http://doi.org/10.35784/iapgos.7321 *received: 21.02.2025 | revised: 08.04.2025 | accepted: 30.05.2025 | available online: 27.06.2025*

HETEROGENEOUS ENSEMBLE NEURAL NETWORK FOR FORECASTING THE STATE OF MULTI-ZONE HEATING FACILITIES

Maria Yukhymchuk[1], Volodymyr Dubovoi[1], Zhanna Harbar[2], Bakhyt Yeraliyeva[3]
[1]Vinnytsia National Technical University, Department of Computer Control Systems, Vinnytsia, Ukraine, [2]Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University, [3]M. Kh. Dulaty Taraz Regional University, Taraz, Kazakhstan

Abstract. *The research is aimed at increasing the accuracy of forecasting the state of multi-zone thermal facilities. Such facilities include multi-room premises, multi-zone greenhouses, tunnel kilns for brick production, and others. The high inertia of such facilities reduces the effectiveness of "ad hoc control". Modern proactive control systems based on forecasting are mainly based on using neural network training. However, to forecast the state of a specific multi-zone thermal facility, training the network requires a very large dataset, which is difficult to create and use. A combined neuro-structural method for forecasting the state of multi-zone thermal facilities is proposed, in which the structure of the neural model reflects the structure of the mutual influence of the facility zones. The research of the method has shown the possibility of ensuring sufficiently high forecast accuracy with a smaller size of the training dataset.*

Keywords: forecasting, multi-zone facility, neuro-structural method, training, dataset

HETEROGENICZNA SIEĆ NEURONOWA DO PROGNOZOWANIA STANU WIELOSTREFOWYCH OBIEKTÓW GRZEWCZYCH

Streszczenie. *Badania mają na celu zwiększenie dokładności prognozowania stanu wielostrefowych obiektów cieplnych. Obiekty takie obejmują obiekty wielopokojowe, wielostrefowe szklarnie, piece tunelowe do produkcji cegieł i inne. Duża bezwładność takich obiektów zmniejsza skuteczność "sterowania ad hoc". Nowoczesne proaktywne systemy sterowania oparte na prognozowaniu opierają się głównie na szkoleniu sieci neuronowych. Jednak w celu prognozowania stanu konkretnego wielostrefowego obiektu termicznego, szkolenie sieci wymaga bardzo dużego zbioru danych, który jest trudny do utworzenia i wykorzystania. Zaproponowano połączoną neurostrukturalną metodę prognozowania stanu wielostrefowych obiektów cieplnych, w której struktura modelu neuronowego odzwierciedla strukturę wzajemnego wpływu stref obiektu. Badania metody wykazały możliwość zapewnienia wystarczająco wysokiej dokładności prognozy przy mniejszym rozmiarze zbioru danych treningowych.*

Słowa kluczowe: prognozowanie, obiekt wielostrefowy, metoda neurostrukturalna, trening, zbiór danych

Introduction

The problem of various processes forecasting has many methods and examples of solutions. Most of them are devoted to problems of a non-technical nature, mainly socio-economic and political. In technical applications, forecasting is necessary for proactive control (preemptive control) of the state, resources, safety, etc. In particular, forecasting the thermal state is important for thermal facilities due to their high inertia. Such forecasting allows to reduce the average cost of thermal energy.

Most works on the problem of forecasting consider methods that use expert assessments and conclusions in one way or another. This mainly concerns socio-economic processes, forecasts in medicine, education, etc. [2, 3]. Our work considers forecasts for multi-zone thermal facilities, which are cyber-physical systems, and although models of physical processes and formal forecasting methods play a much greater role here, expert assessments are also used at certain stages, in particular regarding the ranges of possible changes in parameters, the list of influential factors, etc.

External influences on technical systems in forecasting problems are represented as time series. Time series forecasting can be carried out on the basis of deterministic, stochastic and fuzzy models.

There are many time series forecasting models, but most of them are based on a limited set of basic methods and approaches and their combinations. Methods of time series analysis and forecasting are implemented in various data analysis packages, specialized programming languages, and are described both in theoretical studies and in documentation.

In many cases, the concept of "forecast" of a time series is replaced by the concept of "extrapolation" [1, 5]. Extrapolation as a forecasting method is a mathematical procedure that allows you to calculate the next value of a process based on previous values. A distinction is made between process extrapolation and trend extrapolation. Process extrapolation is carried out on the basis of a small number of the last members of the time series. Trend extrapolation is carried out by averaging a larger number of data.

The moving average method [12, 14] is one of the most common methods of smoothing time series. It allows to reduce random fluctuations and obtain values corresponding to the influence of the main factors. The forecast value is obtained as the average of a certain number of the latest values of the time series.

Exponential smoothing [9] is most commonly used in practical applications and has been described in many papers. The model is based on the calculation of the so-called exponential average. The model differs from the moving average in that it reduces the weight of previous data in proportion to their distance in time from the forecast interval.

Most statistical forecasting methods are based on linear univariate or multivariate regression.

