UDC 004.001; 007; 658.6.002; 621.397.13
DOI: 10.15587/1729-4061.2024.310326

SYNTHESIS OF EXPERT MATRICES IN INDUCTIVE SYSTEM-ANALYTICAL RESEARCH BASED ON FUZZY LOGIC ALGORITHM

Volodymyr Osypenko
Doctor of Technical Sciences, Professor*
Hanna Korohod
Corresponding author
PhD, Associate Professor*
E-mail: 5618607@ukr.net
Borys Zlotenko
Doctor of Technical Sciences, Professor
Dean of Faculty of Mechatronics and Computer Technologies
Department of Computer Engineering and Electromechanics**
Nataliia Chuprynka
PhD, Associate Professor, Head of Department*
Volodymyr Yakhno
PhD, Senior Lecturer*
*Department of Computer Science**
**Kyiv National University of Technologies and Design
Mala Shyianovska (Nemyrovycha-Danchenka) str., 2, Kyiv, Ukraine, 01011

The object of research is the process of inductive modeling of complex systems. The studies that were conducted related to the application of algorithms of the fuzzy logic theory were to coordinate the conclusions of top-level experts in system information-analytical research (SIAR) in the tasks of innovative design. The possibilities of constructing elements of expert matrices of results, as well as evaluating the effectiveness of such applications, are defined. Thanks to this, obtaining formal expert evaluations in numerical form became possible. Experimental studies have confirmed that the proposed approach to the application of fuzzy logic algorithms to the construction of matrices of expert evaluations of SIAR results is quite effective and simple to implement. In addition, this approach fits well into the general paradigm of the Group Method of Data Handling (GMDH). In particular, it was established that the possibility of «retraining» such a block without significant efforts of professional experts can have a positive result, as well as have a good effect on the economic and time parameters of the research project. The main calculation formulas for the algorithm for building a fuzzy system using a neural network in a system with two rules are given. The construction of a fuzzy information output system trained on expert evaluations in the Matlab system is shown. As a result, a technologically acceptable standard deviation of 0.28268 mg/l was obtained. It has been established that by accumulating a database (knowledge) and/or using an information monitoring system, it is possible to «additionally train» a fuzzy system periodically or according to the established quality criterion in the program mode, without involving experts in this process. Thus, there are reasons to assert the importance of using a fuzzy system as one of the tools in inductive SIAR procedures

Keywords: inductive approach, fuzzy logic, criterion of relevance, criterion of corelevance, expert evaluations

Received date 05.06.2024
Accepted date 13.08.2024
Published date 30.08.2024

***How to Cite:** Osypenko, V., Korohod, H., Zlotenko, B., Chuprynka, N., Yakhno, V. (2024). Synthesis of expert matrices in inductive system-analytical research based on fuzzy logic algorithm. Eastern-European Journal of Enterprise Technologies, 4 (4 (130)), 54–62. https://doi.org/10.15587/1729-4061.2024.310326*

1. Introduction

The apparatus of inductive modeling of complex systems (IMSS) is quite well known among specialists in mathematical modeling in a wide range of applications. This is facilitated by its rather developed algorithmic base in the part known as the inductive methodology of the Group Method of Data Handling (GMDH) [1, 2].

One of the latest areas of application of IMCS principles was the inductive technology of system information-analytical research (IT SIAR) [3, 4]. Systematic information and analytical research in the innovative or reengineering direction is intended for the purposeful search for new thorough solutions in a certain subject area during the development of financially intensive investment and innovation projects. Examples of such SIARs are, in particular, the development of complex projects for the strategic development of a certain industry, which require large capital investments, complex systems engineering projects in large companies, the implementation of large-scale socio-economic studies, etc. Application of inductive GMDH principles in SIAR provides for the construction and use of special criteria that have the properties of external addition [1, 2]. from separate independent results of a lower order, which are called elementary. Each elementary result is expressed in the form of a certain number, as an agreed assessment of the members of the expert commission. Such a system of external criteria includes, first of all, the criterion of relevance, the criterion of correlation, and the criterion of balance of information requests. Such criteria are already known from relevant scientific sources, for example [3, 4]. In such studies, general and intermediate research results are presented in the form of certain special matrices, the elements of which are estimates of elementary results, presented by experts who are part of the expert commission. Such commissions are at the highest level in the decision-making process and are defined as expert commission of the highest level (ECHL). Since at SIAR, such estimates are obtained with the help of experts, appropriate mathematical methods and algorithms are used to construct them. Such methods include mathematical statistics, special methods of so-called non-numerical statistics, qualitative assessment, etc.

Given the need for continuous improvement of system information and analytical research technologies, research aimed at applying new approaches to the assessment of intermediate and final results developed by groups of researchers can be considered relevant.

2. Literature review and problem statement

In system information-analytical research the assessments of intermediate and final results (conclusions) developed by ECHL experts are accompanied by agreement according to accepted criteria. The most used criteria (estimates, statistics, etc.) are the well-known mathematical statistics tools given in the literature [5, 6].

Among such tools, it is worth highlighting the following:

– average value;
– weighted arithmetic mean value;
– mode;
– median;
– root mean square deviation;
– Spearman's criterion;
– Kendel's concordance criterion and others.

Options for solving problems with the participation of high-level experts in evaluating the results of complex multi-aspect system-analytical research are the use of the Delphi method [7], the method of paired comparisons [8, 9], and many others. A significant number of publications are devoted to the application of the Kemeny median [10–12]. This mathematical apparatus of the so-called «non-numerical» statistics has proven itself well in the tasks of harmonizing the conclusions of experts. In [10] a certain attempt was made to apply the Kemeny metric in inductive research technologies of complex systems of various natures. In this work, the Kemeny median acted as the value of one of the elements of the matrix of the reference result of the form (E) and the matrices of intermediate results (W^A) and (W^B). It should be noted that such a metric is quite effective for problems in which the specified matrices have a small dimension. When the size of such matrices increases, the effectiveness of using the Kemeny median may tend to decrease. The paper [11] describes some problems in choosing a group judgment algorithm for expert opinions presented in the form of orders of preference. At the same time, special attention is paid to the Kemeny median and its modifications. The use of a heuristic approach to calculating the Kemeny median is special. To solve the problem of the dependence of aggregate results on contextual assessment in the aggregation of ratings, a unified structure is proposed in [12]. The presented framework includes various generalizations of Kemeny estimator for incomplete rankings with ties, and an algorithmic approach is presented that includes exact methods and heuristic algorithms. Algorithms are based on the theory of fuzzy logic, the foundations of which are laid in primary sources [13, 14]. A large number of researchers' works are devoted to the application of such a theory, for example [15–19] and many others. In [15], the basics of the theory of fuzzy sets and some areas of computational intelligence are outlined. It is known that rule-based interpretation of a fuzzy system is one of its advantages. However, for high-dimensional data, this can also be a problem, and to solve it, a certain non-iterative integrated learning mechanism based on the Takagi-Sugeno algorithm is proposed in [16]. From the point of view of practical applications of IT SIAR, the works [17, 18] are interesting. This means that in the early stages of the development process, a huge amount of information is generated, much of which is subjective and remains unused. These articles present a formal structure for the collection of primary information and its use in the assessment of a significant increase in the credibility of expert opinions.

