DOI 10.1007/s10559-021-00385-3
Cybernetics and Systems Analysis, Vol. 57, No. 4, July, 2021

FUZZY COGNITIVE MAP VS REGRESSION

A. P. Rotshtein[1] and D. I. Katielnikov[2] UDC 681.5.015:007

Abstract. *A fuzzy cognitive map is considered as an alternative to regression analysis, i.e., tools for modeling the inputs-output dependence based on expert-experimental information. To calculate the output value at the given input values, increments of variables are used. The optimal values of the weights of the arcs are found using the genetic algorithm in which the chromosomes are generated from the intervals of their feasible values and the selection criterion is the sum of the squared deviations between the model and the observed output values.*

Keywords: *fuzzy cognitive map, regression, approximation, unknown parameters, tuning, genetic algorithm.*

INTRODUCTION

Regression analysis is one of the most common methods of empirical [1] (inductive as defined by A. G. Ivakhnenko) modeling aimed at extracting patterns from observations. The regression equation models the dependence of a value (output or effect) on influencing factors (inputs or causes). To obtain the regression equation, we need:

— to define output and input variables;

— to present experimental data in the form of an inputs-output table;

— to choose an inputs-output dependence model with unknown parameters;

— to find the parameter values that minimize the sum of the squares of the deviations between the calculated and experimental values of the output variable.

The limitations of classical regression analysis [1] are as follows:

— the quantitative nature of the input and output variables is assumed. The method is not adapted to the direct processing of expert linguistic statements that play an important role in modern control systems [2];

— to approximate the "inputs-output" dependence, a polynomial expansion of a multidimensional function in Taylor series is used, which allows, using a change of variables, to reduce the problem of finding unknown parameters to solving a system of linear equations. However, accounting the nonlinear part of the polynomial with an increase in the number of inputs leads to such a cumbersome model that is impossible to use;

— unknown parameters of the regression equation allow meaningful interpretation only for the linear part of the polynomial. Difficulties in interpretation of parameters for nonlinear terms of the polynomial make it impossible to estimate these parameters expertly, i.e., without tedious experiments.

The emergence of fuzzy logic [3] stimulated the development of methods of empirical modeling based on the processing of natural language expressions. Fuzzy rules "if – then" and fuzzy relations can be used to approximate nonlinear dependences [3]. The application of fuzzy rules is described in [4, 5]. The construction of nonlinear dependences based on the integrated use of fuzzy logic, genetic algorithms, and neural networks was first studied in [6, 7]. The modeling elements here are fuzzy terms (low, high, etc.) included in the rules "if – then", and the parameters of the

[1]Jerusalem College of Technology, Machon Lev, Israel, and Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine, *rothstei@g.jct.ac.il*. [2]Vinnytsia National Technical University, Vinnytsia, Ukraine, *fuzzy2dik@gmail.com*.
Translated from Kibernetyka ta Systemnyi Analiz, No. 4, July–August, 2021, pp. 118–130. Original article submitted May 27, 2020.

membership functions of fuzzy terms and the weights of the rules are subject to adjustment. The principal advantage of the fuzzy rules over regression models is the ability to construct pure inputs-output relationships where insufficient experimental data are compensated by expert knowledge. The disadvantage of fuzzy rules is the need to completely restructure them when introducing additional input variables.

Fuzzy relations [3] are implemented in fuzzy cognitive maps (FCM), which became widespread after publishing [8, 9]. Such FCMs use expert information on the interaction of variables. If the number of input variables increases, then the FCM requires fewer expert judgments than a system of fuzzy rules. In contrast to the compositional rule of inference from the theory of fuzzy sets [3], the FCM uses a recurrence relation with a threshold function from the theory of neural networks [9]. In the well-known publications (see, e.g., [10]), the FCM is considered as a model of a dynamic system that allows tracing a step-by-step changes in the values of variables based on the initial vector and the influence matrices. In this sense, the FCM is similar to a Markov chain. The principal difference between these models is determined by the fundamental difference between fuzzy sets and probabilities [11].

In [12], the FCM was used to rank the input factors according to the degree of their influence on the output variable. This work is a continuation of [12] where FCM is considered as an alternative to regression, i.e., a tool for modeling the inputs-output dependence based on expert-experimental information. The vertices of the FCM graph are input and output variables, and the weights of the arcs are unknown parameters that are set by experts and are adjusted based on the results of observations. Variable increments are used to calculate the output value for given input values. The optimal values of the weights of the arcs are found using genetic algorithm.

BASIC CONCEPTS

General Information. According to [8], the FCM is a directed graph whose arcs are weighted by fuzzy terms [3]. The vertices of the graph, called concepts, correspond to the variables accepted in the model, and the weights of the arcs reflect the impacts of changes in the cause variables on changes in effect variables. The use of the term "cognitive" means that the initial data for modeling are subjective opinions of an expert about influences, expressed in terms of "increases" or "decreases." The term "fuzzy" means that the FCM uses different levels of increase and decrease, which are given by values from the intervals [0, 1] and $[-1, 0]$ and correspond to the terms weakly, moderately, strongly, and others from the theory of fuzzy sets [3].

