DOI: 10.17323/2587-814X.2023.3.7.23

Short-term forecasting of electricity prices using generative neural networks

Andrej S. Kaukin
E-mail: kaukin@iep.ru

Pavel N. Pavlov
E-mail: pavlov@ranepa.ru

Vladimir S. Kosarev
E-mail: kosarev-vs@ranepa.ru

Russian Presidential Academy of National Economy and Public Administration
Address: 82, Vernadskogo Prospect, Moscow, 119571, Russia

Abstract

This article studies the predictive abilities of the generative-adversarial neural network approach in relation to time series using the example of price forecasting for the nodes of the Russian free electricity market for the day ahead. As a result of a series of experiments, we came to the conclusion that a generative adversarial network, consisting of two models (generator and discriminator), allows one to achieve a minimum of the error function with a greater generalizing ability than, all other things being equal, is achieved as a result of optimizing the static analogue of the generative model – recurrent neural network. Our own empirical results show that with a near-zero mean square error on the training set, which is demonstrated simultaneously by the recurrent and generative models, the error of the latter on the test set is lower. The adversarial approach also outperformed alternative reference models in out-of-sample forecasting accuracy: a convolutional neural network adapted for time series forecasting and an autoregressive linear model. Application of the proposed approach has shown that a generative-adversarial model with a given universal architecture and a limited number of explanatory factors, subject to additional training on data specific to the target node of the power system, can be used to predict prices in market nodes for the day ahead without significant deviations.

Keywords: electricity market, day-ahead market, time series, generative neural network, recurrent neural network

Citation: Kaukin A.S., Pavlov P.N., Kosarev V.S. (2023) Short-term forecasting of electricity prices using generative neural networks. *Business Informatics*, vol. 17, no. 3, pp. 7–23. DOI: 10.17323/2587-814X.2023.3.7.23

Introduction

A two-level electricity and capacity market operates on the territory of Russia. The day-ahead market is the main platform where industrial producers purchase electricity and capacity. The Day Ahead Market (DAM) is a system where competitive selection of price bids from suppliers and buyers of electricity is carried out a day before its actual delivery to consumers with the determination of prices and supply volumes for each hour of the day[1]. Marginal pricing is carried out on the DAM. In other words, the price is determined on the basis of the balance of supply and demand and applies to all participants in this market.

The level of electricity prices directly affects the costs of companies. The availability of an accurate forecast of electricity prices is necessary for planning the operating activities of enterprises representing energy-intensive industries, as well as for building financial models. The non-linear nature of electricity price dynamics makes their forecasting a difficult empirical task[2].

There are many empirical works by both domestic and foreign researchers where neural network methods were used to predict the prices of the free electricity market and demonstrated their superiority over linear models. However, recently there is evidence that the use of a generative adversarial networks (GAN[3]) approach can further improve the accuracy of a neural network built and optimized specifically for solving a specific problem.

GAN is both a representative of statistical forecasting methods and simulation methods[4]. An inexperienced expert using a set of available predictors (a generative neural network at the beginning of training) makes predictions about the price of electricity in a

[1] Competitive selection of applications is carried out by a commercial operator ATS. The price of DAM is determined for each node of two price zones: the first price zone includes the territories of the European part of Russia and the Urals (Central, North-Western with the exception of territories belonging to non-price zones, Southern, North Caucasus, Volga and Ural Federal districts); the second is the territory of Siberia (Siberian Federal District).

[2] Significant volatility, a complex structure of seasonality (annual, intra-weekly, intraday seasonality), as well as the presence of frequent emissions are a characteristic feature of the prices of the free electricity market. The elections can be explained by both abnormal situations in the energy system and the conjuncture of consumer industries, which makes it impossible to build a complete structural model of the market

[3] The symbiosis of two neural networks: a generator and a discriminator. The generator is designed directly for making forecasts. The discriminator is designed to distinguish real data from the generator forecast and stimulate the generator to learn more efficiently.

[4] Simulation modeling methods are usually used to represent the structure and connections of elements of complex economic systems, where it is explicitly necessary to take into account the interactions of many agents [1–3].

particular node, and another neural network (discriminator) learns to distinguish them from real data. Networks stimulate each other in the process of learning to more and more accurately solve the individual task assigned to them. Training continues until the expert (generator) forecasts reach the required accuracy in accordance with the selected metric.

The purpose of this work is to test the methodology for constructing generative adversarial networks (GAN) to predict per-node electricity prices in Russia on the day-ahead market, as well as to conduct a comparative analysis of the quality of forecasts based on the GAN methodology and forecasts obtained by alternative (reference) methods.

The following tasks were solved in the work presented here to achieve the goal of the study. First, approaches to electricity price forecasting were summarized, including those using generative adversarial neural networks. Secondly, a description of the data characterizing the wholesale electricity market in Russia was given. Thirdly, methodological approaches to the development of a generative-adversarial neural network model for predicting the price of electricity at the level of individual nodes of the energy system in Russia were characterized. Fourth, an overview of the empirical results of the study was provided, comparing the results of electricity price forecasting using the Generative Adversarial Neural Network (GAN) methodology and the main benchmarks, including the basic recurrent network, convolutional neural network and autoregressive model (ARIMA), widely used in modeling time sequences in the field of economics and finance [4–6]. A summary of the results of the study is given in the conclusion.

1. Analysis of the literature

1.1. Classical approaches to forecasting electricity prices on DAM

Various economic-mathematical and simulation models were used to solve the problem of forecasting electricity prices. For example, domestic scientists in [1] used a simulation model based on CGE (general economic equilibrium model) to search for optimal rates of tariff growth in the electric power industry in the regions of the Russian Federation in the regulated market segment. The system of equations developed by the authors took into account the interaction of many economic agents: consumer, producer, importers and exporters of electricity and the state.

The system of simultaneous equations was used in [7] to predict free market prices. The authors modeled spot electricity prices in Scandinavia using a model based on 29 equations, which took into account functional relationships between climatic factors, snow cover development, river water content (hydroelectric power plants are the main source of electricity in Scandinavia) and power system parameters. In [8], a structural model of spot electricity prices for New England in the United States was developed, taking into account functional relationships between fuel prices, as well as electricity demand and the availability of generating capacities.