In [19], a new interesting approach to the approximation of any known function using the fuzzy Takagi-Sugeno-Kanga system with a guaranteed upper bound on the approximation error is proposed. In [20], the authors proposed a fuzzy association using classification methods in high-dimensional problems to obtain an accurate and compact fuzzy classifier rule with low computational cost. From the point of view of applications in IT SIAR, the work is also interesting [21]. This paper examines the application of fuzzy logic to support decision-making in ecosystem management. In particular, the identification, optimization, and uncertainty aspects of fuzzy rule-based models for decision support in ecosystem management are discussed.

Despite the good results of applications of the fuzzy logic toolkit in various applied fields, this toolkit has not been widely used in system-analytical search projects in general and in the procedures for coordinating experts' conclusions.

Hence, research aimed at applying the theory of fuzzy logic in SIAR is appropriate.

3. The aim and objectives of the study

The aim of the study is to obtain formal evaluations in numerical expression for the synthesis of expert matrices in inductive system-analytical studies using a fuzzy logic algorithm. This will make it possible to build criteria for selecting the best results in the processes of performing complex system-analytical research.

To achieve the aim, it is necessary to solve the following objectives:

– to develop a procedure for applying an algorithm based on systems of fuzzy logical inference to build an expert matrix element in SIAR;
– to create an expert matrix element based on fuzzy neural networks;
– to conduct a computer simulation of a fuzzy logical inference system based on the example of determining the initial concentration of suspended substances in particle water.

4. Research materials and methods

4. 1. Object and research hypothesis

The object of research is the process of inductive modeling of complex systems of various application areas.

The hypothesis of the research is the possibility of applying the mathematical apparatus of fuzzy logic for the construction of elements of the matrices of results W^A and W^B, as well as the evaluation of the beneficial or ineffective sides of such application in inductive SIAR.

Some simplifications are adopted in the work. First of all, this concerns the use of the term and essence of «expert matrix», which in this context reflects the adequacy of the achieved result to the main requirements of the project customer. However, in real design, the elements of such a matrix can be further refined if necessary, for which additional time and funds must be allocated. That is, here a certain error is allowed in the estimation of design parameters (elements of matrix E). The same applies to the results obtained by two independent groups and recorded in the W^A and W^B matrices.

4. 2. Subject and tools of modeling

During the study, the software environment MatLab 2016 (USA) for Windows, Linex I OS X operating systems was chosen as a software tool for mathematical modeling and computer analysis of the obtained results.

The demonstration of the proposed algorithm was carried out on the example of the synthesis of the ANFIS expert system for determining the initial concentration of suspended substances in the water of particles (C, mg/l) after the electrocoagulator during the purification of wastewater of a poultry farm.

4. 3. The method of applying the SIAR inductive technology

4. 3. 1. SIAR inductive technologies

In the inductive technology of SIAR [3, 4], the concept of the so-called reference result $R^0\left(I_b^0\right)$ is introduced, as a formalized and agreed conclusion of some higher-level expert commission regarding each of its elements e_{ij} formed in the form of a matrix $E\left(R^0\left(I_b^0\right)\right)$ dimension of $n{\times}m$. This matrix $E\left(R^0\left(I_b^0\right)\right)$ characterizes the prototype of some specific element of the final result and has the following form:

$$E\left(R^0\left(I_b^0\right)\right)=\left(e_{ij}\right)=\begin{pmatrix} e_{11} & \cdots & e_{1n} \\ \vdots & \ddots & \vdots \\ e_{m1} & \cdots & e_{mn} \end{pmatrix}. \qquad (1)$$

where the i-th line, $i=1,..., m$, reflects one of the undeniable types of requirements for the final result (but not the result itself) of the study from the agreed positions of independent experts; j-th column is the justified subdivisions of the i-th element.

In the creation of such matrices in SIAR inductive technologies, the expert commission is directly involved and is a structural external independent subject of the project, the main task of which is, in addition to creating a benchmark [3, 4], the researching the performance of analysts as well. ECHL consists of independent experts who have successfully passed a special competitive selection. It is worth noting that the selection of members of the expert commission is also determined in a special way, using pattern recognition algorithms with the possibility of building ensembles of informative features, for example [22].

Therefore, such a matrix $E^0=E\left(R^0\left(I_b^0\right)\right)$ reflects only the reference form of the final or customer-acceptable result $R^*\left(I_b^*\right)$ and the corresponding final document $D\left(R^*\left(I_b^*\right)\right)$. The elements of such a matrix are formalized and in a certain way agreed by experts' qualitative assessments of its components (sections, subsections, items, paragraphs, appendices, etc.). In addition, the objective interpretation of matrix (1) is generally agreed with the customer of the study and must satisfy its unquestionable requirements. This means that $R^0\left(I_b^0\right)$ is the formal reference prototype of the result for the final document, that can solve the problem in the best way for the customer, and $R^*\left(I_b^*\right)$ is the actual result close to the reference $R^0\left(I_b^0\right)$. The importance of this coordination of the formal presentation with the customer's position and requirements is obvious and most modern researchers pay attention to it.

4. 3. 2. Criteria for the selection of solutions in inductive technology of SIAR

The inductive approach to the execution of SIAR involves the gradual expansion of the initial information base I_b^s, which is at the disposal of analysts at the beginning of the work, through the creation of intermediate results $R_k\left(I_b^s\right)$ and the receipt of additional blocks I_{bs}^+ from the external environment of information to the level I_b^0 and is a multistage procedure $n{\times}m$. It should be noted that in this case $s=1,..., S$ – stages and $k=1,..., K$ – the number of options for intermediate results at the s-th step). Therefore, the results of $R_k\left(I_b^s\right)$ must be evaluated for compliance with their etalon benchmark $R^0\left(I_b^0\right)$ for which the experts from ECHL have already created a matrix $E^0=E\left(R^0\left(I_b^0\right)\right)$.