Concepts. Let $C = \{C_1, C_2, \ldots, C_n\}$ be a known set of concepts, i.e., variables used in the model. According to [8], each concept $C_i \in C$ is estimated by the value $A_i \in [0, 1]$, which determines the level of the concept and is given expertly. To obtain the value of A_i, it is convenient to use the function of fuzzy perfection [12, 13], which characterizes the degree of proximity of the concept value $C_i \in C$ to a certain ideal: 0 is the least perfection, 1 is the greatest perfection.

Connections between Concepts. The weight w_{ij} of the arc connecting the concepts C_i and C_j indicates the impact of C_i on C_j. Let the concepts C_i and C_j be characterized by the variables x_i and x_j, and the dependence $x_j = \varphi(x_i)$ can be constructed as a result of the experiment. Then the weight w_{ij} is defined as a derivative $w_{ij} = dx_j / dx_i$, which can be one of the three types:

— $w_{ij} > 0$ if the increase (decrease) in the value x_i leads to an increase (decrease) in the value x_j (positive influence of C_i on C_j);

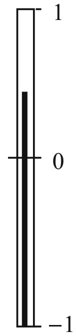
— $w_{ij} < 0$ if the increase (decrease) in the value x_i leads to a decrease (increase) in the value x_j (negative influence of C_i on C_j);

— $w_{ij} = 0$ if the value x_j does not depend on the value x_i (no influence of C_i on C_j).

We will estimate expertly the power of influence w_{ij} using linguistic terms and thermometer scale (Table 1). If the opinions of several experts are taken into account, then the value w_{ij} is estimated as a weighted average of each expert estimate.

The applicability of expert estimate w_{ij} is directly related to the remarkable ability of the human eye to catch linear dependencies in extrapolation problems, as V. M. Glushkov noted [14].

TABLE 1. Methods for Estimating the Power of Influence

Thermometer Scale	Linguistic Ratings	Quantitative Values
	Positive maximum	1
	Positive above average	0.75
	Positive average	0.5
	Positive below average	0.25
	Does not exist	0
	Negative below average	– 0.25
	Negative average	– 0.5
	Negative above average	– 0.75
	Negative maximum	– 1

Recurrent Relation with a Threshold Function. The dynamics of changes in the concept values in the FCM is determined by the relation [15]

$$A_i^{k+1} = f\left(\sum_{\substack{j=1 \\ j \neq i}}^{n} A_j^k w_{ji} + cA_i^k \right), \ k = 0, 1, 2, \ldots, \tag{1}$$

where A_i^{k+1} is the value of the concept C_i at the $(k+1)$th step; A_i^k and A_j^k are the values of concepts C_i and C_j at the kth step, respectively; w_{ji} is a power of influence of the concept C_j on the concept C_i; c is a parameter responsible for the prehistory, i.e., a contribution of a concept value at the previous step, $c \in [0, 1]$; f is the threshold function due to which a concept value does not exceed 1.

The most widely used is a sigmoid threshold function

$$f(x) = \frac{1}{1 + e^{-\lambda x}}, \ \lambda > 0, \tag{2}$$

and the positive part of the hyperbolic tangent

$$f(x) = \begin{cases} \tanh(x) & \text{for } x \geq 0, \\ 0 & \text{for } x < 0, \end{cases} \tag{3}$$

where $\tanh(x) = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$.

Recurrent relation (1) can be represented in a matrix form

$$\boldsymbol{A}^{k+1} = f(\boldsymbol{A}^k \boldsymbol{W}_0 + c\boldsymbol{A}^k), \ k = 0, 1, 2, \ldots, \tag{4}$$

where $\boldsymbol{A}^{k+1}$ and $\boldsymbol{A}^k$, $k = 0, 1, 2, \ldots$, are $(1 \times n)$-vectors of FCM state, whose elements give the concept values at the $(k+1)$th and kth steps, respectively;

$$\boldsymbol{W}_0 = \begin{bmatrix} 0 & w_{12} & \ldots & w_{1n} \\ w_{21} & 0 & \ldots & w_{2n} \\ \cdot & \cdot & \ldots & \cdot \\ w_{n1} & w_{n2} & \ldots & 0 \end{bmatrix}. \tag{5}$$

Formula (5) is an $(n \times n)$-matrix of the powers of concept C_i influences, whose diagonal elements are equal to zero.

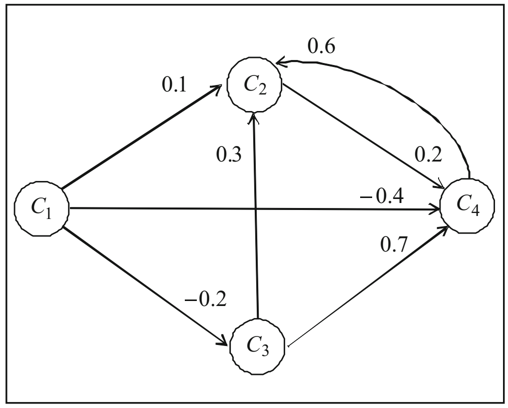


Fig. 1. FCM for computer experiment.