The simulation approach is an effective tool for modeling the economy based on an analytical representation of the interaction of various agents and takes into account the physical processes that affect the economy. For example, in [2], a simulation model was presented for predicting the dynamics of oil production by wells, taking into account the implemented and planned geological and technical measures at each well. In another case, a simulation approach was used to develop a digital twin of a TV production plant [3]. However, the use of a simulation approach for forecasting, in the case of node prices for DAM, seems difficult, since it requires detailed information about the topology of the energy network, its parameters and operating conditions.

Economic and mathematical approaches in terms of time series models are more common for free market price forecasting. They can be generalized into two groups: statistical and machine learning methods. Statistical methods usually include additive econometric models [9]. For example, in [5, 6, 10], various variations of autoregressive models (ARIMA, ARMAX, AGARCH) were used to predict free market prices. More and more works began to appear with the development and popularization of machine learning methods, where linear econometric models are compared with such representatives of machine learning methods as support vector models, gradient boosting and neural networks [9, 11, 12] , turned out to be less accurate in short-term price forecasting.

Many researchers have turned to neural networks to predict the prices of the free electricity market [12–14]. Most researchers, as of 2020 [15], preferred the multilayer perceptron[5] (MLP). Domestic researchers have also repeatedly used this architecture to predict DAM prices. For example, Maryasin and Lukashova used MLP with two hidden layers to forecast free electricity prices in the Yaroslavl region. Zolotova and Dvorkin [16] in their study proposed to use a perceptron with 8 neurons in the hidden layer to predict the hourly equilibrium price index of the first price zone.

Many works exist in the foreign literature using other architectures that have proven themselves in the problems of time series forecasting in other areas. For example, a combination of convolutional and recurrent neural networks was used to predict prices and demand for electricity in [14, 17]. Article [14] shows that this architecture has justified itself in many areas where forecasting required the extraction of both temporal and spatial characteristics of time series. The authors of the study [12] proposed using a convolutional neural network with extended convolutions[6] to predict prices on the wholesale electricity market in the Canadian city of Ontario. However, recently there is evidence that the use of a generative adversarial approach can improve the result of a network of any architecture, if this neural network is used as a generator in a GAN [18].

1.2. Generative-adversarial neural networks

1.2.1. General characteristics

The basic theory of generative adversarial networks with examples of practical use can be found in Nikolenko's monograph [19]. A simple generative adversarial network consists of two artificial neural networks that interact with each other in turn. The first is a generator. It spawns objects in the data space. The second is the discriminator. It learns to distinguish objects generated by the generator from real examples from the training sample.

The generator must learn to trick the discriminator, and the discriminator must correctly distinguish between generated examples and real ones. This is the adversarial component in the interaction of two networks. The above description in terms of game theory is a minimax optimization problem, which can be written as Equation 1:

$$\min_G \max_D V(D,G), \text{ where } \max V(D,G) = \\ = E_{x\sim p_{data}(x)}\left[\log D(x)\right] + E_{z\sim p_z(z)} \log\left(1 - D\left(G(z)\right)\right), \quad (1)$$

where $D(x)$ – functional form of the discriminator;

$G(z)$ – functional form of the generator;

$p_z(z)$ – data distribution generated by the generator;

$p_{data}(x)$ – distribution of actual data.

In practice, the functional forms of the discriminator and generator can be any architecture of neural networks. The solution of the minimax problem provides alternate training of the generator with fixed weights of the discriminator and the discriminator with fixed weights of the generator.

Goodfellow and Benji [20] first described and put into practice adversarial networks in 2014. Sub-

[5] Multilayer perceptron is an artificial neural network that is characterized by several layers of input nodes connected in the form of a directed graph between the input and output layers.

[6] Adaptation of a convolutional network for time series forecasting, which allows taking into account a wide range of history when forecasting.

sequently, their idea was widely applied in practice. Adversarial networks have made it possible to achieve significant results in such areas as image generation from a text description [21], the creation of drugs [22], the generation of pseudo-realistic time sequences with the preservation of distribution highlights [23], etc.

For example, in the case of generating an image by description, the text is converted into a numerical feature space using a recurrent encoder, and then these features are used as a condition in the GAN that generates the image. As a result, for example, a person's face in a photograph can first be displayed in the feature space, and then the age feature can be changed and a new image generated. Thus, it is possible to artificially "age" or "rejuvenate" a person.

To create a new drug, the researchers from [22] used an adversarial autoencoder to generate molecules that can be promising candidates for creating new drugs based on them.

Time series are a unique object for generative modeling. In [23], it is noted that time series forecasting models, such as classical or neural network autoregressions, are inherently deterministic. Generative models, in turn, allow us to add an element of randomness to the neural network output.

1.2.2. Application of GAN for forecasting the electricity market

There are a number of examples of the use of generative-adversarial networks for forecasting the electricity market in foreign research practice. For example, in [24], the authors tested the generative-adversarial network model on two data sets: electricity consumption at the level of an individual household and the dynamics of the exchange rate. As a result, the generative-adversarial model in both experiments surpassed in accuracy its deterministic equivalent – a generative neural network that was trained independently.

In [25], researchers propose a model based on generative-adversarial networks for predicting node-by-node prices of a part of the US energy system. The neural network model uses spatial-temporal correlations between historical prices at nodes and accepts historical prices ordered into a three-dimensional tensor as input data. This tensor consists of a series of time-ordered matrices. In turn, each matrix is actually a map of node prices while preserving the spatial location of nodes. The task of the generative model in this case was to generate a new matrix with forecasted node-by-node prices for electrical energy. The basic model was trained to make a forecast for an hour ahead.

In [26], a generative adversarial network is used to forecast wholesale electricity prices with an interval of 30 minutes for the Australian energy market. Unlike previous works, the authors do not build a point estimate of the price, but an interval one. The generative network allowed the authors to obtain predictive intervals covering rare and extreme observations more accurately than alternative stochastic models.

Following the approach described in [24], we use a two-step procedure to develop our own generative-adversarial model. At the first step, we develop and optimize a recurrent neural network to solve the problem of predicting the price of electricity in a random node of the power system. In the second step, we incorporate the resulting neural network into the GAN as a generator and check the stability of the model on a subset of nodes. This approach allows you to narrow the search space of the GAN architecture to a discriminator and makes it possible to test the hypothesis that it is possible to improve the performance of the underlying neural network by including it in the GAN architecture.

2. Data

The analysis is carried out on the basis of hourly reports of the Trading System Administrator on equilibrium prices in the largest nodes of the energy system [27]. The database covers the period from April 13, 2019 to December 31, 2022 and contains information about 7215 nodes in 66 regions of the Russian Federation. The dynamics of averaged prices for all nodes of the energy system of the Russian Federation is shown in *Fig. 1*.