The Kalman filter is an efficient recursive statistical filter that estimates the state vector of a dynamic system using a series of incomplete and noisy measurements [8,13]. The filter operation is divided into two stages: extrapolation and adjustment. The presence of the extrapolation stage allows, under certain assumptions, to use it to predict the state of a stationary object (process).

The ARMA(p, q) model is one of the most frequently used for forecasting [9]. It consists of two parts: autoregressive (AR); moving average (MA), where p is the order of the regression part, and q is the order of the moving average.

Vector autoregression (VAR) [10] is a model of the dynamics of multiple time series in which the current values of each series depend on the previous values of this and other time series of the vector.

Support Vector Machine (SVM) [15] uses a linear classifier that is trained on support data sequences (vectors). The basis of using the SVM algorithm for prediction is the transfer of the original data to a higher-dimensional space, where each input vector is formed as a sequence of output points with time delays. These vectors are then used as support samples, from the combination of which a regression hyperplane is calculated that describes the data distribution with a given accuracy.

Forecasting the state of multi-zone thermal facilities (MTF) is complicated compared to common forecasting problems by the large dimensionality of the problem. First, the state of individual objects of MTF is influenced by the temperature of the external environment, the power of internal heat sources, the state control system, the heat balance of internal technological processes, the state of neighboring objects, etc. Secondly, since MTF consists of several objects, its state is a vector process, and its components influence each other. The inertia of the MTF also has an impact on the complexity of the problem, as a result of which, when forecasting, it is necessary to take into account $k = n_{\max} \cdot T_{\max} / T_{\min}$ previous data, where $n_{\max}$ is the maximum order of the transfer function of the MTF object; $T_{\max}$, $T_{\min}$ are the maximum and minimum time constants of the MTF objects. Therefore, when using regression forecasting methods, as, for example, in [4], the correlation matrix will contain $0.5(k \cdot n \cdot m + n - 1)^2$ independent pairwise correlation coefficients.

The current trend of using neural networks for process forecasting allows simplifying the task [7]. However, the task of training neural predictors remains difficult due to the lack of datasets for each MTF. In a previous article [6], the authors considered the issue of effective training of neural coordinators of MTF control systems, which also have a forecasting function. But the problem of choosing the optimal structure of the predictor and the MTF state monitoring system requires further research.

The aim of this work is to choose an effective structure of the predictor and the MTF state monitoring system for increasing the accuracy of forecasting the state of multi-zone thermal facilities.

1. Problem statement

The method of predicting the state of the MTF depends on the spatiotemporal characteristics of the impacts on the facility in the $[\mathbf{Z}, t]$ coordinate system. Most facilities are simultaneously exposed to several impacts with different spatial distributions. Fig. 1a shows the main impacts on the MTF object: $u(\mathbf{Z_0}, t)$ is the state of the external environment, where $\mathbf{Z_0}$ is the spatial coordinates of the MTF object; x is the flow of raw materials for the technological process in the MTF object; $p_0(\mathbf{Z_0}, F, t)$ is the impact of the local control system (LCS), where F is the specified state of the object; $\mathbf{P}_r(\mathbf{Z_0}, t)$ is the vector of impacts from neighboring MTF objects. Fig. 1b shows an example of the mutual impacts of objects for the structure of their arrangement 2×2. The MTF object produces the result of the internal technological process and affects neighboring objects with intensity p_f^-.

These influences change the dynamic processes in the distributed facility. The LCS maintains the specified state of the MTF object. The predictor predicts the change in the state of the object under certain control.

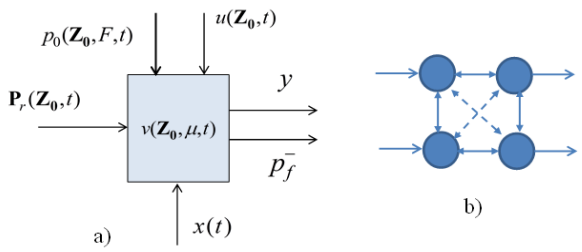


Fig. 1. MTF structure: a) MTF object; b) Example of interaction of the MTF objects

Taking into account the described properties of MTF, the authors used the Heterogeneous Ensemble Neural Network (HENN) to predict the state of its objects, the structure of which is shown in Fig. 2.

In this structure, the prediction of independent influences *F(t+τ)*, *u(t+τ)*, *x(t+τ)* is performed by separate predictors as the prediction of scalar time series. For the experiment, the authors chose an LSTM network, which allows predicting both random processes without trends and taking into account long time trends. To predict *F(t+τ)* and *x(t+τ)*, which do not have pronounced trends, a single LSTM layer with $3 \cdot k$ neurons was used. To predict *u(t+τ)* (ambient temperature), which has pronounced daily and seasonal trends, a 3-layer LSTM was used.

The prediction of the initial state of the MTF object is carried out using NN Elman, which models the dynamics of the object controlled by LCS. LCS of a thermal object with a PID controller has a 3rd order, so NN Elman must have 3 hidden layers.

