Simulation of operation of neural network with purpose of utilisation of hydraulic actuators in complicated technical conditions

Baglan B. Togizbayeva[1*], Bayan T. Sazambayeva[1], Abdikarim A. Karazhanov[1],
Anuar B. Kenesbek[1], Mihaela Cocoşilă[2]

(1. *Department of Transport, Transport Equipment and Technology, L.N. Gumilyov Eurasian National University, Astana 010008, Republic of Kazakhstan*;
2. *Faculty of Juridical, Economical and Administrative Sciences, Spiru Haret University, Craiova 030045, Romania*)

Abstract: Application of neural network technologies is becoming more extensive every year, especially in the economic field. New methods are being discovered; new models of neural networks are being created. Currently, the main application of neural networks in the economy is the prediction of markets, optimization of commodity-money flows, analysis and generalization of various social surveys, prediction of the dynamics of political ratings, optimization of the production process, comprehensive diagnostics of product quality, etc. Utilisation of the hydraulic equipment makes it possible to design systems of automatic operation in the conditions, where minimal human participation and maximum speed of response are required. It is possible to state that such conditions exist in the regions with the permanent access of repair teams or technicians, who perform monitoring of the complex technical objects. Therefore, it is necessary to develop automated systems of operation and monitoring of various equipment components, which are intended for operation in the complicated technical conditions. Authors of this article have selected the hydraulic equipment as the object under investigation due to the fact that it is widely distributed equipment, as well as due to possibilities of this equipment to function or to be adapted for operation in practically any environmental conditions. At the same time, quantity of the state-of-the-art equipment, which is used, as well as complexity of this equipment increase very quickly, therefore process of making decision concerning utilisation of this equipment must be made very quickly. Authors analyse the sphere of automation, where utilisation of a human decision is required. The novelty of this article is connected with the assumption that further direction of operation of such equipment in the complicated technical conditions must be implemented in the sphere of guessing of the user's actions. Authors review neural networks as the toolkit, and they believe that these networks can make decisions in the proactive mode, practically without participation of a user. This article includes description of the model, which can be used as the basis for the system, which is planned to design. In addition, methodological toolkit for assessment of efficiency of this model is proposed.
Keywords: equipment, neural network, complicated technical conditions, functioning, model, economic efficiency
DOI: 10.25165/j.ijabe.20201301.3965

1 Introduction

A specific feature of the development of information technology is the continuous qualitative complication of their fields of application. We can highlight technologies for automating the process of solving complex computing issues; management of facilities and production processes; collection, accumulation, storage and use of large volumes of information; formation of managerial decisions to ensure a level of security not lower than a given military-political state, etc. The use of neural network technologies is becoming more extensive every year, especially in the economic field. New methods are being discovered; new models of neural networks are being created. Currently, the main application of neural networks in the economy is the prediction of markets, optimization of commodity-money flows, analysis and generalization of various social surveys, prediction of the dynamics of political ratings, optimization of the production process, comprehensive diagnostics of product quality and much more.

In the majority of cases, designing of devices includes the stage of experimental adjustment of initial characteristics of these devices. First and foremost, it is connected with a large quantity of those factors, which are not taken into account in the course of the classical synthesis. Other methods of designing with utilisation of the speciality application-dependent software (for example, MicroWave Studio) do not allow to take these factors into account to the full extent as well, because of, particularly, due to the essential growth of the required computational capabilities.

Relevant articles[1] have demonstrated possibility of utilisation of the artificial neural networks as the computational device, which accumulates knowledge and makes it possible to include elements of the artificial intelligence into the process of designing. Neural networks provide a universal way of solving many non-standard problems, they can comprehend complex nonlinear dependencies and self-improvement, learning in the process of use. The use of systems built on their basis is efficient, relatively simple and cost-effective. Therefore, it is not surprising that the growing interest in this technology has led to a significant expansion of the scope of its application, the emergence of a multitude of different approaches, training algorithms, software products, application options. Specific features and main difficulties were considered in respect of their utilisation in the problems, which differ from the problems, which are considered as classical problems for their solving with the help of the neural networks[2]. Prospects of such an approach are determined by timeliness, applicability and cost effectiveness of the focused attention, investigation in this direction[3].

Special interest[4] is connected with utilisation of the neural networks in the course of solving the problems, which are connected with operation of devices in the complicated technical conditions. In respect of these devices, it is worth to say the following: they are characterised by the essential range of change in the output characteristics. In addition, their output characteristics essential differences from the computational characteristics within the filter rejection band. What is more, as of today it is difficult to ensure mathematical forecasting of these differences (such forecasting can be only performed with significant errors), and it is very difficult to adapt conditions of operation of such devices to the state-of-the-art conditions of economic management.

Accelerated development of high technologies has caused an essential growth of the actuality of those spheres of special scientific interests, which are connected with mathematical Simulation of processes and phenomena[5]. As a rule, simulation of real-life objects of the outside world is accompanied by the by essential by difficulties, which occur at the stage of the problem statement already[6]. For the most part, these difficulties are caused by imperfections of the computational methods and facilities for the implementation of these methods. Therefore, these difficulties are the reason of existence of the great quantity of the unsolved problems, which occur in the course of: investigation of aerodynamic models; investigation of hydromagnetics and astrophysics problems; simulation of processes of the solid-state physics on the condition of presence of phase transitions taking into account of quantum effects, and so on[7]. The problems, which have incorrect statements and no optimal algorithms for solving them, become the particularly topical problems. Strictly speaking, the artificial neural networks are used in the course of solving precisely these problems, for the most part[8].

