Assessing natural resources using knowledge-based information processing tools

V V Nosov\textsuperscript{1}, M G Tindova\textsuperscript{2}, I A Ramazanov\textsuperscript{3}, L P Poletaeva\textsuperscript{1} and V P Avdotin\textsuperscript{4}

\textsuperscript{1} Department of Economics and Management, K.G. Razumovsky Moscow State University of technologies and management, 73 Zemlyanoy val, Moscow 109004, Russia
\textsuperscript{2} Economics, Applied Mathematics and System Analysis Department, Yuri Gagarin Saratov State Technical University, 77 Polytechnicheskaya street, Saratov 410054, Russia
\textsuperscript{3} Basic Department of Trade Policy, Plekhanov Russian University of Economics, 36 Strelnosny lane, Moscow 117997, Russia
\textsuperscript{4} Agrarian Technological Institute, Peoples' Friendship University, 6 Miklouho-Maclay street, Moscow 117198, Russia

E-mail: novlal@list.ru

Abstract. To date, other than the land market, the Russian Federation still does not have an established natural resources market due to restrictions on transactions in the natural resources segment; therefore, the market value of natural resources cannot be calculated using classical evaluation methods, and only cadastral appraisal can be conducted. We propose to assess the resources as a whole, using the evaluation of nature reserves, assessment of the resource’s ecological functions and the environmental damage. However, the existing cadastral appraisal methods take no account of resources functions and consider them separately. Moreover, the information on natural systems used for the assessment of resources is incomplete and inaccurate. Such uncertainty makes it difficult to model using classical statistical tools. Therefore, there is a need to use information processing tools based on data mining. We propose to use the fuzzy modelling techniques in natural resources assessment, as they will allow for the evaluation of nature reserves and environmental monitoring at the same time. This paper contains fuzzy models for a land plot evaluation that include linguistic variables - factors that affect the land plot value, and fuzzy rule bases that define relationships between linguistic variables. The results show that the cadastral value of land plots is 2.5 times lower than the one calculated using the fuzzy model, with regard to the ecological factor. The error between the market value of those land plots and the one calculated using the fuzzy model is 11%. Therefore, the use of proposed tools allows for the calculation of the market value of resources based on their cadastral value with an account of various factors using the comparative approach, and also for adjustments to the market value in order to increase the efficiency of natural resources use.

1. Introduction
To date, other than the land market, the Russian Federation still does not have an established natural resources market due to restrictions for transactions in the natural resources segment. Therefore, the market value of natural resources cannot be calculated using classical evaluation methods, and only
cadastral appraisal can be conducted. A significant amount of studies by Russian scholars and scholars abroad has been devoted to the issue of natural resources assessment. In these studies, the assessment of natural resources is conducted separately. Some authors only evaluate the cost of natural resources [1-10], while the others assess the resource’s ecological functions or the environmental damage [11-19]. In our view, there is a need to assess the resources as a whole. However, as can be seen, the existing cadastral appraisal methods take no account of resources functions and consider them separately. Moreover, the information on natural systems used for the assessment of resources is incomplete and inaccurate. Such uncertainty makes it difficult to model using classical statistical tools. Therefore, there is a need to use information processing tools based on data mining. We propose to use the fuzzy modelling techniques in natural resources assessment, as they will allow for the evaluation of nature reserves and environmental monitoring at the same time.

The term “fuzzy” means vague or imprecise or uncertain or inexact [20]. Fuzzy sets enable us to accept the vagueness and lack of precision. Fuzzy sets are used when the classical crisp representation cannot provide a decision for a problem [21-22].

When assessing natural resources, a need often arises to deal with linguistic variables and inaccurate numerical data. There is a comprehensive amount of studies on linguistic variables and their modelling using fuzzy sets [23-28].

2. Methods and materials
A linguistic variable is a variable whose values are words or sentences in a natural or artificial language [23]. This variable is described by the following set \( L=(T, U, G, M) \), where \( T \) is the term-set of \( L \), that is, the collection of its linguistic values, \( U \) is a universe of discourse, \( G \) is a syntactic rule which associates with each linguistic value its meaning, \( M (M \subset U) \).