Hence the need to evaluate and formalize intermediate results so that such independent formalized evaluations can perform the role of commensurate objects of comparison in evaluation criteria. For this, another special object is introduced, which in the work received the name «matrix of intermediate results» $W\left(R_{sk}\left(I_{sb}^j\right)\right)^{A/B}$. Such a matrix, which is formed for each of the two groups of analysts A and B, has the same dimension ($n{\times}m$), as well as the standard matrix $E^0=E\left(R^0\left(I_b^0\right)\right)$, and also has a form similar to it:

$$W\left(R_k\left(I_b^j\right)\right)^A=\begin{pmatrix} w_{11}^A & \cdots & w_{1n}^A \\ \vdots & \ddots & \vdots \\ w_{m1}^A & \cdots & w_{mn}^A \end{pmatrix}; \qquad (2)$$

$$W\left(R_k\left(I_b^j\right)\right)^B=\begin{pmatrix} w_{11}^B & \cdots & w_{1n}^B \\ \vdots & \ddots & \vdots \\ w_{m1}^B & \cdots & w_{mn}^B \end{pmatrix}. \qquad (3)$$

The elements w_{ij} are formed repeatedly according to the same procedures during the execution of IT SIAR, as the elements e_{ij} are formed once.

Therefore, in contrast to IMCS, in the selection criteria of IT SIAR, the objects of comparison are no longer individual numbers (for example, model outputs), but the specified matrices – the matrix of the standard $E^0=E\left(R^0\left(I_b^0\right)\right)$ and matrices of intermediate results $W\left(R_{sk}\left(I_{sb}^j\right)\right)^A$ and $W\left(R_{sk}\left(I_{sb}^j\right)\right)^B$. Moreover, the elements of the matrices of intermediate results are in a certain way formalized assessment of ECHL experts.

Since in IT SIAR, it is necessary to evaluate the conformity of one object to another, namely the conformity of the generated intermediate or final research results with the reference one, the evaluation criteria are naturally based on the concept of relevance. Let's introduce the Δ_{rel}^2 and Δ_{corel}^2 matrices of differences of intermediate results $W\left(R_{sk}\left(I_{sb}^s\right)\right)^A$, $W\left(R_{sk}\left(I_{sb}^s\right)\right)^B$ among themselves and with the reference $E^0=E\left(R^0\left(I_b^0\right)\right)$. These matrices are as follows:

$$\Delta_{rel}^2=\left(\delta_{ij}^2\right)_{WE}=\begin{pmatrix} \delta_{11}^{2(we)} & \cdots & \delta_{1n}^{2(we)} \\ \vdots & \ddots & \vdots \\ \delta_{1m}^{2(we)} & \cdots & \delta_{mn}^{2(we)} \end{pmatrix}, \qquad (4)$$

$$i=1,...,m,\ j=1,...,n.$$

In equation (4), the elements $\left(\delta_{ij}^2\right)_{WE}$ are equal to the squares of the differences between the elements of the matrices $W\left(R_{sk}\left(I_{sb}^s\right)\right)^A$ and $W\left(R_{sk}\left(I_{sb}^s\right)\right)^B$ respectively, and elements of the reference matrix $E\left(R^*\left(I_b^*\right)\right)$:

$$\left(\delta_{ij}^2\right)_{W^A E} = \left(w_{ij}^A - e_{ij}\right)^2, \qquad (5)$$

$$\left(\delta_{ij}^2\right)_{W^B E} = \left(w_{ij}^B - e_{ij}\right)^2. \qquad (6)$$

The relevance of SR_{rel} in IT SIAR is called a measure of the form:

$$SR_{rel}\left(W^A, E^0\right)_{sk} = \left\| W\left(R_{sk}\left(I_b^s\right)\right)^A - E\left(R^0\left(I_b^0\right)\right)\right\|, \qquad (7)$$

$$s = 1,\dots,S,\ k = 1,\dots,K,$$

for the results of group A and

$$SR_{rel}\left(W^B, E^0\right)_{sk} = \left\| W\left(R_{sk}\left(I_b^s\right)\right)^B - E\left(R^0\left(I_b^0\right)\right)\right\|, \qquad (8)$$

$$s = 1,\dots,S,\ k = 1,\dots,K.$$

The system criterion of relevance in IT SIAR is called expressions:

$$CR_{rel}^A\left(W^A, E^0\right)_{sk} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} \delta_{ij}^2\left(W^A, E^0\right)_{R_{sk}}}, \qquad (9)$$

$$CR_{rel}^A\left(W^A, E^0\right)_{sk} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} \delta_{ij}^2\left(W^A, E^0\right)_{R_{sk}}} \qquad (10)$$

or the results of groups A and B, respectively, where for the sake of simplification $W\left(R_{sk}\left(I_{sb}^s\right)\right)^A$ and $W\left(R_{sk}\left(I_{sb}^s\right)\right)^B$, $s = 1,\dots,S,\ k = 1,\dots,K$.

Thus, the expression of the system criterion of relevance in inductive SIAR has a simple and clear meaning. Here the matrices $W\left(R_{sk}\left(I_{sb}^s\right)\right)^A$ and $W\left(R_{sk}\left(I_{sb}^s\right)\right)^B$ can be formed both for any intermediate and final results $R_{sk}\left(I_{sb}^s\right)$ depending on the result number k at the s-th step, and the index of the analytical group (A or B), etc. At the same time, the matrix is unchanged $E^0 = E\left(R^0\left(I_b^0\right)\right)$ throughout SIAR.

The application of the principle of independence of research by two groups and external addition in IT SIAR [3, 4] entails the need to construct one more criterion – an analogue of the criterion of consistency from the IMCS methodology for comparing intermediate results obtained by two independent research groups A and B. For this, matrices $W\left(R_{sk}\left(I_{sb}^s\right)\right)^A$ are formed for intermediate results obtained by the group of analysts A, and matrices $W\left(R_{sk}\left(I_{sb}^s\right)\right)^B$ for results, obtained by a group of analysts B. Procedures for constructing matrices $W\left(R_{sk}\left(I_{sb}^s\right)\right)^A$ and $W\left(R_{sk}\left(I_{sb}^s\right)\right)^B$ for intermediate results $R_{sk}\left(I_{sb}^s\right)^A$ and $R_{sk}\left(I_{sb}^s\right)^B$ respectively, must be the same according to which the and the matrix $E^0 = E\left(R^0\left(I_b^0\right)\right)$.

