Predicting Electric Energy Consumption for a Jerky Enterprise

Elena Kapustina*, Eugene Shutov, Anna Barskaya, Agata Kalganova

National Research Tomsk Polytechnic University, Tomsk, Russia
Email: *lena.k-98@yandex.ru

Abstract

Wholesale and retail markets for electricity and power require consumers to forecast electricity consumption at different time intervals. The study aims to increase economic efficiency of the enterprise through the introduction of algorithm for forecasting electric energy consumption unchanged in technological process. Qualitative forecast allows you to essentially reduce costs of electrical energy, because power cannot be stockpiled. Therefore, when buying excess electrical power, costs can increase either by selling it on the balancing energy market or by maintaining reserve capacity. If the purchased power is insufficient, the costs increase is due to the purchase of additional capacity. This paper illustrates three methods of forecasting electric energy consumption: autoregressive integrated moving average method, artificial neural networks and classification and regression trees. Actual data from consuming of electrical energy was used to make day, week and month ahead prediction. The prediction effect of prediction model was proved in Statistica simulation environment. Analysis of estimation of the economic efficiency of prediction methods demonstrated that the use of the artificial neural networks method for short-term forecast allowed reducing the cost of electricity more efficiently. However, for mid-range predictions, the classification and regression tree was the most efficient method for a Jerky Enterprise. The results indicate that calculation error reduction allows decreases expenses for the purchase of electric energy.

Keywords

Autoregressive Integrated Moving Average Method, Artificial Neural Networks, Classification and Regression Trees, Electricity Consumption, Energy Forecasting

1. Introduction

The amount of electricity consumed is increasing every year around. It is im-

portant for energy producers to generate the amount of electricity that will be consumed by enterprises, because power cannot be stockpiled. Consequently, an accurate forecast of electricity demand is necessary for a stable power supply. The electricity market makes heavy demands on consumers and forces companies to plan energy consumption in order to reduce energy costs. Many works were devoted to resolving predicting energy consumption [1] [2] [3]. Implementation of the forecasting system is considered by the example of a Russian enterprise.

The reliability of predicting energy consumption for large consumers (over 670 kW) plays a significant part, because of the emerging of the Russian wholesale market for electricity and power (WMEP) [4] [5]. For an enterprise to entry into the WMEP, the market players shall submit a qualitative request for the maximum hourly energy capacity to the trading system (pool) administrator and the System Operator of the Unified Power System (SOUPS) no later than 24 hours before starting energy supply [6]. The total cost (Figure 1) of electric energy includes not only the actual consumed energy capacity, but also the payment for a deviation from the energy nomination (the transition of the enterprise to balancing energy market) [4]. An extra charge shall be paid for the purchase (supply) of missing (surplus) power [4]. Therefore, the best correct prediction shall be made to avoid the transition of the enterprise to balancing energy market (BEM) and to reduce extra energy expenses. Reducing the prediction error could make possible the enterprise to substantially reduce the operational expenses without purchasing large-ticket equipment.

There available six price ranges (rates). Enterprises have the possibility of choosing a price range. It is not advisable for large enterprises to use the 1st and the 2nd price ranges (PRs) (integrative price ranges) since they may not prove cost-efficient. If the total capacity of an entity exceeds 670 kW, it is advisable to use the 4th and the 6th PRs [7]. Since the total connected capacity of the power receivers of the enterprise under consideration exceeds 670 kW and hourly planning is intended, the 6th PR is used.

The capacity of the energy consumed by a specific enterprise is governed by features of its operation, such as enterprise load, temperature conditions, illumination level, etc. Therefore, the prediction process, which can help to minimize the expenses, shall be determined on a case-by-case basis. Prediction makes it possible both to perform a cost-effectiveness analysis and to correct equipment operating modes.