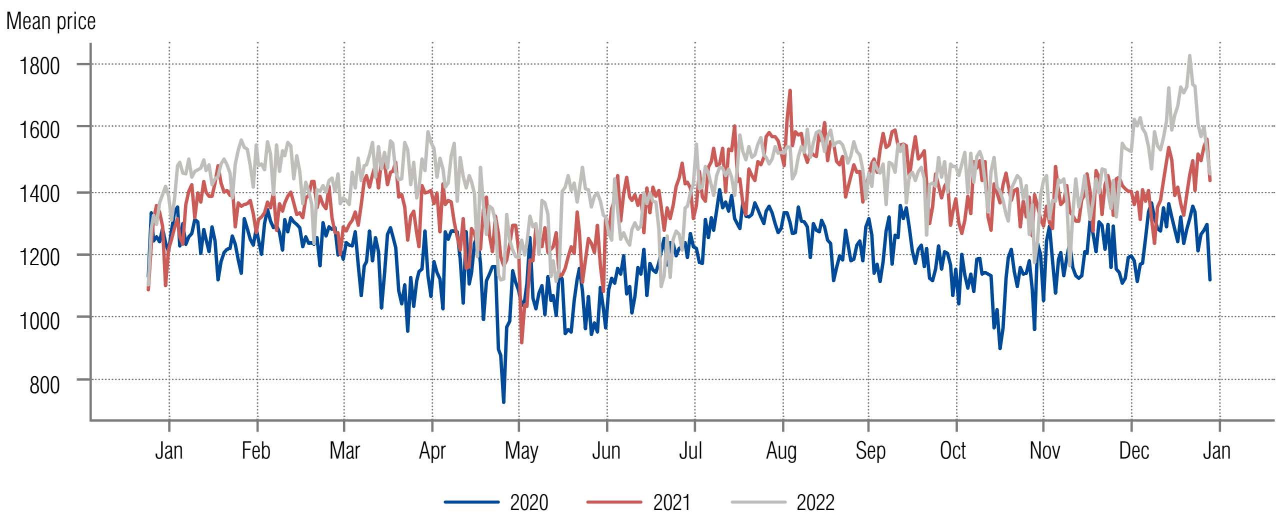

Fig. 1. Average prices for DAM for all nodes of the energy system of the Russian Federation.

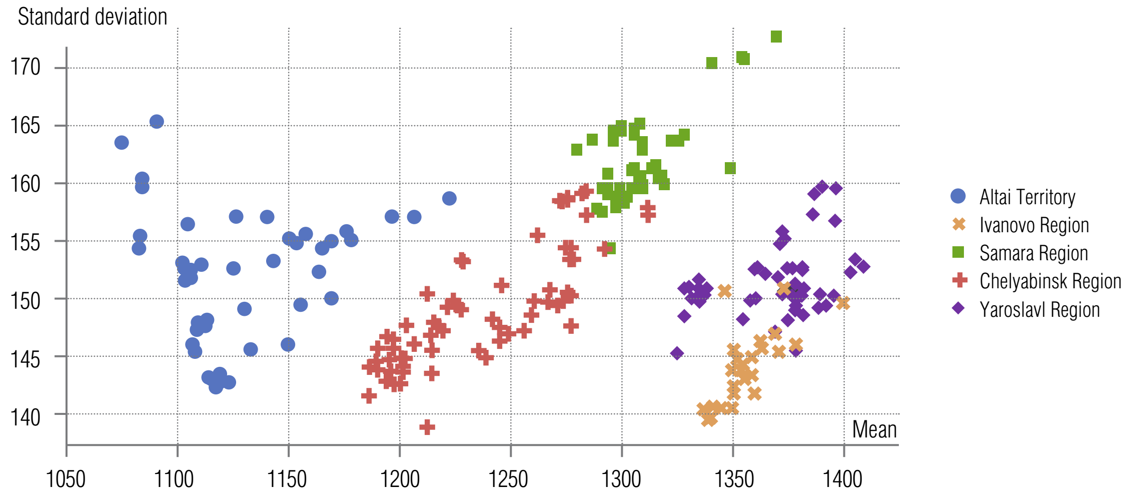

Fig. 2. Clustering of statistical characteristics of nodes within one region.

It is possible to note a tendency to spatial correlation within one region despite the stochastic behavior of each of the node price series. An illustration of this is the clustering of statistical characteristics of prices: the average and standard deviation for the period under review on the example of some subjects of the Russian Federation (*Fig. 2*).

There are a number of features that need to be taken into account when forming a list of explanatory variables:

1. Significant autocorrelation of prices. This phenomenon is described in detail in the work of Zolotova and Dvorkin [16] at the level of price zones. Our own analysis showed that in particular cases, at the node level, the picture is generally similar.

2. Annual seasonality. The seasonal component changes synchronously in most regions of Russia. Price growth is observed in the third quarter, which is associated with the dynamics of the all-Russian production cycle and the beginning of the heating period (the reasons for the annual seasonality are analyzed in more detail in the work of Prokhorova et al. [28]). The exceptions are the Krasnoyarsk Territory and the Irkutsk Region (see *Fig. 3*), in which, on the contrary, there is a decrease in prices in the 3rd quarter, which may be due to the availability of relatively cheap electricity from hydroelectric power plants in the regions.

3. The behavior of prices within the week at the regional level has some relatively constant frequency: in most regions on Monday and Friday prices are at the maximum level, on Saturday – at the minimum. The exceptions are the Irkutsk region and the Republic of Buryatia, where prices on Saturday are on average the highest. However, during the working week, the dynamics may differ from region to region. On public holidays, there is a significant decrease in prices in all regions, with the exception of a number of regions of the North Caucasian Federal District (*Fig. 4*).

4. Nonlinear dependence on temperature. The above-mentioned study by Prokhorova et al. [28] noted the need to take into account both steady changes in temperature dynamics throughout the year affecting the annual seasonality in electricity consumption, and weather factors that take into account deviations from the norm.

3. Research methodology

Analysis of price dynamics in the nodes of the energy system of the Russian Federation made it possible to reveal the presence of spatial autocorrelation. The statistical characteristics of prices differ more and more as the nodes move away, which is associated with changing conditions of supply and demand. For this reason, a predictive model can be specific to:

a) node and take into account the spatial lag (spatial autocorrelation),

b) region and have a multiple output – a forecast simultaneously in all nodes of the regional energy system, where the format of the input data involves taking into account geographic connectivity.