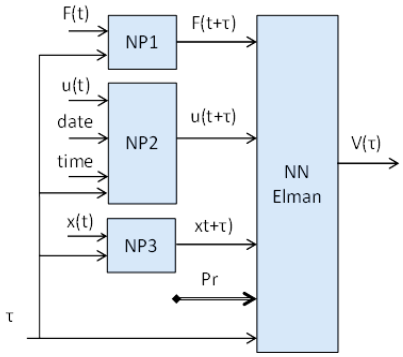


Fig. 2. Predictor structure for state of an object

Taking into account the mutual influences of MTF objects is carried out by combining predictors according to the Hopfield network scheme, as shown in Fig. 3.

With such a connection $\mathbf{P}_r = \mathbf{V}_\varepsilon$, where $\mathbf{V}_\varepsilon$ is the vector of states ε-environment, i.e. neighboring objects that significantly affect the state of the MTF object [13]. To determine the environment of the objects and the corresponding predictor structure, we will construct a spatiotemporal model of the statistical characteristics of the MTF states.

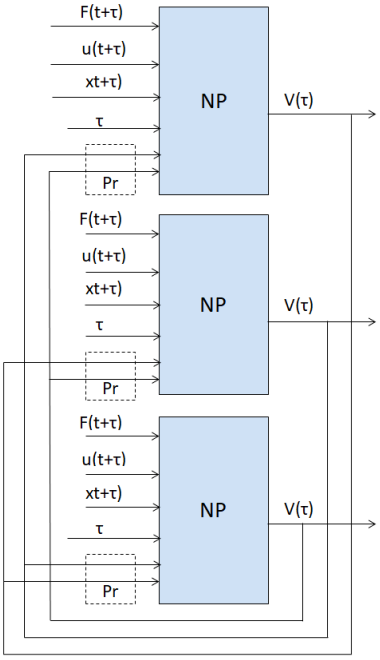


Fig. 3. Model structure for predicting the state of MTF

2. MTF spatio-temporal model

To determine the parameters of the neural model (number of inputs, number of hidden layers of the Elman network, number of neurons in the recursive layers), we will build a spatiotemporal model of the MTF.

Let us consider the dynamic processes in the MTF as the interaction of three fields:

- field of object parameters $v(\mathbf{Z}_0,t)$, where v is a state of the facility's object; $\mathbf{Z}=\{z_k\}$ is a vector of object coordinates;
- field of control influences $\mathbf{p_0}(\mathbf{Z}_0,t)$;
- field of disturbances $\mathbf{u}(\mathbf{Z_0},t)$.

In addition to the impacts, the state of the object also depends on the volume of production, which is determined by the amount of raw materials x and the state v.

In Fig. 1a, in addition to the mentioned fields, heat costs μ for production processes are shown, as well as the interaction of the object with other objects of the facility.

In [13], the MTF model was developed and investigated.

The magnitude of the resource flow between two objects with coordinates $(\mathbf{Z_0},\mathbf{Z_j})$

$$p_j(\mathbf{Z_0},\mathbf{Z_j},\tau_{0j})=\frac{v(\mathbf{Z_0})-v(\mathbf{Z_j})}{8\left(\pi\lambda\tau_{0j}\right)^{3/2}}e^{-\frac{|\mathbf{Z_0}-\mathbf{Z_j}|^2}{4\lambda\tau_{0j}}} \tag{1}$$

where the parameter designations are deciphered in [7].

Since the prediction is based on data coming from sensors, the discrete form is more convenient. In linear facilities, the flows are additive, so

$$P_f(\mathbf{Z_0})=\sum_{j:\mathbf{Z_j}\in\varepsilon(\mathbf{Z_0})}\frac{v(\mathbf{Z_0})-v(\mathbf{Z_j})}{8\left(\pi\lambda\tau_{0j}\right)^{3/2}}e^{-\frac{|\mathbf{Z_0}-\mathbf{Z_j}|^2}{4\lambda\tau_{0j}}} \tag{2}$$

The equivalent structural diagram of the object is shown in Fig. 4.

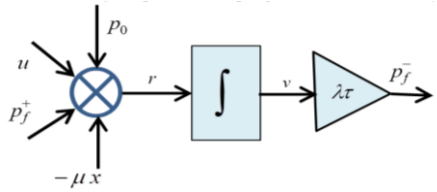


Fig. 4. Equivalent structural diagram of the object

The uncertainty of the state of the object is due to the following factors:

- uncertainty of the influence of the external environment;
- uncertainty of the state of the surrounding objects that influence the object under consideration;
- uncertainty of the parameters of the environment of influence distribution;
- uncertainty of production program and unit costs μ.

The influence of the external environment and surrounding objects is carried out through the heat flow. Let's estimate the dispersion of the state of the object object.

$$\sigma_v^2(\mathbf{Z_0},t)=\sigma_{v(\mathbf{Z_0},0)}^2+\int_0^\infty\int_0^\infty G_{rr}(\omega,\boldsymbol{\Psi})e^{i\boldsymbol{\Psi}Z_0}d\boldsymbol{\Psi}d\omega \tag{3}$$

where G_{RR} is the auto-spectral density of the heat flux entering or leaving the object; ω, $\boldsymbol{\Psi}$ are temporal and spatial circular frequencies, respectively.

