Qualitative measure of the environment risk level for the fuzzy control systems of environmental safety

S V Novikova¹, Y A Tunakova¹, A R Shagidullin² and K N Novikova¹

¹ Kazan National Research Technical University named after A.N. Tupolev (KNRTU–KAI), Kazan, 420111, Russia
² Institute for Environmental Problems and Subsoil Use of the Academy of Sciences of Tatarstan, Kazan, 420087, Russia

E-mail: SVNovikova@kai.ru

Abstract. When creating an environmental management system that should generate a control actions set for regulating the quality of natural objects, it is necessary to be able to "calculate" the qualitative value of the environment state. A qualitative assessment of the environment state is a generalized measure of natural risk. For these purposes, it is very convenient to use the so-called "soft" models that take into account the fuzzy nature of information - a fuzzy inference system and neuro-fuzzy systems. The main advantage of the neuro-fuzzy model over the regression one is the consideration of the ambiguity and fuzziness of the concepts “Acceptable”, “Medium”, “High”, “Unacceptable” and “Catastrophic” risks, as well as the possibility of adaptive correction due to the training of a fuzzy neural network on experimental data.

The calculation of the qualitative values for degree of danger to the environment is important not only in itself, but also as a criterion for making management decisions to normalize or prevent dangerous situations. To develop environmental recommendations on management impacts, the concept of “Risk Measure” is considered in the work, which corresponds to the management decisions made. The criterion “Risk Measure” takes values in the range from 1 to the number of tuples of the experimental sample, and the higher the Measure value, the more global the measures are recommended. Each value can be associated with either a single action or a set of actions. Thus, greater flexibility in the use of the system is achieved, and the range of recommended measures that can compensate for the emerging risks is significantly expanded.

The article describes an algorithm for calculating a quality risk measure. The theory is confirmed by computational experiments using in fuzzy control systems.

1. Introduction

To make decisions on managing the environmental situation, specialists use various assessments of the current and possible future state of the environment. At the same time, they make subjective conclusions regarding the degree of danger and the reasons that cause it. Further, experts determine the necessary measures to eliminate or minimize such a hazard [1].

This approach does not allow automating the decision-making process on the implementation of certain environmental protection measures, introduces significant subjectivity into the process, slows down the development and implementation of the necessary measures.
To automate and accelerate the environmental decision-making process, the authors have developed an automatic decision support system. The main part of the system is the calculation of a specially created qualitative measure of the level of environmental risk. Such a “Risk Measure” will be a numerical expression of the qualitative assessment, which would normally be given by experts based on their own experience. Consequently, the Risk Measure can be applied to select one or more control actions to reduce the hazard in the natural environment.

2. Building regression models of risk measure

The degree of danger to humans is determined on the basis of two fundamental parameters - the overall risk $R_{\text{integral}}$, and the level of risk of metal content in the surface layer of atmospheric air $R_{\text{air}}$. These two factors were chosen because the first criterion expresses a long-term negative anthropogenic impact on the natural environment, and takes into account the mutual influence of its components [2-4] and the second criterion reflects the dynamic component of the impact, since atmospheric air is the most variable environment in the urban ecosystem and has a direct impact on its state at the current time [5-7].

The level of general hazard, subsequently compared with the Risk Measure, was determined as follows: the values of the integral risk $R_{\text{integral}}$ are numbered in ascending order from 1 to the number of tuples of experimental data (in our research – 66), then values of the air risk $R_{\text{air}}$ are also numbered in ascending order. Thus, we obtained paired data tuples with ranked values $R_{\text{air}}$ and $R_{\text{integral}}$. Then the arithmetic mean is calculated for each tuple. The resulting codes rounded to integer values will be the generalized Risk Measure.

The resulting hazard level value is written as an output field in the original datasets of pairs $R_{\text{air}}$ and $R_{\text{integral}}$ without transcoding.

