Economic factors of electricity transport based on energy consumption forecasting

Anna Grabar1*, Darya Starkova2, Olga Soboleva2 and Tatyana Kondratyeva3

1Vyatka State University, Moskovskaya, 36, Kirov, 610000, Russia
2Vyatka State Agricultural Academy, October Avenue, 133, Kirov, 610017, Russia
3Moscow State University of Civil Engineering, 26, Yaroslavskoye Shosse, 129337, Moscow, Russia

Abstract. Forecasting significance in the energy market is extremely high. Demand for electricity determines the key decisions on its purchase and production, load transfer and transmission control. Over the past few decades, several methods have been developed to accurately predict the future of energy consumption. This article discusses various methods for forecasting energy demand. Three blocks of methods are considered: statistical, methods using artificial intelligence and hybrid. Authors defined the metrics that show the quality of the models and help to compare the results of the models: mean absolute error (MAE), mean absolute percentage error (MAPE), root-mean-square deviation (RMSE), minimum and maximum errors on the test sample. A comparative analysis of forecasting methods has been launched on the open data set. The best result is obtained using a combined model based on the Lasso regression method. The accuracy and speed of predictions helps to get an economic effect from regulating generation by selling electricity at the peak of consumption.

1 Introduction

Electricity transport is an essential component of sustainable economic development. Electricity is necessary for the needs of households and production, but also it is an important factor in economic security. Sustainable energy ensures the smooth operation of information systems. The energy system is a complex mechanism with a high level of mutual lock in the generation of transmission and sale of electricity. The transmission and sale of electricity depends on the volume of its generation. The generation of thermal power plants and hydroelectric power plants can be regulated. The number of nuclear power plants and alternative sources of production in Europe and Russia is insignificant, the amount of energy received from them cannot have a serious impact on the energy market, and their generation can be neglected. Using models to accurately predict future energy consumption in individual territories is an important goal for power generation and distribution systems.

Energy management is an important process in the production and consumption of energy, which is widely used to minimize demand, costs and pollutant emissions. On the basis of this process, energy management should seek effective solutions to emerging

* Corresponding author: a.a.grabar@gmail.com
problems in an ever-changing international energy economy, which are caused by global liberalization of the energy market, limited by limited resources and rising prices. Information systems used today represent the basis for monitoring and making such decisions. Energy management using information and computer technologies combines applications from mathematics and computer science to optimize the process of energy production and consumption.

Due to the large volumes of relevant information in management decision support systems, it is necessary to use effective data management. Therefore, the methods of mathematical analysis and optimization should be combined with energy databases and data management of the energy production process. A detailed analysis of the main input and output data of the power system is necessary to increase its efficiency. Improving the efficiency of energy systems or developing cleaner and more efficient energy systems will slow down the growth in energy demand, lead to a deep reduction in the use of fossil fuels and reduce pollutant emissions.

Energy consumption depends on many different influence factors (Figure 1). Usually the demand for energy is influenced by the periods of the year, climatic parameters and economic conditions. The heat demand of the district heating system is highly dependent on the outside temperature, as well as on additional climatic factors such as wind speed, sunlight and humidity. On the other hand, seasonal factors influence energy consumption. Typically, the demand for electricity and heat on weekdays is higher than on weekends. In addition, holidays and holidays have a significant impact on energy consumption. Last but not least, the demand for heat and electricity in the delivery area is affected by the operational parameters of enterprises with high demand for energy and consumer behavior. In addition, electricity and heat requirements follow a daily cycle with low periods at night and with peaks at different hours of the day.

The quality of the energy demand forecast largely depends on the availability of historical data on consumption and on the knowledge of the main parameters for influencing energy demand. Functional relationships are non-linear and there are complex interactions between different types of data. Due to the large number of influence factors and their uncertainty, it is impossible to create an “exact” physical model of energy demand. Therefore, the need for energy is calculated on the basis of mathematical models that simplify real relationships. There are no simple deterministic laws that relate predictive variables (seasonal data, meteorological data and economic factors) on the one hand, and energy demand as a target variable on the other, it is necessary to use statistical models.

