Software implementation of fuzzy logic algorithms for environmental risk assessment

I I Pivovarova¹, R D Terekhin², S V Sarkisov¹, A A Sorokin¹ and V I Musatov¹

¹Military Institute (engineering) of the Military Academy logistics support named after Army General A. V. Hrulev, 22 Zakharyevskaya, St. Petersburg, 196121, Russian Federation
²Saint Petersburg Mining University, 2, 21 Line, Vasilievsky Island, St. Petersburg, 199026, Russia

E-mail: inivpivovarova@gmail.com

Abstract. The article proposes a developed expert information system for assessing environmental risks for complexly structured and uncertain source data. The result of using the program is the desubjectivization of the decision-making process in the presence of a large amount of heterogeneous information. It is supposed to use the system as a technical means of monitoring and assessing the harmful effects of military installations on the environment. The program should serve to increase the effectiveness of environmental protection measures at special facilities of the Ministry of Defense of the Russian Federation. The mathematical algorithm of the developed software application is based on the theory of fuzzy sets (Mamdani's logical inference diagram). The solution involves finding the degree of truth and bringing to clarity the initially fuzzy input data by the centroid method and, as a result, a score assessment of the quality of environmental characteristics.

1. Introduction
Currently, various types of mathematical modeling are widely used in the engineering practice of environmental research. The main ones are deterministic approaches to the study of anthropogenic environmental impact and probabilistic-statistical models [1]. The most common standardized methods include, for example, impact identification matrices, mapping and geographical information systems (GIS), deterministic-stochastic models of environmental risk assessment for receptor regions based on forward and reverse modeling [2-4].

The concept of risks in environmental research is of fundamental importance. The specificity of this class of problems is that it is necessary to consider a wide range of interacting processes over long time intervals in areas of various scales. The description and prediction of complexly structured data volumes is most easily handled by fuzzy logic mechanisms that are widely used in artificial intelligence systems [5].

The proposed approach to the assessment of environmental risks allows taking into account the uncertainty in the external and internal sources of disturbances. The software package makes it possible to determine the extent of the occurrence of environmental disasters and draw conclusions about the causes. In general, the developed tools should help to improve the level of environmental safety, which is especially important for military infrastructure facilities.
2. Materials and methods

There are a large number of fuzzy inference algorithms, among them the most famous are the Mamdani and Sugeno-Takago algorithms. The mechanisms of activation and defuzzification differ in them, since the rules of fuzzy inference are set differently. The first uses the centroid method (calculating the center of mass of the output variables graph), and the second calculates the output value using a number of clear values obtained in a certain way according to the rules specified in the system design [6].

In our project, we used the Mamdani algorithm, when all the stages of fuzzy inference were performed sequentially. All values obtained in the previous step were used in the following:

2.1 Formation of the rule base

A rule base is a set of rules where each sub-conclusion is assigned a specific weight factor. By default, the weight factor is set to one. The linguistic variables in the conditions are called input, and those in the conclusions are output.

2.2 Fuzzification of input variables

This step is often called fuzziness reduction. The formed rule base and the input data array \( A = \{a_1, \ldots, a_m\} \) arrive at the input. The purpose of this step is to obtain truth values for all sub-conditions from the rule base. This happens as follows: for each of the sub-conditions, the value \( b_i = \mu(a_i) \) is found. Thus, a set of \( b_i \) values is obtained.

2.3 Aggregation of sub-conditions

The purpose of this stage is to determine the degree of truth of the conditions for each rule of the fuzzy inference system. Depending on the rule, it is necessary to apply the operation of conjunction and / or disjunction to the sub-conditions. For each of the rules, a set of \( c_i \) is obtained.

2.4 Activation of sub-conclusions

At this stage, there is a transition from conditions to sub-conclusions. For each connection, degree of truth \( d_i = c_i \ast F_i \) is found. Modification of the membership function is made from the sub-conclusion. Its new value is determined as the minimum of \( d_i \) and the initial value of the membership function. This method is called min-activation, which is formally written as follows:

\[
\mu_i'(x) = \min\{d_i, \mu_i(x)\}.
\]

So, the purpose of this stage is to obtain a set of “activated” fuzzy sets for each of the sub-conclusion in the rule base.

2.5 Accumulation of conclusions

The purpose of this step is to obtain a fuzzy set (or their combination) for each of the output variables. It is performed as follows: the i-th output variable matches the combination of sets \( E_i = \bigcup D_j \). J - the numbers of the sub-conclusions in which the i-th output variable is involved. The combination of two fuzzy sets is the third fuzzy set with the following membership function:

\[
\mu_i'(x) = \max\{\mu_1(x), \mu_2(x)\},
\]

where \( \mu_1(x), \mu_2(x) \) are the membership functions of the combined sets.

2.6 Defuzzification of output variables

The purpose of defuzzification is to obtain a quantitative value for each of the output linguistic variables. Formally, this happens as follows. Firstly, it is necessary to consider the i-th output variable and the set \( E_i \) related to it. Then, using the defuzzification method, the final quantitative value of the output variable
is found. This algorithm uses the center of gravity method, in which the value of the i-th output variable is calculated by the formula:

\[ y_i = \frac{\int x \cdot \mu_i(x) \, dx}{\int \mu_i(x) \, dx}, \]

where: \( \mu_i(x) \) – membership function of the corresponding fuzzy set \( E_i \); \( y_i \) – the result of defuzzification [7-9].

3. Results and discussion

The developed software implementation of the described fuzzy classification algorithm allows both to accelerate the decision-making process for environmental monitoring and to move away from a bipolar conformity / non-conformity assessment to a wide range of options intermediate between the two extremes.