In the second case (option b), a significant number of observations over time is required to obtain a mod-

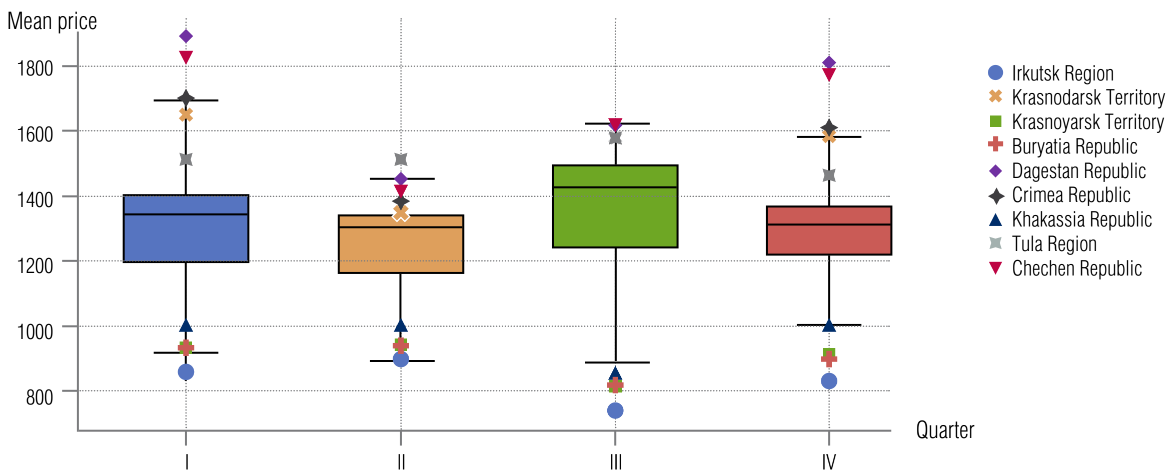

Fig. 3. Diagram of the range of average DAM prices depending on the quarter (the calculation was made for all regions in the sample, emissions are shown for individual regions of the Russian Federation).

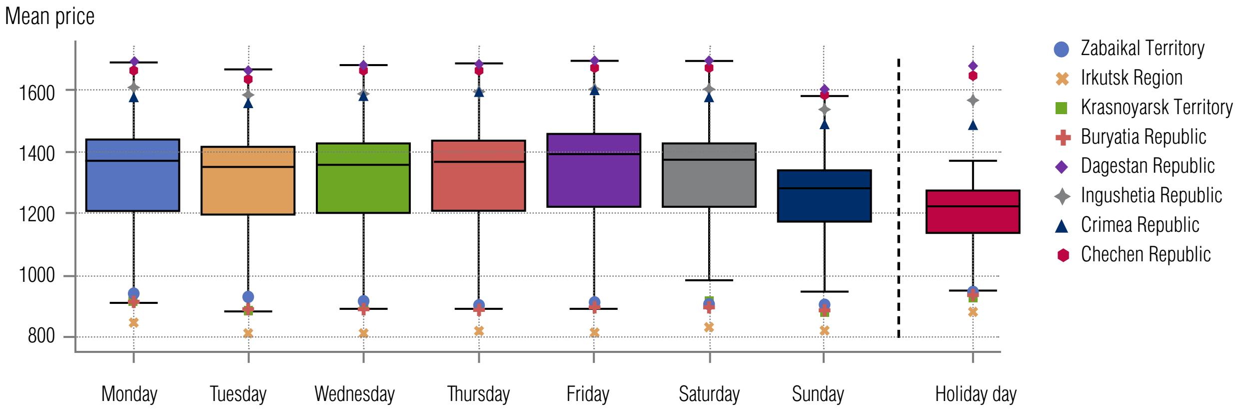

Fig. 4. Diagram of the range of averaged DAM prices by days of the week and holidays (the calculation is carried out for all regions in the sample; emissions are shown for individual regions of the Russian Federation).

el with a high generalizing ability. Unfortunately, the limitations of the available data retrospective on the website of the Trading System Administrator do not allow training such a model. Therefore, in our study, we will focus on the first option (option a), which involves building a universal neural network architecture, which, if trained on node-specific data will allow us to predict prices for any individual node of the energy system without significant error variance.

Taking into account the above data features, the following factors will be used among the explanatory variables:

- lagged values of electricity prices in the target node;
- lag values of electricity prices of the three nearest nodes within the regional network;
- dummy variables for weekends and holidays;
- average daily temperature in the region;
- temperature deviation from moving average.

It is important to note that the node model can also include locality-specific factors, for example, the water level of rivers, which is relevant for the regions of the Urals and Siberia [16], the capacities of industrial consumers, and others. However, the lack of statistics in the required context does not provide this opportunity.

As noted above, at the first step of modeling, a static recurrent neural network is formed which is optimized, and its stability is checked at random nodes of the power system. At the second step of the modeling, a GAN is formed based on the existing architecture of a static recurrent neural network, which is used as a generator. After optimizing the discriminator architecture and GAN hyperparameters, the accuracy of the prediction results of the constructed model is compared with alternative benchmarks.

Initially, all models are trained in parallel on 5 random nodes in each region[7]. Thus, each model is sequentially trained at 330 nodes[8], and the average error and its variance obtained from the test data are used to compare the quality of the models. Accordingly, data preprocessing for all nodes is unified: the sample is divided into training, validation and test samples in the ratio of 80%, 10% and 10%; data standardization and tensor transfor

[7] In order to save computing resources, a limited number of nodes are used for testing the model. This is because each region has a different number of nodes with full data coverage (up to 546 nodes).

[8] 5 nodes · 66 regions.

mation in the format of sliding windows from the original time series (the dimension of the time window is a hyperparameter).

3.1. Building a static generator model (the first step of modeling)

The static model is a two-layer recurrent neural network GRU[1] with 55 cells in the first layer and 20 in the second[2]. The first layer at the output preserves the dimension of the data in time for the next recurrent layer. The second layer transmits a vector of dimension 3 as output data. Thus, the initial matrix of input data of dimension (7 x 8), where 7 is the size of the time window, 8 is the number of explanatory variables, the model maps into a space of dimension (1 x 3). The sum of the elements of this vector is a forecast of the price of electricity for one day ahead (the elements of the vector are summed taking into account the weights, the values of which are selected during the training of the model). The hyperparameters of the model (the size of the time window that determines the number of time lags for all variables, the number of elements of the training sample used to calculate one iteration of gradient descent) were determined empirically as a result of iteration. The criterion was the standard error on the test sample[3].