As of today, there are no any generally accepted definitions of the artificial neural networks, however, it is very important to understand differences between artificial and biological neural networks, which are immeasurably more complicated networks as compared with the mathematical models, on the basis of which the artificial neural networks are created[9]. The state-of-the-art artificial neural networks consist of the great quantity of simple processor elements along with the certain quantity of the local memory elements (of neurons), which are combined with each other with the help of the discrete or continuous communication channels[10]. The problems, which are usually solved with the help of the artificial neural networks, are subjected to decomposition; they must be divided into a great number of local problems, each of which can be solved with the help of the special neuron through implementation of the certain algorithm of the local data processing[11]. Among the problems solved by neural network packages in the field of economics and finance, the following can be mentioned: price forecasting; property valuation; forecast of futures quotations; forecasting currency exchange rates; research of demand factors; insurance risk analysis; optimization of catalog trading, etc. In the field of economics and finance, the following neural network packages are mainly used: Brain Maker, Neuroforester, The AI Trilogy ("Trilogy of Artificial Intelligence"). As concerns the artificial networks, there also exists such a concept as training, which is the special process that ensures adjustment of the network parameters in order to increase the efficiency of its functioning[12]. There exist two main approaches to the training of the artificial neural networks in accordance with the following names: supervised training (learning with a teacher) and non-supervised training (learning without a teacher)[13]. Learning with a teacher is connected with the existence of the learning sample, that is, a certain set of solutions[14].

Another well-known approach is connected with utilisation of the neural network as the universal approximator of functions. Therefore, certain types of splines, for example, B1-splines can be obtained through superimposition of the piecewise-linear activation functions[15]. If one will use such splines as the basis functions, it is possible to find a solution for the system of the linear algebraic equations through the selection of the relevant coefficients of the data of splines[16]. With the purpose of implementation of this approach, the feed-forward artificial neural networks along with the relevant piecewise-linear activation functions are usually utilised[17].

If one would waive restrictions in respect of the form of the neuron activation function, then it would be possible to find out a more state-of-the-art approach to the solution of the boundary-value problems. This approach would be based on the property of approximation of functions, which is typical for the feed-forward artificial neural networks[18,13]. Architecture of the neural network must ensure availability of the parameter, which must be close to the analytical parameter and correct value of which is usually defined through the training, process of which is to be implemented in accordance with the criterion of minimisation of the error function[19].

Major part of the boundary-value problems can be described with the specified degree of accuracy with the help of a certain set of solutions of local problem, each of which describes the dynamics of processes in the fixed point of the certain area[20-22]. Cellular neural networks are based on utilisation of such an approach, within which each neuron is combined with the fixed quantity of the neighboring neurons with the help of the communication channels.

Interest in the use of artificial neural networks in the economy is growing every day. They have proven themselves in solving many applied financial and economic problems. Artificial neural networks are irreplaceable for the high-quality processing of colossal data streams, without which it is very difficult, and sometimes even impossible to adequately evaluate the market situation and make the right decision. All this necessitates further study, development and implementation of the apparatus of artificial neural networks in practice[23,24].

2   Materials and methods

Specification of the experimental investigations was approved in the following manner:

The goal of the investigation: to investigate quality of operation of the selected method of forecasting depending on the parameters of experiments with help of standard samples of data from 5 min, 10 min, 30 min, 1 h, and daily reports of equipment.

Standard samples of the data are used in order to have a possibility for comparison of the investigation results in accordance with the selected method of forecasting with results in accordance with other methods.

Parameters of the experiment: quantity of inputs of the forecasting model (from 2 to 6); quantity of the forecast steps (from 1 to 10); method of forecasting (training of the model by several steps ahead at once or training by one step ahead, and then substitution of forecasts as inputs (in the course of the predetermined number of times)); investigation of parameters of the equipment functioning (parameters as such) or investigation of increments of these parameters.

Specific features of samples: all samples are randomised ones with the purpose of inclusion of the higher quantity of patterns to the training process (within the training sample).

As usual, both direct, and iterative methods, as well as their various combinations (hybrid methods) are used for solving the boundary-value problems with the help of the artificial neural networks (ANN). In principle, major part of these methods is reduced to the solving the system of the linear algebraic equations, which are defined as the result of discretisation of the relevant area. It is worth noting that utilisation of direct methods in this sphere is the traditional operation, which (for the most part) is implemented as mapping of the obtained system of the linear algebraic equations into the architecture of Hopfield type. Then process of solving is based on the minimisation of the energy function of the neural network[25].

In this article, we have been used the methods, which are usually implemented in order to ensure formation of the objective function of designing of the stable technological system. Such a system can be used as the model for implementation of the self-diagnosis principles, as well as for formation of the comprehensive understanding of possibilities of the technological equipment in respect of the autonomous decision-making.

3   Results and discussion

The network is based on the Takagi-Sugeno fuzzy inference rules, but trigonometric polynomials are used in the consequents of rules. The choice of the trigonometric polynomials as approximation functions is caused by their high approximation capacity.

Within each such rule, membership functions correspond to clear inputs of the fuzzy system. Values of membership of inputs to the terms of fuzzy rules are calculated with the help of these membership functions. Then, values of membership of the population of inputs of the fuzzy system to all rules are calculated with the help of the T-norm (triangular norm) operation in the form of the product. In parallel to this process, values of the (pre-trained) trigonometric polynomials are calculated in the consequents of the fuzzy rules for each rule depending on the values of the system inputs. Then, all values of functions, which were obtained from each of these rules, are summed as weighted values in accordance with the values of membership, which were obtained following calculation of values of the term sections of these rules.

The following conclusion was proved with the help of the universal approximation theorem (by fuzzy systems): training of the TSK neural network (Takagi-Sugeno-Kang network) for the retrospective forecasting may be reduced to the problems of clustering of several quadratic optimization problems, as well as to the problems of the independent training of consequents of fuzzy rules.

New method of training of the TSK network was developed on the basis of the above-stated conclusion.

