$; if user is estimated relative importance of criteria C_i and C_j *as* great importance so element $b_{ij} = 7$.

Relative importance of criteria C_i and C_j	Element b_{ij}
Equilibrium	1
Very of little importance	3
Of little importance	5
Importance	7
Great importance	9
Intermediate value	2, 4, 6, 8

Table 1. Scale of evaluation of relative importance of criteria

Next, the self-vector of the matrix *B* is determined from the solution of equation (5):

$$Bw = \lambda_{max} w \qquad (5)$$

where λ_{max} is maximum of self-number of the matrix. The solution sought is given by $\alpha_i = n\, w_i$, where n is a predefined number of criteria.

The mathematical operation of intersection of fuzzy sets is in agreement with operation of the search for minimum of the membership functions of these fuzzy sets. In this specific problem, preference is given to land plots that are characterized by the greatest value of the membership function.

Let us consider an example of multiple assessment of soil conservation actions by this approach. We will proceed to analyze the territory located in the suburbs of St. Petersburg. The plots submitted for analysis are shown on Figure 4 with membership function = 1. They differ in types of soil type, soil hydrology, microclimate, processes of soil degradation, processes of soil contamination, crop rotation, etc.

Multiple assessment in this example was carried out by four criteria:

Technological criteria (TC),

Economical criteria (EC),

Ecological criteria (ECC),

Social criteria (SC).

With the TC we try to evaluate soil conservation actions and the feasibility of re-making the area. TC =1, if the action is fully suitable; TC = 0, if it is not.

With the EC we evaluate the economical efficiency of soil conservation actions. EC =1, if the action is economically efficient; EC = 0, if it is not.

With the ECC we evaluate the environmental additional negative effect of the soil conservation actions. ECC =1, if the environmental additional negative effect is absent, ECC = 0, if it is not.

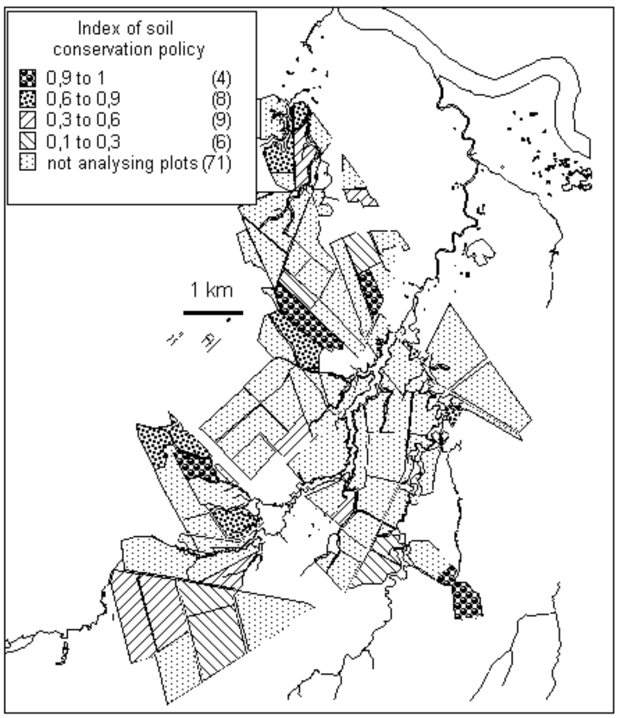

Figure 5: Integral index of evaluation of soil conservation action for selected land plots.

SC provides an estimate of the social factors. SC =1, if the human response is positive; SC = 0, if the human response is negative.

In practice, the values of the four criteria would be assigned by an expert panel. In this example we use the values given in Table 2.

Number of land plot	Criteria			
	1	2	3	4
4	1	0,95	1	0,9
7	0,5	0,4	1	0,7
8	0,8	0,8	1	0,7
9	0,7	0,6	1	0,6

Table 2. Values of the 4 criteria assigned by an expert panel.

The coefficients of relative significance α_i are found by comparison of pair of criteria. In particular, matrix *B* was reconfigured, as shown in (6)

$$B = \begin{Vmatrix} 1 & 5 & 6 & 7 \\ 1/5 & 1 & 4 & 6 \\ 1/6 & 1/4 & 1 & 4 \\ 1/7 & 1/6 & 1/4 & 1 \end{Vmatrix} \quad (6)$$

Components of self-vector of the matrix with λ_{max} = 4,390 are: w_1 = 0,619; w_2 = 0,235; w_3 = 0,101; w_4 = 0,045. Coefficients of relative significance of criteria are: α_1 = 2.48; α_2 = 0,94; α_3 = 0,4; α_4 = 0,18.