2. The Practical Implementation of Predicting Electric Energy Consumption

Information on actual amount consumption of electric energy was received by electric energy metering devices installed on the territory of the enterprise. Figure 2 shows hourly energy consumption trend. The data are well formalized and represent the dynamics of the production process. Figure 3 shows daily energy
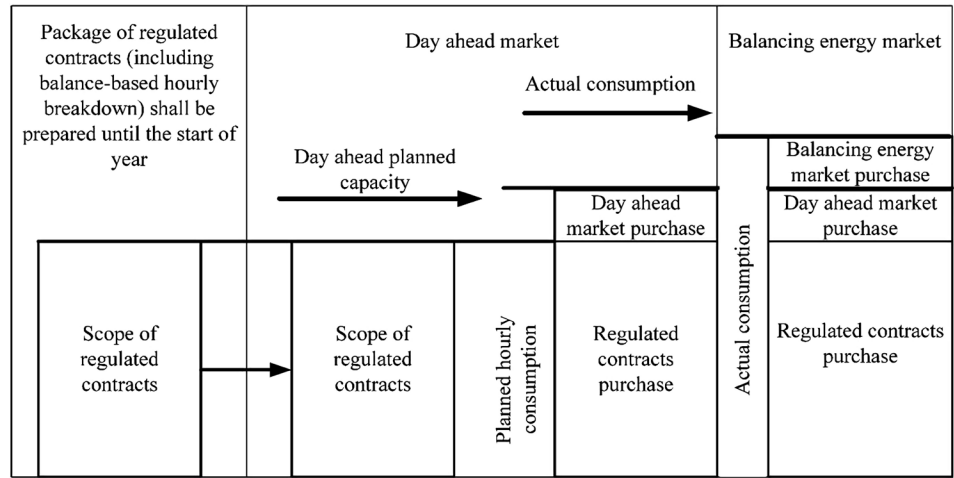
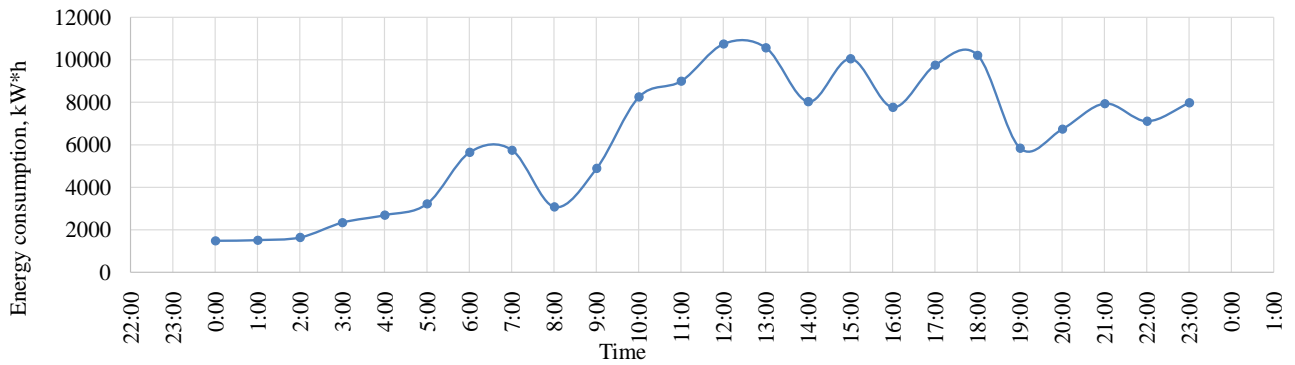**Figure 1.** Diagram of payment for deviation of actual consumed energy capacity from energy nomination.