In IMCS, the criterion of non-contradiction requires that models of the same structure, the coefficients of which are calculated on parts A and B on a given information base, would give the most similar initial results, that is, that the initial values of such models should have a minimum discrepancy.

It is possible to use this property of this criterion in application to IT SIAR and introduce a special new concept for this correlation.

To do this, similarly to (4), it is possible to define the matrix Δ_{corel}^2:

$$\Delta_{corel}^2 = \left(\delta_{ij}^2\right)_{WW} = \begin{pmatrix} \delta_{11}^{2(ww)} & \cdots & \delta_{1n}^{2(ww)} \\ \vdots & \ddots & \vdots \\ \delta_{1m}^{2(ww)} & \cdots & \delta_{mn}^{2(ww)} \end{pmatrix}, \qquad (11)$$

$$i = 1,\dots,m,\ j = 1,\dots,n,$$

where $\left(\delta_{ij}^2\right)_{WW}$ – the elements, which are equal to the squares of the differences of the elements of the matrices $W\left(R_{sk}\left(I_{sb}^s\right)\right)^A$ and $W\left(R_{sk}\left(I_{sb}^s\right)\right)^B$, that is, element by element:

$$\left(\delta_{ij}^2\right)_{WW} = \left(w_{ij}^A - w_{ij}^B\right)^2. \qquad (12)$$

Co-relevance of results in the IT SIAR inductive procedure is called a measure of the form:

$$SR_{corel}\left(W^A, W^B\right)_{sk} = \left\| W\left(R_{sk}\left(I_b^s\right)\right)^A - W\left(R_{sk}\left(I_b^s\right)\right)^B\right\|, \qquad (13)$$

$$s = 1,\dots,S,\ k = 1,\dots,K.$$

The system criterion of co-relevance in IT SIAR is the expression:

$$CR_{corel}\left(W^A, W^B\right)_{sk} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} \delta_{ij}^2\left(W^A, W^B\right)_{sk}}, \qquad (14)$$

where, as in (9) and (10), is introduced: the notation $W^A = W\left(R_{sk}\left(I_{sb}^s\right)\right)^A$, $W^B = W\left(R_{sk}\left(I_{sb}^s\right)\right)^B$, $\delta_{ij}^2\left(W^A, W^B\right)_{sk}$ are the δ_{ij}^2 elements of the matrix Δ_{corel}^2 of the discrepancies of the results $W\left(R_{sk}\left(I_b^s\right)\right)^A$ and $W\left(R_{sk}\left(I_b^s\right)\right)^B$, formalized and agreed by the ECHL in the matrices of the discrepancies of the results $W\left(R_{sk}\left(I_b^s\right)\right)^A$ and $W\left(R_{sk}\left(I_b^s\right)\right)^B$, $s = 1,\dots,S,\ k = 1,\dots,K$.

From the standpoint of the theory of fuzzy sets, each row of the matrix E can be interpreted as a linguistic variable, that is, term A, and the constituent parts of this row – as elements a_{ij} from a set of possible values, $a_{ij} \in A$. For example, in the now classic example [23], an expert assessment of the answer to the question of whether water is very hot can have the following form for $n = 10$:

$$E = \left\{\frac{0,0}{10};\frac{0,0}{20};\frac{0,0}{30};\frac{0,1}{40};\frac{0,2}{50};\frac{0,5}{60};\frac{0,8}{70};\frac{0,9}{80};\frac{1,0}{90};\frac{1,0}{100}\right\}. \qquad (15)$$

In equation (15), the denominator is the water temperature (°C), and the numerator is the number of positive evaluations in the sense that it is really very hot water. That is, the last and penultimate element corresponds to the term «boiling water».

Therefore, the prototype of the expert matrix (for example only) should have the form:

$$E = \left\{e_{ij} / v_{ij}\right\}, \qquad (16)$$

where v_{ij} – all possible values of estimates of gradations of expert conclusions, which may contain the result for the i-th element.

Let's pay attention to the obvious fact that for all positive research results, the denominators of all elements of one row of matrix E will be the same, and, thus, the line of expert assessments from the above example may look like this:

$$E = \left\{0.0; 0.0; 0.0; 0.1; 0.2; 0.5; 0.8; 0.9; 1.0; 1.0\right\}. \qquad (17)$$

The elements of line (17) are the ordinates of the membership function $g(x)$ [20] (x is the input signal), and the expert matrix will take on a simpler form (1).

5. Synthesis of the expert matrix based on the fuzzy logic algorithm

5. 1. The procedure for creating an unstructured system for building an expert matrix

Below is an algorithm for the creation of a practical non-structured system for constructing an expert matrix for evaluating results, oriented towards the development of one value e_{ij} of the formalized matrix E in the range of values [0, 1].

For the practical implementation of such a fuzzy system, it is possible to look at a number of popular algorithms [13, 14]: Mamdani, Tsukamoto, Sugeno, Larsen, a simplified fuzzy inference algorithm. However, with the assurance of expertly tested software products and the fact that ECHL, when prompted by the basis of the rules of unimportant products, will provide numerical values for estimates, for the prompt-based system, the possibility of using Sugeno algorithms has been adopted o or due to the algorithm of incorrect output.

With a simplified algorithm, the rules of our task will look like this:

$$R_1: \text{if } x \in A_1 \text{ then } z_1 = \alpha_1 c_1,$$
$$\ldots\ldots$$
$$R_N: \text{if } x \in A_N \text{ then } z_N = \alpha_1 c_N,$$

where x, z – input and output changes; A – specification of the term-set (needs to be agreed upon in advance with all experts); c_1, c_2 – expert assessments (pure numbers in the range [0, 1]); N – number of rules (equal to the number of experts multiplied by the number of terms, however, if the experts' assessments of the specified parameter are equal, one rule is left).

The simplified fuzzy inference algorithm will take the form:

1. The first stage is to determine the stages of truth for the implementation of the skin rule.

2. Second stage – additional numbers:

$$\alpha_1 = A_1(x_0),\quad \ldots\ldots,\quad \alpha_N = A_N(x_0). \tag{18}$$

3. The third stage is to be clear about the significance of assessing the characteristics of the object of investigation (process, phenomenon):

$$z_0 = \frac{\sum_{i=1}^{N} \alpha_i C_i}{\sum_{i=1}^{N} \alpha_i}. \tag{19}$$

The rules for the Sugeno algorithm will take the form:

$$R_1: \text{if } x \in A_1 \text{ then } z_1 = \alpha_1 c_1,$$
$$\ldots\ldots$$
$$R_N: \text{if } x \in A_N \text{ then } z_N = \alpha_1 c_N,$$

where α_1, α_N – the power of the truthfulness for specific rules.

