Development of mathematical models for the short-term forecasting of daily consumption schedules of active power by Moscow

S A Vyalkova¹, O Kornykova² and I Nadtoka¹
¹Department of Energy, Platov South-Russian State Polytechnic University (NPI), Novocherkassk 346428, Russia
²Department of Information Technology And Management, Platov South-Russian State Polytechnic University (NPI), Novocherkassk 346428, Russia

Email: ii_nadtoka@mail.ru

Abstract. The present article presents the results of research into solving the problem of increasing the accuracy of forecasting power consumption. The purpose of these studies is to develop mathematical models for short-term forecasting of daily schedules of active power consumption in Moscow, taking into account meteorological factors. Research has been carried out on four predictive models based on singular spectral analysis (SSA), least-squares method, trigonometric interpolation, neural and neural fuzzy networks (NFN). It is shown that the NFN and hybrid model based on MSSA and NFN has the smallest error.

1. Introduction

One of the main functions of the System Operator is planning the optimal daily work schedules for power plants and electric grids of the Unified Energy System of Russia. Improving the accuracy of the forecast of electricity consumption is a very urgent and significant task, since the reliability of the operating modes of power systems directly depends on the forecasting accuracy.

Meteorological factors have a significant impact on power consumption in the power system [1-10], which is why, when forecasting in real conditions, dispatch technologists use the predicted values of air temperature, cloud cover, precipitation and natural light.

In this article, four models are proposed for short-term forecasting of daily active power graphs and their comparative analysis is carried out. When carrying out the research, the archived data of daily active power graphs, actual and forecast data on air temperature, cloudiness, precipitation and natural illumination for the territory of Moscow were used. These data were obtained at the Moscow Branch of SO UES for the period from 1 to 29 February 2016. The choice of additional parameters of the mathematical forecasting model is based on the results of the authors [1-3], which show the effect of air temperature and illumination on the power by hours of the daily schedule. Software implemented in the high-level C# language was used as a modeling tool.

2. Problem

Currently, the solution to the problem of forecasting power consumption has a number of features associated with taking into account the actual hourly values of active power \( P(t) \), air temperature \( \theta(t) \), clouds \( O(t) \), precipitation \( \Psi(t) \) and natural light \( E(t) \) current day and predicted temperature data \( \theta_F(t) \), clouds \( O_F(t) \), precipitation \( \Psi_F(t) \) and illumination \( E_F(t) \) for the next day:
Taking into account the actual data of the current day \( P(t) \), \( \theta(t) \), \( O(t) \), \( \Psi(t) \) and \( E(t) \) is carried out from the midnight until the moment of forecasting \( t \), while for the previous days, information about which should be taken into account when building the forecast model, 24-hourly measurements of the actual values of these quantities are used;

- the sampling frequency of the predicted data, and received from the hardware and software system "Meteo" is six times lower than the required one: for each forecast day there are only four measurements: for 3.00 AM, 9.00 AM, 15.00 PM and 21.00 PM hours instead of 24 hourly measurements;
- due to the current lack of forecast data on natural illumination \( E_F(t) \), to determine the values of this quantity, a separate forecast model has to be used, the arguments of which are the actual and forecast values of cloudiness and precipitation;

Taking into account the input data for the forecast model significantly affects the quality of forecasting, increasing its error.

The inputs were taken into account as follows in the developed model. Source data were archives \( P(t) \), \( \theta(t) \), \( O(t) \), \( \Psi(t) \) and \( E(t) \) for to build a forecast. The time interval of the initial data is \( N \in [X - q - 2; X] \) in 2015 and 2016, and \( N \in [X - q; X + q] \) in 2015, where \( X \) - forecast date, \( q \) - number of days in the input interval.

The input interval length is \( q \) in forecasting model algorithms \( P(t) \) for different types of days was:
- \( q = 15 \) for working days;
- \( q = 60 \) for weekends, post-holiday, pre-holiday, irregular days;
- \( q \) for holidays is equal to the number of holidays in the forecast month for the three previous years.

The amount of source data of active energy consumption \( P(t) \) is equal to \( N = 24 \times d \) h, here \( d \) is the number of selected days, taking into account the type of natural day from the previously given interval of source data. The data for the current natural day is also taken into account, i.e. day \( H + X \), where \( H \) is the missing hours of the current natural day.

Accounting for air temperature and natural illumination when predicting active power in the input layer of the described neural fuzzy network (NFN) was carried out by taking into account the data: \( P(t) \), \( \theta(t) \), \( E(t) \), \( \bar{E}_F(t) \), \( E_F(t) \) and the average values of these quantities for four hours \( t = 3, 9, 15, 21h \) [11].

