UDC 621.316.72
DOI: 10.15587/1729-4061.2022.254477

BUILDING AN ADAPTIVE HYBRID MODEL FOR SHORT-TERM PREDICTION OF POWER CONSUMPTION USING A NEURAL NETWORK

Gulnara Ibrayeva
Corresponding author
Master of Engineering Science, Doctoral Student*
E-mail: gulnara388@gmail.com
Yuliya Bulatbayeva
PhD, Acting Assistant Professor*
Yermek Sarsikeyev
PhD, Head of Department
Department of Electrical Equipment Operation
S. Seifullin Kazakh Agro Technical University
Zhenis ave., 62, Nur-Sultan,
Republic of Kazakhstan, 010011
*Department of Automation of Manufacturing Processes
Karaganda Technical University
N. Nazarbayev ave., 56, Karaganda,
Republic of Kazakhstan, 100027

This paper proposes a step-by-step technique for combining basic models that forecast electricity consumption in an artificial neural network by the method of preliminary selection and further hybridization. The reported experiments were conducted using data on hourly electricity consumption at the metallurgical plant AO ArcelorMittal Temirtau in the period from January 1, 2019, to November 30, 2021. The current research is related to the planned introduction of a balancing electricity market. 96 combinations of basic models were compiled, differing in the type of neural network, the set of initial data, the order of lag, the learning algorithm, and the number of neurons in the hidden layer. It has been determined that the NARX-type network is the most optimal architecture to forecast electricity consumption. Based on experimental studies, the number of hidden neurons needed to form a planned daily profile should equal 3 or 4; it is recommended to use the conjugate gradient method as a learning algorithm. When selecting models from three groups, it was revealed that the conjugate gradient method produces better results compared to the Levenberg-Marquardt algorithm. It is determined that the values of the selected RMSE error indicator take values of 23.17, 22.54, and 22.56, respectively, for the first, second, and third data groups. The adaptive hybridization method has been shown to reduce the RMSE error rate to 21.73. However, the weights of the best models with values of 0.327 for the first group of data, and 0.336 for the second and third ones, show that the individual use of a separate combination of models is also applicable. The devised forecasting electricity consumption model can be integrated into an automated electricity metering system

Keywords: short-term forecasting, weighted average forecast, hybrid model, neural network, electrical load

Received date 07.03.2022
Accepted date 12.04.2022
Published date 29.04.2022

How to Cite: *Ibrayeva, G., Bulatbayeva, Y., Sarsikeyev, Y. (2022). Building an adaptive hybrid model for short-term prediction of power consumption using a neural network. Eastern-European Journal of Enterprise Technologies, 2 (8 (116)), 6–12. doi: https://doi.org/10.15587/1729-4061.2022.254477*

1. Introduction

The issue of electricity shortage and the low share of alternative sources of electricity in the energy system of the Republic of Kazakhstan require transformations of the electricity market. The electric power market of the Republic of Kazakhstan is a mechanism for the decentralized purchase and sale of electrical energy, which operates based on bilateral contracts between market participants [1]. The system operator of the unified energy network JSC Kazakhstan Electricity Grid Operating Company(KEGOC) currently provides system services in the form of technical dispatching, transportation, and balancing of electricity production and consumption. Such a mechanism for the functioning of the electric power market does not stimulate market participants to plan and thereby save their energy resources. According to the Concept for the Development of the Fuel and Energy Complex of the Republic of Kazakhstan, in the second half of 2022, it is planned to introduce a balancing electricity market (BEM) [2]. The purpose of BEM is to encourage participants to plan their consumption qualitatively and to encourage manufacturers to follow the commands of the System Operator [3].

The task of the current study is related to changes in legislation in the field of electric power and the launch of BEM in 2022, which works as follows: enterprises submit an application to the System Operator of the Unified Energy System JSC KEGOC for a day ahead. Applications are a planned daily profile of the electrical load of the enterprise, which is compared with the actual one for the next day. Suppose that at a certain hour a company planned to consume 100 MW while the actual load reached 101 MW. A surplus to other consumers at a lower price, and, conversely, in the event of a shortage, the company would buy 1 MW at a higher price. It can be concluded that both excess and shortage are not profitable for the enterprise; there is a need for analysis and forecasting [4].

For a forecasting problem, intelligent methods of analyzing temporal data, such as neural networks, are widely used [5]. To plan a daily profile for the day ahead, BEM consumers can use commercial solutions based on neural networks but ready-made software has a template set of forecasting model settings. Limited control capabilities of the neural network do not make it possible to take into consideration the specificity of a particular production, its electrical load profile, shift schedule, and other factors. In this regard, a relevantarea of research under the conditions of the planned launch of BEM is to build an adaptive model for short-term forecasting of electricity consumption.