On the basis of the fuzzy model, a software oriented for use with MapInfo Professional GIS software, version 4,0 has been designed. By this, software values of integral index of multiple assessment of soil conservation action for selected land plots are calculated and mapped automatically (Figure 5). In particular from Figure 5 it follows that priority should be given to the land plot where the index of multiple assessment > 0,9.

4.3 Analysis of soil acidity

It is well known, that soil acidity (pH) is of one of most important characteristic of soil. Also there is information that agricultural plants may be classified roughly according to their dependence to *pH* according to three classes (first approximation):

(Class 1) $6 < pH < 7$;

(Class 2) $5 < pH < 6$; (6)

(Class 3) $4 < pH < 5$.

The methodological framework proposed on this paper for land use management was applied for calculation and mapping of the correlation index between acidity of land plots and requirements of agricultural plants.

The GIS knowledge management fuzzy model is based on the theory developed by Bogardi et al. [2]. Let *R* and *L* denote the *pH* of plot and *pH* for one of these three classes respectively, described by triangular fuzzy numbers (Figure 6).

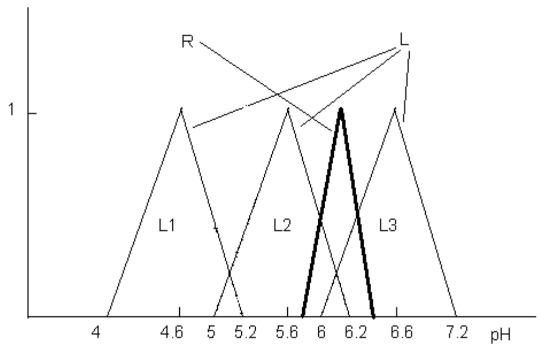

Figure 6: Fuzzy number for plots (*R*) and 3 fuzzy numbers for 3 class of plants.

The fuzzy number *Z* is obtained by the subtraction operation between *R* and *L*:

$$Z(\alpha) = R(\alpha) - L(\alpha), \forall \alpha \in [0,1]. \quad (7)$$

Using fuzzy arithmetic the membership function of the fuzzy number $Z(\alpha) = [Z_1(\alpha), Z_2(\alpha)]$ can be computed from the level sets using the formulas in (8):

$$Z_1(\alpha) = R_1(\alpha) - L_2(\alpha),$$
$$Z_2(\alpha) = R_2(\alpha) - L_1(\alpha), \forall \alpha \in [0,1] \quad (8)$$

The conformity between *pH* of a plot and *pH* of one of classes occurs when $R(\alpha) > L(\alpha)$, i. e. $Z(\alpha)>0$. The event $Z(\alpha) > 0$ implies compatibility. To define a measure to determine the compatibility under such circumstances, we propose a fuzzy compatibility index *Co* as follows in equation (9):

$$Co = \frac{\int_{z>o} \mu_z(z)dz}{\int_z \mu_z(z)dz} \quad (9)$$

Here $\mu_z(z)$ is membership function of fuzzy number *Z.*

On the basis of the model, a software oriented for use with MapInfo Professional GIS software, version 4,0 has been designed.

The approach was used for management of agrarian territories located in the suburbs of St. Petersburg. In particular, values of index of correlation between acidity of land plots and requirements of agricultural plants are calculated and mapped automatically (Figures 7 – 8). In particular, from Figure 7 it follows that only some plots of the second group offer favorable conditions for agricultural plants (plots where index of correlation > 0,9). Most of the territory is suitable for agricultural plants of the first group (Figure 8).

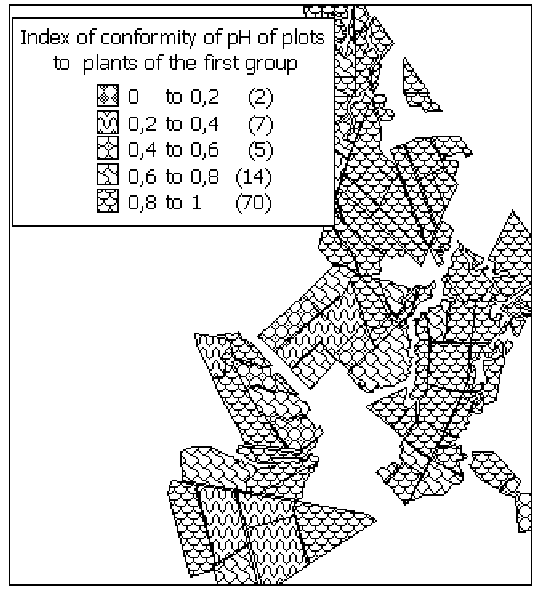

Figure 7: Thematic map describing index of correlation between acidity of land plots and requirements of agricultural plants for the first group.