To solve the problem of forecasting illumination \( \bar{E}_F(t) \), forecasting of natural illumination was carried out taking into account the average daily actual and forecast data of cloudiness and precipitation. The predictive model for natural light is based on a multilayer neural network. To implement the predictive model, the R-Studio software was used [12-13].

When predicting daily graphs of active power, the following models were studied:
1) Hybrid predictive model based on one-dimensional singular spectral decomposition and forecasting by trigonometric interpolation, least squares approximation method (SSA and TI, LS);
2) Neural network;
3) Neural fuzzy network;
4) Hybrid predictive model (MSSA and NFN).

In Fig. 1 shows a diagram of a system of models for short-term forecasting of daily active power \( P(t) \) graphs.
In this article, when forecasting using hybrid models based on SSA - Singular Spectrum Analysis and MSSA - Multi-Channel Singular Spectrum Analysis [12-15], at the first stage, the series was divided into three components: trend, harmonics and noise component. At the second stage of forecasting the selected trend and harmonic components. Prediction was also performed using a multilayer perceptron neural network and a Takagi-Sugeno-Kanga neural fuzzy network (TSK) [14-17]. In works [4-6], strong relationships were established between active power and air temperature during the day, as well as complex nonlinear relationships between active power and natural illumination during daylight hours for the period from 2012 - 2018. Seasonal component in the investigated hybrid models, was obtained on the basis of the input data interval for 2015-2016.

3. Some theoretical basic

3.1. Hybrid forecasting model SSA, TI and LSM

The method of one-dimensional singular spectral analysis used to study the time series of daily graphs of active power is described in detail in [5, 7].

A number of daily graphs of active power $P_N = (p_0, ..., p_{N-1})$ is presented as the sum of additive components:

$$ P = P_T + P_h + P_t $$

(1)

here $P_T$ - trend, $P_h$ - harmonic, and $P_t$ - noise.

The difference between the methods of one-dimensional and multivariate singular spectral analysis is in the construction of the trajectory matrix.

Trajectory matrix $X$ has the form for one-dimensional series $P_N$:

$$ X = [p_1 : : : p_K] = X^{(1)} $$

here $P_N$ - active power range of the current year, $1 \leq i \leq K$, $N$ - the length of the original row.

The nesting procedure forms $K = N - L + 1$ vectors $L$ - attachments:

$$ P_i = (p_{i-L-1}, ..., p_{i+L-2})^T $$

where $1 \leq i \leq K$. 

![Figure 1. Structure of short-term forecasting of daily active power graphs](image-url)
The basic SSA algorithm consists of several complementary stages: embedding, decomposition, grouping and renewal [14].

After applying the SSA method to predict individual additive components $P_T, P_h$, the least squares method (LSM) was used for the trend component, and for the harmonic method, trigonometric interpolation (TI).

3.2. Hybrid forecasting model MSSA and NFN

When studying internal regularities by the MSSA method, the time series of active power $P_N = (p_0, ..., p_{N-1})$ is presented similarly to formula (1).

Trajectory matrix $X$ of a multidimensional series $(\tilde{P}_N, P_N)$ has the form:

$$X = \left[ \begin{array}{c} \tilde{P}_1 \ldots \tilde{P}_K \ P_1 \ldots P_K \end{array} \right] = \left[ x^{(1)}, x^{(2)} \right],$$

where $P_N$ is the row of active power of the current year; $\tilde{P}_N$ – power range of the previous year; $N$ – is the length of the original row [14-16].

The nesting procedure forms $K = N - L + 1$ vectors $L$-attachments for each of the time series:

$$\tilde{P}_i = (\tilde{P}_{i-1}, ..., \tilde{P}_{i+L-2})^T,$$

$$P_i = (p_{i-1}, ..., p_{i+L-2})^T,$$

where $1 \leq i \leq K$.

$$X^{(1)} = \begin{bmatrix} \tilde{P}_0 & \tilde{P}_1 & \tilde{P}_2 & \ldots & \tilde{P}_{K-1} \\ \tilde{P}_1 & \tilde{P}_2 & \tilde{P}_3 & \ldots & \tilde{P}_{K} \\ \tilde{P}_2 & \tilde{P}_3 & \tilde{P}_4 & \ldots & \tilde{P}_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{P}_{L-1} & \tilde{P}_{L} & \tilde{P}_{L+1} & \ldots & \tilde{P}_{N-1} \end{bmatrix},$$

$$X^{(2)} = \begin{bmatrix} p_0 & p_1 & p_2 & \ldots & p_{K-1} \\ p_1 & p_2 & p_3 & \ldots & p_K \\ p_2 & p_3 & p_4 & \ldots & p_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{L-1} & p_L & p_{L+1} & \ldots & p_{N-1} \end{bmatrix}.$$

The composition and sequence of stages of the basic MSSA algorithm is similar to the SSA method [12].