2. Literature review and problem statement

Paper [6] reports a combined model with variable weight optimization for the short-term prediction of electrical load. This model makes it possible to combine the advantages of statistical and intellectual methods of data analysis. The Autoregressive Integrated Moving Average (ARIMA) method was used as a statistical technique. The advantages of the ARIMA model include a clear mathematical justification and a formalized step-by-step construction methodology [7]. However, when external factors change, it is necessary to re-identify the parameters of the ARIMA model, which makes its use unsuitable for the tasks of daily forecasting of electricity consumption due to resource and time costs. This circumstance can be excluded by using only intelligent methods of data analysis, as described in work [8]. It is proposed to use various types of neural networks: stable with back error propagation, Powell-Beal-conjugate gradient, cascade, with gradient descent, linear, and multi-level ones. However, this method is most applicable for medium- and long-term forecasting of electricity consumption with the calculation of the average absolute error. The experience of using neural networks using one for the problem of short-term forecasting is described in work [9]. The cited study used the NARX (Nonlinear Autoregressive with External Input) network to predict network traffic. The prediction model is built on a neural network with 52 neurons in the hidden layer and 1 neuron in the output layer using the Levenberg-Marquardt learning algorithm. However, the model used has a rigid structure that may not be suitable for short-term forecasting of other processes. In addition, other features of the NTS application are not studied, such as the algorithm of learning by the smoothed gradient method, the possibility of changing the architecture of the neural network, and the justification of the selected settings.

Paper [10] proposes a hybrid methodology for forecasting electricity prices for the energy market. The essence of hybridization is as follows. In the first stage, all data are divided into clusters, differing in the selected time lag order. In the second stage, the initial data are divided into the initial training (80 %) and the final testing (20 %) samples. The final test sample is used to verify the adequacy of the finished model. The initial training sample is divided into intermediate training (80 %) and intermediate testing (20 %) samples. Thus, with the help of such a division of the initial sample, a two-stage selection of data is carried out for training the final model, which increases its predictive properties. This hybridization technique is not possible to implement since information on electricity consumption and hourly prices for electrical load, which were not approved before the introduction of BEM in the Republic of Kazakhstan, are used as initial data.

Work [11] describes a procedure for determining the optimal predictive model by selecting the dominant models and combining them to obtain a final model, taking into consideration the weight of each forecast. The main results of the application of the methodology are a decrease in the average absolute error and volatility of the forecast. The only obstacle to the use of a giventechnique is the criterion for selecting dominant models: to forecast forBEM, it is better to use the standard error of the forecast.

Paper [12] reports a technique for predicting the production by an alternative power supply source. As initial data, weather data such as the average hourly temperature, relative humidity, and the level of solar radiation are used. However, such accuracy matters when studying the influence of weather phenomena on the generation of electricity by a solar panel and would be an unnecessary factor for building a model for predicting the electricity consumption by a full-cycle metallurgical complex that works non-stop all year round.

Work [13] demonstrates the effectiveness of the hybrid prediction method, which combines the combination of different types of neural networks and the mechanism for dividing the initial data into subsamples.

Having studied the adaptive forecasting methods that are close in essence,as well as the features of their implementation, we can conclude that the degree of influence of each input variable on the forecast value has not been sufficiently investigated. In addition, it should be noted that the accuracy of the forecast is affected by the architecture of an artificial neural network (ANN), which should be selected taking into consideration the specificity of the subject of study. In this regard, it became necessary to study the schemes of hybrid models with a different set of initial data and ways to combine them.

3. The aim and objectives of the study

The purpose of this study is to construct an adaptive hybrid model for forecasting electricity consumption based on basic models on ANNs, to the input of which various initial data are supplied.

To accomplish the aim, the following tasks have been set:

– to select neural network settings for building basic predictive models;

– to determine the way to combine basic models on neural networks;

– to evaluate the predictive capabilities of the hybrid model.