Procedure of formation of the antecedents is classical one. As concerns the sample of the data (provided that this sample corresponds to the inputs of the network), method of the differential grouping is used, while then (in accordance with the results of the differential grouping) k-means method is used. This results in occurrence of the necessary quantity of the input rules, as well as in occurrence of the centers of membership functions of the terms. At this stage, parameters of the functions of membership-deviation are selected the same. At the next step, these parameters are adjusted with the help of the methods of the quadratic programming, as well as with the help of the considerations, which are presented below. Functions of membership are selected as the Gaussian functions:

$$\mu_{kj}(x_j) = e^{-\frac{(x_j - c_{kj})^2}{2\sigma_{kj}^2}} \tag{1}$$

where, μ is microparticle; k is cluster; x is function argument vector; j is local parameter equal 1, n; e is Euler number; c is fundamental physical constant; σ is root-mean-square deviation.

Let us assume that A_k (where $k = 1$, K) is the collection of points of the data sample, which are included to the cluster k, while K is the total quantity of the clusters, which comply with the rules of the fuzzy system.

In order to adjust the rules in such a manner that in the cases, where a point moves to the input of the fuzzy system, then the antecedent of the rule, to which this point belongs, would show the degree of membership at the level "1". As concerns other rules, then value of membership of the antecedent would be changed in inverse proportion to the increase in the distance of this point from the centers of membership functions of the antecedents. Therefore, in these cases it would be appropriate to perform the following approximation:

$$\begin{cases} \prod_{j=1}^{n} \mu_{kj}(x_{ij}) = 1 \\ (x_{i1}, x_{i2}, ..., x_{in}) \in A_k \\ \prod_{j=1}^{n} \mu_{kj}(x_{ij}) = 1 - \frac{\|x_i - c_k\|}{\max_i \|x_i - c_k\|} \end{cases} \tag{2}$$

where, n is an integer; i is local parameter equal 1, L; A_k is the set of data points belonging to cluster k.

Upon completion taking logarithms of the left-hand and the right-hand members of the equalities ($k=1$, K) and after introduction of the designations (3), (4), (5):

$$h_{kij} = -\frac{(x_{ij} - c_{kj})^2}{2} \tag{3}$$

$$p_{kj} = -\frac{1}{\sigma_{kj}^2} \tag{4}$$

$$\begin{cases} g_{ki} = 0 \\ (x_{i1}, x_{i2}, ..., x_{in}) \in A_k \\ g_{ki} = \ln(1 - \dfrac{\| x_i - c_k \|}{\max\limits_{i} \| x_i - c_k \|}) \end{cases} \qquad (5)$$

where, h is Planck's constant; p is the number of neurons in the neural network; g is Landég-factor; n is an integer.

We will obtain the systems (6):

$$\sum\nolimits_{j=1}^{n} h_{kij} p_{kj} = g_{ki}$$

$$i = 1, \, L \qquad (6)$$

$$j = 1, \, n$$

While after introduction of the designations (7)-(9):

$$H_k = \| h_{kij} \|_{i=1,L}^{j=1,n} \qquad (7)$$

$$p_k = \| p_{kj} \|_{j=1,n} \qquad (8)$$

$$g_k = \| g_{ki} \|_{i=1,L} \qquad (9)$$

where, H is magnetic field strength.

It would be possible to rewrite the systems (6) in the matrix form (10):

$$H_k p_k = g_k \qquad k=1, \, K \qquad (10)$$

Therefore, in order to find out the parameters σ_{kj}, it is necessary to solve the problem of the quadratic programming in Equations (11)-(12):

$$\min\limits_{p_{kj}} (H_k p_k - g_k)^T (H_k p_k - g_k) \qquad (11)$$

$$P_{kj} \geq 0, \quad j=1, \, n \qquad (12)$$

where, T is the period of oscillation.

In the course of training of consequents, each consequent has its separate training course within its own multitude A_k with the purpose of minimisation of its own criterion E_{Ak}. In order to ensure such training, it is possible to apply the principle of the neural network cell.

The radial basis network is the fuzzy neural network within the class of the networks, which utilise the fuzzy logic inference (Takagi and Sugeno). Let us use the following designations: x_1,..., x_n – inputs of the neural network; f_i, $i =1,..., k$, – radial basis functions, values of which depend on the inputs of the neural network. In this article, we have used functions of the Gaussian form in Equation (13):

$$f_i(x) = \exp\left\{ -\frac{\| x - c_i \|^2}{2\sigma_i^2} \right\} \qquad (13)$$

The following designations are used in Equation (13): $x =(x_1$,..., $x_n)$ – vector of the function arguments (of the neural network inputs); $c_i =(c_{1i},...,c_{ni})$ – center of the radial basis function, and σ_i – "scatter" of the radial basis function.

The values of the network inputs are transferred to the calculation units of the radial basis functions. Then, the values, which would be obtained at the outputs, will be multiplied by the relevant weights and will be added to each other. Therefore, we would obtain the output of the radial basis network.

It is usual practice to select the input values of the learning sample of the data as the "centers" of radial basis functions. At the same time, "scatters" are selected by the expert as the values that are equal to each other.

Let us denote the output of the neural network as "y". Then this output will be determined by Equation (14):

$$y = \sum\nolimits_{i=1}^{k} w_i f_i(x) \qquad (14)$$

where, y is the vector of desired network outputs; w is corresponding weight vector.

Let us assume that there exists the training sample of the data $(x_{1j}, x_{2j},..., x_{nj}, y_j)$, $j =1,..., L$. Let us introduce the following matrix and vector notations: $y=(y_1,..., y_L)$ – vector of the desired outputs of the network; $F=\|f_i(x_{1j},...,x_{nj})$, $j=1,...,L$, $i=1,...,k$ – matrix of the relevant outputs of units of the radial basis functions; $w =(w_1,...,w_k)$ – relevant weight vector, then it is possible to rewrite Equation (14) into Equation (15):

$$y = Fw \qquad (15)$$

where, F is the square matrix; w is corresponding weight vector.

Because of it is possible to denote $L = k$, then, in this case, F square matrix, for which (in accordance with the Michelli's theorem) relevant inverse matrix will always exist. If we will fix the value of σ_i, then problem of training of the neural network is reduced to inversion of the matrix F and multiplication of this inverse matrix by vector "y". However, if $L \neq k$, then the problem of adjustment of weights is reduced to the process of finding the pseudo-equal matrix F.