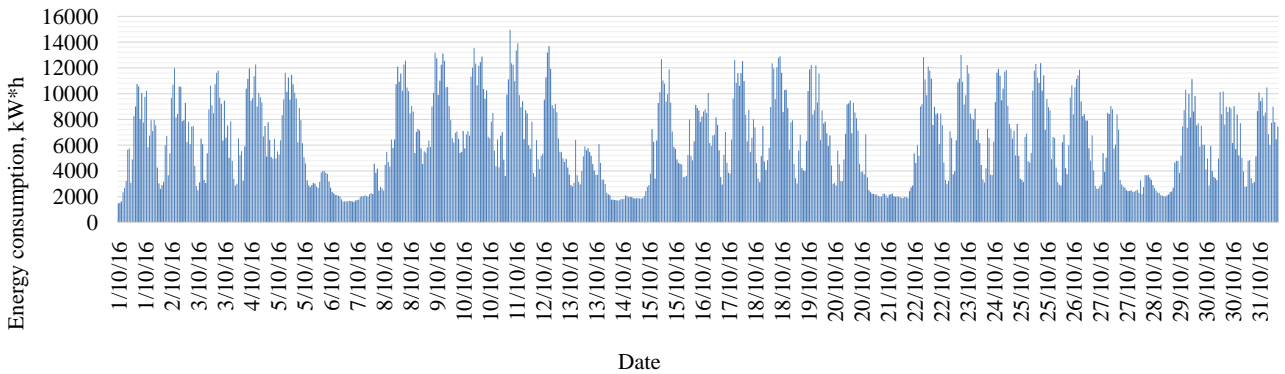**Figure 2.** Daily electricity consumption trend.

Figure 3. Monthly electricity consumption chart.

consumption for October 2016 in which weekly cycle (load variations on the weekend) can be clearly seen.

We can determine seasonal component using a periodogram (**Figure 4**). There are two peaks, which equal to 24 and 168 hours (day and week) and one smaller peak, equal to 84 hours, that means that the process of electric energy consumption is cyclical. It should be taken into account when making a short-term electricity
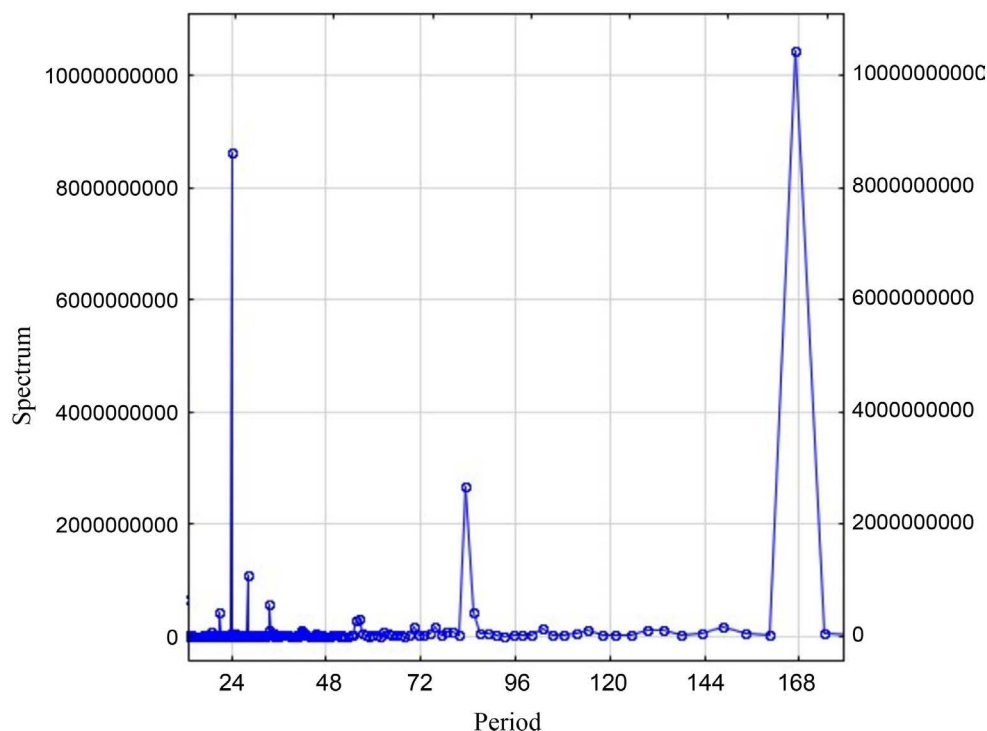
Figure 4. Periodogram.

consumption forecast. Typical days with similar daily load curve and operating schedule were selected to make day ahead prediction. The data for week ahead and month ahead predictions have been selected in the same manner.