4. The study materials and methods

To conduct experimental studies, data on the hourly electricity consumption at the metallurgical plant of JSC ArcelorMittal Temirtau from January 1, 2019, to November 30, 2021, were taken on the basis of the regional electric load profile (RLP) of the Karaganda oblast, which includes 25560 values of the electric load of the enterprise. RLP is the share of consumption byan enterprise calculated according to the approved norms in the context of the load of the entire region.Data on the electrical load are given in MW. It should be noted that, along with RLP, the company operates an automated system for commercial electricity metering. However, for mutual settlements with the System Operator under the production-consumption balancing agreement, it is the RLP that is used. Data from RLP for the construction of basic forecasting models are taken on Greenwich Time, which entails a shift in the load profile by 5 hours back from the local time of Nur-Sultan.Also, in addition to the hourly electrical load (L), the day of the week (W), the indicators of the day off or working day (*E*), and the indicators of the studied hour (*H*) are used as initial data. It is expected that the combined model based on selection and hybridization of the initial data could improve the predictive capabilities compared to the baseline models with separate initial data.

Our experiments were conducted in the Neural Net Time Series (NTS) application, which is part of the MATLAB 2018a software package (9.4.0.813654). This application can be initiatedusing the *ntstool* command.

5. Results of studying an adaptive hybrid model for forecasting electricity consumption on an artificial neural network

5. 1. Selection of neural network settings for building basic predictive models

The choice of input data for network training and their processing is the most difficult stage in building a model on ANN. The initial data for forecasting are selected taking into consideration the type and objectives of forecasting. Analysis of the electrical load profile of the object under study significantly helps in solving the problem of correlation of the predicted parameter and the hypothetical predictor [14].

For this study, the annual, monthly, weekly, and daily load profiles of the metallurgical plant were analyzed [15]. It was found that historical data on energy consumption, weather data, and the number of days of downtime and repairs for previous years can be used for long-term forecasting.For short-term forecasting, historical data on energy consumption for previous days, as well as data on seasonality and cyclicality, should be taken into consideration. Our study considers the problem of forecasting in order to form a planned daily profile of the electrical load for a day ahead for the effective operation of the enterprise on BEM, that is, short-term forecasting is necessary [16].

The training data collection phase is followed by the data preparation and normalization phase. Normalization refers to the process of equalizing the original data for the interval from 0 to 1. Unprepared data affectANN, which leads to incorrect learning [17].

The NTS application proposes to build NAR (Nonlinear Autoregressive), NARX, and nonlinear input-output models. A NAR network is a closed-type model that makes it possible to predict the future values of a variable from a certain number of previous values of the same variable fed to the network input. The number of previous values of variables that affect its current value determines the backlog order. The NARX network has a semi-closed architecture for building models taking into consideration the influence of the previous values of the variable and external parameters on the variable under study. A nonlinear input-output network builds an open-type model, to the input of which external parameters are supplied. Only NAR and NARX networks are applicable for the tasks of our study since the nonlinear input-output network does not take into consideration the impact of retrospective data on electricity consumption.

After selecting the network structure, NTS requests the loading of input and target variables, which are divided into the training (70 %), checking (15 %), and test (15 %) samples.

The next step is to refine the network architecture. The maximum number of hidden neutrons is recommended to be determined by the sum of neurons in the input and output layer. Thus, the number of hidden neurons is proposed to alternate from 1 to 4, which includes data on electricity consumption, an indicator of the hour, week, and weekend or working day. At the same stage, it is proposed to choose the order of lagging behind. For the problem of forecasting on BEM, it is optimal to consider the effect of electrical load for the previous day, two days ago, and three days ago [18].

The last step is to choose a learning algorithm. In the present study, experiments were conducted using the Levenberg-Marquardt algorithm and the conjugate gradient method [19].

Based on a large number of basic models, there is a need to select working models and find a further way to combine them [20].

5. 2. Determine how to combine basic models on neural networks

According to [11], simple basic models can be combined by the selective or hybrid method. The selective method involves the continuous selection of the model with the best indicator for the selected criterion and switches to the appropriate model. The hybrid model averages the forecast taking into consideration the weight of each individual forecast, thereby making it possibleto smoothly switch from one model to another. It has been empirically revealed that it is optimal to use three models when building a hybrid forecasting model among many combinations of basic models. Thus, it is proposed to combine the methods of selection and hybridization as follows.

In the first stage, three models should be selected according to a pre-selected criterion for the accuracy of forecasts among the 96 models presented. The NTS application calculates the Mean Squared Error (MSE) but the Root Mean Squared Error (RMSE) should be selected as the main criterion for selecting the best models. The latter criterion (1) demonstrates the impact of the error of each hour for BEM tasks where the consumer must pay a fine for each discrepancy in the plan with the fact.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{\left(X_i - Y_i\right)^2}{n}}, \quad (1)$$

where i is the current hour of study,

n is the total number of study hours,

X_i is the actual electrical load, MW,

Y_i is the predicted electrical load, MW.