The risk measure is used to form decisions as part of the DSS-neuro-fuzzy recommender control system. Neural fuzzy systems allow to imitate the way of human thinking in their work, in particular, they operate with qualitative concepts for computations [8]. To form the right-hand sides of such a system, it is necessary to present the value of the Measure in the form of a linear polynomial, for which, according to the resulting table, 5 regression models were constructed, corresponding to five qualitative values of the state of the urban ecosystem: “Acceptable”, “Medium”, “High”, “Unacceptable” and “Catastrophic”. Since the level of risk of metal content in the atmospheric air is an integral part of the integral risk, the samples for constructing regressions were selected according to the gradations of the integral risk:

- $<0.021$ acceptable (0.25 quantile)
- $0.021-0.028$ - medium (0.5 quantile)
- $0.028-0.035$ - high (0.75 quantile)
- $>0.035$ - unacceptable
- $>0.041$ - catastrophic (0.90 quantile)

In the case when the level of risk due to the state of atmospheric air is high, its regression coefficients in the corresponding equations increase in proportion to the gradations of risk in the air (0.01-0.015-0.02-0.025), which smoothly corrects the measure.

The resulting models are shown in Table 1.

| State of the urban ecosystem | Values of Risk measure |
|-----------------------------|------------------------|
| Acceptable                  | Measure = 786* $R_{\text{integral}} + 827* R_{\text{air}}$ |
| Medium                      | Measure = 1234* $R_{\text{integral}} + 745* R_{\text{air}}$ |
| High                        | Measure = 1383* $R_{\text{integral}} + 694* R_{\text{air}}$ |
| Unacceptable                | Measure = 1572* $R_{\text{integral}} + 660* R_{\text{air}}$ |
| Catastrophic                | Measure = 1696* $R_{\text{integral}} + 633* R_{\text{air}}$ |
3. Creation of neuro-fuzzy recommender control system

While creating the recommendation system, the linear dependencies of Table 1 are used as the initial values of the right-hand sides (consequences) of the linguistic inference rules. The left parts (prerequisites) are formed as all kinds of combinations of linguistic values of variables Rair and Rintegral. In total, the rule base thus contains 20 statements, shown in Table 2.

Table 2. A complete set of inference rules for determining action code for control actions.

| If Rair... | And Rintegral | Then Risk measure (action code)... |
|------------|---------------|-----------------------------------|
| Acceptable | Acceptable    | Measure = 786 * Rintegral + 827 * Rair |
| Acceptable | Medium        | Measure = 1234 * Rintegral + 745 * Rair |
| Acceptable | High          | Measure = 1383 * Rintegral + 694 * Rair |
| Acceptable | Unacceptable  | Measure = 1572 * Rintegral + 660 * Rair |
| Acceptable | Catastrophic  | Measure = 1696 * Rintegral + 633 * Rair |
| Medium     | Acceptable    | Measure = 786 * Rintegral + 1240 * Rair |
| Medium     | Medium        | Measure = 1234 * Rintegral + 1117 * Rair |
| Medium     | High          | Measure = 1383 * Rintegral + 1041 * Rair |
| Medium     | Unacceptable  | Measure = 1572 * Rintegral + 990 * Rair |
| Medium     | Catastrophic  | Measure = 1696 * Rintegral + 950 * Rair |
| High       | Acceptable    | Measure = 786 * Rintegral + 1654 * Rair |
| High       | Medium        | Measure = 1234 * Rintegral + 1490 * Rair |
| High       | High          | Measure = 1383 * Rintegral + 1388 * Rair |
| High       | Unacceptable  | Measure = 1572 * Rintegral + 1320 * Rair |
| High       | Catastrophic  | Measure = 1696 * Rintegral + 1266 * Rair |
| High       | Acceptable    | Measure = 786 * Rintegral + 2067 * Rair |
| Unacceptable| Medium       | Measure = 1234 * Rintegral + 1862 * Rair |
| Unacceptable| High         | Measure = 1383 * Rintegral + 1735 * Rair |
| Unacceptable| Unacceptable | Measure = 1572 * Rintegral + 1650 * Rair |
| Unacceptable| Catastrophic | Measure = 1696 * Rintegral + 1582 * Rair |

Membership functions are set in the form of Gaussians with centers and spreads as parameters adjusted in the process of training a neural fuzzy network.