In linear facilities, the power spectral densities are additive, taking into account mutual correlation. Therefore, the balance of the heat flux spectral densities

$$\begin{cases}
G_{p_0v_0}(\omega,\boldsymbol{\Psi})=G_{v_0v_0}(\omega,\boldsymbol{\Psi})\cdot W_{v_0p_0}(\omega)\\
G_{rr}(\omega,\boldsymbol{\Psi})=G_{p_0v_0}(\omega,\boldsymbol{\Psi})+G_{uu}(\omega,\boldsymbol{\Psi})+G_{xx}(\omega,\boldsymbol{\Psi})+\\
+\sum_{f\in\varepsilon}G_{p_fp_f}+2G_{p_0u}(\omega,\boldsymbol{\Psi})+2G_{p_0x}(\omega,\boldsymbol{\Psi})+\\
+2\sum_{f\in\varepsilon}G_{p_0p_f}(\omega,\boldsymbol{\Psi})-2G_{ux}(\omega,\boldsymbol{\Psi})+\\
+2\sum_{f\in\varepsilon}G_{up_f}(\omega,\boldsymbol{\Psi})-2\sum_{f\in\varepsilon}G_{p_fx}(\omega,\boldsymbol{\Psi})\\
G_{vv}(\omega,\boldsymbol{\Psi})=G_{rr}(\omega,\boldsymbol{\Psi})\cdot\left(\frac{1}{T\omega}\right)^2\\
G_{p_fp_f}(\omega,\boldsymbol{\Psi})=\int_0^\infty\left\{\iiint_\Omega\left[R_{p_fp_f}(\tau,\mathbf{Z})e^{-j\boldsymbol{\Psi}\mathbf{Z}}\right]e^{-j\omega\tau}d\mathbf{Z}\right\}d\tau
\end{cases} \tag{4}$$

where G_{rr} is the auto-spectral power density of the total input heat flow; $G_{p_0p_0}$ is the auto-spectral power density of the control influence (external heat flow); G_{uu} is the auto-spectral power density of disturbances; G_{vv} is the auto-spectral power density of the state (stored heat); G_{xx} is the auto-spectral power density of heat use and/or dissipation; $G_{p_{f0}p_{f0}}$ is the auto-spectral power density of the input heat flow taking into account the distribution; G_{ux} is the mutual spectral power density of the disturbance and raw materials.

The influence of the external environment and the consumption of raw materials on production can be considered independent, therefore $G_{ux}=0$.

The mutual power spectral density of the control influence and the input resource coming from neighboring objects $G_{p_0p_f}$ is determined by the LCS algorithm and parameters. In the simplest case, the controller calculates the influence

$$p_0=\mu x-k\left(\sum_{f\in\varepsilon}p_f-u\right) \tag{5}$$

where $p_f=p_f^+-p_f^-$, $k=\dfrac{W_0W_{LCS}}{1+W_0W_{LCS}}$ is the coefficient of attenuation of the external environment influence by the influence of LCS, after which it solves the optimization problem. The presence in the calculation algorithm of a large number of factors affecting the result leads to a decrease in the mutual pair correlation, therefore we can consider $G_{p_fx}\approx 0$ and $G_{up_f}\approx 0$.

To obtain the time dependence $R_{p_0x}(\tau,Z_0)$, let we consider the structural diagram of the local control.

If a PID controller with time constants T_1 and T_2 is used in the LCS and the object has an integrator transfer function with time constant T_0, then the power spectral densities are related by the relation

$$\begin{aligned}G_{p_0x}(\omega,Z_0)&=G_{xx}(\omega,Z_0)\cdot W_{x_0\to p_0}(j\omega)=\\&=G_{xx}(\omega,Z_0)\cdot\frac{1-T_2T_0\omega^2+T_0j\omega}{(1+T_1j\omega)(1+T_2j\omega)T_0j\omega+(1+T_1j\omega)+1}\end{aligned} \tag{6}$$

The spatial component of spectral densities is determined by the geometric characteristics of the MTF and the location of the points of application of the control influence. The influence of external disturbance can be considered as white noise in time, the amplitude of which is maximum at the boundaries of the facility and exponentially decreases when separated from the boundaries.

The results of modeling of the MTF control process shown in Fig. 5, show that the spatial distribution of the resource is the sum of exponentially decreasing functions with modes at the points of application of the influence, i.e.

$$G_{uu}(\omega,\Psi)=G_{uu}(\omega,\Psi\notin S)\cdot e^{-\frac{1}{4\lambda\tau}\left[\min\limits_{\forall\Psi_\Omega}(\Psi_\Omega-\Psi)\right]}= \\ =g_u e^{-\frac{1}{4\lambda\tau}\left[\min\limits_{\forall\Psi_\Omega}(\Psi_\Omega-\Psi)\right]} \tag{7}$$

where Ω is a boundary of the distributed facility (S area); g_u is the amplitude of external noise;

$$G_{p_0p_0}(\omega,\Psi)=\sum_i G_{p_0p_0}(\omega,\Psi_{0i})\cdot e^{-\frac{1}{4\lambda\tau}(\Psi_{0i}-\Psi)} \tag{8}$$

where Ψ_{0i} is the coordinate vector of the *i*-th point of application of the control influence.