One of the defining provisions in SSA and MSSA are: the choice of the length of the window $L$ (in hours), as well as the interval of the initial power data $(P(t) = (P_0(t), ..., P_i(t), ..., P_{N-1}(t)), \text{ where } N = \text{number of days}, i = \text{day number}, t = 1, \ldots, 24 = \text{times of day}).$

The window length for both methods (SSA and MSSA) was chosen equal to the length of the daily power graph $L = 24$ hours on the basis of research presented in the works [4-7].

For automatic grouping of vector components, the algorithm proposed in [15] was used, the essence of which is to select the required component from a series by identifying eigenvectors using the low-frequency method.

After obtaining the additive components by the MSSA method, the trend prediction $P_T(t), P_h(t)$ was carried out separately using the NFN [16, 17].

3.3. Prediction model based on neural networks

Studies of the predictive model of daily graphs of active power consumption in Moscow were carried out on the basis of an artificial neural network of a multilayer perceptron and a neural fuzzy network (TSK).

A multilayer perceptron consists of several layers of neurons: an input layer, an output layer, and one “hidden” layer. The neurons of each layer are not connected with each other and interact only with the neurons of the previous and subsequent layers [8-10, 18].
Initial data of power consumption and meteorological factors are received at the inputs of the NN. The output of the NN is the forecast of the power vector $\mathbf{P}_{t+k}$. All input data are pre-normalized. In addition to the input data for the neural network, the following parameters are set:

- the number of neurons in the input layer is equal to the amount of power data and meteorological factors, taking into account the type of day;
- the number of hidden layers is equal to one; the number of neurons in the hidden layer – 30; the number of neurons in the output layer – 24;
- constant characterizing the learning rate $\eta = 0.7$; experimentally found parameter, $\alpha = 0.01$. These parameters were found as a result of 50 experiments;
- bias $\Theta_j$ – unit vector in the $j$-layer.

The activation function was found by the formula:

$$net^{(K)}_j = \sum_{i=1}^{n} w_{ij} x_i + \Theta_j,$$

To activate the neuron, a sigmoidal function equal to [10] was used:

$$f(net^{(K)}_j) = \frac{1}{1+e^{-\alpha \cdot f(net^{(K)}_j)}}.$$

The mathematical task was to find the weighting coefficients at which the error in the functioning of the network for the learning sample would be minimal:

$$\min_{1} \sum_{p} E_p \rightarrow \min,$$

where $E_p$ an error, the value of which must satisfy the condition [11] $E_p > \zeta$, $\zeta = 10^{-5}$.

$$E_p = \frac{1}{2} \sum_{j=1}^{n} (net^{(K)}_j - d_j)^2,$$

where $d_j$ – desired network output, $p$ – index of the sample in the training set, $K$ – layer number, $j=1..n$, $n$ – number of neurons in the previous layer [11].

3.4. Forecasting model based on fuzzy neural networks

The Takagi-Sugeno-Kanga (TSK) fuzzy neural network algorithm consist of six layers: input data, determining the degree of membership $\mu_A^{(k)}(x_i)$ for each variable $x_i (i=1,2,...,N)$, for vector $x$ calculating the value of the membership coefficient $w_k = \prod_{j=1}^{N} \mu_A^{(k)}(x_j)$, calculating TCK functions:

$$y_k = p_{k0} + \sum_{j=1}^{N} p_{kj} x_j,$$

here $p_{kj}$ linear weights $(k=1,2,...,M), (j=1,2,...,N)$, calculation of the weighted sum of signals $y_k(x)$ and the sum of the weights $\sum_{k=1}^{M} w_k$ and calculating the output neuron.

The theory of this TSK method is described in detail in [11].

At the beginning of the neural fuzzy network algorithm, the number of input variables is determined $N \times M$ rules. Each rule forms $N+1$ variables $p_{j}^{(k)}$ linear dependence TSK. As a result, we find $M(N+1)$ linear network parameters. [11].

He learning process is divided into two stages.