Thus, Sugeno algorithm for our main task can be presented in the form of an ongoing procedure:

1. The first stage is like in the simplified algorithm.

2. The second stage – initially, similar to the simplified algorithm, additional numbers α are calculated, after individual rule outputs:

$$z_1^* = \alpha_1 x_0,\quad \ldots\ldots,\quad z_N^* = \alpha_N x_0. \tag{20}$$

3. The third stage is finding a clear way out:

$$z_0 = \frac{\alpha_1 z_1^* + \ldots + \alpha_N z_N^*}{\alpha_1 + \ldots + \alpha_N}. \tag{21}$$

By creating an unvalued system (simplifying algorithm or Sugeno algorithm), it is possible to calculate the residual estimate for one element of the matrix E.

5. 2. The procedure for creating an expert matrix based on fuzzy neural networks

However, one of the main disadvantages of fuzzy inference systems is their inability to self-learn [23–25], and to adjust them, it is necessary to re-involve experts at a complete functional stop. To solve this problem, the ability to self-adapt when changing expert assessments or receiving additional information about a certain element (for example, a paragraph of the source document $D\left(R^*\left(I_b^*\right)\right)$) of the matrix E is necessary.

Therefore, it makes sense to use the apparatus of fuzzy neural networks - neural networks with clear signals, weights and activation functions, but combining them using t-norm, t-conorm, or other operations [26].

In this case, the inputs, outputs and weights of the fuzzy neural network, for example, are real numbers from the necessary range [0, 1]. The fuzzy system training algorithm integrated into applied software products (for example, MatLab) is quite easy to use.

The main calculation formulas for the algorithm for constructing a fuzzy system using a neural network on a system with two rules (the form of the entry corresponds to a simplified algorithm of fuzzy logical inference) have the form:

$$R_1: \text{if } x \in A_1 \text{ then } y = z_1,$$
$$\ldots\ldots$$
$$R_N: \text{if } x \in A_2 \text{ then } y = z_2.$$

The calculations were performed for the case when the fuzzy concepts A_1 and A_2 have sigmoid membership functions:

$$A_1(x) = \frac{1}{1 + e^{b_1(x - a_1)}}, \tag{22}$$

$$A_2(x) = \frac{1}{1 + e^{b_2(x - a_2)}}, \tag{23}$$

where a_1, a_2, b_1 and b_2 – coefficients.

Then the degrees of truth of the rules are determined by the relations:

$$\alpha_1 = A_1(x) = \frac{1}{1 + e^{b_1(x - a_1)}}, \tag{24}$$

$$\alpha_2 = A_2(x) = \frac{1}{1 + e^{b_2(x - a_2)}}, \tag{25}$$

and system output:

$$o = \frac{\alpha_1 z_1 + \alpha_2 z_2}{\alpha_1 + \alpha_2} = \frac{A_1(x) z_1 + A_2(x) z_2}{A_1(x) + A_2(x)}. \tag{26}$$

The error function in general looks like this:

$$\Delta_k = \Delta_k\left(a_1, b_1, a_2, b_2, z_1, z_2\right) = \\ = \frac{1}{2}\left(o^k\left(a_1, b_1, a_2, b_2, z_1, z_2\right) - y^k\right)^2, \quad (27)$$

where $k=1,..., N$.

Using the approaches of the backpropagation algorithm [26], it is possible to obtain the required correction formulas for partial conclusions:

$$z_1 := z_1 - \eta\frac{\partial E_k}{\partial z_1} = z_1 - \eta\left(o^k - y^k\right)\frac{\alpha_1}{\alpha_1+\alpha_2} = \\ = z_1 - \eta\left(o^k - y^k\right)\frac{A_1\left(x^k\right)}{A_1\left(x^k\right)+A_2\left(x^k\right)}, \quad (28)$$

$$z_2 := z_2 - \eta\frac{\partial E_k}{\partial z_2} = z_2 - \eta\left(o^k - y^k\right)\frac{\alpha_2}{\alpha_1+\alpha_2} = \\ = z_2 - \eta\left(o^k - y^k\right)\frac{A_2\left(x^k\right)}{A_1\left(x^k\right)+A_2\left(x^k\right)}, \quad (29)$$

where η – the learning rate constant.

5. 3. Software implementation of the ANFIS expert system

The result of the software implementation of ANFIS (adaptive neuro-fuzzy inference system) in the Matlab system will be the construction of a fuzzy information output system trained on expert assessments (Sugeno algorithm).

As an example, for the verification of the proposed procedure for the formation of matrix elements, the determination of the initial concentration of suspended substances in the water after the electrocoagulator during the purification of wastewater of a poultry farm was taken [27].

This example can be interpreted as the task of obtaining one element $e_{ij}\in E$ of the integral standard matrix E of the reference result $R^0\left(I_b^0\right)$, i.e. considered as a small fragment of a large investment project element. Similar tasks are often encountered in complex information and analytical studies assessing the possible long-term consequences of the influence of agro-industrial processing complexes on the environment and human health. In our terms, the result of such a study will be of an «elementary» (private) nature, i.e. enter into the system expert matrix as an element. But our task is to use this particular example to verify the effectiveness of the procedure proposed above.

The task of the expert system is to calculate the concentration of particles suspended in water (C, mg/l) for possible combinations of selected technological parameters in the established ranges of their values: current density at the electrodes (a) – 0.1–1.2 A/dm²; pH of water supplied for purification (3–7); flow velocity in the interelectrode space (V) – 2–6 m/hour. The controlled parameter of this process is the concentration of suspended particles in the water supplied for cleaning (C), which is 0.5–7 g/l.

Technological parameters and their ranges of change were selected based on the recommendations of experts (technologist, chief power engineer, chief mechanic, instrumentation and automation engineer). Statistical data for electrocoagulation treatment of poultry farm wastewater were obtained experimentally [27] and are presented in Tables 1–3.

In particular, Table 1 presents the results of studies that were included in the Training subset of data, Table 2 shows the data of the Testing data of the general sample, and Table 3 contains the Checking data. The presented statistical data reflect the real situation in the poultry farm in the process of research.