In the second step, one needs to average the selected three predictions, taking into consideration the impact of a more accurate forecast of each base model on the final forecast of the final model. The weights ω_1, ω_2, and $\omega_3$of the predictions Y_1, Y_2, and $Y_3$are respectively defined as follows (2):

$$\omega_1 = \frac{RSME_2 \cdot RSME_3}{RSME_1 \cdot RSME_2 + RSME_2 \cdot RSME_3 + RSME_1 \cdot RSME_3},$$

$$\omega_2 = \frac{RSME_1 \cdot RSME_3}{RSME_1 \cdot RSME_2 + RSME_2 \cdot RSME_3 + RSME_1 \cdot RSME_3},$$

$$\omega_3 = \frac{RSME_1 \cdot RSME_2}{RSME_1 \cdot RSME_2 + RSME_2 \cdot RSME_3 + RSME_1 \cdot RSME_3}, \quad (2)$$

where $RMSE_1$, $RMSE_2$, and $RMSE_3$ are the $RMSE$ values of the Y_1, Y_2, and Y_3 predictions, respectively.

Thus, the final forecast Y of the adaptive hybrid model should be determined from (3):

$$Y = \omega_1 \cdot Y_1 + \omega_2 \cdot Y_2 + \omega_3 \cdot Y_3. \quad (3)$$

In addition, the NTS application, in addition to the MSE indicator, calculates the regression coefficient R, which determines the degree of correlation between the target and output data under test; it can take values from 0 to 1.

5. 3. Evaluating the predictive capabilities of a hybrid model

After collecting and processing the initial data, 96 basic models with different parameters were investigated, differing in the type and architecture of the neural network, the learning algorithm, the lag order, the number of neurons in the hidden layer, and the set of initial data. The results of testing models with a lag of 24, 48, and 96 hours are given in Tables 1–3, respectively.

Table 1

Results of experiments with the lag indicator n=24

ANN type	Number of hidden neurons	Learning algorithm	Initial data	MSE	R
NAR	4	Levenberg-Marquardt	L	643	84
NAR	4	Conjugate gradient	L	649	83
NAR	3	Levenberg-Marquardt	L	642	84
NAR	3	Conjugate gradient	L	674	82
NAR	2	Levenberg-Marquardt	L	633	86
NAR	2	Conjugate gradient	L	838	81
NAR	1	Levenberg-Marquardt	L	647	86
NAR	1	Conjugate gradient	L	776	80
NARX	4	Levenberg-Marquardt	H	542	86
NARX	4	Conjugate gradient	H	608	84
NARX	3	Levenberg-Marquardt	H	560	87
NARX	3	Conjugate gradient	H	698	85
NARX	2	Levenberg-Marquardt	H	673	84
NARX	2	Conjugate gradient	H	809	80
NARX	1	Levenberg-Marquardt	H	592	85
NARX	1	Conjugate gradient	H	719	81
NARX	4	Levenberg-Marquardt	E	547	86
NARX	4	Conjugate gradient	E	609	86
NARX	3	Levenberg-Marquardt	E	656	85
NARX	3	Conjugate gradient	E	583	86
NARX	2	Levenberg-Marquardt	E	612	86
NARX	2	Conjugate gradient	E	716	83
NARX	1	Levenberg-Marquardt	E	636	84
NARX	1	Conjugate gradient	E	613	85
NARX	4	Levenberg-Marquardt	W	605	84
NARX	4	Conjugate gradient	W	537	86
NARX	3	Levenberg-Marquardt	W	672	83
NARX	3	Conjugate gradient	W	688	82
NARX	2	Levenberg-Marquardt	W	547	85
NARX	2	Conjugate gradient	W	539	87
NARX	1	Levenberg-Marquardt	W	639	85
NARX	1	Conjugate gradient	W	619	84

Fig. 1 shows the architecture and test results of such ANN with a lag order of 96 hours, with 3 hidden neurons, an indicator of the day of the week at the input, and a learning algorithm by the conjugate gradient method.