3.2. Building a GAN (second modeling step)

The generative-adversarial model is a generator and discriminator connected in series (*Fig. 5*). A previously defined model of a recurrent neural network is used as a generator. A noise vector is added to the input data matrix as a separate factor, which prevents the generator from retraining and allows adding a stochastic element to the output of the model.

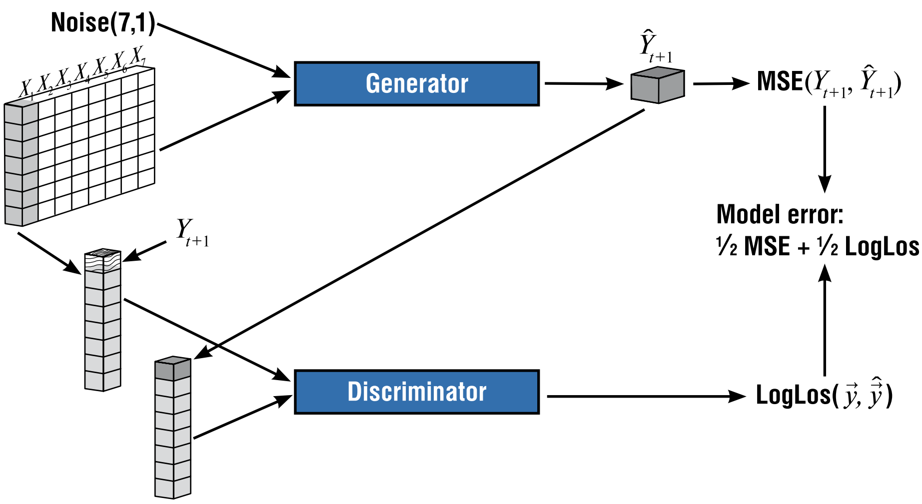

Fig. 5. The device of a generative-adversarial model used to build a forecast of node-by-node electricity prices on the DAM.

[1] Gated Recurrent Units is a type of recurrent neural network designed to model time sequences.

[2] The number of neurons in each layer, the type of activation function and the gradient descent step were determined as a result of enumeration on a given set using the KerasTuner neural network optimization package.

[3] Experiments to determine the underlying architecture were made on the data of a randomly selected node. Later, the model was further trained for each node individually.

The discriminator architecture is a convolutional neural network with one hidden layer that accepts dimension 7 as input. The first 6 elements of the vector are the actual price values in retrospect, and the last element is the generator forecast or the actual price, depending on the template being submitted to the input. The discriminator learns to classify incoming data: (X, Y) – "True" or 1 and $(X, G(X))$ – "False" or 0. The discriminator learns using the error function (Equation 2), which is based on cross-entropy (BCA – binary cross entropy), Equation 3:

$$L_D = l_{bce}(D(X,Y), 1) + l_{bce}(D(X,G(X)), 0), \quad (2)$$

$$l_{bce} = -(y \cdot \log(p) + (1 - y) \cdot \log(1 - p). \quad (3)$$

To train the generator, a two-component error function is used, consisting of cross–entropy and mean squared error (MSE – mean squared error), Equation 4.

$$L_G = \lambda_1 l_{bce}(D(X,G(X)), 1) + \lambda_2 \sum (Y - G(X))^2. \quad (4)$$

The cross-entropy metric is calculated by assigning the "True" label to the vector with the predicted price value. This is necessary to mislead the discriminator and allows the generator to pick up weights in the learning process that complicate the task of training the discriminator in the subsequent iteration, which introduces an element of "competition" into the learning process. It is important to note that theoretically adding the MSE component to the error function is not strictly necessary [29]. We are adding MSE for faster generator convergence and shorter training time.

As mentioned above, the discriminator and generator are trained alternately until the specified accuracy of the price forecast is achieved. This provides a solution to the original minimax problem. A general description of the entire learning process is shown in pseudocode (*Table 1*).

More detailed information about the algorithm of generative adversarial network learning is contained in [30, 31].

Thus, the discriminator learns to distinguish generated samples from real data in the learning process. At the next iteration of training, the generator strives to improve its forecast so that the probabilistic output of the model is closer to the "Truth" with fixed discriminator weights. This approach allows you to change the gradient trajectory in the learning process and come to a different optimum in comparison with the static model, where only the root-mean-square error is used [32].

The results of the GAN model are compared with "reference" alternative models: ARIMA autoregression, basic recurrent neural network, two-layer convolutional neural network (CNN). The architecture of the latter is based on the model proposed in [33].

4. Empirical results

We trained the discriminator five times more than the generator during each training iteration. This allows the discriminator to be somewhat ahead of the generator in accuracy and distinguish the generated data from the real ones. *Fig. 6* shows the process of learning a generative-adversarial model for some random node.

The left axis shows the total error of the model, which is the weighted sum of the error of the discriminator and the generator (mse + logloss). The value of the discriminator error at each training iteration (logloss) is shown on the right axis.

In order to test the stability of the proposed model, the following experiment was conducted: 5 nodes of the power system were randomly selected in each region. The only selection criterion is the absence of gaps in the data on the analyzed time horizon. Next, a set of explanatory variables was formed for all nodes in accordance with a unified procedure, data was preprocessed, and then the previously saved generative model was further trained on the data of each of the 5 nodes. Based on the results of the experiment, the error of the electricity price forecast was calculated on a test sample (in parallel, training and error calculation were carried out for a static generator model and other benchmarks).

Table 1.

General algorithm of generative-adversarial neural network training

Setting conditions: gradient descent step ρ_D, ρ_G; error function weights parameters for the generative model λ_1, λ_2; random initialization of weights in discriminator and generator models.