The statistical model studies quantitative relationships from historical data. During this learning process, the quantitative relationships between target variables (variables that should be predicted) and predictor variables are determined from historical data. Training datasets should be provided for known predictor target variables. According to this

![Fig. 1. Factors Affecting Electricity Consumption.](image-url)
example, a mathematical model is determined. This model can then be used to calculate the values of the target variables as a function of predictor variables for periods for which only predictor variables are known. When using meteorological data as predictive variables, accurate forecasts are also needed for them.

An analysis of the relationship between energy consumption and climate factors includes the following activities:
- balancing energy (distribution of demand);
- analysis of the main factors of influence (Figure 2);
- development of a mathematical model;
- analysis and modeling of typical demand profiles.

The working hypothesis of this study is that the grid operators in the Russian Federation and over the world have the opportunity to regulate the generation of electricity and to obtain economic benefits through interregional flows. The purpose of the article is to test the hypothesis put forward, evaluate and select tools for predicting the generation and transportation of electricity.

The goal achievement requires solving the following two tasks:
- analysis of existing models to accurately predict future trends in energy consumption, including using non-linear data;
- development of a methodology for choosing a method for forecasting the demand for electric energy within the territory.

2 Materials and method

Models for predicting the electrical load and, therefore, the balance of energy flows can be divided into three main categories [1] models of time series, models of artificial intelligence and hybrid models.

Authors consider statistical methods for modeling time series. Their main purpose is to collect and analyze chronological observations of a certain phenomenon to build a model that has the properties and structure inherent in a series of observations [2]. Statistical time series models, compared to other models, are easy to use and develop. Basically, statistical methods use historical data to produce forecasts for short periods of time. As the forecast period increases, the forecast error increases. The following statistical time series models were used to predict the electrical load: linear regression [3], seasonal autoregression [4], threshold autoregression [5], Kalman filter [6], the integrated ARIMA moving average model of autoregression [7], and the seasonal integrated autoregression model - Moving average SARIMA [8], etc.

Due to the short-term forecast and many factors affecting the size of the balance of flows in power networks, the use of statistical models alone is inefficient [9].

Next, we consider artificial intelligence models that can be used to confirm the hypothesis put forward. The following models / methods of artificial intelligence are used to predict electric load:
- expert systems [10];
- support vector method [11],
- support vectors regression method (SVR) [12];
- methods of fuzzy logic [13];
- artificial neural networks (ANN) [13];
- echo state networks (ESN) [14].

Using artificial intelligence models, a larger number of influence factors can be considered, including those that are not explicitly defined. In addition, during the construction and training of artificial intelligence models, an explicit relationship is established between the initial data and the forecast results. Neural networks, as well as
statistical models, can use the forecast error - the difference between the predicted and the actual value of the flow balance to clarify the parameters of the model (training). Artificial intelligence methods are not without drawbacks, such as convergence to a local extremum and the effect of retraining for neural networks.

And finally, consider hybrid models. Two approaches to the construction of a hybrid model can be distinguished. The first approach is based on the idea of averaging the results of applying a series of models using the following algorithm:

1. Obtaining a forecast using a set of models;
2. Calculation of the weight of each model;
3. The weighted sum of the calculation results of all models is taken as the final forecast [15, 16, 17].

The main problem of the first approach is the complexity of calculating the weight coefficients of the components, which significantly affect the accuracy of forecasting. However, if this problem is resolved, the effectiveness of the hybrid model is superior to the effectiveness of each of its components.

The second approach to building a hybrid model is to decompose the process under study into a series of subprocesses, each of which uses a method that considers the a priori properties of the subprocess. The end result of forecasting is the sum of the results of the simulation of subprocesses.

We distinguish three main ways of signal decomposition: Fourier transform, wavelet transform and empirical mode decomposition (EMD).

The classic signal analysis method is the discrete Fourier transform. In this case, the time series of the balance change can be considered as a linear combination of signals with different frequency ranges. Fourier transform is an important tool for analyzing stationary signals. In the Fourier spectrum is used to analyze and predict seasonal fluctuations in electricity demand, while trend forecasting is performed using neural networks [18]. However, the Fourier spectrum is not effective for the analysis of short-term phenomena [19].