Table 2

Results of experiments with the lag indicator n=48

ANN type	Number of hidden neurons	Learning algorithm	Initial data	MSE	R
NAR	4	Levenberg-Marquardt	L	575	86
NAR	4	Conjugate gradient	L	660	82
NAR	3	Levenberg-Marquardt	L	602	85
NAR	3	Conjugate gradient	L	651	86
NAR	2	Levenberg-Marquardt	L	599	87
NAR	2	Conjugate gradient	L	765	82
NAR	1	Levenberg-Marquardt	L	651	84
NAR	1	Conjugate gradient	L	802	80
NARX	4	Levenberg-Marquardt	H	539	86
NARX	4	Conjugate gradient	H	548	87
NARX	3	Levenberg-Marquardt	H	538	86
NARX	3	Conjugate gradient	H	614	86
NARX	2	Levenberg-Marquardt	H	625	85
NARX	2	Conjugate gradient	H	628	83
NARX	1	Levenberg-Marquardt	H	654	84
NARX	1	Conjugate gradient	H	580	86
NARX	4	Levenberg-Marquardt	E	561	85
NARX	4	Conjugate gradient	E	522	87
NARX	3	Levenberg-Marquardt	E	670	84
NARX	3	Conjugate gradient	E	542	85
NARX	2	Levenberg-Marquardt	E	601	85
NARX	2	Conjugate gradient	E	581	84
NARX	1	Levenberg-Marquardt	E	673	84
NARX	1	Conjugate gradient	E	646	83
NARX	4	Levenberg-Marquardt	W	622	86
NARX	4	Conjugate gradient	W	510	86
NARX	3	Levenberg-Marquardt	W	636	86
NARX	3	Conjugate gradient	W	508	87
NARX	2	Levenberg-Marquardt	W	611	86
NARX	2	Conjugate gradient	W	627	85
NARX	1	Levenberg-Marquardt	W	536	87
NARX	1	Conjugate gradient	W	679	83

Of the three groups of data fromTables 1–3, the following ANN models were selected for the following parameters:

– the NARX network with 4 hidden neurons, a conjugate gradient learning algorithm, 24-hour lag order, and data on the day of the week at the input: MSE_1=537 and R=86;

– the NARX network with 3 hidden neurons, a conjugate gradient learning algorithm, a lag order of 48 hours, and data on the day of the week at the input: MSE_2=508 and R=87;

– the NARX network with 1 hidden neuron, a conjugate gradient learning algorithm, a lag order of 96 hours, and input day of the week data: MSE_3=509 and R=87.

However to combine the models,one must recalculate the MSE to RMSE:

$$RMSE_1 = \sqrt{MSE_1} = 23.17,$$

$$RMSE_1 = \sqrt{MSE_1} = 22.54,$$

$$RMSE_1 = \sqrt{MSE_1} = 22.56.$$

Table 3

Results of experiments with the lag indicator n=96

ANN type	Number of hidden neurons	Learning algorithm	Initial data	MSE	*R*
NAR	4	Levenberg-Marquardt	L	582	87
NAR	4	Conjugate gradient	L	645	83
NAR	3	Levenberg-Marquardt	L	571	86
NAR	3	Conjugate gradient	L	625	84
NAR	2	Levenberg-Marquardt	L	633	85
NAR	2	Conjugate gradient	L	638	84
NAR	1	Levenberg-Marquardt	L	652	84
NAR	1	Conjugate gradient	L	759	83
NARX	4	Levenberg-Marquardt	H	642	84
NARX	4	Conjugate gradient	H	646	84
NARX	3	Levenberg-Marquardt	H	517	86
NARX	3	Conjugate gradient	H	566	86
NARX	2	Levenberg-Marquardt	H	572	84
NARX	2	Conjugate gradient	H	623	85
NARX	1	Levenberg-Marquardt	H	582	86
NARX	1	Conjugate gradient	H	568	86
NARX	4	Levenberg-Marquardt	E	583	86
NARX	4	Conjugate gradient	E	655	85
NARX	3	Levenberg-Marquardt	E	546	85
NARX	3	Conjugate gradient	E	549	85
NARX	2	Levenberg-Marquardt	E	654	86
NARX	2	Conjugate gradient	E	571	86
NARX	1	Levenberg-Marquardt	E	600	86
NARX	1	Conjugate gradient	E	694	83
NARX	4	Levenberg-Marquardt	W	648	86
NARX	4	Conjugate gradient	W	593	83
NARX	3	Levenberg-Marquardt	W	621	84
NARX	3	Conjugate gradient	W	604	84
NARX	2	Levenberg-Marquardt	W	672	83
NARX	2	Conjugate gradient	W	621	85
NARX	1	Levenberg-Marquardt	W	509	87
NARX	1	Conjugate gradient	W	646	85

Then the weight of each forecast can be calculated as follows:

$$\omega_1 = \frac{22.54 \cdot 22.56}{22.54 \cdot 22.56 + 23.17 \cdot 22.56 + 23.17 \cdot 22.54} = 0.327,$$

$$\omega_2 = \frac{23.17 \cdot 22.56}{22.54 \cdot 22.56 + 23.17 \cdot 22.56 + 23.17 \cdot 22.54} = 0.336,$$

$$\omega_3 = \frac{23.17 \cdot 22.54}{22.54 \cdot 22.56 + 23.17 \cdot 22.56 + 23.17 \cdot 22.54} = 0.336.$$

When calculating the final forecast using (3), we obtained $Y$with RMSE=21.73. Thus, the hybrid model produces better results compared to the base models chosen according to the selected criterion, which makes it possible to reduce the price of electricity for a specific operating hour.