While (until the algorithm converges):

Discriminator Training (*D*):

Getting M samples from a training sample: X^{train}:

$(X, Y) = (X^1, Y^1), \ldots, (X^m, Y^m) \subset X^{train}$

Stochastic gradient descent step and updating of weights D at fixed weights G::

$$W_D = W_d - \rho_D \sum_{i=1}^{M} \frac{\partial L_D\left(X^{(i)}, Y^{(i)}\right)}{\partial W_D}$$

Generator Training (*G*):

Getting M new samples from X^{train}:

$(X, Y) = (X^1, Y^1), \ldots, (X^m, Y^m) \subset X^{train}$

Stochastic gradient descent step and updating of weights G at fixed weights D:

$$W_G = W_G - \rho_G \sum_{i=1}^{M} \frac{\partial L_G\left(X^{(i)}, Y^{(i)}\right)}{\partial W_G}$$

End While

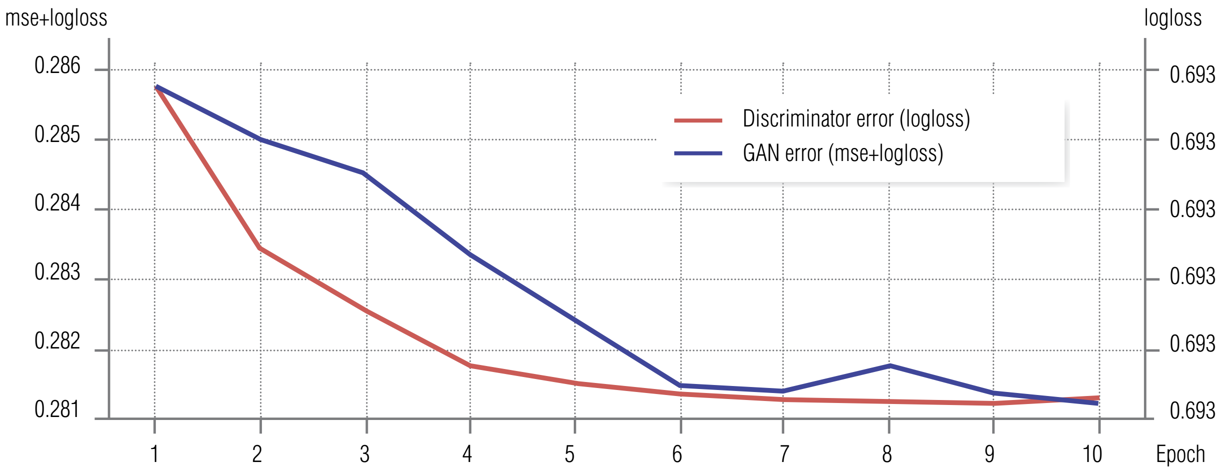

Fig. 6. The value of errors of the generative-adversarial model in the learning process on a training sample.

The results of the experiment showed that the generative-adversarial model demonstrates the smallest average error on the test sample and the minimum spread of error values for individual nodes of the power system. *Table 2* shows the average errors on the test sample of the models under consideration and their standard deviations.

In [25], the authors associate the superiority of GAN with the ability to provide the required gradient for optimizing the generator during training: the gradient directed by the discriminator allows one to achieve a wider minimum than, other things being equal, is achieved as a result of optimizing the static analogue of the generative model. Our own empirical results also support this thesis, since with a near-zero error on the training set achieved simultaneously by a static generator model and a GAN, the error of the latter on the test set is lower.

The average errors of the GAN model and their standard deviations are visualized in *Fig. 7*: the color gradation on the conditional map of regions corresponds to the value of the average error calculated for five randomly selected nodes in the corresponding region of the Russian Federation (see the left part of *Fig. 7*) and the variance of average errors (see the right part of *Fig. 7*).

The model demonstrates the best results in most of the regions of the Central Federal District, the Northwestern Federal District and the Volga Federal District. In turn, the model demonstrates the largest average error in such regions as the Republic of Tatarstan, Bashkiria, a number of regions of Siberia and the Caucasus. The high variance of the average error is also manifested mainly in these regions. The unsatisfactory performance of the model in these regions seems to be due to the insufficiency of a set of explanatory factors, which in all cases was standard. It is possible that energy bridges with other countries connected with the domestic energy system are important for pricing in border regions; for the central regions of Siberia, the water level of rivers is an important factor, since a significant share of generation falls on hydroelectric power plants.

It is also important that the test part of the sample (from August 19, 2022 to December 31, 2022) accounts for a period with a clear structural shift – a sharp increase in electricity prices at the end of 2022 in regions such as Irkutsk, Tyumen and Tomsk Regions, Krasnoyarsk Krai, Stavropol Krai, Krasnodar Krai. The sharp increase in the price trend in this case was due to conjunctural factors, the influence of which could not be present in the training sample.

Table 2.

Comparative table of averaged errors and their spread when applying the model to five random nodes in the analyzed regions

Indicator		Error on the test sample			
		GAN	GRU	CNN	ARIMA
MAE	mean	0.0497	0.062	0.0681	0.0724
	std	0.0002	0.0009	0.0063	0.0092
MAPE	mean	0.092	0.0973	0.1055	0.1114
	std	0.0005	0.0018	0.0113	0.0242
MSE	mean	0.0046	0.0059	0.0073	0.0082
	std	0	0.0002	0.0013	0.0034