TABLE 2. Experimental Results Using Relation (1)

Threshold Function	**Initial State**				**Steady State**			
	A_1^0	A_2^0	A_3^0	A_4^0	A_1^l	A_2^l	A_3^l	A_4^l
$f(x) = \dfrac{1}{1+e^{-x}}$	0.8	0.6	0.4	0.2	0.66	0.82	0.62	0.75
	0.1	0.3	0.5	0.7	0.66	0.82	0.62	0.75
	0.1	0.1	0.1	0.0	0.66	0.82	0.62	0.75
	0.5	0.5	0.5	0.0	0.66	0.82	0.62	0.75
$f(x) = \begin{cases} \tanh(x) & \text{for } x \ge 0, \\ 0 & \text{for } x < 0 \end{cases}$	0.8	0.6	0.4	0.2	0.00	0.86	0.00	0.70
	0.1	0.3	0.5	0.7	0.00	0.86	0.00	0.70
	0.1	0.1	0.1	0.0	0.00	0.86	0.00	0.70
	0.5	0.5	0.5	0.0	0.00	0.86	0.00	0.70

The initial state of the FCM is defined by the vector

$$\boldsymbol{A}^0 = [A_1^0, A_2^0, \dots, A_n^0], \tag{6}$$

whose elements reflect the concept values at the step $k = 0$.

A steady state of the FCM is defined by the vector

$$\boldsymbol{A}^l = [A_1^l, A_2^l, \dots, A_n^l] \tag{7}$$

at the lth step, such that as a result of the interaction of the concepts, FCM enters the steady mode in which $|A_i^l - A_i^{l-1}| < \varepsilon$, where ε is a small positive value and $i = 1, 2, \dots, n$.

Experiment Using Relation (1). The use of this relation to approximate the inputs-output dependence was tested experimentally on a simple FCM (Fig. 1). Various vectors (6) were used as inputs into FCM, and the corresponding vectors (7) were calculated in the steady state mode using relation (1). It was assumed that $c = 1$. The experiment was carried out with threshold functions (2) and (3). Sigmoid function (2) was used for the parameter values $\lambda = 0.5$, $\lambda = 1$, and $\lambda = 2$. A fragment of the experimental results is shown in Table 2. In all the experiments, it turned out that for each threshold function, the vector of concept values in the steady state mode does not depend on the corresponding vector in the initial state. Thus, relation (1) does not provide the output sensitivity to changes in input variables.

INPUTS-OUTPUT DEPENDENCE APPROXIMATION

Interrelation of Variable Increments. To approximate the inputs-output dependence using the FCM, we will use the increments of the concept values as they interact at the kth steps, $k = 0, 1, 2, \dots$. The value of the concept C_i at the $(k+1)$th step depends on the values of the concepts C_j $(j = 1, 2, \dots, n)$ at the previous kth step. Let us denote this dependence as

$$A_i^{k+1} = \psi(A_1^k, \dots, A_j^k, \dots, A_n^k). \tag{8}$$

From (8), the relationship between the increments (Δ) of concept values on the neighboring $(k+1)$th and kth steps follows:

$$\Delta A_i^{k+1} = \frac{\partial A_i^{k+1}}{\partial A_1^k} \Delta A_1^k + \dots + \frac{\partial A_i^{k+1}}{\partial A_j^k} \Delta A_j^k + \dots + \frac{\partial A_i^{k+1}}{\partial A_n^k} \Delta A_n^k. \tag{9}$$

Partial derivatives in (9) correspond to the mutual interactions of concepts: $\frac{\partial A_i^{k+1}}{\partial A_j^k} = w_{ji}$. Hence, relation (9) can be written in the form

$$\Delta A_i^{k+1} = \sum_{j=0}^{n} \Delta A_j^k w_{ji}, \; i = 1, 2, \dots, n, \tag{10}$$

where

$$\Delta A_i^{k+1} = A_i^{k+1} - A_i^k, \; \Delta A_j^k = A_j^k - A_j^{k-1}. \tag{11}$$

Considering (10) and (11), we obtain the equation for the dynamics of a step-by-step change in the concept values

$$A_i^{k+1} = A_i^k + \sum_{j=1}^{n} (A_j^k - A_j^{k-1}) w_{ji}. \tag{12}$$

Similarly to (4), relation (12) can be represented in the matrix form

$$\boldsymbol{A}^{k+1} = \boldsymbol{A}^k \oplus (\boldsymbol{A}^k \ominus \boldsymbol{A}^{k-1}) \boldsymbol{W}_0, \tag{13}$$

where $\oplus$ and $\ominus$ are the operations of elementwise addition and subtraction of vectors, performed according to the scheme

$$[a, b] \oplus [c, d] = [a + c, b + d],$$

$$[a, b] \ominus [c, d] = [a - c, b - d].$$

In (13), we assume that, for $k = 0$, the equality

$$\boldsymbol{A}^1 = \boldsymbol{A}^0 \oplus \boldsymbol{A}^0 \boldsymbol{W}_0$$

holds.

Output Variable Prediction. The output value corresponding to the fixed values of the input variables can be calculated by the following algorithm.

Step 1. Set the initial state of the FCM (6) by the vector

$$\boldsymbol{A}^0 = [A_1^0, A_2^0, \dots, A_{n-1}^0, A_n^0 = 0], \; A_i^0 \in [0, 1], \; i = 1, 2, \dots, n-1. \tag{14}$$

Step 2. Using relation (13), calculate the vector (7) of concept values in a steady state. Fix the value A_n^l of the output concept C_n.

Step 3. Find the initial vectors (14) that correspond to the highest ($\overline{A}_n$) and the lowest ($\underline{A}_n$) values of the output concept C_n in a steady state. The algorithm for solving the necessary optimization problems is considered in the next section.