We used the *Statistica* software to test the hypothesis on normality of distribution of the background data. Using the Kolmogorov-Smirnov test ($p < 0.1$) and the Shapiro-Wilk test ($p = 0.000$), the conclusion has been drawn that the time series is nonstationary, which complicates the prediction making process [8].

Currently, there are many methods for short-term forecasting of electricity consumption, which are used by industrial enterprises [9]-[14]. Since the time series is nonstationary and there are seasonal components, autoregressive integrated moving average method (ARIMA) can be applied. ARIMA is one of the most often-used forecasting methods which allows for a stationary series by differentiating the time series. ARIMA is used quite often to analyze initial data in which distinct and stable periods of operation can be seen. That is why this method requires typical time frames to be selected [9].

Recently, methods using artificial neural networks (ANNs) have been applied not only to make short-term forecasting of electricity consumption but also in the construction of medium- and long-term forecasts. Such networks consist of many input/output neurons that interact with each other through nerve endings (synapses). The application of ANNs allows one to process data with a larger degree of stochasticity. These networks can be learned, *i.e.* it is possible to reveal regularities and integrate background information [15]. The advantages of the neural network are fast learning algorithms and the ability to operate with noisy

input signals.

A multivariate prediction method of making a forecast has also been chosen, because in univariate models are built only on the basis of a one-dimensional series and does not take into account other indicators, such as the amount of output. The use of classification and regression trees (CART) enables to make predictions based on data that affect the dependent variable, which allows one to explore the model in more detail and to reveal the factors that have the maximum effect on electric energy consumption [16].

The efficiency of ARIMA method has been assessed. Because of the high degree of stochasticity of the background data, the following typical sequence of data analysis has been implemented: making an assessment of the behavior of the autocorrelation function (ACF) and the partial autocorrelation function of the load curve and selecting a model for prediction making, as well as adding a seasonal component (definite cyclicality). Taking into account these transformations, a day ahead prediction has been made. Then, the hypothesis on normality of distribution of the data obtained has been tested. It has been shown that the distribution is normal. To perform an analysis of the prediction, an analysis of the normal probability graph, the normal distribution of the random variable and the remainder series has been made, as well as the mean absolute percent error of the prediction (MAPE) has been calculated.

The mean absolute percentage error can be determined using the following formula [17] [18]:

$$\text{MAPE} = \frac{100\%}{L} \sum_{t=1}^{L} \left| \frac{X_t - \hat{X}_t}{X_t} \right|$$

where X_t—actual value,

$\hat{X}_t$—predicted value,

L—number of steps.

If MAPE is <10%, the precision accuracy is high, if MAPE is <20%, the precision accuracy is well and if MAPE is <50%, the precision accuracy is satisfactory.

The mean absolute percent error for the prediction under consideration is equal to 13.66%.

Week ahead and month ahead predictions have been also made. The mean absolute percent error is equal to 23.35% for the week ahead prediction of the enterprise load and to 37.19% for the month ahead prediction. The trend of increasing the mean absolute percent error of the prediction with increasing prediction interval has been revealed. Therefore, this type of prediction is not suitable for forward planning because of the excessive error value.

The efficiency of ANN method has been assessed. The assessment has been made using the regression analysis of time series. Upon completing the procedure of learning of the multilayer perceptron (MP), the network with the highest performance has been chosen. To assess the prediction quality, the remainder distribution histogram, the remainder normal probability graph and the normal probability graph have been analyzed. The performance of the model is equal to

99.41%.