For the linguistic variable Rair, the parameters of the membership functions are transferred from the neuro-fuzzy system for calculating Rintegral after training (Table 3).

Table 3. Characteristics of the membership functions of the linguistic variable Rair.

| Membership function | c   | σ    |
|---------------------|-----|------|
| Acceptable          | 0.000 | 0.0051 |
| Medium              | 0.013 | 0.0028 |
| High                | 0.018 | 0.0016 |
| Unacceptable        | 0.030 | 0.0063 |

To determine the characteristics of the membership functions of the variable Rintegral, the following procedure is applied: Integral risk is associated with five linguistic variables that determine the degree
of hazard: Acceptable”, “Medium”, “High”, “Unacceptable” and “Catastrophic”. The association is made on the basis of ranking and allocation of 4 basic quantiles corresponding to the concepts of “Acceptable”, “Medium”, “High”, “Unacceptable”. Concept of “Catastrophic” is defined as the risk value that exceeds the right border of the last of the selected quantiles. The characteristics of the resulting Gaussians are shown in Table 4.

Table 4. Characteristics of the membership functions of the linguistic variable $R_{\text{integral}}$

| Membership function | $c$     | $\sigma$  |
|---------------------|---------|-----------|
| Acceptable          | 0.010556| 0.005364  |
| Medium              | 0.024708| 0.005137  |
| High                | 0.031677| 0.005318  |
| Unacceptable        | 0.038111| 0.004333  |
| Catastrophic        | 0.049072| 0.004355  |

4. Construction of a fuzzy inference system of the TSK type for calculating the action code for the control action

Based on the found characteristics, a TSK system was built in the MatLab Fuzzy Tool Box [9]. Figure 1 shows a graphical representation of the TSK fuzzy system for determining the environmental protection code based on the air risk levels $R_{\text{air}}$ and the integral risk $R_{\text{integral}}$.

![Figure 1. TSK fuzzy system for determining action code for control action.](image)

This system is capable of making recommendations by calculating the code of the required control action (example for a pair of input parameters $R_{\text{air}} = 0.0151$ and $R_{\text{integral}} = 0.0292$) – Figure 2. However, the calculation accuracy can be increased by retraining (adjusting) the system parameters.
5. Building a neuro-fuzzy recommender system

To increase the adequacy of the control system developed by the neural net, it was presented and retrained in the form of a neural fuzzy net. The neuro-fuzzy recommender system was created in the MatLab ANFIS package.

A dataset of 53 tuples was used for training. A hybrid learning algorithm was applied for seven eras. The resulting error on the training set was 5% (0.51 in absolute values).

The training efficiency was tested on a test sample (13 data tuples). The error on the test set was about 15%.

The neuro-fuzzy recommender system is a fuzzy neural network of the ANFIS type (adaptive network-based fuzzy inference system) – Figure 3.
After training, the recommendation system began to give softer recommendations. So, for the considered example (Figure 2), the system BEFORE training recommended using Activity 41. After training, for the same input data (example for a pair of input parameters $R_{\text{air}} = 0.0151$ and $R_{\text{integral}} = 0.0292$), the system recommends using Activity 35 (Figure 4), which according to the gradation of measures corresponds to actions of a local nature.

![Figure 4. Calculation of the action code (system AFTER training).](image)

### 6. Interpretation of responses by the recommender neuro-fuzzy control system

For the practical application of the recommender system built by the neuro-fuzzy, it is necessary to be able to interpret the answers calculated by the system (the code of the required environmental protection measure) into the name of the action itself.

The absence of the initial binding of codes to actions makes it easy to modify and build up the system, changing the interpretation of the codes. In addition, this approach allows you to localize the system, adjusting the interpretation for each specific territory according to its own algorithm. In this case, in different territories, the same code can mean different measures reflecting the peculiarities of the natural and anthropogenic conditions of a particular territory.

An example of constructing the correspondence of the measures taken to the calculated value of the Risk Measure is presented in Table 5.