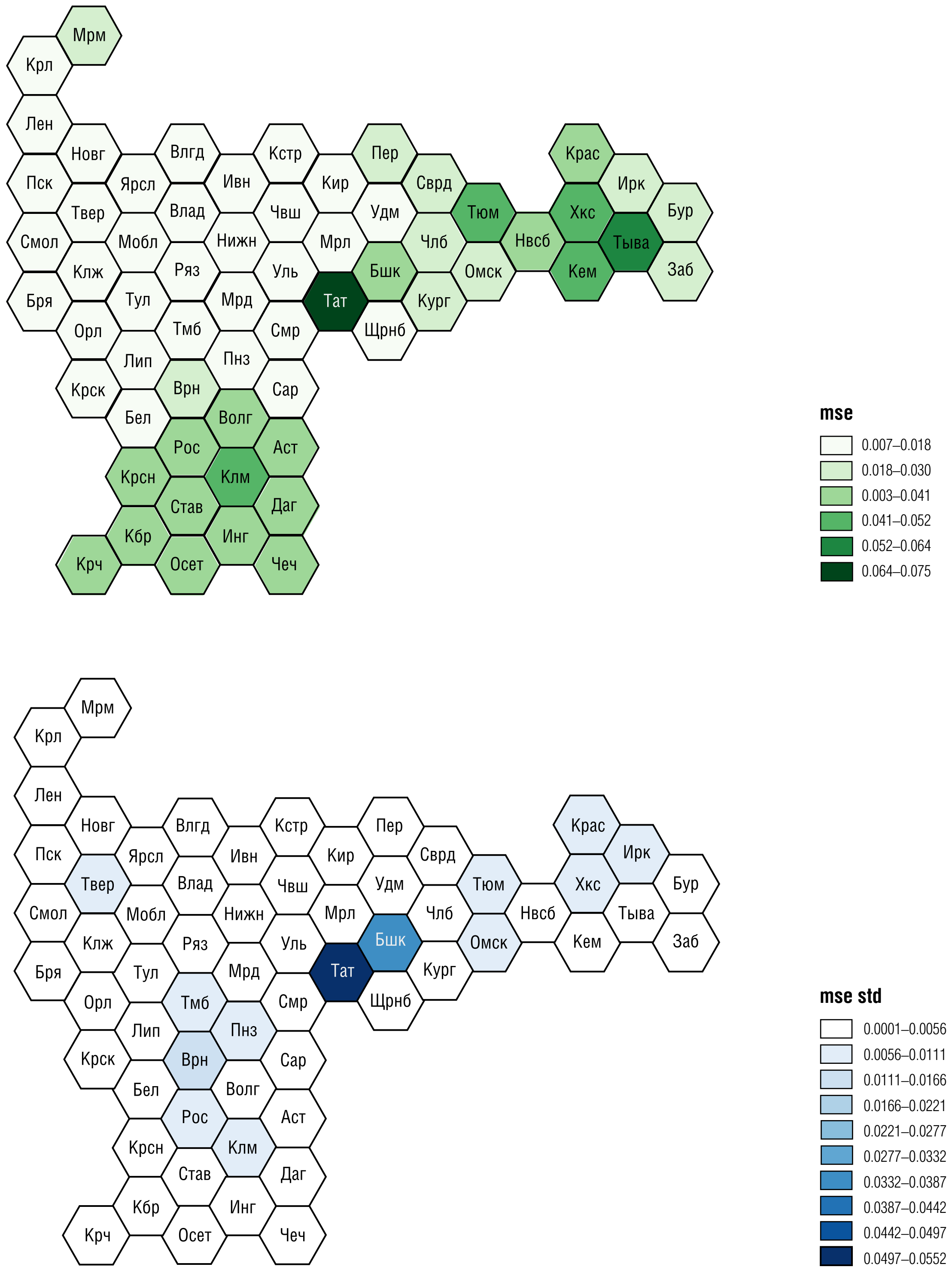

Fig. 7. The average error of the generative-adversarial model in the test sample (graph below) and its variance (right graph) for five random nodes in the region.

The hypothesis of the group equality of the Levene variances for the root-mean-square errors obtained at randomly selected nodes of the power system within a single region. The null hypothesis of the test is that "all subsamples have equal variances". When outliers are excluded (12 regions with the largest standard deviation of error), the null hypothesis of the test is not rejected for the remaining 54 regions. Thus, for most regions, the proposed architecture of the generative model is universal and is able to provide a relatively low error on the test sample for various nodes of the country's power system.

Conclusion

The results of the study demonstrate that the proposed generative-adversarial network can be used to predict prices in DAM nodes for the day ahead without significant deviations in accuracy for 54 of the 66 regions of Russia under consideration. The network architecture is universal (it does not change during the transition from region to region of the Russian Federation) and uses a limited number of explanatory factors. The network needs to be retrained on data specific to the target node. The neural network model included the following set of variables: historical price values in the target and geographically close to it nodes of the power system (node-by-node prices in the electricity market are correlated both spatially and in time), ambient temperature and seasonal factors.

The proposed generative-adversarial model reduced the mean square error by 22% on a test sample of a static generator model based on a recurrent neural network, and also surpassed the quality of alternative benchmark models: convolutional neural network and autoregressive linear model (ARIMA). ■

References

1. Akopov A.S., Beklarjan G.L. (2005) Analysis of the effectiveness of the regulatory policy of the state using the regional CGE model of the behavior of natural monopolies (on the example of the electric power industry). *Jekonomicheskaja nauka sovremennoj Rossii*, no. 4(31), pp. 130–139 (in Russian).
2. Akopov A.S., Beklarjan A.L., Hachatrjan N.K., Fomin A.V. (2017) Oil production dynamics forecasting system using simulation modeling. *Informacionnye tehnologii*, vol. 23, no. 6, pp. 431–436 (in Russian).
3. Makarov V.L., Bahtizin A.R., Beklarjan G.L., Akopov A.S. (2021) Digital plant: methods of discrete-event modeling and optimization of production characteristics. *Business Informatics*, vol. 15, no. 2, pp. 7–20. https://doi.org/10.17323/2587-814X.2021.2.7.20
4. Hachatrjan N.K., Kravchenko T.K., Akopov A.S., Uvarova O.M. (2016) Forecasting of the main indicators of the Russian stock market by autoregressive models with distributed lags. *Audit and financial analysis*, vol. 3, pp. 128–133 (in Russian).
5. Gonzalez J.P., San Roque A.M., Perez E.A. (2017) Forecasting functional time series with a new Hilbertian ARMAX model: Application to electricity price forecasting. *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 545–556. https://doi.org/10.1109/TPWRS.2017
6. Zhao Z., Wang C., Nokleby M., Miller C.J. (2017) Improving short-term electricity price forecasting using day-ahead LMP with ARIMA models. *2017 IEEE Power & Energy Society General Meeting*, pp. 1–5. https://doi.org/10.1109/PESGM.2017.8274124
7. Vehviläinen I., Pyykkönen T. (2005) Stochastic factor model for electricity spot price – the case of the Nordic market. *Energy Economics*, no. 2, pp. 351–367.

8. Howison S., Coulon M. (2009) Stochastic behaviour of the electricity bid stack: from fundamental drivers to power prices. *The Journal of Energy Markets*, pp. 29–69.

9. Weron R. (2014) Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081. https://doi.org/10.1016/j.ijforecast.2014.08.008

10. Garcia R.C., Contreras J., Akkeren M., Garcia J.B. (2005) A GARCH forecasting model to predict day-ahead electricity prices. *IEEE Transactions on Power Systems,* vol. 20, no. 2, pp. 867–874. https://doi.org/10.1109/TPWRS.2005.846044

11. Lynch C., O'Leary C., Sandareshan P.G., Akin Y. (2021) Experimental analysis of GBM to expand the time horizon of Irish electricity price forecasts. *Energies*, vol. 14, no. 22, 7587. https://doi.org/10.3390/en14227587

12. Deng Z., Liu C., Zhu Z. (2021) Inter-hours rolling scheduling of behind-the-meter storage operating systems using electricity price forecasting based on deep convolutional neural network. *International Journal of Electrical Power & Energy Systems*, vol. 125, 106499. https://doi.org/10.1016/j.ijepes.2020.106499

13. Zhang J., Tan Z., Yang S. (2012) Day-ahead electricity price forecasting by a new hybrid method. *Computers and Industrial Engineering*, vol. 63, pp. 695–701. https://doi.org/10.1016/j.cie.2012.03.016

14. Kim T.Y., Cho S.B. (2019) Predicting residential energy consumption using CNN-LSTM neural networks. Energy, vol. 182, pp. 72–81. https://doi.org/10.1016/j.energy.2019.05.230

15. Mar'yasin O.Yu., Lukashov A.I. (2020) Prediction of free electricity prices using neural networks. *Neiroinformatika-2020. Sbornik nauchnykh trudov. XXII mezhdunarodnaya nauchno-tekhnicheskaya konferentsiya*, pp. 294–303 (in Russian).