A comparison of the mean absolute percent errors of the prediction for the two methods has been made (**Figure 5**). For the univariate prediction, ANN has been found to be the most suitable method since it is characterized by the lower error of 1.88%, which is smaller than the error of 13.66% for ARIMA method. Therefore, ANN is the most cost-efficient method.

It is known that the sum of predictions is more efficient than a univariate prediction. Therefore, to reduce the calculation error the multiple prediction method has been used. A few prediction models have been built in each interval and then the arithmetic mean has been found [19]. A comparison of the two multiple prediction methods has been made (**Figure 6**) and the mean absolute percent error of the prediction has been determined.

The multiple univariate prediction for the ARIMA model yields a higher accuracy than the single prediction (the mean absolute percent error of the prediction is equal to 4.35%). The mean absolute percent error of the multiple prediction according to the ANN model is equal to 5.89%, which is higher than the error of the single prediction according to the ANN model and the error of the multiple prediction according to the ARIMA model. The increase in the error can be explained by the absence of multitude of learning time frames, which could be helpful for the further design of a prediction model.

Univariate models possess a serious drawback, *i.e.* it is impossible to take into account other factors that affect energy consumption and make a prediction using only preceding values. The method of constructing classification and regression trees, namely, the *CART* algorithm for constructing an extended prediction, is free of this drawback [16]. "The consumed power, kW" was used as the dependent variable and the energy-consuming factors, such as "Casting of SCh-15, t", "Pig iron, t", "Casting, t" and "Number of heats", were selected as continuous factors. Using the *Statistica* software, the classification and regression tree has been constructed (**Figure 7**) for which the quality of the prediction was assessed using the remainder distribution histogram and the normal probability graph. The mean absolute percent error over a month is equal to 1.49% (**Figure 8**).
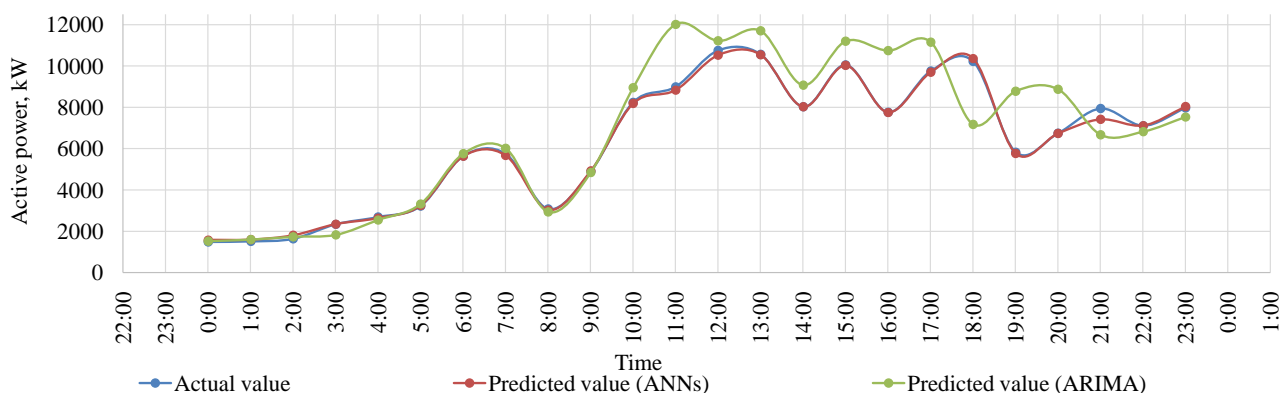