It is usual practice to select "scatters" as the values that are equal to each other. However, the following disadvantage of such an approach is obvious at once: selection of "scatters" depends on the specific problem. In addition, inversion of the matrix is connected with the essential computational efforts. The article, which is devoted to the training of the network, proposes to use the genetic algorithm in order to solve this problem. It is solution of the problem, but such a solution cancels the main advantage of the radial basis network: high learning rate.

In order to ensure learning of the weights, it is proposed to fix the value "of the scatter" for clarity of the computations (subsequently this restriction will be cancelled). In this case, learning is reduced to utilisation of the least-square method in order to restore the linear dependence on the single variable. In this article, we have used precisely this method of learning of the weights of the radial basis network. In addition, we have proposed the method of the learning of the values "of scatters". This method is also reduced to utilisation of the small neural network in order to restore dependence on the single variable.

Hereinafter the following designation will be used (16):

$$k_i = \frac{1}{2\sigma_i^2} \qquad (16)$$

Value k_i from Equation (16) will be called the coefficient of contraction within this article.

Therefore, this article proposes to select two implementations of the data sample provided that these implementations would be situated as far apart as possible in accordance with the Euclidean metric. It will be assumed that these two implementations are the centers of two radial basis functions, which create a new network. Values of the coefficients of contraction for the Gaussian functions are selected as the values that are equal to each other. Then, with the help of the approach on the basis of the Michelli's theorem, it is possible to restore both weights of the radial basis function. Then all implementations of the learning sample of the data are transmitted through the constructed neural network. If the common error of operation of the network would be lesser than the acceptable value, then synthesis and training of the network is completed, otherwise – the following step. At each step, values of the weights, which were obtained earlier, are not changed, while each new weight minimises the inaccuracy of the network at the previous level.

Let us assume that "p" weights are already restored. Now we have to select the implementation with the greatest inaccuracy from those implementations of the data sample, did not take part yet in the synthesis of the network and in the learning of weights. Input values of the selected implementation of the sample initiate the center of a new neuron of the radial basis network (of a new unit for calculation of values of the radial basis function). The coefficient of contraction is selected as the value that is equal to the coefficients of other neurons (hereinafter this restriction will be cancelled). The entire input part of the implementations of the learning data sample, which took part already in the synthesis and learning of the network, as well as new implementation, are transmitted through the network, which was synthesised and learned in the course of the previous step. Therefore, we will obtain "$p+1$" deviations of outputs of the network from the desired values of outputs:

$$e_1 = y_1 - d_1, \ldots, e_{p+1} = y_{p+1} - d_{p+1} \qquad (17)$$

where, d is deutron.

Designations within Equation (17) are as follows: e_i – deviation of the obtained value of the neural network output y_i from the real-life value d_i.

The goal of the learning is to find such value w_{p+1}, which will ensure the existence of the system of Equation (18):

$$f_{p+1}(x_i)w_{p+1} = -e_i, \quad i=1,\ldots,p+1 \qquad (18)$$

It is worth noting once again that subscript "i" does not define the sequential number of any implementation within the learning sample of the data. This subscript defines the sequential number of the neuron of the radial basis network, while x_i (in accordance with the relevant agreement) defines the center of the relevant neuron.

With the help of the least-square method, we will obtain the following equation:

$$w_{p+1} = -\frac{\sum_{i=1}^{p+1} f_{p+1}(x_i)e_i}{(f_{p+1}(x_i))^2} \qquad (19)$$

The implementations of the data sample, which did not take part in the synthesis and learning, "are transmitted" through the network. If common error would be lesser than the predetermined acceptable value, then process of the network synthesis is completed, otherwise, we have to go to the following step.

The above-discussed algorithm of the neural network learning makes it possible to ensure quick learning of the neural network without wasting time for inversion of great matrices. In addition, this algorithm makes it possible to waive of the learning methods, which ensure convergence of the process of learning during infinite quantity of the learning iterations.

However, this method has an imperfection: similarity of values of the coefficients of contraction. Below we propose a relevant idea in order to solve this problem. Firstly, general output of the radial basis neural network has an inessential influence upon the coefficients. Therefore, in the course of learning k_i it is not necessary to take into account values of the weights w_i. In addition, (in accordance with the statement), part of the radial basis network, with the multitude of which the weights are connected, performs classification of the vector, which was transmitted to the input of the network. Therefore, in an ideal scenario, we would like to ensure that in the case of transmission of the certain vector from the learning data sample to the network input, the neuron, which corresponds to this input, would provide the biggest value "1". At the same time, other neurons would ensure other values provided that these values would depend on the distance between

the point, which was transmitted to the network input, and the centers of the radial basis functions. Taking into account the Equations (13) and (16), as well as taking into account the considerations, which are described in this paragraph, it is possible to create the system of Equation (20):

$$\exp\{-k_i \| x_i - c_j \|^2\} = \begin{cases} 1, \; i = j \\ 1 - \dfrac{\| x_i - c_j \|}{\max\limits_{j} \| x_i - c_j \|} \\ \quad i \neq j \end{cases} \qquad (20)$$

where, k_i is the contraction ratio.

Within this system, subscripts "i" and "j" are changed from one to the values of the neuron quantity in the neural network.

Now, if we would introduce designations $h_{ij} = \|x_i - c_j\|^2$, take natural logarithms in respect of each equation in the system (20), as well as exclude identical equations from this system, it is possible to find out the system (21):

$$-k_i h_{ij} = \ln\left(1 - \frac{\| x_i - c_j \|}{\max\limits_{j} \| x_i - c_j \|} \right), \; i \neq j \qquad (21)$$

It is possible to define values of the coefficients contraction with the help of the small neural network:

$$k_i = -\frac{\sum_{z=1}^{p} h_{ij} \ln\left(1 - \frac{\| x_i - c_j \|}{\max\limits_{j} \| x_i - c_j \|} \right)}{h_{ij}^2} \qquad (22)$$

If we will average all coefficients, then it would be possible to find out the classical case, where coefficients of contraction of all neurons are on the same level.