![Table 5. Examples of correspondence of actions codes to specific control actions.](image)

| Action code | Examples of a list of environmental protection measures |
|-------------|-------------------------------------------------------|
| 1-13        | Environmental monitoring in the zones of action of mobile sources of pollution |
| 14-26       | Environmental monitoring in the zones of operation of stationary emission sources  
Increase in the area of green spaces for general use at the expense of urban forests and forest parks. |
| 27-39       | Environmental monitoring in the zones of operation of stationary and mobile sources of emissions.  
Increase in the area of green spaces for general use at the expense of urban forests and forest parks. |
Formation of green sanitary protection zones between residential areas and industrial enterprises, taking into account data on the levels of integral risk.

40-52 Environmental monitoring in the areas of stationary and mobile emission sources and residential areas.
Increase in the area of green spaces for general use at the expense of urban forests and forest parks.
Formation of green sanitary protection zones between residential areas and industrial enterprises, taking into account data on the levels of integral risk.
Construction of district highways, highways to reduce traffic flows within the city – reducing the traffic load.

>52 Environmental monitoring in the areas of operation of stationary and mobile sources of emissions and areas of residence of the population, as well as residential areas.
Increase in the area of green spaces for general use at the expense of urban forests and forest parks.
Formation of green sanitary protection zones between residential areas and industrial enterprises, taking into account data on the levels of integral risk.
Construction of district highways, highways to reduce traffic flows within the city – reducing the traffic load.
Removal of enterprises with harmful and hazardous industries from the zone – exclusion of industrial production.

As can be seen from the presented table, the list of environmental protection measures increases, including all the previous items, that is, if the risk measure increases, then it is obvious that it includes control actions for lower order risk measures. Block measure ranges can be subdivided into smaller ones, which provides flexibility in the proposed approach.

7. Conclusion
The proposed algorithm for calculating a qualitative measure of the level of risk to the environment makes it possible to build automatic fuzzy control systems of environmental safety. Based on the Risk Measure, such systems automatically recommend a set of environmental protection measures that most adequately meet the current situation.

In this case, each value of the Risk Measure can be associated with both a single measure and a set of measures. A set of such events can be selected individually for any territory. This allows for greater flexibility in the use of the system and significantly expands the range of recommended measures that can compensate for emerging risks.

Funding
The research was carried out with the financial support of the Ministry of education and science of the Russian Federation in the framework of fulfilling obligations under Agreement number 075-03-2020-051/3 dated 09.06.2020 (topic number fzu-2020-0021).

References
[1] Federal Law of January 10, 2002 N 7-FL “On Environmental Protection” (in Russian)
[2] Sornette D, Maillart T and Kroger W 2013 Exploring the limits of safety analysis in complex technological systems Journal of Risk Reduction 6 59-66
[3] Tunakova Y, Novikova S, Shagidullin A, Valiev V and Morais A 2019 Methodology for Assessing the Level of Territorial Environmental Risk for the Planned Management of Environmental Safety of the Urban Environment Ecology and Industry of Russia 23(10) 44-
49 (in Russian)

[4] Janssen C 2012 Putting ecological realism in environmental risk assessment *Evolution or Revolution – Research priorities for future risk assessment* 14-15

[5] Tunakova Y A, Novikova S V and Shagidullin A R 2017 Calculation of the probability integral indicator of the level of air pollution *Environmental Engineering* 7

[6] Tunakova Y, Novikova S, Shagidullin A and Shagidullin R 2017 Model development for the determination of PM2.5 and PM10 fine particles in ground layer of atmospheric air *Ecology and Industry of Russia* 21(12) 38-43 (in Russian)

[7] Dominici F, McDermott A, Zeger S L and Samet J M 2002 On the use of generalized additive models in time-series studies of air pollution and health *American Journal of Epidemiology* 156(3) 193-203

[8] Zadeh L A 1965 Fuzzy sets *Information and Control* 8(3) 338-53

[9] Shtovba S D 2007 Designing fuzzy systems using MatLab (Telecom) 288 (in Russian)