**Stage 1.** Calculation of linear parameters $p_{kj}$ polynomial TSK by solving a system of linear equations by Interval Greville Algorithm for find the “Pseudo-inverse” matrix [11].
Stage 2. Calculation of output signals \( y(i) \) for \( i = 1, \ldots, p \) and error vectors. Error signals are routed through the connected network towards the network input (backpropagation) up to the first layer where the parameters are calculated \( c_{ij}^{(k)}, \sigma_{ij}^{(k)}, b_i^{(k)} \) [11].

After the refinement of the nonlinear parameters, the process of adaptation of the function parameters is started again TSK (Stage 1) and nonlinear parameters (Stage 2). This cycle was repeated until the criterion for minimizing the forecast error was met [11,19].

A fuzzy network with learning rate coefficients \( \alpha_1 = \alpha_2 = 0.2, \alpha_3 = 0.05 \) was used for forecast. These parameters were found empirically as a result of 50 experiments.

3.5. Assessment of the forecast quality

To assess the quality of the forecasts obtained, the values of the error per day were calculated as a percentage using the formula (2) [20].

\[
\varepsilon_{ij} = \frac{|p_{\text{Fact}_i} - p_{F_i}|}{p_{\text{Fact}_i}} \cdot 100\% \tag{2}
\]

where \( i = 1, 2, \ldots, 24 \) is the number of the hour; \( p_{\text{Fact}_i} \) — the actual value of active power for the \( i \)-th hour; \( p_{F_i} \) — is the value of active power for the \( i \)-th hour, obtained using the forecast model, \( j = 1, 2, \ldots, n \) is the number of the day.

The calculation \( E_{Fkj} \) — of the absolute percentage error per day was performed for the entire test set [20]:

\[
E_{Fkj} = \frac{1}{24} \sum_{i=1}^{24} \varepsilon_{ij} \tag{3}
\]

where \( k \) is the number of the month (= 1,2,…, 12).

The final average monthly forecast error for active power was calculated using the

\[
E_{Fk}^{\text{Mec}} = \frac{1}{N_F} \sum_{j=1}^{N_F} E_{Fkj} \tag{4}
\]

where \( N_F \) — is the number of days in a month [20].

4. Experimental results

The models described above predicted the active power values from February 01 to February 29, 2016 for Moscow. In Fig. 2, bar plots of the relative errors (\( E_{Fkj} \)) of these forecasts are shown for February 2016.
5. Discussions of results

Table 1 shows the results of calculating the forecasting errors of the daily active power graphs of Moscow in February 2016 with a lead time of $24 + H$ hours, where $H$ is the missing hours until the end of the current day. The table shows the maximum and minimum daily mean values of errors determined by formula (3) and monthly mean values determined by formula (4) for each of the studied models.

| Model                               | Average month values, % | Average daily Max, % | Average daily Min, % |
|-------------------------------------|--------------------------|----------------------|----------------------|
| Hybrid model (SSA and TI, LS)       | 2.45                     | 7.21                 | 0.33                 |
| Neural network (NS)                 | 1.46                     | 5.30                 | 0.39                 |
| Neural fuzzy network (NFS)          | 1.34                     | 6.52                 | 0.53                 |
| Hybrid model (MSSA and NFN)         | 1.30                     | 3.01                 | 0.42                 |

6. Conclusions

The hybrid model (MSSA and HNS), as well as the HNS model as a whole, give more accurate forecasting results than other forecasting models. Also during the research, it was determined that when predicting weekends and holidays, the forecast gave unacceptable errors. on some days can increase for the hybrid model (MSSA and HNS) up to 3.01%, and for the HNS model up to 6.52%. It is necessary to conduct additional research to improve the structure and selection of experimentally tunable parameters of the considered neural fuzzy network in order to reduce the forecast error.

The scientific results of present article can be recommended for using in other megalopolis of Russia or other countries. However, in each specific case, it is necessary to additionally analyze the initial data, since power consumption depends on the influence of various meteorological factors. Namely: the geographical latitude of the area, local weather conditions and the characteristics of the electricity consumer of the megapole. These factors require additional study in order to determine effective methods of their application.
7. References

[1] Shumilova G P, Gottman N E, Startseva T B 2008 “Prediction of electrical loads in the operational control of electric power systems based on neural network structures”. (Yekaterinburg) pp 89

[2] Nadtoka I I, Sukhominlova O A, Maltseva A A 2002 “Analysis of the dependence of power consumption in the power system on air temperature”// Methods and algorithms of applied mathematics in engineering, medicine and economics: materials of the III scientific-practical. Conf. (Novocherkassk) vol. 1 pp 9–13