In the course of practical application, it is proposed to introduce neurons selectively. If the quality of functioning of the network is increased essentially as the result of addition of a neuron, then it is necessary to apply the above-described algorithm. If this is not the case (increase in quality of functioning of the network is inessential as the result of addition of a neuron), it is necessary to do the following: either to retrofit the network with the help of the RLSM (recursive least-square method) without creation of a new neuron, or ignore this information completely.

In the course of planning activities of the system at the level of its formation approach to the forecasting with the help of the fuzzy logics and the Hurst exponent is usually implemented. Figure 1 shows the process of implementation of the method schematically.
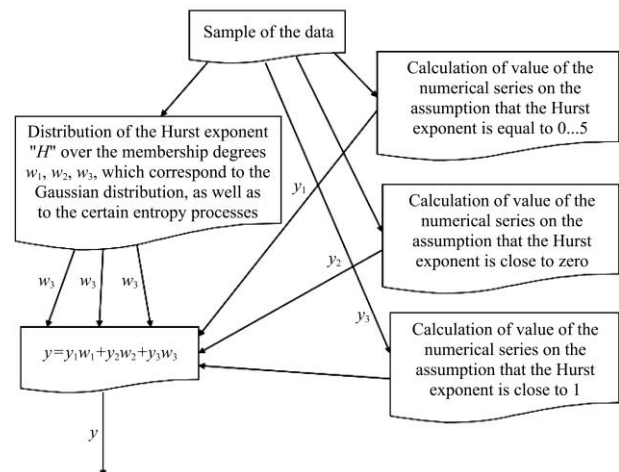

Figure 1　Process of forecasting with the help of the approach, which is based on the fuzzy logics and Hurst exponent

As it is evident from the above Figure, three independent models are constructed on the basis of the existing data sample from the very beginning of implementation of this method. Each of these models is based on its own assumption concerning degree of persistence of the process. Parameters of models are restored and outputs of these units are transferred to the unit of the result aggregation in accordance with these assumptions. In parallel to the above-described process, another process is performed: calculation of the Hurst exponent of the temporal series of the relevant process. The calculated value is subject to fuzzification. Fuzzification is to be implemented on the basis of the assumption that any process includes locally non-deterministic components, in respect of which it is not possible to state unambiguously that they are black, white, or pink noises.

Upon completion of the fuzzification procedure, final output is calculated as the weighted sum of the models, which were restored earlier in accordance with various assumptions in respect of the composition of the locally non-deterministic components. The degrees of membership of the process to one of three above-described kinds are precisely weights of the models.

Complexity of this approach is caused by the fact that this process (in the course of its functioning) can change its membership to one of three varieties in a gradual manner depending on various factors (which are frequently unknown). Therefore, in the case of change in membership of the process, it is necessary to analyse another kind of model and construct this model from the very beginning.

At the same time, on the basis of the assumption concerning fuzziness of the process type, it is possible to construct three "clean" models and then adapt them in a gradual manner (in the course of time). In this case, degrees of membership will demonstrate, to which of three models this process is "closer". Therefore, this method would have greater possibilities in the course of selection of models.

In order to ensure forecasting, this approach can be implemented in the following manner. The Hurst exponent is calculated on the basis of the temporal series of this process. Then, the model temporal series are constructed with the help of the inverse fast Fourier transform (as it was proposed by Crownover). They are constructed on three assumptions: the Hurst exponent is equal to 0; if it is equal to 0.5, then the forecasted value of the temporal series trend is assumed; if it is equal to 1, then the previous value is assumed as the forecasted value. At this stage, it is necessary to add the normalized first values, which were obtained from the model temporal series, to the current value of the process, which is subject to forecasting. Then fuzzy weighting is performed in accordance with the manner, which is already described above. At the next step, new real-life value is taken and in order to add next normalized model values, etc. Therefore, it is possible to call such an approach to forecasting as the adaptive method of forecasting, which is based on the decomposition of the noise component in accordance with the noise bases.

The investigations with the help of the above-described methods were performed in order to confirm one of the mutually exclusive hypotheses: is there a universal method of forecasting or each method is efficient in its own specific situation?

It was proved that the neural network (according to Mamdani model) is the best in accordance with value of the criterion, which estimates quality "on the average". The adaptive linear estimator (estimation algorithm) was the best in accordance with the values

of the criterion, which estimates maximum deviation.

Below we describe the process of the designing and implementation of the automatic system of the equipment monitoring and making decisions on the basis of the neural networks, which utilise the multiagent approach, as well as the theory of the fuzzy optimization. We have performed the comparative analysis of these systems in order to determine their horizons of maximum efficiency. Then we compare these automatic systems and the systems, which are known already. Let us call this system as the automatic system of making technological solutions (ASMTS). As a rule, the ASMTS consists of the units, which are presented in Figure 2.
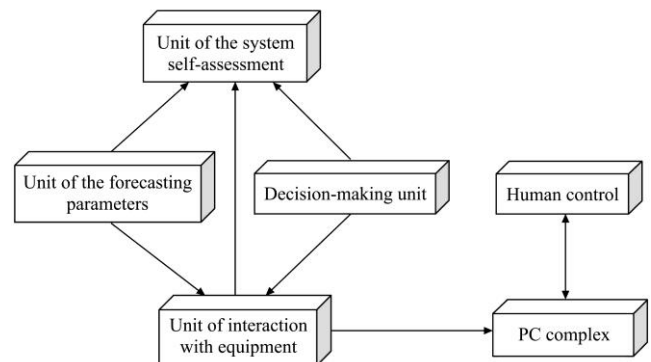

Figure 2   Generalized structural and functional diagram of the ASMTS

As it can be seen from the above Figure, the ASMTS must perform three actions: to forecast meaningfulness of the technological tools, with which this system works (Unit of analysis of parameters of the technological tools); to distribute available information concerning abnormal situations between the technological tools, with which this system works (Decision-making unit); as well as to estimate efficiency of operation of its components (Unit of the system self-assessment).