Membership functions for fuzzy sets should satisfy the following conditions:

- order: \( \mu_{T_i}(u_{min}, \mu_{T_n}(u_{max}) \)
- completeness and consistency: \( 0 < \max_{u \in U} \mu_{T_i}(u) < 1 \)
- normality: \( \forall u, \exists i : \mu_{T_i}(u) = 1 \)
- boundedness: \( 0 < \int_{0}^{\infty} \mu_{T_i}(u)du < +\infty \)

The process of creating a linguistic variable using a method of fuzzy clustering of experimental data involves the following steps:

1. Define a name and a number of terms of every linguistic variable.
2. Determine the bounds of the universe of discourse \( U \) using the minimum of the variable (Umin) as a lower bound and the maximum of the variable (Umax) as an upper bound.
3. Perform fuzzy clustering and determine the bounds of every term's membership functions using the number of terms and the values of the upper and lower bounds of the variable that belongs to the set.
4. Determine the mean squared deviation of each cluster and build a triangular membership function for every term by plotting its value left and right from the peak (centre) of the membership function. Build a Gaussian function for terms \( T_i \) and \( T_n \) in a similar way by setting single values of the membership function for the extreme terms.

Introduced linguistic variables satisfy all the requirements.

Using clusters as terms of the linguistic variable, we can obtain the bounded universe of discourse \( U \) as the trained network provides a limited number of clusters.

For example, in accordance with the proposed approach, terms \( T_1=\{\text{«low»}, \text{«medium»}, \text{«high»} \} \) were chosen for the linguistic variable \( L_1 \) – price in the fuzzy model of land assessment.

The universe of discourse of the database analysis is defined by the interval \( X_1=[1, 50] \). By following the third and fourth steps of the algorithm, we obtain the membership functions of terms presented in figure 1.
Figure 1. Membership functions for the terms of a linguistic variable $L_1$.

Analytically, the membership functions of terms of the linguistic variable $L_1$ can be described as follows:

$$
\mu_{T_1} = \begin{cases} 
1, & t \leq 1 \\
\frac{5-t}{4}, & 1 \leq t \leq 5 \\
0, & t \geq 5 
\end{cases}
; 
\mu_{T_2} = \begin{cases} 
0, & t \leq 3 \\
\frac{t-3}{17}, & 3 \leq t \leq 20 \\
1 - \frac{t}{10}, & 20 \leq t \leq 30 \\
0, & t \geq 30 
\end{cases}
; 
\mu_{T_3} = \begin{cases} 
0, & t \leq 25 \\
\frac{t}{25} - 1, & 25 \leq t \leq 50 \\
1, & t \geq 50 
\end{cases}
$$

After choosing the linguistic variables, the next step in the fuzzy model construction is to compose a rule base (linguistic model) in terms of linguistic «If $X$ is $A$ then $Y$ is $B$» rules.

The completeness requirement for the «if $A_i$ then $B_i$» system means that for every current state of the process there is at least one rule and that for every term of an input variable there is at least one rule in which this term is used as a premise (the left-hand side of the rules).

The next step after formulating logical rules is their processing, which consists of:

1. Calculation of the truth degrees of the left-hand side of the rules, i.e. determination of the degree of membership of input data to fuzzy subsets, shown in the left-hand side of the rules.
2. Modification of fuzzy subsets, shown in the right-hand side of the rules, according to truth values obtained in the first stage.
3. Grouping (superposition) of modifies subsets.
4. Scaling of superposition results.

3. Results and discussion

Assessment of the value of natural resources through the fuzzy logic inference system can be represented as an algorithm (the flowchart is presented in figure 2). Realization of the algorithm can be as follows:

1. Consider the assessed resource as the land plot and improvements.
2. Based on the land plot category and size, calculate the cadastral value using the fuzzy model for land assessment (FML).
3. Assess the improvements of land using one of the following: a fuzzy model for forest resources assessment (FMFR), a fuzzy model for water resources assessment (FMWR), a fuzzy model for minerals assessment (FMM).
4. Determine the environmental damage using the fuzzy model for environmental damage assessment (FMED).
5. Calculate the total value of the natural resource as the sum of steps 2, 3 and 4.
Taking into account that the value of land depends on its category, we grouped all fuzzy rules in the rule base. To find the base price, we implement the rules that define the location and size of land plots. Next, we define the category of land and implement the rules that define features essential for each category. Thus, by following the described algorithm, we can find the final price of the land plot.

To determine the value of improvements (forest, water and mineral resources) of the land plot, we constructed fuzzy models for assessment of forest, water and mineral resources, which also consist of
linguistic variables – factors that affect the value of these resources – and fuzzy rule bases that define the logical relationship among the linguistic variables.