TABLE 3. Experimental Results Using Relation (13)

Initial State				Steady State				
A_1^0	A_2^0	A_3^0	A_4^0	A_1^l	A_2^l	A_3^l	A_4^l	$\hat{A}_4$
0.5	0.5	0.5	0.0	0.50	0.82	0.40	0.24	0.50
0.1	0.4	0.8	0.0	0.10	1.08	0.78	0.72	0.78
0.9	0.5	0.3	0.0	0.90	0.52	0.12	−0.17	0.26
0.1	0.2	0.7	0.0	0.10	0.77	0.68	0.59	0.70
0.2	0.7	0.4	0.0	0.20	1.06	0.36	0.38	0.58

Step 4. Calculate the normalized value of A_n^l obtained at Step 2 by the formula

$$\hat{A}_n = \frac{A_n^l - \underline{A}_n}{\overline{A}_n - \underline{A}_n}. \tag{15}$$

Consider the value $\hat{A}_n$ as a prediction for the output value that corresponds to the given input vector (14). We need to normalize (15) because there is no threshold function in (13), and, due to this fact, the value of A_n^l may be greater than one.

Experiment Using Relation (13). Table 3 shows a fragment of the results of the experiment with FCM (see Fig. 1) using relation (13). The steady state was observed for $l = 6$. The highest and lowest values $\overline{A}_4 = 1.09$ and $\underline{A}_4 = -0.61$ were attained at the initial vectors $\boldsymbol{A}_0 = [0.0, 1.0, 1.0, 0.0]$ and $\boldsymbol{A}_0 = [1.0, 0.0, 0.0, 0.0]$, respectively.

Consider, e.g., the first row in Table 3. The initial vector $\boldsymbol{A}_0 = [0.5, 0.5, 0.5, 0.0]$ leads to the steady state vector $\boldsymbol{A}_0 = [0.5, 0.82, 0.40, 0.24]$, whence $A_4^6 = 0.24$. As a result of normalization of (15), we obtain

$$\hat{A}_n = \frac{0.24 + 0.61}{1.09 + 0.61} = 0.5.$$

ADJUSTING THE PARAMETERS OF THE INPUTS-OUTPUT MODEL

Data for Adjusting. Let us assume that as a result of the observations we manage to collect the data, presented in Table 4, where $A_{ip} \in [0, 1]$ is a concept C_i level in the observation p, $i = 1, 2, \ldots, n$, $p = 1, 2, \ldots, N$, N is the number of observations. It is assumed that the values of A_{ip} in Table 4 can be obtained in two ways: experimentally or by expert assessments. The experimental method makes it possible to accurately measure the parameter x_{ip} associated with the concept C_i and the transformation of x_{ip} into the level $A_{ip} \in [0, 1]$ using membership functions for the fuzzy perfection [12]. The expert method uses qualitative estimates for levels A_{ip} in the form of linguistic (fuzzy) terms, which correspond to numerical values (Table 5). The formal way of transition from linguistic terms to numerical values is the defuzzification procedure [4], which is not considered in this paper.

Arc Weight Intervals. The FCM parameters adjusted with the allowance for the observation results, are the weights of the arcs $w_{ij} \in [\underline{w}_{ij}, \overline{w}_{ij}]$, where $\underline{w}_{ij} \in (\overline{w}_{ij})$ is a lower (upper) limit of the interval of feasible values w_{ij}. We will base the choice of the intervals $[\underline{w}_{ij}, \overline{w}_{ij}]$ on the following assumptions:

— the type of influence $w_{ij} > 0$, $w_{ij} < 0$, or $w_{ij} = 0$ is determined expertly and does not change when adjusting the FCM;

— the power of influence w_{ij} is estimated by an expert to the accuracy of one linguistic term (see Table 1), i.e., ± 0.2;

— the power of positive $w_{ij} > 0$ and negative $w_{ij} < 0$ influences change within the intervals $[0.05, 0.95]$ and $[-0.95, -0.05]$, respectively.

TABLE 4. Observational Data for the Inputs-Output Model

Number	Inputs				Output
	C_1	C_2	$\dots$	C_{n-1}	C_n
1	A_{11}	A_{21}	$\dots$	$A_{n-1,1}$	A_{n1}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
p	A_{1p}	A_{2p}	$\dots$	$A_{n-1,p}$	A_{np}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
N	A_{1N}	A_{2N}	$\dots$	$A_{n-1,N}$	A_{nN}

TABLE 5. Expert Estimation of Concept Level

Term	Designation	Numerical Value
Exceptionally low	*Exxl*	0.045
Extremely low	*Exl*	0.135
Very low	*Vl*	0.255
Low	*L*	0.335
Below average	*Ba*	0.410
Average	*A*	0.500
Above average	*Aa*	0.590
High	*H*	0.665
Very high	*Vh*	0.745
Extremely high	*Exh*	0.865
Exceptionally high	*Exxh*	0.955