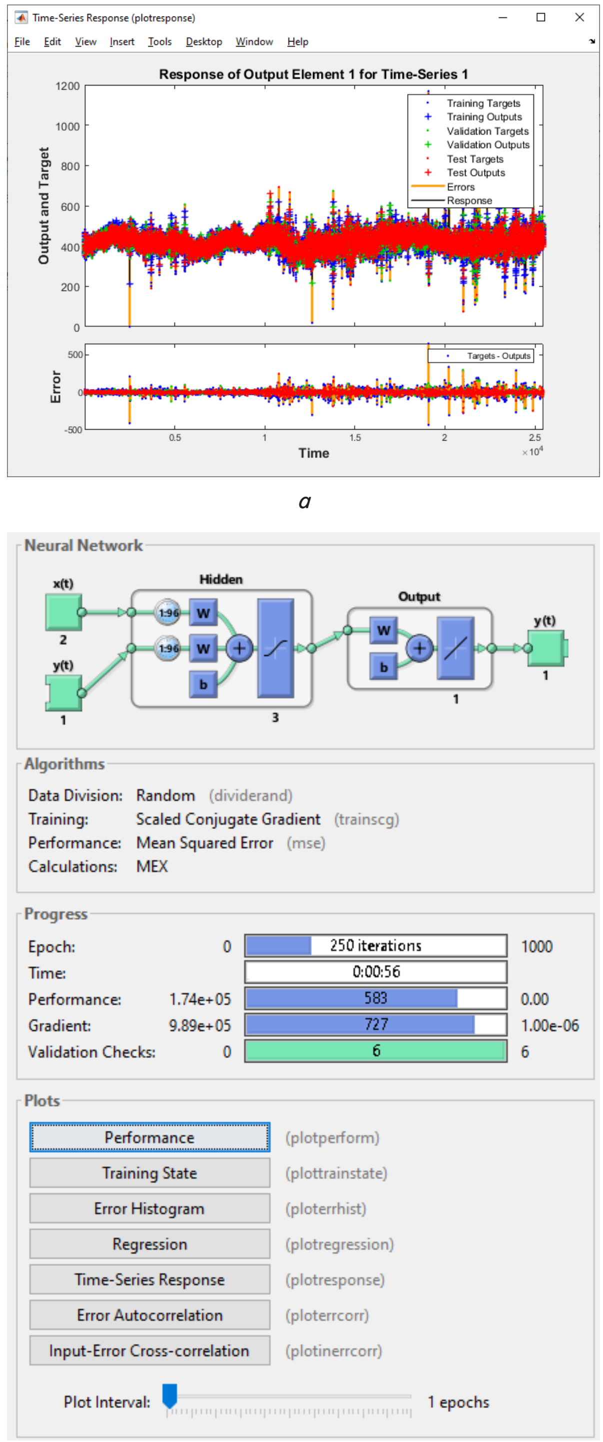

Fig. 1. The results of testing the basic model on an artificial neural network with a lag order of 96 hours, 3 hidden neurons, a day of the week indicator at the input, and a learning algorithm by the conjugate gradient method: *a* – a plot for comparing the target and output values of the electrical load; *b* – the final window of ANN modeling results

6. Discussion of results of studying the adaptive hybrid model

Based on the capabilities of the NTS application, various configurations of the neural network were compiled. In the first stage, the models were divided into three groups: in Table 1 – with a backlog of 24 hours, in Table 2 – 48 hours, and in Table 3 – 96 hours. This separation was made for the convenience of working with a large set of setting parameters. The backlog order determines the degree of influence of the load for the previous days on electricity consumption for the current day.Next, the type of neural network is defined. Thus, in each of the three groups of models with the NAR-type network, the highest MSE value was observed, equal to 838, 802, and 752 in the first, second, and third groups, respectively. This is explained by the fact that this type of network does not take into consideration the influence of other factors on the amount of electricity consumption. In the next step, the three large groups were divided into subgroups depending on the parameter taken into consideration as an external factor. It was found that the best results in all three groups were shown by a subgroup of NARX models with such a parameter of the cyclic component as the day of the week. In work [9], the model on the NARX network also demonstrated high predictive capabilities, where the value of the regression coefficient was 0.97134.