In the stationary mode of using a multi-zone facility, the production task does not change over time, therefore $G_{xx}(\omega,\Psi)=G_{xx}(\omega=0,\Psi)$.

Let us find the spatiotemporal correlation function of the resource. Based on the Wiener-Khinchin theorem, we write

$$K_{vv}(\tau,\mathbf{Z})=2\int_0^\infty\int_0^\infty G_{vv}(\omega,\Psi)\cdot e^{i\omega\tau}\cdot e^{i\boldsymbol{\Psi}\mathbf{Z}}d\omega d\Psi \tag{9}$$

Similarly, other spectral power densities of the model are transformed into correlation functions.

The correlation function (9) allows us to estimate the spatial interval of correlation of the states of the MTF objects. It is the basis for determining the structure of the Hopfield model connections in Fig. 3.

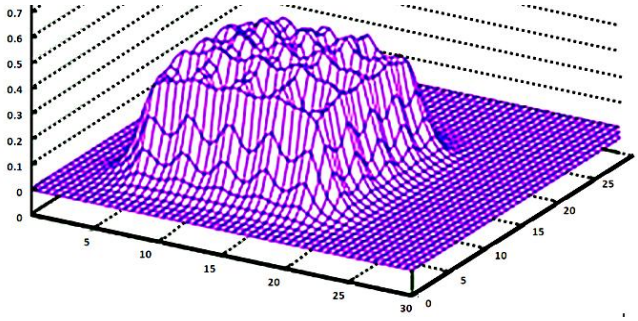


Fig. 5. Result of modeling the MTF control process

3. Numerical experiments and results

Forecasting was studied on a model in the Scilab environment. MTF with four linearly arranged controlled elements was simulated.

The parameters for setting up the simulation model are:

- coefficients of connection between elements,
- coefficient of influence of the external environment,
- time constants of integrator objects,
- specific resource costs μ.

External influence $u(t)$ is common to all elements. The process graphs of the model are shown in Fig. 6: 1 – change in the state of the external environment $u(t)$; 2 – change in the input of raw materials, the processing of which requires resources, for element $i=1$; 3 – change in the input of raw materials for element $i=3$; 4 – change in the input of raw materials for element $i=4$; 5 – change in the input of raw materials for element $i=2$; 6 – state of element $i=2$.

As a basis for comparison, the forecast by the method of spatiotemporal regression for element $i=2$, surrounded by elements $i=1$ and $i=3$, is used. There is also a remote element $i=4$.

In Fig. 6, graph 7 is the forecast of the state of the element $i=2$; 8 is the forecast error.

Taking into account the linearity of the dependence of the state of elements on the influencing factors in model (5), as well as the correlation between the amount of raw materials x and the controlling influence p_0, we write the predicted value of the state of the i-th element over a period of time τ in the form of a multivariate regression equation on independent factors

$$v(\tau,\mathbf{Z}_i)=v(0,\mathbf{Z}_i)+K_{vu}(\tau,\mathbf{Z}_i)\sqrt{\frac{K_{vv}(0,\mathbf{Z}_i)}{K_{uu}(0,\mathbf{Z}_i)}}\cdot \\ \left[u(0,\mathbf{Z}_i)-m_u(\mathbf{Z}_i)\right]+K_{vx}(\tau,\mathbf{Z}_i)\sqrt{\frac{K_{vv}(0,\mathbf{Z}_i)}{K_{xx}(0,\mathbf{Z}_i)}}\cdot \\ \left[x(0,\mathbf{Z}_i)-m_x(\mathbf{Z}_i)\right]+\sum_{j\in\varepsilon(\mathbf{Z}_i)}\left\{\gamma(\mathbf{Z}_i-\mathbf{Z_j})K_{vu}(\tau+\tau_{0j},\mathbf{Z}_j)\cdot\right. \\ \left.\cdot\sqrt{\frac{K_{vv}(0,\mathbf{Z}_j)}{K_{uu}(0,\mathbf{Z}_j)}}\left[u(\tau-\tau_{0j},\mathbf{Z}_j)-m_u(\mathbf{Z}_j)\right]\right\}+ \\ +\sum_{j\in\varepsilon(\mathbf{Z}_i)}\left\{\gamma(\mathbf{Z}_i-\mathbf{Z_j})K_{vx}(\tau+\tau_{0j},\mathbf{Z}_j)\sqrt{\frac{K_{vv}(0,\mathbf{Z}_j)}{K_{xx}(0,\mathbf{Z}_j)}}\cdot\right. \\ \left.\cdot\left[x(\tau-\tau_{0j},\mathbf{Z}_j)-m_x(\mathbf{Z}_j)\right]\right\} \tag{10}$$