Table 1

Experimental data: *A* – training input data subset

I, A/dm²	pH	*V*, m/h	*C*, g/l	*C*, mg/l
1.2	3	2	0.5	4.54
0.1	7	2	0.5	0.13
0.1	3	6	0.5	13.69
1.2	7	6	0.5	8.18
1.2	3	2	3	25.47
0.1	7	2	3	27.29
0.1	3	6	3	41.86
1.2	7	6	3	44.64
0.65	5	4	1.75	18.97
0.65	7	2	3	27.26
0.1	5	2	0.5	2.78
1.2	5	6	3	43.49
0.1	7	2	1.75	12.11
0.1	5	6	3	43.58
0.65	7	6	0.5	8.72
1.2	3	6	1.75	25.84
1.2	3	4	3	32.88
0.65	3	6	0.5	13.11
1.0	4	3	0.8	7.35
0.7	6	5	2	25.13

Table 2

Experimental data: *B* – testing input data subset

I, A/dm²	pH	*V*, m/h	*C*, g/l	*C*, mg/l
0.2	3.1	2.1	0.6	6.45
0.3	3.2	2.2	0.7	7.11
0.4	3.3	2.3	0.8	7.82
0.5	3.4	2.4	0.9	8.59
0.6	3.5	2.5	1	9.41
0.7	3.6	2.6	1.1	10.29
0.8	3.7	2.7	1.2	11.22
0.9	3.8	2.8	1.3	12.21
1	3.9	2.9	1.4	13.25
1.1	4	3	1.5	14.34
1.2	4.1	3.1	1.6	15.49
1.2	7	6	2.9	42.84
1.1	6.9	5.9	2.9	42.37
1	6.7	5.8	2.8	40.25
0.9	6.8	5.7	2.7	38.24
0.8	6.7	5.6	2.6	36.25
0.7	6.6	5.5	2.5	34.30
0.6	6.5	5.4	2.4	32.42
0.5	6.4	5.3	2.3	30.59
0.4	6.3	5.2	2.2	28.81

Table 3

Experimental data: *C* – checking input data subset

I, A/dm^2	pH	*V*, m/h	*C*, g/l	*C*, mg/l
0.2	7	2	1.75	12.12
0.3	6.9	2.1	3	27.52
0.4	6.8	2.2	1.74	12.59
0.5	6.7	2.3	2	15.81
0.6	6.6	2.4	1.73	13.06
0.7	6.5	2.5	1.78	13.91
0.8	6.4	2.6	1.72	13.53
0.9	6.3	2.7	1.79	14.69
1	6.2	2.8	1.71	14.08
1.1	6.1	2.9	1.8	15.39
1.2	6	3	1.7	14.55
1.2	5.9	3.1	1.81	16.16
1.1	5.8	3.2	1.69	15.20
1	5.7	3.3	1.82	17.08
0.9	5.6	3.4	1.68	15.90
0.8	5.5	3.5	1.83	18.00
0.7	5.6	3.6	2	20.32
0.6	5.4	3.7	1.84	18.88
0.1	5.3	3.8	3	33.73
0.2	5.2	3.9	1.85	19.86

When training an expert system in ANFIS-Editor (MatLab software environment) from a training sample of data (Table 1), it is possible to select the default settings and defined 30 «running» cycles (E_{nox}). As a result, a technologically acceptable standard deviation was obtained 3.9788e^{-005} mg/l (Fig. 1).

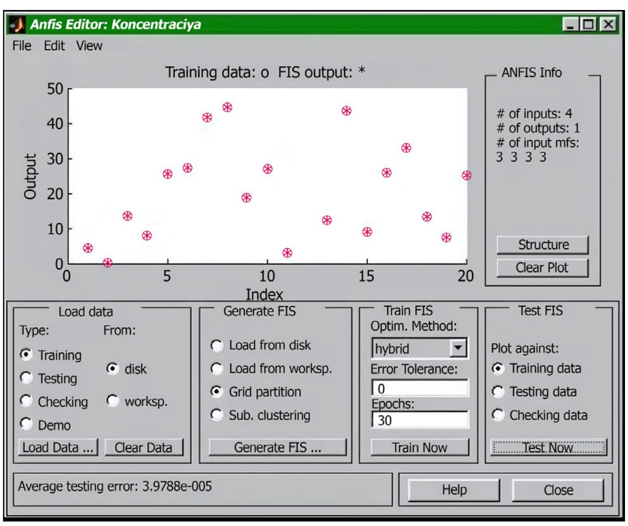

Fig. 1. Setting up ANFIS on training experimental data

But after entering the values of the control sample (Testing data), the quality of work obtained using the default settings turned out to be unsatisfactory – the root mean square error was 12.6569 mg/l (Fig. 2).

In the process of expert adjustment in the structure of the fuzzy neural network, the number of neurons in each ball was reduced by one and the parameters of the membership functions were varied. As a result of 64 stages of iterative learning (Fig. 3), a root-mean-square error of 0.28268 mg/l was obtained. This value, according to the technological aspects of electrocoagulation treatment of poultry farm wastewater from suspended solids, is accepted as satisfactory.

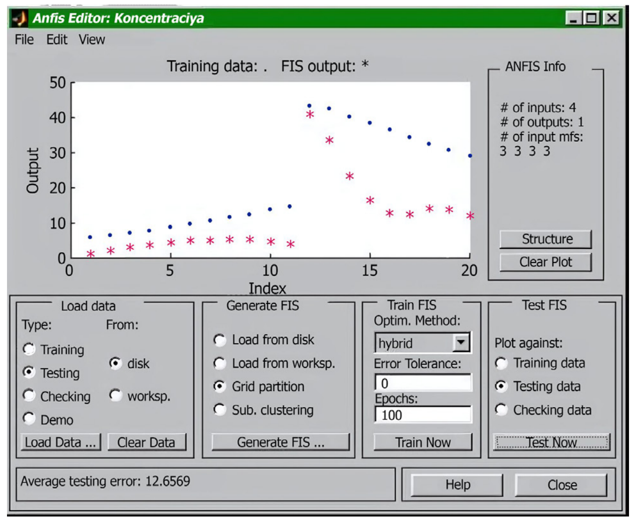

Fig. 2. Setting up ANFIS to control experimental data (default)

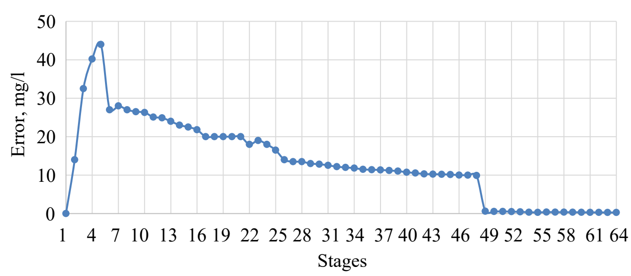

Fig. 3. Results of expert iterative training of ANFIS

The quality of the ANFIS system was finally assessed after the input of checking data. The root mean square error was 0.34339 mg/l. This indicated that the system was not «overtrained» and could be used for further research.