After compiling various sets of network configurations and obtaining simulation results, a technique was determined to discard the non-productive and combine the remaining selected models. The decrease in the value of the root mean square error (1) of the final weighted average forecast (3) is explained by a smooth switch between the best selected baseline models according to a pre-selected forecast accuracy criterion, taking into consideration the weight of each forecast (2). The value (1), in relation to the average value of the absolute error of the forecast described in [8], makes it possible to take into consideration the maximum magnitude of the error without averaging it. Such an approach to the choice of the forecast accuracy criterion is most appropriate for forecasting tasks on BEM.

The proposed method of combining basic models makes it possible to change the structure of the model, responding to the slightest changes in external factors. The current study was conducted for BEM on the day ahead and cannot be used for the task of dynamic forecasting of electricity consumption. The complexity of applying our method is in a large number of basic models at the initial stage. This circumstance can be overcome by revising the tuning parameters of the basic models to reduce their number, which would reduce the time for data processing. Unlike [6], the adaptive hybrid model has a more flexible structure where additional parameters can be integrated. For example, with the direct introduction of BEM and the further approval of prices for services, prices for balancing electricity can also be entered into the initial data.

7. Conclusions

1. To build 96 basic models, networks such as NAR or NARX with the number of neurons in the hidden layer from 1 to 4, trained by the Levenberg-Marquardt algorithm or the conjugate gradient method, were used. As input data, data on electricity consumption, day of the week, information about the day-off or working day, and the hour indicator with a lag of 24, 48, and 96 hours were applied. The NARX-type model makes it possible to consider the influence of third-party factors. In our experiments, the weekly cyclic component had the greatest impact in all three groups of base models. Among the learning algorithms, the best results were shown by the conjugate gradient method, namely in the first and second data groups. Additional advantages of the conjugate gradient method are speed and memory savings, although it requires the calculation of first-order derivatives.In practice, it was proved that the number of neurons of the hidden layer should tend to the sum of the neurons of the input and output layer: in the first group of data – 4 neurons, in the second – 3 neurons and, as an exception, in the third – 1 neuron.

2. To build the final model, selection and hybridization methods were used. At the selection stage, 3 basic models were selected, chosen according to the criterion of the smallest root mean square error. The resulting model can be attributed to the category of adaptive or self-adjusting. The RMSE values for the selected model predictions take close values, which is reflected in the weight of each prediction. The weights of the best predictions of the three different groups of data demonstrate that the selected combinations work equally effectively and produce an approximately similar result. This means that each model can be applied separately but it is the hybrid model that will make it possible to make a smooth switch, monitor changes, and take them into consideration when forming a planned daily load profile.

3. The weighted average final forecast makes it possible to reduce the RMSE value from 23.17 to 21.73 and, thereby, reduce the company's penalty payments for exceeding the planned daily electrical load profile.

Acknowledgments

We express our gratitude to Doctor of Technical Sciences, Professor Iosif Breido, and the management of the Central Plant Electrotechnical Laboratory at AO ArcelorMittal Temirtau.

References

1. On Electric Power Industry. Law of the Republic of Kazakhstan dated 9 June 2004 No. 588. Available at: https://adilet.zan.kz/eng/docs/Z040000588_
2. Ob utverzhdenii Kontseptsii razvitiya toplivno-energeticheskogo kompleksa Respubliki Kazakhstan do 2030 goda. Postanovlenie Pravitel'stva Respubliki Kazakhstan ot 28 iyunya 2014 goda No. 724. Available at: https://adilet.zan.kz/rus/docs/P1400000724
3. Ob utverzhdenii Pravil organizatsii i funktsionirovaniya optovogo rynka elektricheskoy energii. Prikaz Ministra energetiki Respubliki Kazakhstan ot 20 fevralya 2015 goda No. 106. Zaregistrirovan v Ministerstve yustitsii Respubliki Kazakhstan 26 marta 2015 goda No. 10531. Available at: https://adilet.zan.kz/rus/docs/V1500010531
4. Ob utverzhdenii Pravil funktsionirovaniya balansiruyuschego rynka elektricheskoy energii. Prikaz Ministra energetiki Respubliki Kazakhstan ot 20 fevralya 2015 goda No. 112. Zaregistrirovan v Ministerstve yustitsii Respubliki Kazakhstan 26 marta 2015 goda No. 10532. Available at: https://adilet.zan.kz/rus/docs/V1500010532