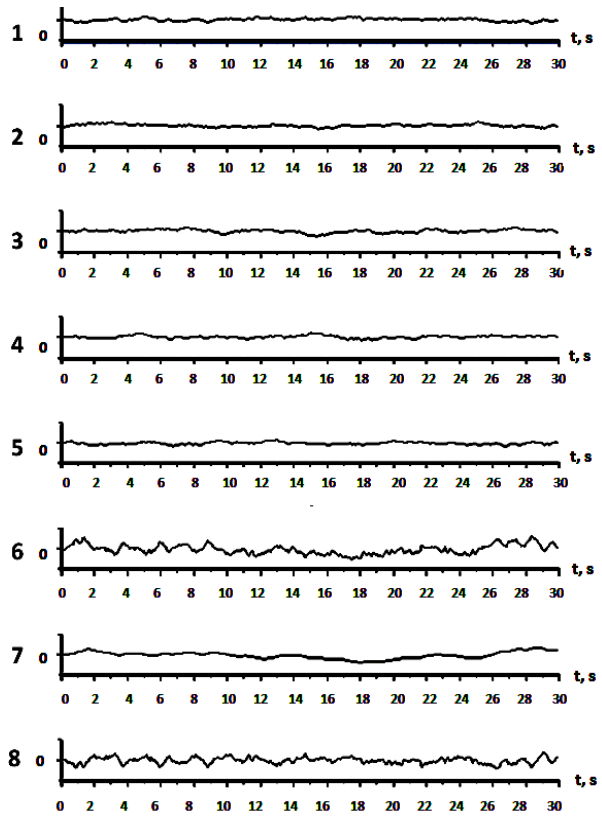


Fig. 6. MTF simulation results

The results of the RMSE study using the spatiotemporal regression method for different sets of input data $\{x,p,u\}$ are shown in Fig. 7. The dependence of RMSE on the scatter of the values of the given state σ_F for the combination of parameters u – state of the external environment; k – coefficient of mutual influence in formula (5) is shown in Fig. 8. The significant influence of input data on RMSE leads to the conclusion that in order to obtain a reliable forecast, tuning on a large sample of input data is necessary with subsequent averaging of the parameters.

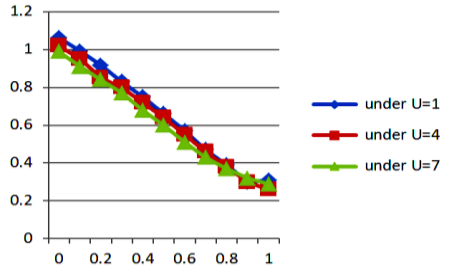


Fig. 7. Dependence of RMSE on attenuation during influence propagation

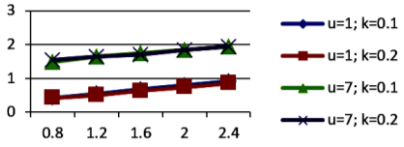


Fig. 8. Dependence of RMSE on the spread of values of a given state σ_F

4. Determining the appropriate depth of forecasting

As the depth (interval) of the forecast increases, the uncertainty of the forecast results increases. At the same time, time and computational resources are spent on forecasting. Gradually, a situation arises when the positive effect of forecasting becomes less than the costs of its implementation. This determines the appropriate maximum depth of forecasting.

It is also necessary to take into account the duration of transient processes in the system when predicting transitions to other operating modes of the system. Fig. 9 shows transient processes when changing the mode: Fig. 9a – when changing the state of the external environment, which acts on all external elements; Fig. 9b – when changing the amount of raw materials supplied to one internal element. Simulation shows that the more elements are affected by the change in operating mode, the more time is required to achieve a stationary mode. This time determines the appropriate minimum depth of prediction.

Let us estimate the appropriate depth of the forecast when forecasting by the spatiotemporal regression method.

Residual variance of linear forecast

$$D_T = D_0\left(1-k_{vv}^2(T)\right) \tag{11}$$

where $k_{vv}(T)=\dfrac{\int_S K_{vv}\left(\tau=T,\boldsymbol{\Psi}\right)d\boldsymbol{\Psi}}{\int_S K_{vv}\left(\tau=0,\boldsymbol{\Psi}\right)d\boldsymbol{\Psi}}$ is the correlation coefficient; $D_0=\int_S K_{vv}\left(\tau=0,\boldsymbol{\Psi}\right)d\boldsymbol{\Psi}$ is the a priori variance at the beginning of the forecast interval.

If the predictor is configured correctly and the forecast error γ has a zero mean, then the loss from the forecast error is

$$q(T)=\frac{1}{\sqrt{2\pi D_T}}\int_{-3\sqrt{D_T}}^{3\sqrt{D_T}} g(\gamma)e^{-\frac{\gamma^2}{2D_T}}d\gamma \tag{12}$$

where $g(\gamma)$ is the lost function.

Accordingly, the profit from forecasting is

$$e_T=q_0-q(T)=\frac{1}{\sqrt{2\pi D_0}}\int_{-3\sqrt{D_0}}^{3\sqrt{D_0}} g(\gamma)e^{-\frac{\gamma^2}{2D_0}}d\gamma- \\ -\frac{1}{\sqrt{2\pi D_T}}\int_{-3\sqrt{D_T}}^{3\sqrt{D_T}} g(\gamma)e^{-\frac{\gamma^2}{2D_T}}d\gamma \tag{13}$$

On the other hand, the cost of forecasting is a linear function

$$q_p(T)=c_1T+c_0 \tag{14}$$

where c_0 is a fixed costs associated with obtaining, inputting, outputting and using forecast data; c_1 is an operating costs of the forecasting system. In particular, the costs depend on the method of implementing the forecasting function.