An alternative to the Fourier transform is the wavelet transform, which is better suited for transient analysis. So, the wavelet transform was used to filter high-frequency components [20], and to decompose the signal into components [21], for which hybrid models based on ARIMA and the support vector method NLSSVM were constructed (finite Newton method for Lagrangian support vector machine). M. Besseca and J.L. Fouquaub built a hybrid model combining wavelet transforms with neural network methods to predict short-term power consumption [22]. Although the wavelet transform demonstrates efficiency in decomposing the time series of the balance of electricity into more regular components, it is impossible to obtain comprehensive information about the signal with its help [23].

To increase the efficiency of signal decomposition, the EMD (Empirical Mode Decomposition) method is actively used in conjunction with regression methods [24] and a neural network [25].

The EMD transform can provide better decomposition results than the wavelet transform, but its drawback is the boundary effect — the error at the ends of the observation interval that accumulates during the selection of empirical modes. The authors compensate for this effect by artificially adding new boundary points to the study interval — a local minimum on the left and a local maximum on the right [26].

In this case, there is only one empirical mode and a residual term characterizing the signal trend. The ARIMA method is applied to the residual term, and the wavelet-neural network (WNN) optimized using the FOA algorithm (Fruit Fly Optimization Algorithm) for finding the extremum is applied to the empirical mode, a high-frequency signal depending on calendar events.
In the study and preparation of the signal, it is proposed to use the methods of expansion in the Fourier spectrum and EMD conversion. The first one helps to clarify the optimal discreteness of BP of changes in TI and highlight seasonal processes, and the second leads to explore the trend and identify noise in BP. Empirical EMD transform modes are used as features for training the LSTM neural network in a prediction model.

As a data source for research, we used a data set on energy consumption PJM with a clock discrete. PJM Interconnection LLC (PJM) is a regional broadcasting organization (RTO) in the United States. It is part of an Eastern Accession network operating a transmission system serving all or parts of Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia, and the county. Colombia. Hourly energy consumption data comes from the PJM website and is presented in megawatts (MW).

3 Results

A number of results were obtained in research and construction of the predictive model:
- the periodic properties of the time series with periods of 12 hours, 24 hours, 6 days were confirmed and considered in the forecasting model, seasonal fluctuations with 5 seasons per year were identified and considered in the calculation of the training period of the predictive model, the difference in daily behavior of the series was confirmed and taken into account in the model and night time;
- trend changes in the values of the time series during the day and during the week are highlighted, while we note that only the daily trend can be predicted;
- the absence in the time series of additive components with the properties of white noise has been established;
- the absence of correlation of high-frequency components with both the work schedule and the duration of daylight hours and weather conditions has been established, the source of this component should be considered significant differences in the characteristics of the control equipment and the human factor in the work of energy network operators.

The methodology for constructing a predictive model of the problem is based on an empirical-pragmatic approach and consists in empirical testing of methods for predicting time series of their analysis and subsequent adaptation to solving the problem of forecasting power consumption by consumers of the energy network. To predict the magnitude of such power, the study considered regression, autoregressive, and neural network methods for solving the forecasting problem.

Authors have used the next quality metrics:
- mean absolute error (MAE);
- mean absolute percentage deviation (MAPE);
- root-mean-square deviation (RMSE);
- minimum and maximum errors on the test sample;
- distribution parameters: 1.3 quartiles and median.

The results of predictive models are shown in table 1.

| Model                              | Study          | Test           | MAE | MAP E | RMS E | Min | Max | p25 % | p50 % | p75 % |
|------------------------------------|----------------|----------------|-----|-------|-------|-----|-----|-------|-------|-------|
| Random forest (500 trees)          | 120 days       | 180 hours      | 364 | 11.63 % | 2835  | 1   | 1473| 1523  | 3098  | 5233  |
| Gradient boost over decisive trees | 120 days       | 180 hours      | 350 | 11.13 % | 2668  | 0   | 1436| 2958  | 4993  |
Among the investigated methods are highlighted:

- The Lasso-regression method, which showed a large error in its original form, but considers the behavior of the time series, while having a high learning speed and allowing the use of many features (order of several hundred).