Figure 5. Actual electric energy capacity and electric energy capacities predicted using ANN and ARIMA (single prediction).
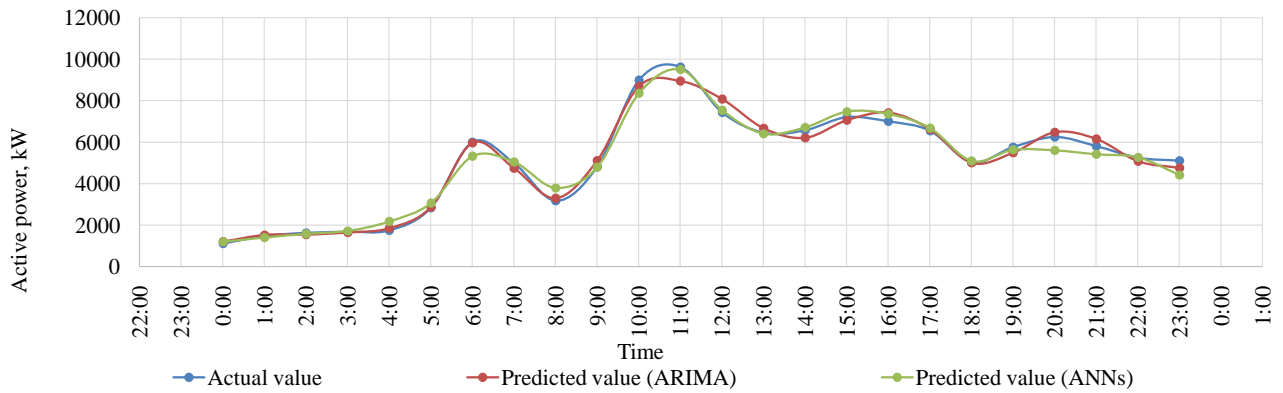
Figure 6. Actual electric energy capacity and electric energy capacities predicted using ANN and ARIMA (multiple prediction).
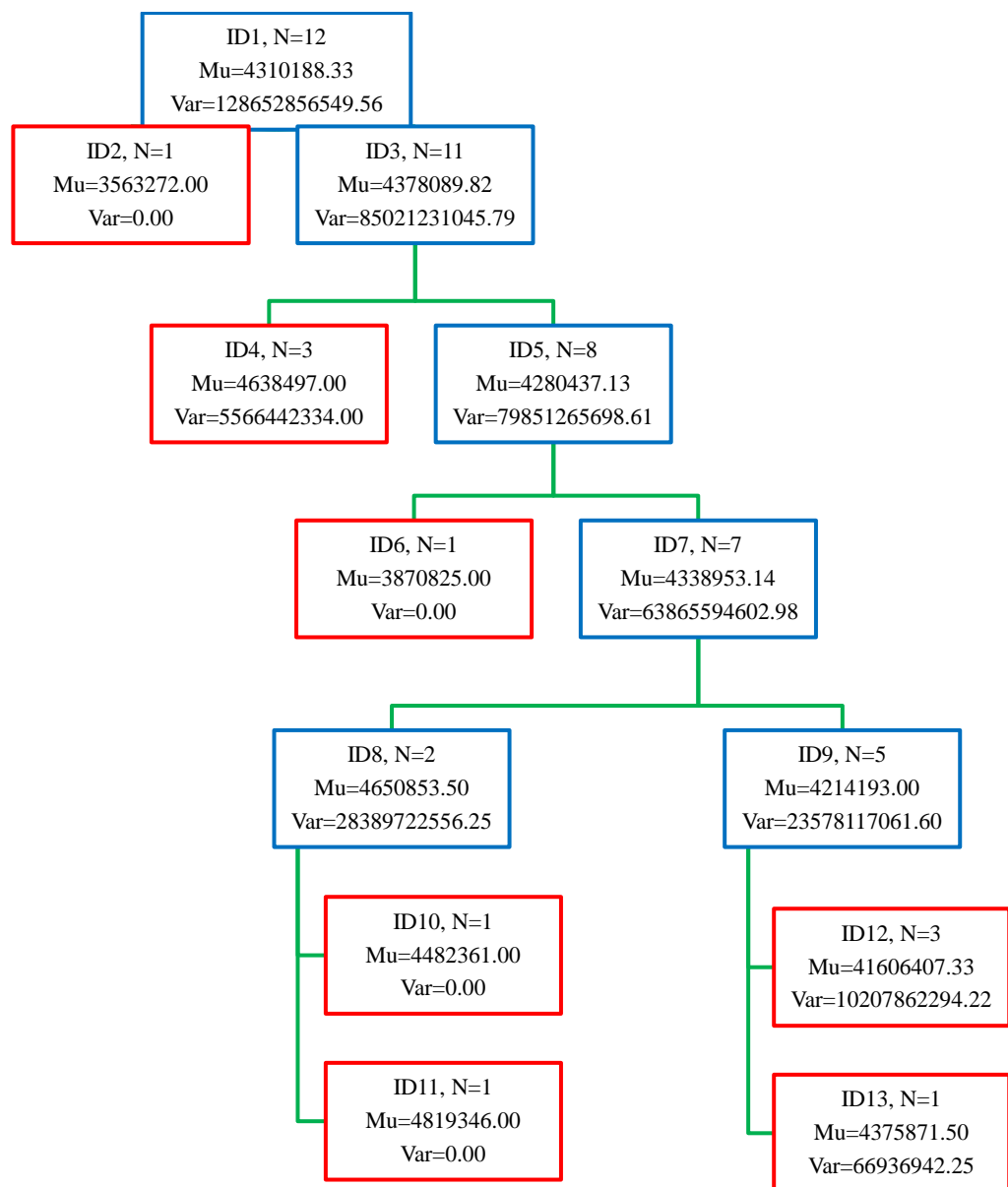