The form of dependencies (13) and (14) is shown in Fig. 10. Therefore, the condition for choosing the depth of the forecast is $q_T > q_p$.

In periodic control, it is advisable to limit the depth of the forecast to the control period, i.e. the condition for evaluating the depth of the forecast will be as follows:

$$(T<\tau_w)\cap(q_T>q_p) \tag{15}$$

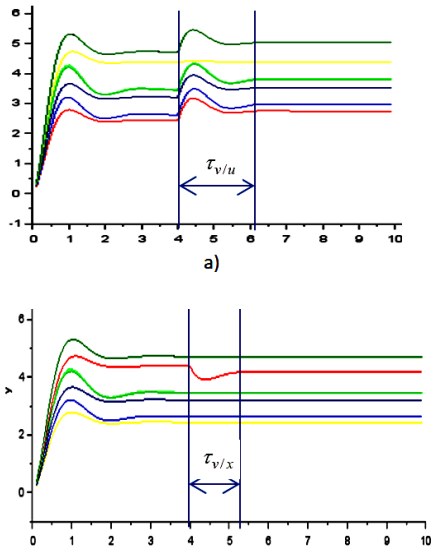


Fig. 9. Transient processes when changing the MTF operating mode

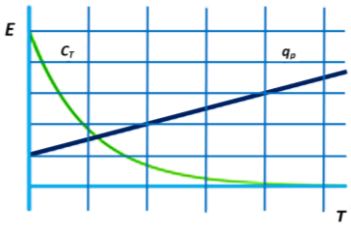


Fig. 10. Profit and cost of forecasting

5. Predicting with Heterogeneous Ensemble Neural Network

This work compared statistical methods that are characterized by ease of implementation and setup, and the proposed forecasting using Heterogeneous Ensemble Neural Network, which takes into account the architectural and nonlinear features of MTF.

There are no publicly available datasets for training and testing predictors for real MTFs. Since the creation of such datasets requires a lot of time and effort, NNs were trained on data from a simulation model. The methodology of such model-oriented training was investigated by the authors in [6].

The results of the study are presented in table 1. Exponential smoothing for 10 cycles, vector autoregression, Heterogeneous Ensemble Neural Network, spatiotemporal regression, and prediction using an unsegmented LSTM network were compared.

Input processes were generated with a normal distribution and an exponential correlation function with a time constant equal to the time of one prediction cycle. The number of discrete states of MTF $\|Y\|=20$ elements. In addition to inputs and element states, the input data of the predictors are the state of the environment and time.

To ensure the independence of the results from computer performance, the data are presented in normalized form. The normalizing value is taken as the exponential smoothing performance for 10 cycles for the MTF 1×4 architecture (four MTF control elements arranged in one line). The setup time, prediction time, and prediction error were studied for two values of the prediction depth: minimum (1 simulation cycle) and optimal 3 simulation cycles).

The long training time of the Heterogeneous Ensemble Neural Network is due to the need to separately train all NNs of Ensemble, but it provides higher prediction accuracy.

The prediction time using the RNN neural network was not recorded because the neural network was implemented on a serial processor, which is not typical. It is obvious that a parallel RNN can provide ultra-fast prediction.

The tuning time was determined by the size of the training sample, which in turn was determined from general considerations regarding the number of predictor parameters to be tuned and the sample size that provides a significance level according to the Pearson criterion of 5% with an acceptable error of also 5%. The exponential growth of the training sample and, accordingly, the tuning (training) time of the predictor with an increase in the number of elements and connections of the MTF elements is due to a corresponding increase in the number of estimated parameters. The required size of the training sample can be reduced provided that the system is homogeneous.

The experiments have shown that under given conditions, the neural network predictor has the highest accuracy, but requires long training and quite significant costs for parallel implementation.

Exponential smoothing gives the worst results in terms of forecast accuracy, but it requires minimal training of the predictor.

The optimal approach is to use a simulation model for forecasting with one cycle of active training and subsequent gradual improvement of results using passive identification.

Table 1 Results of the study of the characteristics of predictors of MTF status

	Training time (settings)	Forecast time	RMSE (forecast for 1 cycle)	RMSE (3-cycle forecast)
Exponential smoothing in 10 cycles	1	1	1	2.3
Vector autoregression	3.2	1.3	0.67	1.63
Spatiotemporal regression	59	54	0.41	1.41
Heterogeneous Ensemble Neural Network	800	0.2	0.22	0.96
LSTM	1200	0.1	0.34	1.25

6. Conclusion

The research is aimed at increasing the accuracy of forecasting the state of multi-zone thermal objects. A combined neuro-structural method for forecasting the state of multi-zone thermal objects is proposed, in which the structure of the neural model reflects the structure of the mutual influence of the object's zones. The method is implemented in the form of a Heterogeneous Ensemble Neural Network. The study of the method has shown the possibility of providing higher forecast accuracy with a relatively smaller size of the training data set.