- SARIMAX method, demonstrating the highest accuracy, but demanding on computing equipment and each time the forecast is needed to refine the hyperparameters of the autoregressive model. The latter property is caused by the presence in a number of processes that are initiated by automatic systems with different characteristics in response to the actions of a small number of operators. Such an effect cannot be predicted with high accuracy based on statistical data, but it cannot be considered noise.

- A neural network based on LSTM cells, showing high accuracy, but not able to quickly learn when changing the behavior of a series.

It was decided to use the abandonment of the concept of a neural network in favor of a combined model based on the Lasso regression method to obtain a forecast for periods of up to 12 hours (half a daily cycle) and SARIMAX for a forecast of more than 12 hours. The resulting calculation model allows predicting power consumption with an accuracy at which the MAE of forecast values obtained no less than 1 hour before the actual delivery begins, on average per day, is in the range of 800-1000 MW.

A table showing the comparison of different methods:

| Method       | Days | Hours | Error (%) | MAE (MW) |
|--------------|------|-------|-----------|----------|
| Lasso regression | 120  | 168   | 10.08     | 2564     |
| Ride regression  | 120  | 168   | 10.94     | 2627     |
| SARIMAX       | 30   | 72    | 4.18      | 1408     |
| LSTM (3 layers BiLSTM and 2 fully connected layers) | 420  | 144   | 5.16      | 1278     |

An example of the forecast is shown in Figure 2.

![Figure 2](image_url)

**Fig. 2.** Example of the forecast of electricity consumption in the network.
4 Discussion

In addition to the proposed regular methods for predicting energy consumption, peak load forecasting is currently an important issue. Accurate forecasting of peak demand for electricity is critical for the safe and reliable planning, supply and transportation of energy. Accurate peak time prediction provides the ability to efficiently distribute the load between the transmission substations, ensure reliable power supply in the distribution network, analyze the flow of energy and schedule the start time of the generation.

Most studies have shown an increase in energy consumption in recent years and believe that energy consumption will increase in subsequent years, but this trend cannot continue forever. Studying factors such as industry saturation, economic downturns, climate change, and environmental issues that could trigger a downward trend in energy consumption is an interesting area of research. This area of research can detect when energy consumption begins to decrease or what changes may occur in other types of energy consumption. Another attractive way of research may be to find the optimal delay in the effect on energy consumption for effective factors. A change in the effective factor can affect energy consumption after one or more periods.

As a further research route, we can study the influence of environmental parameters, with the exception of temperature, such as wind, cloud cover and humidity, on energy consumption and the reliability of forecast variables. Another area of future research includes the development of hybrid methods. Literature shows that classical methods can no longer lead to dominant results. Another area of future research may focus on new measures to evaluate the effectiveness of forecasting methods. Energy forecasting methods apply the same estimation criteria based on errors that are used for other forecasting tasks, such as forecasting water consumption, forecasting travel demand and forecasting demand for telecommunication devices. Therefore, it would be interesting if it were possible to determine some evaluation criteria that are not based on errors in order to find out the effectiveness of energy prediction methods.

The results of the study confirmed the working hypothesis and proved that with the help of energy consumption forecasting it is possible to obtain an economic effect if energy is sold at the approach to the peak of consumption, and the purchase is at the decline.

5 Conclusion

The analysis and forecast of energy demand are an important part of management in the energy industry. Energy consumption in the area where the power plant is delivered is affected by seasonal data, climatic parameters, and economic boundary conditions. The quality of demand forecasting tools largely depends on the availability of historical data on consumption, as well as knowledge of the main parameters for influencing energy consumption. Energy data management should provide information for energy control, including all activities for the planning, operation and control of the generation and distribution process. Detailed knowledge of energy requirements in the delivery area is necessary to increase the efficiency of power plants and realize the potentials of optimizing the transport energy network.

In addition to the growing demand for energy in the current century, unbalanced supply and demand for energy at different time periods, as well as in geographical, climatic and even political conditions, increases the importance of accurate forecasts.