Thus, a software implementation of a fuzzy system for determining the quality of electrocoagulation treatment of poultry farm wastewater from suspended matter was obtained (Fig. 4).

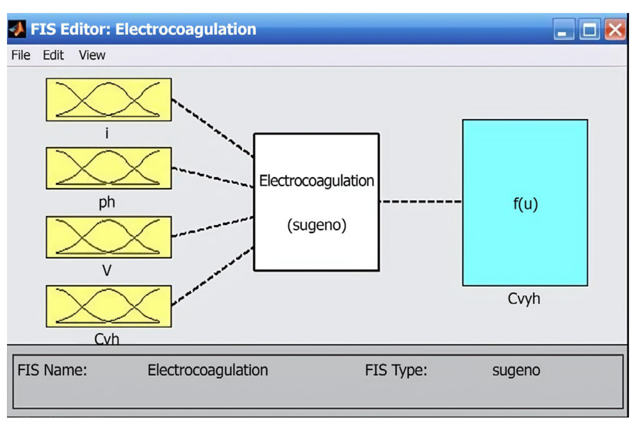

Fig. 4. Software interface of a trained fuzzy system (Sugeno algorithm)

6. Discussion of the results of the synthesis of matrix elements of expert evaluations of results in SIAR

For the synthesis of matrixes of agreement of expert evaluations (elements of such matrices) of results in system information-analytical research, a procedure of application based on a fuzzy system was proposed. The degrees of truth of the rules were determined by relations (24), (25), and the output of the system was determined by (26). At the same time, the error function was generally described by equation (27). This does not differ from the approaches that were used in works [13, 14], where the use of a simplified fuzzy inference algorithm was also assumed.

Thus, using the apparatus of fuzzy neural networks, self-adaptation becomes possible when expert assessments are changed. This is not inconsistent with the practical data presented in [25], which authors also used a combination of neural networks with distinct signals, weights, and activation functions using certain operations. In addition, the ability to «retrain» such a unit without significant efforts by professional experts is quite effective and also has a positive effect on the cost and time of the project. Another advantage of the proposed procedure for applying the algorithm based on fuzzy logical inference systems for constructing an element of the expert matrix in SIAR is the following. Accumulating a database (knowledge) and/or using an information monitoring system [3, 4], it is possible to «retrain» a fuzzy system to obtain one element of the matrices W^A, W^B and E^0 periodically or according to a set of criteria. Moreover, the involvement of experts in this process is not mandatory. It is important to emphasize that the proposed procedure for constructing an element of expert matrices fits well into the general paradigm of inductive modeling of complex systems and the group method of data handling, as one of the powerful tools in the field of artificial intelligence. The synthesis of the optimal result in GMDH takes place according to inductive combinatorial or multi-line procedures of gradually increasing the complexity of the model and step-by-step selection of the best samples according to a special system of criteria without human intervention.

The synthesized fuzzy system according to the proposed procedure was used in the research project on electrocoagulation processes without conducting complex experiments in the design of automated control systems for wastewater treatment complexes. Here, the basis was a fuzzy system, and the neural network approach was used only for its adjustment. The developed procedure for the synthesis of elements of the expert matrix based on fuzzy neural networks was tested in the Matlab system. As a result of 64 stages of iterative training (Fig. 3), a root mean square error of 0.28268 mg/l was obtained. This value is considered acceptable according to the requirements of the technological aspects of electrocoagulation purification of wastewater of poultry farms from suspended substances.

For large-scale projects, the application of the described approach can be somewhat overloaded and less efficient in the form presented above. However, for relatively small projects, using fuzzy logic is effective. Therefore, when forming a rather complex matrix of complex information-analytical research, a certain compilation of systems, similar to the one used in the example of electrocoagulation assessment, will probably be necessary. This is due to the diversity of SIAR, and the need to use the knowledge of specialists from various subject areas.

The impossibility of removing the mentioned shortcomings within the framework of this study gives rise to a potentially interesting direction for further research. They, in particular, can be focused on finding out the optimality of documentation detailing, since it is closely related to the dimension of such leading technological elements as the specified matrices.

7. Conclusions

1. The procedure for applying the algorithm based on systems of fuzzy logical inference for the creation of a practical unstructured system of synthesis of elements of the expert matrix of evaluation of results in inductive technologies of system information-analytical research is given. The peculiarity of the proposed approach is that by accumulating a database (knowledge) and/or using an information monitoring system, it is possible to «additionally train» a fuzzy system periodically or according to a set quality criterion. Moreover, the involvement of experts in this process is not mandatory. Based on the Sugeno algorithm, it became possible to calculate the estimate for one element of the expert matrix.

2. The main calculation formulas for the algorithm for constructing the elements e_{ij}, w_{ij}^A and w_{ij}^B of the matrix of the standard (or benchmark) $E^0 = E\left(R^0\left(I_b^0\right)\right)$ and the matrices of intermediate results $W\left(R_{sk}\left(I_{sb}^j\right)\right)^A$ and $W\left(R_{sk}\left(I_{sb}^j\right)\right)^B$ are given. The use of a neural network apparatus with clear signals, weights and activation functions, as well as their combinations, made it possible to implement the self-learning process without the need to re-engage experts.

3. Using the Matlab application, a system for deriving fuzzy information, trained on expert evaluations (Sugeno algorithm), was built. In the process of experimental adjustment of the structure of the fuzzy neural network, as a result of 64 stages of iterative training, a result with a root mean square error of 0.28268 mg/l was obtained. The obtained results from the point of view of technological aspects of electrocoagulation purification of wastewater of poultry farms from suspended substances are accepted as satisfactory.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this study, including financial, personal, authorship, or any other, that could affect the study and its results presented in this article.

Financing

The study was conducted without financial support.