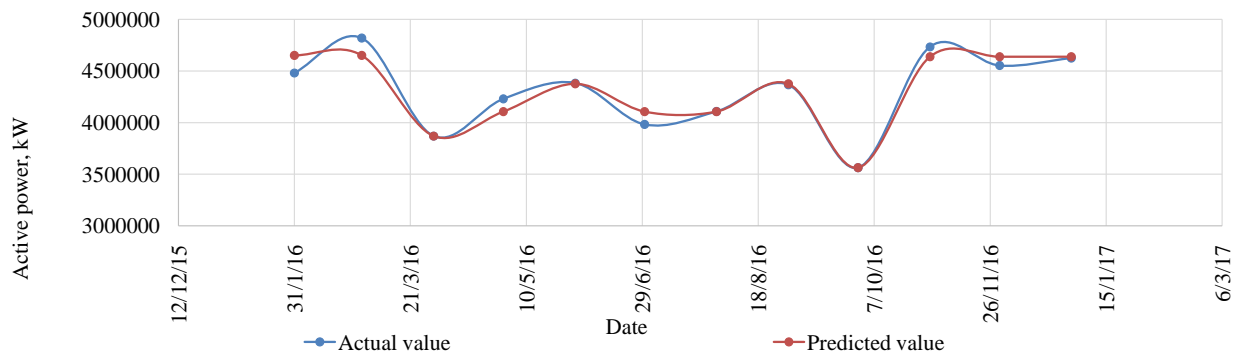**Figure 7.** Classification and regression tree No. 6.

Figure 8. Actual and predicted electric energy capacities (multivariate prediction).

3. Estimation of the Economic Efficiency of Prediction Methods

A comparison of the cost efficiency of the four following day ahead predictions has been made: the single univariate prediction and the multiple univariate prediction made using ARIMA and ANN. The calculation has been performed on the basis of the data on non-regulated electricity price cap and the retail markup calculated by the last resort supplier for an enterprise using the 6th PR and having the maximum capacity of power receivers of 10 MW min. over the period of August and November, 2016 [20].