Forecasting demand is one of the most important factors in the operation and planning of energy production and distribution systems. If energy consumption is overstated, production downtime is planned and the cost of consumption of all subscribers will increase without a good reason. Underestimating energy consumption, on the other hand, has its
negative economic and social consequences, such as blackouts and blackouts. Forecasting models must be reliable to maintain a high level of accuracy at different time periods and in response to different values of input variables. Some researchers believe that accurate forecasting of energy consumption affects weather quality, supply chain security, capital investment, revenue analysis, and market management.

Factors for short-term and long-term energy fluctuations can be different. Economic factors, seasonal variations and climatic conditions significantly affect the daily consumption of electricity. The proposed models can be used by manufacturers, suppliers and regulatory authorities who want to ensure reliable transportation of electricity at a minimum price.

References

1. J. Zhang, Y.M. Wei, D. Li, Z. Tan, J. Zhou, Energy 152, 774 (2018) doi: 10.1016/j.energy.2018.06.012
2. R. Adhikari, R. Agrawal, An Introductory Study on Time Series Modeling and Forecasting (Saarbrücken, Lambert Academic Publishing, 2013)
3. A.D. Papalexopoulos, T.C. Hesterberg, IEEE Trans Power Syst. 4, 1535 (1990)
4. G. Mbamalu, M. El-hawary, IEEE Trans Power Syst. 1, 343 (1993)
5. S. Huang, IEEE Proc-Gener Transm Distrib. 5, 477 (1997)
6. M. Zhang, H. Bao, L. Yan, J.P. Cao, D.U. Jian-Guang Power Syst Technol 10, 1 (2003)
7. J.F. Chen, W.M. Wang, C.M. Huang, Electr. Power Syst. Res. 3, 187 (1995)
8. A. Tarsitano, I.L. Amerise, Energy 133, 108 (2017)
9. J. Che, J. Wang, G. Wang, Energy 37, 657 (2012)
10. S. Rahman, R. Bhatnagar, IEEE Trans. Power Syst. 3, 392 (1988)
11. P.F. Pai, W.C. Hong, Electr. Power Syst. Res. 74, 417 (2005)
12. Y.L. Yang, J.X. Che, Y.Y. Li, Y.J. Zhao, S.L. Zhu, Energy 113, 796 (2017)
13. S. Chenthur Pandian, K. Duraiswamy, C. Christober Asir Rajan, N. Kanagaraj, Electr. Power Syst. Res. 76, 541 (2006)
14. A. Deihimi, H. Showkati, Energy 1, 327 (2012)
15. S.B. Kyung, S.B. Young, H.H. Dug, J. Gilsoo, IEEE Transac. Power Syst. 1, 96 (2005)
16. N. Amjadi, F. Keynia, Energy 34(1), 46 (2009)
17. A. Mottaeva, E3S Web of Conferences 110, 02164 (2019) doi.org/10.1051/e3sconf /201911002164
18. E. González-Romera, M.A. Jaramillo-Morán, D. Carmona-Fernández, Energy Conversion and Management 49(11), 3135 (2008) doi:10.1016/j.enconman.2008.06.004
19. S.J. Yao, Y.H. Song, L.Z. Zhang, X.Y. Cheng, Energy Conversion and Management 41(18), 1975 (2000) doi:10.1016/s0196-8904(00)00035-2
20. A. Mottaeva, A. Ivashchenko, A. Ryattel, E3S Web of Conferences 164, 10038 (2020) https://doi.org/10.1051/e3sconf/202016410038
21. A. Ghasemi, H. Shayeghi, M. Moradzadeh, M. Nooshyar, Appl. Energy 177, 40 (2016)
22. M. Besseca, J. Fouquaub, Eur. J. Oper. Res. 1, 149 (2018)
23. E. Alickovic, J. Kevric, A. Subasi, Biomed Sigmaal Proc Proces 39, 94 (2018)
24. G.F. Fan, L.L. Peng, W.C. Hong, F. Sun, Neurocomputing 173, 958 (2016)
25. X.H. Qiu, Y. Ren, P.N. Suganthan, G.A. Amaratunga, J. Applied Soft Computing 54, 246 (2017)
26. Jinliang Zhang, Yi-Ming Wei, Dezhi Li, Zhongfu Tan, Jianhua Zhou, Energy (2018) doi: 10.1016/j.energy.2018.06.12