Documents regulating the sphere of hydroecological safety in the Russian Federation are presented by SanPiN. However, these documents do not allow an accurate assessment of water quality, since they give only 2 options: compliance with established standards and non-compliance with them. In the field of environmental decision-making may require a broader gradation, taking into account many factors, and making it possible to display even the most minimal changes in these factors in the final assessment. The corresponding opportunity is just implemented in the presented information system.

3.1 Structure

The program is written in accordance with modern programming techniques - an object-oriented approach is used. The program implements the features provided by the Java language: Generics, Collections API, lambda expressions. The Swing-based user interface was created using the visual form designer included with the NetBeans IDE. There was also created a special module (Database) that reads and processes the file with the source data saved using MS Excel.

During the development, the structure of the developed program was modeled and a class diagram of the main mathematical package of the application was created (figure 1).

3.2 Implementation

The Application class has a main () method, which is the entry point to the program. In this procedure, the databases are initialized using the Configurator class, which analyzes the specified file with fuzzy system parameters (the file must be located in the directory from which the program is launched). The
class is written so as to unify the fuzzy system task method with the format used by the MathLab Fuzzy Logic Toolbox. The module returns an array of fuzzy production rules.

Then the MainFrame screen form is displayed, on which the completion of this procedure is ended. When the form display method is called, a separate thread is launched, which processes screen events in a loop and calls methods that have signed up (using the Listener programming pattern).

The user interface allows to call the file selection dialog box. After selecting using the Parser class, the file is read, its format is determined, analysis is performed, and columns with the required data are determined. The program allows to configure file import settings - the user can switch data in columns using the additionally called DataImportFrame dialog box. The source data is displayed on the left side of the main form of the application.

By clicking on the "Calculation" button, the entered values are checked, and if they correspond to the logic of the task, then it accesses the MamdaniAlgorithm class to perform calculations - the execute () method is called, the arguments of which are the initial parameters and fuzzy inference rules. Thus, separation is performed: for the form to work, it does not directly require a module that performs calculations, and vice versa, the calculation part is in no way connected with the implementation of the form.

After calculating all the required values, additional processing of the results is carried out (standardization to obtain a percentage representation of risk). The resulting values are displayed on the right side of the main form (figure 2).

![Figure 2. Program window.](image)

4. Conclusion
The software application was tested on the data of hydrochemical monitoring of water bodies in military towns of the Leningrad Region. The average data for 2018 were processed for the following pollutants: oil products, suspended solids, manganese, chlorides and the indicator of biological oxygen demand (BOD).

The excess of maximum concentration limits for the content of petroleum products and suspended solids was established in 50 and 70 percent of cases, respectively. The content of chlorides and BOD5
is within normal limits. The assessment was carried out both for all factors simultaneously and separately for two identified pollutants.

The problem of wastewater pollution by the technical zones of military camps is quite acute, although over the past decade, the supply of military facilities with treatment facilities has changed for the better - from 77% to 46% of the estimated need. Wastewater from everyday washing of technical equipment makes up 80 ... 85% of the total volume of industrial waste water from technical zones [10]. The main types of pollution of these waters are suspended solids and petroleum products. The concentration of suspended solids in wastewater depends on the type of equipment used, the nature of the pavement, seasonal conditions, soil composition in the area of operation of the equipment, the frequency of washing and the type of washing devices. The main feature of oil products contained in washing wastewater is their low emulsification and a high level of fixation on suspended solids.

The main indicators characterizing the environmental impact of the Armed Forces of the Russian Federation in recent years indicate some stabilization of the ecological situation in the areas where military facilities are located, nevertheless, the problems of ensuring the environmental safety of military activities remain relevant. And the need for quality monitoring and quick decision-making in the field of environmental control is the task of ensuring national security.

Generally speaking, today there is a serious gap in our ability to predict environmental risk with insufficient quantity and quality of hydrochemical monitoring data. Therefore, the most promising method in this case is the use of fuzzy logic algorithms. And the proposed software implementation of the mathematical apparatus will speed up and simplify the classification process and eliminate the human factor in decision making.

References

[1] Łomnicki A 1988 The Place of Modelling in Ecology Oikos 52(2) 139-42
[2] Qian J et al 2016 Multivariate statistical analysis of water chemistry in evaluating groundwater, geochemical evolution and aquifer connectivity near a large coalmine, Anhui, China Environ Earth Sci. 75 747
[3] Ashley R P and Lloyd J W 1978 An example of the use of factor analysis and cluster analysis in groundwater chemistry interpretation. J. Hydrol 39 441–4
[4] Pivovarova I I 2019 Cluster analysis for structuring spatial data International Journal of Mechanical Engineering and Technology 10(2) 1085-109
[5] Prato T 2005 A fuzzy logic approach for evaluating ecosystem sustainability Ecological Modelling 187(2) 361-8
[6] Zade L 1976 The concept of a linguistic variable and its application to making approximate decisions Mir 166
[7] Sandler U and Tsitolovsky L 2008 Neural Cell Behavior and Fuzzy Logic Springer 478
[8] Lu G and Wong D 2008 An adaptive inverse-distance weighting spatial interpolation technique, Computer&Geosciences 34 1044–55
[9] Akhmetshin E M, Barmuta K A, Yakovenko Z. M, Zadorozhnaya L I, Mironov D S and Klochko E N 2017 Advantages of cluster approach in managing the economy of the Russian federation. International Journal of Applied Business and Economic Research 15(23) 355-64
[10] Luzgin B 2013 “Pacification” of high military technologies: risks and consequences Journal Civil protection strategy: Issues and Research 12 11-22