Data availability

Data can be provided upon reasonable request.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the presented work.

References

1. Ivakhnenko, A. G. (1970). Heuristic self-organization in problems of engineering cybernetics. Automatica, 6 (2), 207–219. https://doi.org/10.1016/0005-1098(70)90092-0
2. Madala, H. R., Ivakhnenko, A. G. (2019). Inductive Learning Algorithms for Complex Systems Modeling. CRC Press. https://doi.org/10.1201/9781351073493
3. Osypenko, V. (2013). Info-logical structure of inductive technologies of the searching system-information-analytical researches. Visnyk Natsionalnoho universytetu «Lvivska politekhnika». Ser. «Komp. nauky ta informatsiyni tekhnolohiyi», 751, 315–319.
4. Osypenko, V. (2012). The Results Estimation in the Integrated System-Analytical Investigations Technologies. Control systems and computers, 1, 26–31. Available at: http://usim.org.ua/arch/2012/1/6.pdf
5. Seno, P. S. (2007). Teoriya ymovirnostei ta matematychna statystyka. Kyiv: Znannia, 556.
6. Wackerly, D., Mendenhall, W., Scheaffer, R. (2007). Mathematical Statistics with Applications. Brooks/Cole.
7. Dalkey, N., Helmer, O. (1963). An Experimental Application of the DELPHI Method to the Use of Experts. Management Science, 9 (3), 458–467. https://doi.org/10.1287/mnsc.9.3.458
8. Saaty, T. L. (2008). Relative measurement and its generalization in decision making why pairwise comparisons are central in mathematics for the measurement of intangible factors the analytic hierarchy/network process. Revista de La Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas, 102 (2), 251–318. https://doi.org/10.1007/bf03191825
9. Andrieu, P., Cohen-Boulakia, S., Couceiro, M., Denise, A., Pierrot, A. (2023). A unifying rank aggregation framework to suitably and efficiently aggregate any kind of rankings. International Journal of Approximate Reasoning, 162, 109035. https://doi.org/10.1016/j.ijar.2023.109035
10. Osypenko, V. (2011). Syntez ekspertnoi matrytsi za metrykoiu Kemeni v induktyvnykh tekhnolohiyakh informatsiyno-analitychnykh doslidzhen. Naukovyi visnyk NUBiP Ukrainy: Seriya «Enerhetyka i avtomatyzatsiya v APK», 166 (3), 119–127.
11. Bury, H., Wagner, D. (2003). Application of Kemeny's Median for Group Decision Support. Applied Decision Support with Soft Computing, 235–262. https://doi.org/10.1007/978-3-540-37008-6_10
12. Davenport, A., Kalagnanam, J. (2004). A Computational Study of the Kemeny Rule for Preference Aggregation. AAAI'04, 697–702. Available at: https://cdn.aaai.org/AAAI/2004/AAAI04-110.pdf
13. Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8 (3), 338–353. https://doi.org/10.1016/s0019-9958(65)90241-x
14. Zadeh, L. A. (1994). Fuzzy logic, neural networks, and soft computing. Communications of the ACM, 37 (3), 77–84. https://doi.org/10.1145/175247.175255
15. Bede, B. (2013). Mathematics of Fuzzy Sets and Fuzzy Logic. In Studies in Fuzziness and Soft Computing. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35221-8
16. Pal, N. R., Saha, S. (2008). Simultaneous Structure Identification and Fuzzy Rule Generation for Takagi-Sugeno Models. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 38 (6), 1626–1638. https://doi.org/10.1109/tsmcb.2008.2006367
17. Yadav, O. P., Singh, N., Chinnam, R. B., Goel, P. S. (2003). A fuzzy logic based approach to reliability improvement estimation during product development. Reliability Engineering & System Safety, 80 (1), 63–74. https://doi.org/10.1016/s0951-8320(02)00268-5
18. Mehmanpazir, F., Asadi, S. (2016). Development of an evolutionary fuzzy expert system for estimating future behavior of stock price. Journal of Industrial Engineering International, 13 (1), 29–46. https://doi.org/10.1007/s40092-016-0165-7
19. Sonbol, A. H., Fadali, M. S., Jafarzadeh, S. (2012). TSK Fuzzy Function Approximators: Design and Accuracy Analysis. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 42 (3), 702–712. https://doi.org/10.1109/tsmcb.2011.2174151
20. Alcala-Fdez, J., Alcala, R., Herrera, F. (2011). A Fuzzy Association Rule-Based Classification Model for High-Dimensional Problems With Genetic Rule Selection and Lateral Tuning. IEEE Transactions on Fuzzy Systems, 19 (5), 857–872. https://doi.org/10.1109/tfuzz.2011.2147794
21. Adriaenssens, V., Baets, B. D., Goethals, P. L. M., Pauw, N. D. (2004). Fuzzy rule-based models for decision support in ecosystem management. Science of The Total Environment, 319 (1-3), 1–12. https://doi.org/10.1016/s0048-9697(03)00433-9
22. Osypenko, V. V. (2014). Dva pidkhody do rozviazannia zadachi klasteryzatsiyi u shyrokomu sensi z pozytsiy induktyvnoho modeliuvannia. Energy and Automation, 1, 83–97. Available at: https://journals.nubip.edu.ua/index.php/Energiya/article/view/3433
23. Ross, T. J. (2010). Fuzzy Logic with Engineering Applications. Wiley. https://doi.org/10.1002/9781119994374
24. Passino, K. M., Yurkovich, S. (1997). Fuzzy Control. Addison-Wesley.
25. Shi, Y., Mizumoto, M. (2000). A new approach of neuro-fuzzy learning algorithm for tuning fuzzy rules. Fuzzy Sets and Systems, 112 (1), 99–116. https://doi.org/10.1016/s0165-0114(98)00238-3
26. Osypenko, V. V., Shtepa, V. N. (2010). Alhorytmy syntezu ekspertnoi matrytsi informatsiyno-analitychnykh doslidzhen na osnovi fazilohiky. Systemni tekhnolohiyi, 6 (71), 154–165. Available at: https://journals.nmetau.edu.ua/index.php/st/issue/view/76/51
27. Shtepa, V. N., Donchenko, M. I., Sribnaya, O. G. (2007). Ochistka rastvorov ot dispersnyh primesey metodom elektrokoagulyatsii. 1. Elektrohimicheskoe poluchenie koagulyanta. Vistnyk NTU «KhPI». Khimiya, khimichna tekhnolohiya ta ekolohiya, 9, 86–94.