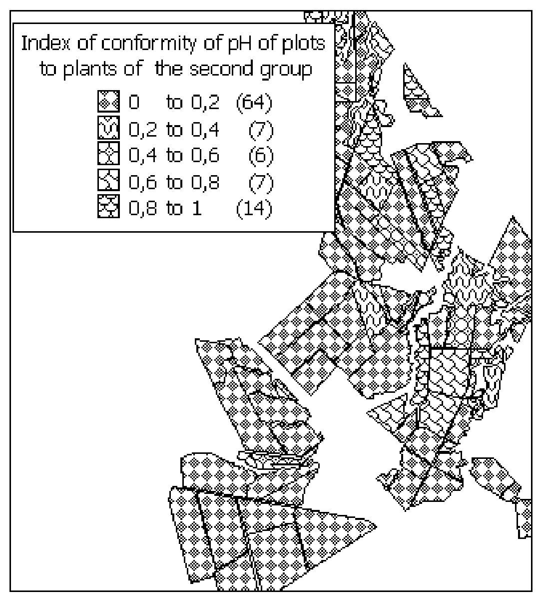

Figure 8: Thematic map describing index of correlation between acidity of land plots and requirements of agricultural plants for the second group.

6 Conclusions

This paper presented a GIS methodological framework based on fuzzy sets theory for land use management.

It has discussed three applications of the GIS methodological framework. Recently we have applied this framework to several other applications. In particular, for evaluation of cumulative influence of consequence of flooding [8] and for multiple assessment of the territorial prophylactic action in the design of strategy of using the health resources [9]. Currently, we have developed an application for the determination of priority for the restoration of burned areas.

The further development of this approach could create a methodological framework for GIS oriented for social, economical and environmental support of decision-making processes.

References

[1] M. Abbott. Introducing hydroinformatics. *Journal of Hydroinformatic.s* 1(1):3-20, 1999.

[2] I. Bogardi, A. Bardossy, M.D. Mays, L. Duckstein. Risk Assessment and Fuzzy Logic as Ralated to Environment. Data Reliability and Risk Assessment in Soil Interpretation. *SSSA Special Publication.* 47:81-97, 1996.

[3] A.N. Borisov,. O.A. Krumberg, I.P. Federov. Decision-making on the base of fuzzy models. Zinatre Publ. Riga. 1990.

[4] A. Kaufmann, M. Gupta Fuzzy mathematical models in engineering and management science. Amsterdam. Elsevier Science, 1988.

[5] D. Kurtener, V. Badenko.. Questions of integration of some ecological models into geoinformation system. *In Proceeding of 21st Urban Data Management Symposium.* CD-ROM, Theme V, 1999.

[6] D. Kurtener, G. Richard, V. Badenko.. Theoretical description of heat and water flow in soil taking into account soil tillage. *In Agrophysical and Ecological Problems Of Agriculture in The 21st Century.* 1:29-53. 1998

[7] D.A. Kurtener, V.L. Badenko. Use of geographical information systems (GIS) in agrophysics. *In Proceeding on Flow and deformation in biology and environment. Prague, pages* 34-39, 1998

[8] D. Kurtener and V. Badenko. Questions of integration of some ecological models into geoinformation system. *In Proceedings of UDMS'99 on Information technology in the service of local government planing and management. Venice, Italy, pages* 21-23, 1999.

[9] Dmitry Kurtener, Viktor Yakushev, Vladimir Badenko, Elaheh Pourabbas. Development of methodology of multiple assessment of landscape parcels on the base fuzzy models integrated into GIS environment. *R. 1*, St. Petersburg, Russia, 23p. 1999.

[10] V.P. Yakishev, D.A.. Kurtener, N.V. Arefiev, V.L. Badenko, N.M. Khimin, T.I. Prokof'eva and L.K. Shvetsova. Use of geographical information systems (GIS) in agrophysics. *Russian Agricultural Science.* 2:52-54. 1999.

[11] L.A. Zadeh. Fuzzy sets. *Inform. Control*. 8(3):338-353, 1965

[12] H.J. Zimmerman. Fuzzy set theory and its application. MA, 1985

[13] X.Y. Chen. A multivatiable analysis of grid data as a soil classification and mapping tool: a case study of a homogeneous plain in Vagia, Viotia, Greece. *Geoderma.* 77(1):63-81, 1997.

[14] Alex McBratney, Inakwu, O.A. Odeh. The application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. *Geoderma.* 77(2-4):85-113. 1997.

[15] R.M. Lark, H.C. Bolam. An uncertainty in the prediction and interpretation of spatially variable data on soils. *Geoderma.* 77(2-4):263-282, 1997.

[16] P. Lagacherie, D.R. Cazemier, P.F.M. van Gaans, P.A. Burrough. Fuzzy k-means clustering of fields in an elementary cathment and extrapolation to a larger area. *Geoderma.* 77(2-4):197-210, 1997