[3] Nadtoka I I, Al-Zihery Balasim M “Mathematical modeling and short-term forecasting of electricity consumption of the power system, with due account of air temperature and natural illumination, based on support vector machine and particle swarm”. International Conference on Industrial Engineering. Engineering 129 (2015) 657 – 663. ICIE, 2016, Chelyabinsk. URL: https://www.sciencedirect.com/science/article/pii/S1877705815039715

[4] Nadtoka I , Vyalkova S 2019 “Hybrid Speculation Model of Energy Consumption Based on Multivariate Singular Spectrum Analysis and Neural Fuzzy Network”. 2019 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), (Sochi, Russia) URL: https://ieeexplore.ieee.org/document/8742971.

[5] Vyalkova S A, Poddubskaya S S, Kupriyanova M A, Kornyukova O A 2011 “Application of a one-dimensional tracked method for modeling daily load graphs on the example of the Rostov region”. Collection of works by the winners of the selection round of the All-Russian competition of scientific research works of students, graduate students and young scientists in several interdisciplinary, (Novocherkassk) pp 297–299

[6] Vyalkova S A, Nadtoka I I 2018 “Application of the MSSA method to the time series of power consumption, air temperature and natural light”. Dispatching and control in the electric power industry: materials of the XIII All-Russia. open young. scientific-practical conf., (Kazan) pp 4–9

[7] Vyalkova S A, Nadtoka I I 2015 “Taking into account the properties of the noise component of the power consumption series in short-term forecasting”. Journal Electric power industry through the eyes of youth: scientific. (tr. V Int. youth. scientific. tech. conf., 9-13 November, Ivanovo) vol. 1 pp 83–86

[8] Nadtoka I, Vyalkova S, 2016 “Short-term forecasting using fuzzy neural network accountability air temperature and natural lighting”. International Conference on Industrial Engineering, (Chelyabinsk) pp 450 – 454, URL: http://ieeexplore.ieee.org/document/7910991.

[9] Shepelev I E, Nadtoka I I, Vyalkova S A, Gubsky S O 2015 Comparing ambient temperature account methods in neural network based city short-term load forecasting. Optical memory and neural networks. July pp 220–229 URL: http://link.springer.com/article/10.3103/S1060992X15030108.

[10] Vyalkova S A, Nadtoka I I 2014 Short-term forecasting of electricity consumption in Moscow using a neural network, Electric power industry through the eyes of youth: scientific. tr. V Int. youth. scientific. tech. conf. (10-14 November, Tomsk, Russia) vol. 1 pp 386–389

[11] Osovsky S, 2004 “Neural networks for information processing”. Finance and statistics, (Moskow) pp 344

[12] Shitikov V K, Mastitsky S E “Classification, regression and other Data Mining algorithms using R”. URL: http://www.iievbras.ru/ecostat/Kiril/R/DM/DM_R.pdf.

[13] Robert I. Kabakov: “Data analysis and graphics with R”. URL: https://www.bioclass.ru/files/R_0_3.pdf

[14] Golyandina N, Zhigljavsky A Singular Spectrum Analysis for Time Series – 2013 URL: http://www.gistatgroup.com/gus/book3/index.html.

[15] Alexandrov P I 2006 “Development of software for automatic allocation and prediction of additive components of time series in the framework of the Caterpillar -SSA”: thesis... Candidate of Physical Mathematics Sciences: 05.13.18. St. Petersburg p 152
[16] Shekalkin V N 2014 “Hybrid models and time series forecasting methods based on the Caterpillar -SSA and Box-Jenkins methods”. Journal Eastern European advanced technology, vol. 5(4) pp 43–62
[17] Schelkalin V N, Teveryashev A D “Trend and decomposition approaches to predict the processes of electricity consumption”. URL: journals.uran.ua/eejet/article/download/1207/1111
[18] Mohsen Hayati, Yazdan Shirvany 2007 Artificial neural network approach for short term load forecasting for Ilam Region. World academy of science, engineering and technology international journal of electrical, robotics, electronics and communications engineering vol: 1 №:4, pp 659–663
[19] Miguel López & Carlos Sans & Sergio Valero & Carolina Senabre, 2019. “Classification of Special Days in Short-Term Load Forecasting: The Spanish Case Study”, Energies, MDPI, Open Access Journal, vol. 12(7), pp 1–31, April.
[20] Technique of control of accuracy of the forecast of consumption: The order of the head of operational staff on improvement of the competitive balancing System Operator of the Unified Energy System market of April 14, Moscow, 2011, № 132.