Given these assumptions, the weight intervals for the FCM arcs are chosen according to the scheme:

$$0.3 \in [0.1, 0.5],\ -0.3 \in [-0.5, -0.1],$$

$$0.1 \in [0.05, 0.3],\ -0.1 \in [-0.3, -0.05],$$

$$0.8 \in [0.6, 0.95],\ -0.8 \in [-0.95, -0.6].$$

Optimization Problem. Let us denote by $\hat{A}_n = F(\boldsymbol{A}_0, \boldsymbol{W}_0)$ an inputs-output dependency model that corresponds to the above prediction algorithm. Using this model and Table 4, let us find the deviations

$$\varepsilon_p = A_{np} - \hat{A}_{np},\ \ p = 1, 2, \dots, N, \tag{16}$$

where A_{np} is an output value in pth observation, $\hat{A}_{np}$ is the output forecast for input values from the pth observation, i.e.,

$$\hat{A}_{np} = F(A_{1p}, A_{2p}, \dots, A_{n-1,p}, A_{np} = 0, \boldsymbol{W}_0).$$

Following the method of least squares from regression analysis, we formulate the problem of adjusting the FCM based on observations as follows: find the matrix $\boldsymbol{W}_0 = [w_{ij},\ i = 1, 2, \dots, n,\ j = 1, 2, \dots, n]$ of powers of influences, such that its elements satisfy the constraints $w_{ij} \in [\underline{w}_{ij}, \overline{w}_{ij}]$ and minimize the sum of squares of deviations (16), i.e.,

$$S(\boldsymbol{W}_0) = \sum_{p=1}^{N} [A_{np} - F(\boldsymbol{A}_p, \boldsymbol{W}_0)]^2 \xrightarrow[\boldsymbol{W}_0]{} \min, \tag{17}$$

where $\boldsymbol{A}_p = [A_{1p}, A_{2p}, \dots, A_{n-1,p}, A_{np} = 0],\ \ p = 1, 2, \dots, N.$

Genetic Algorithm. To solve nonlinear optimization problem (17), a genetic algorithm is proposed based on the following concepts and operations [16]: chromosome — a coded version of the solution; population — the initial set of solutions; fitness function — a criterion for selecting options; crossover — the operation of generating offspring chromosomes from the parent chromosomes; mutation — a random change in chromosome elements.

If $P(t)$ are parent chromosomes and $C(t)$ are offspring chromosomes on the tth iteration, then the general structure of the genetic algorithm has the form

begin

$t := 0$; set initial value $P(t)$;

estimate $P(t)$ using the fitness function;

while (until termination conditions are fulfilled) **do**

cross $P(t)$ to obtain $C(t)$;

mutate $C(t)$;

estimate $C(t)$ by the fitness function;

select $P(t+1)$ from $P(t)$ and $C(t)$; $t := t+1$;

end

end.

A chromosome is defined as a row of non-zero matrix elements $\boldsymbol{W}_0 = [w_{ij}]$, $w_{ij} = R[\underline{w}_{ij}, \overline{w}_{ij}]$, where $R[\underline{x}, \overline{x}]$ is the operation of finding a random number uniformly distributed over the interval $[\underline{x}, \overline{x}]$. For example, for FCM (see Fig. 1), the generator of the initial population of chromosomes is the row $[w_{12}, w_{13}, w_{14}, w_{24}, w_{32}, w_{34}, w_{42}]$, where $w_{12} = R[0.05, 0.3]$, $w_{13} = R[-0.4, -0.05]$, $w_{14} = R[-0.6, -0.2]$, $w_{24} = R[0.05, 0.4]$, $w_{32} = R[0.1, 0.5]$, $w_{34} = R[0.5, 0.9]$, $w_{42} = R[0.4, 0.8]$, whence we have the following examples of chromosomes:

$$[0.196, -0.211, -0.252, 0.143, 0.227, 0.548, 0.776], \tag{18}$$

$$[0.171, -0.262, -0.331, 0.309, 0.308, 0.639, 0.450]. \tag{19}$$

The crossing of a pair of parent chromosomes gives rise to an offspring chromosome. The crossing operation is performed by a random exchange of genes (elements) of the parent chromosomes. To do this, each gene of the offspring chromosome is assigned a random number $\xi_1 = R[1, 0]$. If $\xi_1 \le 0.5$, then this gene is taken from the first parent; otherwise, the gene is taken from the second parent. Let the parent chromosomes be given by rows (18) and (19), and the random numbers ξ_1 correspond to the row [0.19, 0.62, 0.21, 0.71, 0.94, 0.17, 0.33]. Then, as a result of crossing (18) and (19), we obtain the offspring chromosome

$$[0.196, -0.262, -0.252, 0.309, 0.308, 0.548, 0.776]. \tag{20}$$

Each gene in (20) may undergo mutation. To do this, a random number $\xi_2 = R[0, 1]$ and the mutation coefficient q are assigned to each gene (in this case, $q = 0.1$). If $\xi_2 \le q$, then this gene is replaced with a random number from the range of feasible values.

Let a row of random numbers ξ_2 be of the form [0.63, 0.11, 0.49, 0.08, 0.18, 0.74, 0.37]. Hence, in chromosome (20), only the fourth gene undergoes mutation, and, after that, the offspring chromosome takes the form $[0.196, -0.262, -0.252, 0.151, 0.308, 0.548, 0.776]$, where $0.151 = R[0.05, 0.4]$.

The fitness function is criterion (17) with a minus sign, i.e., the better the chromosome meets the optimization criterion, the greater the fitness function.

The selection of parent chromosomes for the crossover is not occasional. The best solutions are prioritized. The greater the fitness functions, the more likely a given chromosome will produce an offspring [16]. When the genetic algorithm is executed, the population size remains constant. Therefore, after the crossovers and mutations from the resulting population, it is necessary to remove the chromosomes that have the worst value of the fitness function.