References

[1] Armstrong J:. Extrapolation for Time-Series and Cross-Sectional Data. Armstrong J. S. (eds.): Principles of Forecasting. International Series in Operations Research & Management Science 30. Springer, Boston 2001 [http://doi.org/10.1007/978-0-306-47630-3_11].

[2] Bisikalo O., et al.: Parameterization of the Stochastic Model for Evaluating Variable Small Data in the Shannon Entropy Basis. Entropy 25(2), 2023, 184.

[3] Collopy F., Adya M., Armstrong J.: Expert Systems for Forecasting. Armstrong J. (eds.): Principles of Forecasting. International Series in Operations Research & Management Science 30. Springer, Boston, 2011 [http://doi.org/10.1007/978-0-306-47630-3_14].

[4] Dias G. F., Kapetanios G.: Estimation and forecasting in vector autoregressive moving average models for rich datasets. Journal of Econometrics 202(1), 2018, 75–91.

[5] Dubovoi V., et al.: Model-Oriented Training of Coordinators of the Decentralized Control System of Technological Facilities with Resource Interaction. IEEE Access 13, 2025, 13414–13426.

[6] Dubovoi V., et al.: Functional safety assessment of one-level coordination of distributed cyber-physical objects. Przegląd Elektrotechniczny 97(9), 2021, 38–41.

[7] Geysena D., et al.: Operational thermal load forecasting in district heating networks using machine learning and expert advice. 2017, arXiv:1710.06134v1.

[8] Gyeera T., Simons A., Stannett M.: Kalman filter based prediction and forecasting of cloud server KPIs. TechRxiv. May 19, 2021 [https://doi.org/10.36227/techrxiv.14583342.v1].

[9] Hasan M., Wathodkar G., Muia M.: ARMA Model Development and Analysis for Global Temperature Uncertainty. arXiv, 2023, 2303.02070v1.

[10] Hou P., et al.: Vector Autoregression Model-Based Forecasting of Reference Evapotranspiration in Malaysia. Sustainability 15, 2023, 3675 [http://doi.org/10.3390/su15043675].

[11] Hyndman R., et al.: Forecasting with Exponential Smoothing. Springer Berlin, Heidelberg 2008 [https://doi.org/10.1007/978-3-540-71918-2].

[12] Kolobrodov V. G., Nguyen Q. A., Tymchik G. S.: The problems of designing coherent spectrum analyzers. Proc. of SPIE 9066, 2013.

[13] Kukharchuk V. V., et al.: Information Conversion in Measuring Channels with Optoelectronic Sensors. Sensors 22(1), 2022, 271 [https://doi.org/10.3390/s22010271].

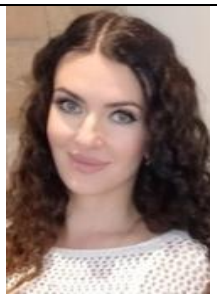
[14] Ivanovski Z., Milenkovski A., Narasanov Z.: Time Series Forecasting Using a Moving Average Model for Extrapolation of Number of Tourist. UTMS Journal of Economics 9(2), 2018, 121–132.

[15] Thorsten J., et al.: Predicting structured objects with support vector machines. Communications of the ACM 52(11), 2024, 97–104 [https://doi.org/10.1145/1592761.1592783].

D.Sc. Eng. Maria Yukhymchuk
e-mail: umcmasha@gmail.com

Professor of the Computer Control Systems Department of the Vinnytsia National Technical University, Ukraine. Received the degree of Doctor of Philosophy. degree in mathematical modeling and computational methods in 2013, Dr. Tech. degree in Control Process Automation in 2023.
Her main research interests include system analysis, information technology, reliability and efficiency of coordinated control of distributed cyber-physical systems with multi-zone thermal facilities.

https://orcid.org/0000-0002-8131-9739



D.Sc. Eng. Volodymyr Dubovoi
e-mail: vmdubovoy@gmail.com

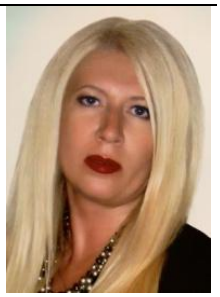
Professor at the Department of Computer Control Systems Vinnytsia National Technical University. Senior Member IEEE. 1989 received the title of associate professor at the Department of Automation and Information Measuring Systems; 1997: doctor of sciences (eng.). Author of more than 300 scientific works, 19 books (monographs, textbooks), 25 patents and inventions.
The main scientific direction: modeling, research and design of complex information systems under uncertainty.

https://orcid.org/0000-0003-0440-3643

D.Sc. Zhanna Harbar
e-mail: garbar_janna@ukr.net

Doctor of economic sciences, associate professor at the Department of Pedagogy and Educational Management, Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University.
Her main research interests include system analysis, information technology, pedagogy.

https://orcid.org/0000-0003-3492-9224



Ph.D. Bakhyt Yeraliyeva
e-mail: yeraliyevabakhyt81@gmail.com

Senior lecturer of the Information Systems Department, Faculty of Information Technology, M. Kh. Dulaty Taraz Regional University, Taraz, Kazakhstan.
Research interests: fiber optic technologies, information systems, internet of things and blockchain technologies.

https://orcid.org/0000-0002-8680-7694