5. Breido, J. V., Kalinin, A. A., Lissitsyn, D. V. (2018). Algorithms of Energy Efficient Control of Electric Technological Complex for Autonomous Heat Supply. EAI Endorsed Transactions on Energy Web, 5 (19), 155040. doi: https://doi.org/10.4108/eai.10-7-2018.155040
6. Li, W.-Q., Chang, L. (2018). A combination model with variable weight optimization for short-term electrical load forecasting. Energy, 164, 575–593. doi: https://doi.org/10.1016/j.energy.2018.09.027
7. Ibrayeva, G., Breido, I., Bulatbayeva, J. (2021). Application of Box-Jenkins Methodology in Electric Consumption Forecasting of a Metallurgical Enterprise. DAAAM Proceedings, 0493–0499. doi: https://doi.org/10.2507/32nd.daaam.proceedings.071
8. Ahmad, T., Zhang, D., Shah, W. A. (2020). Efficient Energy Planning With Decomposition-Based Evolutionary Neural Networks. IEEE Access, 8, 134880–134897. doi: https://doi.org/10.1109/access.2020.3010782
9. Serikov, T., Zhetpisbayeva, A., Mirzakulova, S., Zhetpisbayev, K., Ibrayeva, Z., Soboleva, L. et. al. (2021). Application of the NARX neural network for predicting a one-dimensional time series. Eastern-European Journal of Enterprise Technologies, 5 (4 (113)), 12–19. doi: https://doi.org/10.15587/1729-4061.2021.242442
10. Ghofrani, M., Azimi, R., Najafabadi, F. M., Myers, N. (2017). A new day-ahead hourly electricity price forecasting framework. 2017 North American Power Symposium (NAPS). doi: https://doi.org/10.1109/naps.2017.8107269
11. Mulesa, O., Povkhan, I., Radivilova, T., Baranovskyi, O. (2021). Devising a method for constructing the optimal model of time series forecasting based on the principles of competition. Eastern-European Journal of Enterprise Technologies, 5 (4 (113)), 6–11. doi: https://doi.org/10.15587/1729-4061.2021.240847
12. Georgiou, G. S., Nikolaidis, P., Kalogirou, S. A., Christodoulides, P. (2020). A Hybrid Optimization Approach for Autonomy Enhancement of Nearly-Zero-Energy Buildings Based on Battery Performance and Artificial Neural Networks. Energies, 13 (14), 3680. doi: https://doi.org/10.3390/en13143680
13. Pan, L., Feng, X., Sang, F., Li, L., Leng, M., Chen, X. (2017). An improved back propagation neural network based on complexity decomposition technology and modified flower pollination optimization for short-term load forecasting. Neural Computing and Applications, 31 (7), 2679–2697. doi: https://doi.org/10.1007/s00521-017-3222-2
14. Mosbah, H., El-hawary, M. (2016). Hourly Electricity Price Forecasting for the Next Month Using Multilayer Neural Network. Canadian Journal of Electrical and Computer Engineering, 39 (4), 283–291. doi: https://doi.org/10.1109/cjece.2016.2586939
15. Breido, J. V., Bulatbayeva, J. F., Orazgaleyeva, G. D. (2020). Daily load profile analysis of metallurgical production for determining the forecasting model input data.Vestnik PGU, 2, 88–97. Available at: http://vestnik-energy.tou.edu.kz/storage/journals/80.pdf
16. Kaytez, F., Taplamacioglu, M. C., Cam, E., Hardalac, F. (2015). Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. International Journal of Electrical Power & Energy Systems, 67, 431–438. doi: https://doi.org/10.1016/j.ijepes.2014.12.036
17. Lee, N. C., Leal, V. M. S., Dias, L. C. (2018). Identification of objectives for national energy planning in developing countries. Energy Strategy Reviews, 21, 218–232. doi: https://doi.org/10.1016/j.esr.2018.05.004
18. López, M., Valero, S., Rodriguez, A., Veiras, I., Senabre, C. (2018). New online load forecasting system for the Spanish Transport System Operator. Electric Power Systems Research, 154, 401–412. doi: https://doi.org/10.1016/j.epsr.2017.09.003
19. Jimenez, J., Donado, K., Quintero, C. G. (2017). A Methodology for Short-Term Load Forecasting. IEEE Latin America Transactions, 15 (3), 400–407. doi: https://doi.org/10.1109/tla.2017.7867168
20. Angamuthu Chinnathambi, R., Mukherjee, A., Campion, M., Salehfar, H., Hansen, T., Lin, J., Ranganathan, P. (2019). A Multi-Stage Price Forecasting Model for Day-Ahead Electricity Markets. Forecasting, 1 (1), 26–46. doi: https://doi.org/10.3390/forecast1010003