EXAMPLE OF FCM APPLICATION

Multivariate Analysis of the Reliability. The initial stage of modeling the reliability of the system [17] is its structuring, i.e., selecting the elements with associated probabilities of failures. In cases where structuring is complicated, the system has to be considered as a black box whose output is reliability, and the inputs are influencing factors.

Let us consider modeling the reliability of a car in a driver–car–road system, taking into account the material, technical, ergonomic, organizational, and environmental factors. The information needed to build and adjust the FCM is provided by an expert in the field of vehicle operation and maintenance.

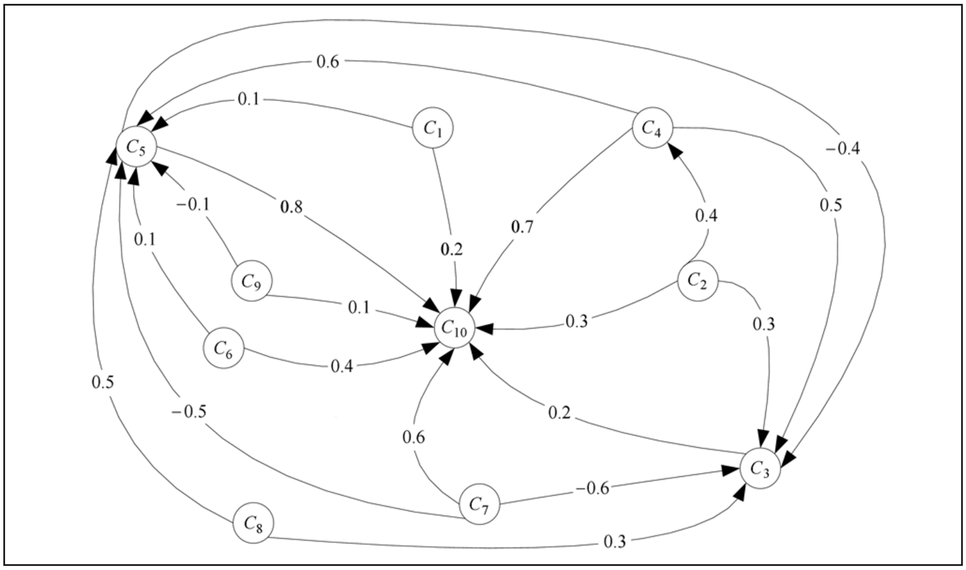


Fig. 2. FCM of the system "driver–car–road."

TABLE 6. Expert Estimate of Inputs-Output

Number	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}
1	*Vh*	*Exh*	*Ba*	*Exh*	*Vh*	*Exh*	*Exh*	*Vh*	*Exh*	*Exh*
2	*Exh*	*Vh*	*Vl*	*Exh*	*Exh*	*Vh*	*Exh*	*Exh*	*H*	*Exh*
3	*Exxh*	*Exh*	*L*	*Exxh*	*Vh*	*Vh*	*Exxh*	*Exxh*	*Exxh*	*Exh*
4	*Vh*	*H*	*A*	*Exh*	*Vh*	*Vh*	*H*	*Vh*	*Exh*	*Vh*
5	*H*	*Vh*	*Vl*	*H*	*H*	*Aa*	*Aa*	*Vh*	*Exh*	*Vh*
6	*A*	*H*	*L*	*Exxh*	*Aa*	*Aa*	*Vh*	*Vh*	*Vh*	*H*
7	*Aa*	*Aa*	*A*	*Ba*	*A*	*H*	*H*	*A*	*Aa*	*Aa*
8	*A*	*Ba*	*A*	*A*	*Aa*	*Aa*	*Vh*	*Vh*	*Aa*	*Aa*
9	*H*	*Aa*	*Aa*	*Vh*	*H*	*Aa*	*Aa*	*Ba*	*Vl*	*A*
10	*A*	*A*	*H*	*Ba*	*A*	*Aa*	*H*	*Ba*	*H*	*A*
11	*A*	*Aa*	*Vh*	*H*	*Aa*	*Aa*	*Aa*	*L*	*Ba*	*A*
12	*L*	*Ba*	*Exh*	*A*	*Aa*	*Ba*	*Ba*	*H*	*H*	*A*
13	*Exl*	*L*	*Ba*	*Ba*	*A*	*L*	*Vl*	*A*	*A*	*Ba*
14	*Vl*	*H*	*Exh*	*L*	*Ba*	*A*	*Vl*	*Exl*	*A*	*L*
15	*Aa*	*Vl*	*Vl*	*Ba*	*L*	*Exl*	*Exl*	*L*	*Exl*	*L*

Approximation. The given FCM is shown in Fig. 2, where C_{10} is an output concept corresponding to the reliability and safety of the vehicle, $C_1, C_2, \ldots, C_9$ are input concepts corresponding to the following influencing factors: C_1 is driver's qualification, C_2 is road conditions, C_3 are unit operating costs, C_4 are terms of use, C_5 is periodicity of maintenance, C_6 is quality of maintenance and repair, C_7 is quality of car's structure, C_8 is quality of operating materials and spare parts, and C_9 is storage conditions.