Linguistic variables of the fuzzy model for water resources assessment (FMWR) are as follows: price per 1 cubic meter; category of water bodies; location (watershed districts); size; water quality; use; protection and restoration activities.

For the fuzzy model for forest resources assessment (FMFR), linguistic variables would be the price per 1 ha; forest category; size; location; distance from traffic arteries; value.

In the fuzzy model for minerals assessment (FMM), we used the following linguistic variables: price; available information; geological features; feasibility; economic efficiency; purpose.

Ecological condition of land is one of the main pricing factors; therefore, we constructed a fuzzy model for environmental damage assessment (FMED). The final price can be adjusted based on the more detailed study of the ecological condition of the land plot.

Thus, the constructed fuzzy model allows for the economic assessment of natural resources as a whole, including the evaluation of nature reserves, assessment of the resource’s ecological functions and the environmental damage to determine the most efficient type of their use.

Data for the Saratov region presented in table 1 were used for testing of the fuzzy model for assessment of land plots, forests and isolated water bodies.

| Category         | Cadastral value, млн. руб/га | Market value, млн. руб/га | FML | FMWR | FMFR | FMM | FMED | σ₁ | σ₂ | σ₃ | σ₄ |
|------------------|-------------------------------|---------------------------|-----|------|------|-----|------|----|----|----|----|
| Agriculture      | 0.103                         | 0.162                     | 0.170 | 0.165 | 0.628 | 0.097 | 0.926 | 0.031 |
| Water            | 3.7                           | 6                         | 5.8  | 6.2  | 6.25 | 12.8 | 0.041 | 7.055 | 0.075 |
| Forest           | 1.06                          | 1                         | 1.086 | 1.1  | 0.031 | 0.1  |      |      |
| Settlements      | 44.1                          | 70.95                     | 60   | 54   | 0.224 | 0.239 | 83.56 | 0.1  |
| Summer houses    | 24.07                         | 19.01                     | 22.1 | 22   | 2.343 | 0.169 | 14.4  | 0.004 |
| Production       | 1                             | 0.812                     | 0.76 | 0.8  | 0.772 | 0.071 | 0.817 | 0.075 |
| Recreation       | 1                             | 2.45                      | 2.2  | 2.21 | 2.42 | 2.58 | 1.58  | 0.053 | 0.91  | 0.172 |
| Reserve          | 1                             | 1.15                      |      |      |      | 1.3  | 0.3   | 0.13  |
| **Mean value**   |                               | **2.571**                 | **0.112** | **17.946** | **0.086** |

With regard to the ecological situation, the relative error σ₁ between the cadastral value and the final price calculated using the fuzzy model is 2.571 (257%), which means that the cadastral value is lower than the final price by a factor of 2.5.

The error σ₂ between the market value and the final price calculated using the fuzzy model with an account of the ecological situation is 0.112 (11%); hence, the use of tools allows for an adjustment of 11% to the market value found using the comparative approach.

The error σ₃ between the market value and the base price calculated using the fuzzy model with an account of location, size and category is 17.95 (17%).

The error σ₄ between the base price found using the fuzzy model with an account of location, size, category and final price (with regard to all factors, including the ecological condition) is 0.086 (0.86%).

4. Conclusion

The results allow for a conclusion that the environmental factor doesn't greatly affect the value of land plots of different categories. In our view, the existing scheme for categorization of lands ignores the best use principle, which in most cases results in the lowered market value of various categories of lands. This, in turn, affects the rental rates for the use of land plots with forest, water and mineral resources and reduces the tax revenues that would otherwise go into the country's fund.
Thus, information processing tools can be used as a basis for management decision-making to increase the efficiency of land use, while the environment assessment module – as a basis for rational area zoning.