16. Zolotova I.Yu., Dvorkin V.V. (2017) Short-term forecasting of prices in the Russian wholesale electricity market based on neural networks. *Problemy prognozirovaniya*, no. 6, pp. 47–57 (in Russian).

17. Guo X., Zhao Q., Zheng D., Ning Y., Gao, Y. (2020) A short-term load forecasting model of multi-scale CNN-LSTM hybrid neural network considering the real-time electricity price. *Energy Reports*, vol. 6, pp. 1046–1053. https://doi.org/10.1016/j.egyr.2020.11.078

18. Wu W., Huang F., Kao Y., Chen Z., Wu Q. (2021) Prediction method of multiple related time series based on generative adversarial networks. *Information*, vol. 12, no. 2, 55. https://doi.org/10.3390/info12020055

19. Nikolenko S., Kadurin A., Arkhangel'skaya E. (2019) *Glubokoe obuchenie*. Sankt-Peterburg, Piter (in Russian).

20. Goodfellow I., Pouget-Abadie I., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. (2014) Generative adversarial networks. *arXiv*:1406.2661. https://doi.org/10.48550/arXiv.1406.2661

21. Zhang H., Zhang H., Yu Koh J., Baldridge J., Lee H., Yang Y. (2021) Cross-modal contrastive learning for text-to-image generation. Proceedings of the *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 833–842.

22. Kadurin A., Nikolenko S., Khrabrov K., Aliper A., Zhavoronkov A. (2017) DruGAN: An advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties in silico. *Molecular Pharmaceutics*, vol. 14, no. 9, pp. 3098–3104. https://doi.org/10.1021/acs.molpharmaceut.7b00346

23. Yoon J., Jarrett D., van der Schaar M. (2019) Time-series generative adversarial networks. *Neural Information Processing Systems*.

24. Koochali A., Dengel A., Ahmed S. (2021) If you like it, GAN it — Probabilistic multivariate times series forecast with GAN. *Engineering Proceedings*, vol. 5, no. 1, 40. https://doi.org/10.3390/engproc2021005040

25. Zhang Z., Wu M. (2020) Real-time locational marginal price forecasting using generative adversarial network. *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids*, pp. 1–6. https://doi.org/10.1109/SmartGridComm47815.2020.9302938

26. Kabir H.M., Khosravi A., Nahavadini S., Kavousi-Fard A. (2019) Partial adversarial training for neural network-based uncertainty quantification. *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 5, no. 4, pp. 595–606. https://doi.org/10.1109/TETCI.2019.2936546

27. *Otchet o ravnovesnykh tsenakh v naibolee krupnykh uzlakh raschetnoi modeli*. Administrator Torgovoi Sistemy. Available at: https://www.atsenergo.ru/nreport?rname=big_nodes_prices_pub&region=eur (accessed 02.03.2023) (in Russian).

28. Porokhova N.V., Rudakov E.N., Saakyan Yu.Z. (2009) Alternative methods for monitoring the dynamics of industrial production. *Problemy Prognozirovaniya*, no. 4, pp. 36–53 (in Russian).

29. Zhang K., Zhong G., Dong J., Wang S., Wang Y. (2019) Stock market prediction based on generative adversarial network. *Procedia Computer Science*, vol. 147, pp. 400–406. https://doi.org/10.1016/j.procs.2019.01.256

30. Mathieu M., Couprie C., LeCun Y. (2015) Deep multi-scale video prediction beyond mean square error. *arXiv*:1511.05440. https://doi.org/10.48550/arXiv.1511.05440

31. Radford A., Metz L., Chintala S. (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv*:1511.06434. https://doi.org/10.48550/arXiv.1511.06434

32. Staffini A. (2022) Stock price forecasting by a deep convolutional generative adversarial network. *Frontiers in Artificial Intelligence*, vol. 5, 837596. https://doi.org/10.3389/frai.2022.837596

33. Liu S., Ji H., Wang M.C. (2020) Nonpooling convolutional neural network forecasting for seasonal time series with trends. *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 8, pp. 2879–2888. https://doi.org/10.1109/tnnls.2019.2934110

34. Izmailov P., Maddox J., Kirichenko P., Garipov T., Vetrov D., Gordon A. (2020) Subspace inference for Bayesian deep learning. Proceedings of the *35th Uncertainty in Artificial Intelligence Conference. Proceedings of Machine Learning Research*, vol. 115, pp. 1169–1179.

About the authors

Andrej S. Kaukin

Cand. Sci. (Econ.);

Head of the Laboratory of Industry Markets and Infrastructure, Gaidar Institute for Economic Policy, 3–5, Gazetnyy Lane, Moscow 125993, Russia;

Head of the Laboratory of System Analysis of Industrial Markets, Institute of Applied Economic Research, Russian Presidential Academy of National Economy and Public Administration, 82, Vernadskogo Prospect, Moscow 119571, Russia;

E-mail: kaukin@iep.ru

ORCID: 0000-0003-2892-5278

Pavel N. Pavlov

Cand. Sci. (Econ.);

Senior researcher, Laboratory of System Analysis of Industrial Markets, Institute of Applied Economic Research, Russian Presidential Academy of National Economy and Public Administration, 82, Vernadskogo Prospect, Moscow 119571, Russia;

E-mail: pavlov@ranepa.ru

ORCID: 0000-0001-6200-6481

Vladimir S. Kosarev

Research Fellow, Laboratory of System Analysis of Industrial Markets, Institute of Applied Economic Research, Russian Presidential Academy of National Economy and Public Administration, 82, Vernadskogo Prospect, Moscow 119571, Russia;

E-mail: kosarev-vs@ranepa.ru

ORCID: 0000-0001-5435-9076