To test the FCM as a forecasting model, we used inputs-output expert estimates (Table 6), obtained with the use of linguistic terms from Table. 5. Each row in Table 6 was defined by an expert as a combination of levels $(A_1, A_2, \ldots, A_9)$ of influencing factors leading to a given level (A_{10}) of reliability and safety of the car.

TABLE 7. Expert Estimate and Simulation Results

Number	A_{10}		**Before Adjusting**			**After Adjusting**		
	Term	Number	A_{10}^{l}	$\hat{A}_{10}$	$\lvert A_{10} - \hat{A}_{10} \rvert$	A_{10}^{l}	$\hat{A}_{10}$	$\lvert A_{10} - \hat{A}_{10} \rvert$
1	*Exh*	0.865	3.474	0.805	0.060	3.482	0.807	0.058
2	*Exh*	0.865	3.448	0.799	0.066	3.582	0.830	0.035
3	*Exh*	0.865	3.672	0.851	0.015	3.814	0.884	0.019
4	*Vh*	0.745	3.240	0.751	0.006	3.294	0.763	0.018
5	*Vh*	0.745	2.852	0.661	0.084	2.924	0.677	0.068
6	*H*	0.665	3.073	0.712	0.047	3.029	0.702	0.037
7	*Aa*	0.590	2.249	0.521	0.069	2.329	0.540	0.050
8	*Aa*	0.590	2.324	0.538	0.052	2.532	0.587	0.003
9	*A*	0.500	2.727	0.632	0.132	2.599	0.602	0.102
10	*A*	0.500	2.110	0.489	0.011	2.192	0.508	0.008
11	*A*	0.500	2.534	0.587	0.087	2.374	0.550	0.050
12	*A*	0.500	2.195	0.509	0.009	2.212	0.512	0.012
13	*Ba*	0.410	1.682	0.390	0.020	1.659	0.384	0.026
14	*L*	0.335	1.855	0.430	0.095	1.550	0.359	0.024
15	*L*	0.335	1.400	0.324	0.011	1.415	0.328	0.007

To use formula (13) for calculations, linguistic terms were replaced by numerical values from Table 5. Predicting the output level ($\hat{A}_{10}$) for different input levels ($A_1, A_2, \ldots, A_9$) was performed for each row of Table 6. The steady state FCM mode was observed at $l = 5$. The values $\underline{A}_{10}$ and $\overline{A}_{10}$ needed for the calculation by formula (15) are the following: $\underline{A}_{10} = 0$ for $A_i = 0$, $i = 1, 2, \ldots, 9$, and $\overline{A}_{10} = 4.317$ for $A_i = 1$, $i = 1, 2, \ldots, 9$.

Comparing expert estimates A_{10} from Table 6 and the simulation results before adjusting (Table 7), we conclude that the mean absolute deviation (MAD) and mean square error (MSE) are

$$MAD = \frac{1}{15} \sum_{p=1}^{15} \lvert A_{10,p} - \hat{A}_{10,p} \rvert = 0.051 \text{ and } MSE = \frac{1}{15} \sum_{p=1}^{15} (A_{10,p} - \hat{A}_{10,p})^2 = 0.004,$$

respectively.

Note that the proximity of the values A_{10} and $\hat{A}_{10}$ prove the high qualification of the expert whose knowledge was used to construct the FCM (see Fig. 2) and inputs-output estimates (see Table 6).

Adjusting. In the genetic algorithm for FCM adjusting (see Fig. 2), 200 chromosomes were used. At each iteration, 20 crossovers were performed with the mutation factor $q = 0.1$. Dynamics of changes in the optimization criterion (17) as the number of iterations (M) increases is shown in Fig. 3. The fact that the optimization algorithm attains a steady state interval with more than 2000 iterations determines the convergence of the genetic algorithm. The change in the weights of the arcs after adjustment is given in Table 8. Application of the FCM with adjusted arc's weights (see Table 7) delivers improved MAD and MSE indices whith values 0.035 and 0.002 after adjustment, respectively.

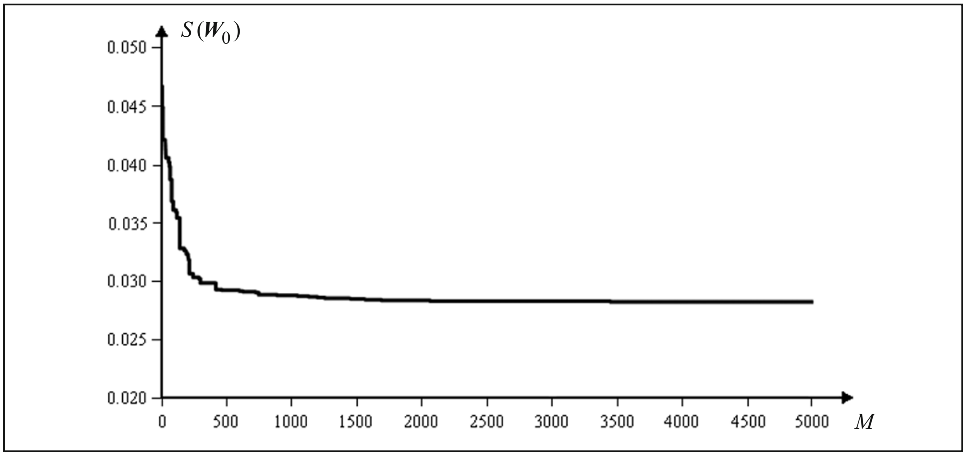


Fig. 3. Dynamics of changes in the optimization criterion.