However, if we now would specify this diagram in detail, then the first ASMTS, which was developed, was the ASMTS on the basis of the fuzzy portfolio theory. Key component of this ASMTS is the diversification unit on the basis of the fuzzy theory. After the moment, when forecasts of values of the equipment components will be obtained (in the form of the fuzzy triangular numbers), the problem of the efficient distribution between the equipment components occurs. Theory of the fuzzy optimization is used with this object in view. The formal meaning of the optimization problem is as follows (23):

$$\sum_{i=1}^{n} r_i x_i \rightarrow \max_{x_i, i=1,...n}, \ \beta(x, r, r^*) \leq \beta_1$$

$$\sum_{i=1}^{n} x_i = 1, \ x \geq 0, \ \forall i = 1,...,n \qquad (23)$$

where, r is expected fault tolerance; β is Beta function; $\forall$ is universal quantifier.

The objective function includes the following components: r_i – expected fault tolerance of the i-th component of equipment in the forecast for the period of time T; x - share of the fault tolerance of the i-th component of equipment; n – the total quantity of the equipment components. Therefore, a criterion is established as follows: maximization of the expected period of operation of the entire equipment.

The first restriction means that there exists the following requirement: risk of the inefficient utilisation must not exceed a certain level. If values have membership functions of the triangular form, then it is possible to write down the risk function

in the following manner (24):

$$\beta(x,r,r^*) = \beta(x_1,\ldots,x_n, r_{min1},\ldots,r_{minn}, r_1,\ldots r_n, r_{max1},\ldots,r_{maxn}, r_{min}^{*^{*^{max}}})$$
(24)

The designations (r_{mini}, r_i, r_{maxi}) in the formula (24) as the fuzzy (triangular) number require special explanations. These designations are minimal value, the most expected value, and maximum value, respectively. (r_{min}^*, r^*, r_{max}^*) is the triangular fuzzy number of the criterial quantity, that is, the value of the fault withstand ability, below which relevant equipment component is considered as inefficient one.

Definition of the risk function is based on the probability of the parameter movement to the zone of the inefficient parameters of functioning.

Figure 9 presents arrangement of the membership functions of the criterial value of the portfolio yield and value of the portfolio yield, which is expected.

If we use $\varphi(\alpha)$ to denote the geometrical probability of the parameter movement to the zone of inefficiency depending on the level α, then it is obvious that the following formula is correct and valid:

$$\varphi(\alpha) = \frac{S_\alpha}{(r_2(\alpha) - r_1(\alpha))(r_2^*(\alpha) - r_1^*(\alpha))}$$
(25)

where, φ is arbitrary function; α is energy level; S is area.

Let us assume that "c" will hereinafter denote a certain constant value. In this case, level of the risk can be determined as the mathematical expectation of the parameter movement to the zone of inefficiency in respect to all levels of α (energy levels) with the help of the formula (26):

$$\beta = \int_0^1 p(\alpha)\varphi(\alpha)d\alpha$$
(26)

where, d is deutron.

In order to ensure that the risk function would be changed within the range from 0 to 1, it is necessary to introduce another normalizing condition (27):

$$\int_0^1 p(\alpha)d\alpha = 1$$
(27)

where, $p(\alpha)$ is probability density of the parameter movement to the energy level α.

Situation, where $p(\alpha)$ = const, is investigated sufficiently well already. In this case, formula (26) will be written as follows:

$$\beta = \int_0^1 p(\alpha)\varphi(\alpha)d\alpha = \int_0^1 \varphi(\alpha)d\alpha$$
(28)

Another two risk functions were proposed. The first of them is connected with the situation, where probability density of the parameter movement to the energy level is changed in proportion to value of the energy level. It is a natural situation, because of the higher speed is required in order to ensure achievement of the higher amplitude of oscillations, and this fact transforms the risk of the parameter movement to the zero level (the level, which requires the higher speed) into the zero risk. Such risk function is determined by the following formula:

$$\beta = \int_0^1 p(\alpha)\varphi(\alpha)d\alpha = \frac{\int_0^1 \alpha\varphi(\alpha)d\alpha}{\int_0^1 tdt} = 2\int_0^1 \alpha\varphi(\alpha)d\alpha$$
(29)

where, t is time.

In addition, the article has proposed the approach, which ensures minimization of operation with non-fuzzy values of the risk function. Essence of this approach is connected with the fact that for the situations, where $\varphi(\alpha)$ = 0 or $\varphi(\alpha)$ = 1, probability density is equal to zero ($p(\alpha)$ =0), that is, these situations are rejected as non-productive ones. In this case, if we will denote the significance level as α_1, where $\varphi(\alpha) \neq 0$ or $\varphi(\alpha) \neq 1$, then formula (26) will be written as follows:

$$\beta = \int_0^1 p(\alpha)\varphi(\alpha)d\alpha = \int_0^{\alpha_1} \frac{c}{\int_0^{\alpha_1} cdt}\varphi(\alpha)d\alpha = \frac{\int_0^{\alpha_1}\varphi(\alpha)d\alpha}{\alpha_1}$$
(30)

In the context of the fuzzy theory, formula (23) is commonly referred to as the direct problem. The problem, which is the dual problem in respect of the direct problem, is formulated as (31):

$$\beta(x,r,r^*) \to \min_x$$

$$\sum_{i=1}^n r_i x_i \geq r^*$$
(31)

$$\sum_{i=1}^n x_i, \ x_i \geq 0, \ \forall i = 1,\ldots,n$$

Specific feature of the formula (23) is connected with the fact that power of the set of the obtained solutions can exceed one, therefore it is inexpedient to apply the methods of the local optimization, which allow only one solution (unambiguous solution). In a number of articles, it was shown that the risk function for the situation, which was described in the present article, has a convex nature. This fact is the basis of the majority of the methods, which are used for optimization of the fuzzy making decision concerning expedience of performing operations with the help of the relevant equipment.