References
[1] Manzoor S, Qureshi T, Liaqat M, Farooq M and Shamaill S 2009 A comparison between cadastre 2014 and cadastral systems of different countries ACM Intern. Conf. Proc. Ser. 293-8
[2] Zhichkin K, Nosov V, Zhichkina L, Zhenzebir V and Sagina O 2020 Cadastral appraisal of lands: agricultural aspect IOP Conf. Ser. Earth Environ. Sci. 421 022066
[3] Polat Z 2018 Evolution and future trends in global research on cadastre: a bibliometric analysis GeoJournal 84 1121-34
[4] Noszczyk T and Hernik J 2019 Understanding the cadastre in rural areas in Poland after the socio-political transformation J. Spat. Sci. 64 73-95
[5] Mjøs L 2019 Cadastral development in Norway: the need for improvement Surv. Rev. 1-12
[6] Mika M 2020 Modernisation of the Cadastre in Poland as a tool to improve the land management and administration process Surv. Rev. 52 224-34
[7] Szturc J and Hybler V 2019 The impact of updated soil properties on the development of land price in selected cadastral area near the strongly urbanized areas J. Ecol. Eng. 20 161-8
[8] Zhichkin K, Nosov V, Zhichkina L, Alborova M and Kuraev A Methodology for cadastral valuation of agricultural lands occupied by water bodies 2020 E3S Web Conf. 175 06014
[9] Kharitonov A, Ershova N and Vikin S 2019 The Improvement of Conceptual and Categorical Framework for the Classification of Objects of Cadastral Registration IOP Conf. Ser. Earth Environ. Sci. 272 022210
[10] Zhichkin K, Nosov V, Lakomiak A and Zhichkina L 2020 Cadastral valuation of lands dedicated to perennial plantings: features and practice E3S Web Conf. 177 04002
[11] Myers J and Thorbjornsen K 2004 Identifying Metals Contamination in Soil: A Geochemical Approach Soil Sediment Contam. 13 1-16
[12] Korobova E 2017 Combined assessment of the ecological and geochemical state of anthropogenically impacted areas Geochem. Int. 55 861-71
[13] Zhichkin K, Nosov V, Andreev V, Kotar O and Zhichkina L 2019 Damage modelling against non-targeted use of agricultural lands IOP Conf. Ser. Earth Environ. Sci. 341 012005
[14] Prostov S and Shabanov E 2019 Substantiation of the method of operational monitoring of soil contamination and oil decontamination processes E3S Web Conf. 105 020131
[15] Poltarykhin A, Nosov V, Poletayeva L, Avdotin V, Grishin V and Babakistyev M 2019 Problems of the ecological system in Russia and directions for their solution based on economic and social development programs J. Environ. Manag. Tour. X pp 508-14
[16] Zhichkin K, Nosov V, Zhichkina L, Tkachev S and Voloshchuk L 2020 Prediction methodology for potential damage from misuse of agricultural lands E3S Web Conf. 161 01060
[17] Shepelev I, Nemerov A, Eskova E, Keller V and Potapova S 2019 Ecological engineering as a mean to reduce the anthropogenic impact of production on biota IOP Conf. Ser. Mater. Sci. Eng. 537 062049
[18] Zhichkina L, Nosov V, Zhichkin K, Starikov P, Vasyukova A and Smirnova Z 2020 Monitoring of technogenic pollution of soil in the region IOP Conf. Ser. Mater. Sci. Eng. 862 062061
[19] Parveen S, Parveen S and Rahman N 2020 Fuzzy Systems: A Human Reasoning Approach Using Linguistic Variables IJICV 2019 33 538-545
[20] Zadeh L 1965 Fuzzy sets Inform. Control 8 338-53.
[21] Kilić J, Rogulj K and Jajac N 2019 Fuzzy expert system for land valuation in land consolidation processes Croat. Oper. Res. Rev. 10 89-103
[22] Zadeh L 1975 The concept of a linguistic variable and its application to approximate reasoning Ing.Ser. 8 301-57
[23] Schwartz D 1988 An alternative semantics for linguistic variables *Lecture Notes in Computer Science* 313 87-92
[24] D’Ambrosio B 1989 *Linguistic Variables Qualitative Process Theory Using Linguistic Variables* (New York: Springer-Verlag)
[25] Romanenko Y, Tsybul'skii A and Romanenko A 2000 Algorithm for determining classes of indistinguishable faults by using linguistic variables *Meas. Tech.* 43 673-4
[26] Bailón A, Blanco A, Delgado M and Fajardo W 2001 A Coding Method to Handle Linguistic Variables *Granular Computing* 70 187-214
[27] Williams J, Steele N, Robinson H 2001 Modelling Non-Numeric Linguistic Variables *Developments in Soft Computing* 9 124-31
[28] Xu Z 2007 Group Decision Making with Triangular Fuzzy Linguistic Variables *IDEAL* 4881 17-26