TABLE 8. Weights of the Arcs before and after Adjustment

Weight	Before Adjusting	After Adjustment	Weight	Before Adjusting	After Adjustment
w_{23}	0.3	0.220	w_{85}	0.5	0.700
w_{43}	0.5	0.301	w_{95}	– 0.1	– 0.051
w_{53}	– 0.4	– 0.311	$w_{1,10}$	0.2	0.359
w_{73}	– 0.6	– 0.711	$w_{2,10}$	0.3	0.103
w_{83}	0.3	0.496	$w_{3,10}$	0.2	0.051
w_{24}	0.4	0.212	$w_{4,10}$	0.7	0.500
w_{15}	0.1	0.245	$w_{5,10}$	0.8	0.868
w_{45}	0.6	0.400	$w_{6,10}$	0.4	0.293
w_{65}	0.1	0.153	$w_{7,10}$	0.6	0.712
w_{75}	– 0.5	– 0.309	$w_{9,10}$	0.1	0.300

CONCLUSIONS

Based on a fuzzy cognitive map, a method for constructing an inputs-output dependence is proposed, which is an alternative to regression analysis. The method is illustrated by the example of multivariate analysis of the reliability of a man–machine system. The idea of the method is to represent the model in the form of a directed graph whose vertices correspond to the input and output variables, and the weights of the arcs are unknown parameters that are provided by experts and adjusted according to the results of observations. We show that the threshold functions, which are traditionally used in works on FCM, do not provide the output sensitivity to variations in the input variables; therefore, to approximate the inputs-output dependence, an algorithm based on variable increments is proposed. The problem of optimal adjustment of the weights of arcs is posed and a genetic algorithm for its solution is developed. Chromosomes are generated from ranges of acceptable values, and the selection criterion is the sum of the squared deviations between the model and observed output values. The fundamental advantages of the method are that, firstly, by simply adding and removing vertices and arcs from the graph, it is possible to change many input variables affecting the objective function, and secondly, due to the convenience of interpreting unknown parameters, it is possible to compensate the lack of experimental data by the use of expert assessments. We recommend to use the method in predicting problems with a large number of input variables, where expert assessments are an important source of information.

REFERENCES

1. D. C. Montgomery, G. C. Runger, and N. F. Hubele, Engineering Statistics, John Wiley & Sons, New York (1998).
2. D. A. Pospelov, Logic-Linguistic Models in Control Systems [in Russian], Energoatomizdat, Moscow (1986).
3. L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning," Inform. Sci., Vol. 8, No. 3, 199–249 (Part I); Vol. 8, No. 4, 301–357 (Part II); Vol. 9, No. 1, 43–80 (Part III) (1975).
4. E. H. Mamdani, "Application of fuzzy algorithms for control of a simple dynamic plant," Proc. of IEE, Vol. 121, Iss. 12, 1585–1588 (1974).
5. M. Takagi and M. Sugeno, "Fuzzy identification of systems and its application to modeling and control," IEEE Trans. on Systems, Man and Cybern., Vol. SMC-15, No. 1, 116–132 (1985).
6. A. P. Rotshtein, Intellectual Technologies of Identification: Fuzzy Sets, Genetic Algorithms, and Neural Networks [in Russian], Universum, Vinnytsia (1999).
7. A. Rotshtein and H. Rakytyanska, Fuzzy Evidence in Identification, Forecasting and Diagnosis, Springer, Berlin–Heidelberg (2012).
8. B. Kosko, "Fuzzy cognitive maps," Intern. J. of Man–Machine Studies, Vol. 24, Iss. 1, 65–75 (1986).
9. B. Kosko, Neural Networks and Fuzzy Systems, Prentice-Hall, Englewood Cliffs, NJ (1992).
10. E. I. Papageorgiu (ed.), Fuzzy Cognitive Maps for Applied Sciences and Engineering. From Fundamentals to Extensions and Learning Algorithms, Springer, Berlin–Heidelberg (2014).
11. A. Kaufmann, Introduction à la Théorie des Sous-Ensembles Flous à l'Usage des Ingénieurs: Fuzzy Sets Theory, Masson, Paris–New York–Barcelone–Milan (1977).
12. A. Rotshtein, D. Katielnikov, and A. Kashkanov, "A fuzzy cognitive approach to ranking of factors affecting the reliability of man–machine systems, " Cybern. Syst. Analysis, Vol. 55, No. 6, 958–966 (2019). https://doi.org/10.1007/s10559-019-00206-8.
13. A. P. Rotshtein, "Selection of human working condition based on fuzzy perfection," J. Comput. Syst. Sci. Int., Vol. 57, No. 6, 927–937 (2018).
14. V. M. Glushkov, An Introduction to Automatic Control Systems [in Russian], Tekhnika, Kyiv (1974).
15. C. D. Stylios and P. P. Groumpos, "Modeling complex systems using fuzzy cognitive maps," IEEE Trans. on Systems, Man, and Cybernetics — Part A: Systems and Humans, Vol. 34, No. 1, 155–162 (2004).
16. M. Gen and R. Cheng, Genetic Algorithms and Engineering Design, John Willey & Sons, New York (1997).
17. R. Barlow and F. Proschan, Statistical Theory of Reliability and Life Testing, Holt, Rinehart and Winston, New York (1975).