The problem (32) is a more practically significant one as compared with the previous problems:

$$\sum_{i=1}^n r_i x_i \to \max_x, \ \beta(x,r,r^*) \to \min_x$$
$$\sum_{i=1}^n x_i = 1, \ x_i \geq 0, \ \forall i = 1,\ldots,n$$
(32)

In order to solve the problem (32), we have used the method of solving the multicriteria nonlinear problems with the help of the linear additive convolution in order to reduce this problem to the single-objective problem. User of the ASMTS before beginning of operation with this system must define the weight for all criteria. Let us assume that w_1 is the weight of the criterion of efficiency, while w_2 is the risk criterion weight, $w_1 \geq 0$, $w_2 \geq 0$, $w_1 + w_2 = 1$. Independences are solved with the help of the optimization problems in the absence of one of the criteria. Let us assume that after solving the optimization problems we have obtained the minimal value of the fault tolerance r_{min}, the maximum value of the fault tolerance r_{max}, the minimal value of the risk level β_{min}, the maximum value of the risk level β_{max}. Then problem of optimization of the fuzzy logics (32) may be reduced to the form (33):

$$\frac{r_{max} - \sum_{i=1}^n r_i x_i}{r_{max} - r_{min}}w_1 + \frac{\beta(x,r,r^*) - \beta_{min}}{\beta_{max} - \beta_{min}}w_2 \to \min_{x_i, i=1,\ldots,n}$$
$$\sum_{i=1}^n x_i = 1, \ x_i \geq 0, \ \forall i = 1,\ldots,n$$
(33)

This problem can be solved with the help of the nonlinear programming methods already. It is solution of the formula (33), which is the basis of the Decision-making unit (Diversification unit) of the ASMTS on the basis of the fuzzy theory.

The ASMTS on the basis of the multiagent approach excludes the centrally-controlled distribution. To be more accurate, from the very beginning of operation of the ASMTS, equipment components are distributed among the agents (such distribution is developed with the help of the certain tool and within a certain

timeframe). In the course of implementation of the next periods, each component of equipment makes its own decision in the course of operation: either to use the data, which it has, autonomously or to transfer these data to the more efficient decision-making center. Therefore, process of the equipment functioning is not distributed by the single decision-making unit in respect of the entire system; such decisions are made by the distribution unit for each specific component of equipment.

In more detail, there exists certain quantity of the decision-making centers within the system, and each center operates at its own horizon and with the help of its tools. They interact with each other with the help of the database (MySQL) and memory areas of Windows are divided between them. They provided each other with the information concerning their forecasted periods of the information provision, their forecasted fault tolerance, as well as the forecasted level of the risk. Then the information in the form of the set of such triads, which was obtained by each equipment component (which plans to make a decision), makes it possible to distribute relevant decision, thus having been solved the optimization formula (34).

$$\max \frac{1}{S_j} \sum_{i=1}^{n} r_i L_i x_i$$

$$\min \frac{1}{S_j} \sum_{i=1}^{n} \delta_i L_i x_i$$

$$\min \frac{1}{S_j} \sum_{i=1}^{n} t_i L_i x_i$$

$$\sum_{i=1}^{n} L_i x_i \le S_j, \ x_i \in N \cup \{0\}, \ i=1,...,n \qquad (34)$$

where, L is length; δ is inexact differential; N is normal response force.

It is this system, which is capable to make decision that not only behavior of certain units or machines will be designed and forecasted, but all decisions (on the whole), which will define potential fault tolerance of equipment will be designed and forecasted as well.

4 Conclusions

These experiments have demonstrated that each of the forecasting methods under investigation has been better than other methods at the certain values of the experimental parameters. Quality of the forecasting systems is formally decreased along with the growth of the horizon number; however, range of the tolerable errors in the course of forecasting is increased at the same time. Therefore, it is possible to say with reasonable confidence that methods of forecasting work at all horizons, which have been investigated, with the same efficiency. Quality of forecasting depends on the specific tool, which is analysed. Reason for such dependence is connected with different volatility of the tools. Forecasting of increments have ensured the better values of the quality criteria as compared with the method of forecasting values as such, and it is possible to explain this fact as follows: this process is closer to the stationary conditions. One-step method of forecasting is more efficient (from the practical point of view) as compared with the multistep method of forecasting, because of difference in the quality of forecasting for both these methods was not essential. For example, utilisation of the network (Mamdani model): the single-step method has ensured value of the absolute criterion at the level of 2.61, at the time when the multistep method has ensured this value at the level of 3.01), while for utilisation of the single-step method it is necessary to have only one neural network after relevant learning (for utilisation of the multistep method it is necessary to ensure learning process several times).

Multiple applications of implementations of the forecasting methods have caused occurrence of the problem of the learning stability for many of these methods. Each method has its own techniques for overcoming this problem. Particularly, as concerns, the radial basis neural network, the necessity of analysis of increments (analysis of values as such) through the problem of oversaturation was experimentally confirmed. Similar problem exists also for the TSK network with trigonometric polynomials, but in accordance with the approach, which is described in the article, the input data are normalised automatically.

As concerns quantity of inputs of the model "on the average", it is possible to state that quantity of the model inputs at the level from 3 to 5 was sufficiently efficient parameter practically for all methods of forecasting (except for the method of forecasting with the help of the model of fuzzy chaotic oscillations).

[References]

[1] Bigné E, Aldas-Manzano J, Küster I, Vila N. Mature market segmentation: a comparison of artificial neural networks and traditional methods. Neural Computing and Applications, 2010; 19(1): 1–11. https://doi.org/10.1007/s00521-008-0226-y.

[2] Kovalevskyy, Sergiy, Olena Kovalevska, and Raul Turmanidze. Acoustic diagnostics of lever mechanisms with subsequent processing of data on neural networks. In New Technologies, Development and Application, ed. Isak Karabegović. Cham: Springer International Publishing, 2019; pp.202–210.

[3] Stapelberg R F. Safety and risk in engineering design. In Handbook of Reliability, Availability, Maintainability and Safety in Engineering Design, London: Springer London, 2009; pp.529–798. https://doi.org/10.1007/978-1-84800-175-6_5.

[4] Gotovtsev P M, Voronov V N, Smetanin D S. Analysis of the coolant condition with the help of artificial neural networks. Thermal Engineering, 2008; 55(7): 552–557. https://doi.org/10.1134/S0040601508070045.

[5] Meshalkin V P, Bol'shakov A A, Petrov D Y, Krainov O A. Algorithms and software system for controlling the quality of glass batch using artificial neural networks." Theoretical Foundations of Chemical Engineering, 2012; 46(3): 284–87. https://doi.org/10.1134/S0040579512030062.

[6] Rubtsov V P, Shcherbakov A V, Solov'eva E V. On the use of fuzzy logic in control systems for technological equipment. Russian Electrical Engineering, 2016; 87(4): 219–222. https://doi.org/10.3103/S1068371216040131.

[7] Guo Y J, Kang Q, Wang L, Wu Q D. Data-based state forecast via multivariate grey RBF neural network model. In Advances in Swarm Intelligence, eds. Ying Tan, Yuhui Shi, and Carlos A Coello Coello. Cham: Springer International Publishing, 2014; pp.294–301.

[8] Wu W, Liu M, Liu Q, Shen W M. A quantum multi-agent based neural network model for failure prediction. Journal of Systems Science and Systems Engineering, 2016; 25(2): 210–228. https://doi.org/10.1007/s11518-016-5308-2.

[9] Bakharev V P, Kulikov M Y, Nechaev D A, Yakovchik E V. Optimizing the diamond machining of ceramics on the basis of systemic analysis using neural networks. Russian Engineering Research, 2008; 28(12): 1183–1187. https://doi.org/10.3103/S1068798X08120071.

[10] Lisboa P J G, Vellido A, Wong H. Outstanding issues for clinical decision support with neural networks. In Artificial Neural Networks in Medicine and Biology, eds. Helge Malmgren, Magnus Borga, and Lars Niklasson. London: Springer London, 2000; pp.63–71.

[11] Arranz N, de Arroyabe J C F. A model to analyse governance structures in technological networks. In Strategy and Governance of Networks: Cooperatives, Franchising, and Strategic Alliances, eds. George Hendrikse, Mika Tuunanen, Josef Windsperger, and Gérard Cliquet. Heidelberg: Physica-Verlag HD, 2008; pp.249–68. https://doi.org/10.1007/978-3-7908-2058-4_14.

[12] He X G, Xu S H.   Application of process neural networks.   In Process Neural Networks: Theory and Applications, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010; pp.195–232.   https://doi.org/10.1007/978-3-540-73762-9_9.

[13] Vulfin A M, Frid A I, Giniyatullin V M.   Neuralbase Model for Detection and Recognition of Technological Situations within the Scope of Data Mining Strategy.   Optical Memory and Neural Networks, 2010; 19(3): 207–12.   https://doi.org/10.3103/S1060992X1003001X.

[14] Gerasimov A V, Vasil'kov Y V.   Imbalance detection and classification system based on wavelet analysis and artificial neural networks.   Automation and Remote Control, 2013; 74(11): 1883–1889.   https://doi.org/10.1134/S0005117913110106.

[15] Sung H-Y, Yeh H-Y, Lin J-K, Chen S-H.   A visualization tool of patent topic evolution using a growing cell structure neural network.   Scientometrics, 2017; 111(3): 1267–1285.   https://doi.org/10.1007/s11192-017-2361-7.

[16] Tugengol'd A K, Luk'yanov E A, Remizov E V, Korotkov O E.   Intelligent control of technological systems.   Russian Engineering Research, 2008; 28(5): 479–484.   https://doi.org/10.3103/S1068798X08050158.

[17] Lütjering G, Williams J C.   Technological Aspects.   In Titanium, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007; pp.53–173.   https://doi.org/10.1007/978-3-540-73036-1_3.

[18] Li X F, Wang L.   The establishment of rough-ANN model fordynamic risk measure of enterprise technological innovation and its application.   In Proceedings of the Sixth International Conference on Management Science and Engineering Management, eds. Jiuping Xu, Masoom Yasinzai, and Benjamin Lev. London: Springer London, 2013; pp.65–75.

[19] Schmitt R.   Self-optimising production systems.   In Integrative Production Technology for High-Wage Countries, ed. Christian Brecher. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012; pp.697–986.   https://doi.org/10.1007/978-3-642-21067-9_6.

[20] Stepanov P, Nikitin Y.   Diagnostics of mechatronic systems on the basis of neural networks with high-performance data collection.   In Mechatronics 2013, eds. Tomáš Bvrezina and Ryszard Jabloński.   Cham: Springer International Publishing, 2014; pp.433–440.

[21] Shevgunov T.   Using artificial neural networks for time difference of arrival target localization based on reduced discrete cosine transform.   Periodico Tche Quimica, 2019; 16(33): 530–40.

[22] Shevgunov T, Efimov E, Kirdyashkin V.   Scattering target identification based on radial basis function artificial neural networks in the presence of non-stationary noise.   Periodico Tche Quimica, 2019; 16(33): 573–582.

[23] Kamenskikh M.   Assessment of cluster and network collaboration influence on regional economy.   Journal of Advanced Research in Law and Economics, 2018; 9(2): 510–515.

[24] Krechetov I V, Skvortsov A A, Poselsky I A, Paltsev S A, Lavrikov P S, Korotkovs V.   Implementation of automated lines for sorting and recycling household waste as an important goal of environmental protection.   Journal of Environmental Management and Tourism, 2018; 9(8): 1805–1812.

[25] Buyankin V M.   Neuroidentification with neuro-self tuning to ensure the operation of the current loop of the electric drive with the desired static and dynamic characteristics.   Periodico Tche Quimica, 2018; 15(30): 513–519.