The following is the example of calculation for the univariate prediction over the period of 00:00-01:00, October 1, 2016, which has been made using ARIMA [21]:

$$
\begin{aligned}
C_e &= \sum_{i=1}^{24} W_i \cdot C_i + P_p \cdot C_p \big/ 30 + P_{max} \cdot C_{max} \big/ 30 + \sum_{i=1}^{24} W_i \cdot C_{other} \\
&\quad + \sum_{i=1}^{24} W_i^+ \cdot C_{iretail}^+ + \sum_{i=1}^{24} W_i^- \cdot C_{iretail}^- + \sum_{i=1}^{24} W_i^+ \cdot C_i^+ + \sum_{i=1}^{24} W_i^- \cdot C_i^- \\
&= 1516.67 \times 0.027 + 283426.00 \times 9.00 \big/ (30 \times 24) + 0.00 \times 0.00 \big/ (30 \times 24) \\
&\quad + 1516.67 \times 0.001 + 34.67 \times 0.00 + 34.67 \times 0.00 \\
&= 47.67 \text{ USD}
\end{aligned}
$$

where W_i is energy consumption per hour, kW*h;

C_i is the rate for the power, USD/kW per hour;

P_p is the power paid on the wholesale market (the peak load charge), kW*h;

C_p is the rate for the power purchased by the consumer at the regulated electricity price cap, RUB/kW per month;

P_{max} is the grid power paid by the consumer as the power consumed for energy transmission services, kW (to be determined as the maximum power over planned peak hours approved by the OAO SOUPS with account for the pricing zone of the consumer), kW*h;

C_{max} is the rate for the power consumed for energy transmission services, RUB/kW per month;

C_{other} is the payment for other services, which are integral part of the process of electric energy and power supply, over the account period, USD/kW*h;

W_i^+ is the increase of the actual electric energy capacity on the planned capacity, kW*h;

$C_{iretail}^+$ is the retail markup differentiated in hours of the account period, which is applied to the non-regulated wholesale market electricity price determined based on the results of the competitive procedure of selecting bids for system balancing in regard to the increase of the actual electric energy capacity on the planned capacity, USD/kW*h;

W_i^- is the increase of the planned electric energy capacity on the actual capacity, kW*h;

$C_{iretail}^-$ is the retail markup differentiated in hours of the account period, which is applied to the non-regulated wholesale market electricity price determined based on the results of the competitive procedure of selecting bids for system balancing in regard to the increase of the planned electric energy capacity on the actual capacity, USD/kW*h;

C_i^+ is the non-regulated wholesale market electricity price differentiated in hours of the account period, which is determined by the commercial operator of the wholesale market based on the results of the competitive procedure of selecting bids for system balancing in regard to the increase of the actual electric energy capacity on the planned capacity per hour (i) of the account period, USD/kW*h;

C_i^- is the non-regulated wholesale market electricity price differentiated in hours of the account period, which is determined by the commercial operator of the wholesale market based on the results of the competitive procedure of selecting bids for system balancing in regard to the increase of the planned electric energy capacity on the actual capacity per hour (i) of the account period, USD/kW*h.

The purchased electric energy for each hour over October 1, 2016 was calculated in the same manner and then the results were summed up (Table 1).

4. Conclusions

Based on the analysis of the results, the following conclusions can be drawn:

- For short-range predictions, the model constructed on the basis of the single univariate prediction using the ANN method is the most efficient because the mean absolute percent error and the purchased electric energy are lower than those obtained using other methods. This result is ensured owing to the possibility of learning the networks when stochastic background data are used;

Table 1. Results of the calculations.

A predictive algorithm	Single univariate prediction		Multiple univariate prediction	
	ARIMA	ANN	ARIMA	ANN
The mean absolute percent error. %	13.66	1.88	4.35	5.89
The electric energy expenses. USD.	1,879,539.60	1,720,153.62	1,335,331.55	1,330,072.21

- Multiple univariate predictions are efficient for jerky enterprises with continually changing order package. In this case, shorter time intervals, which are closest to the prediction period, are used. The averaging of several predictions makes it possible to smooth "jumps" (data that do not fit a definite regularity for the background data);

- For mid-range predictions, the classification and regression tree is the most efficient method because the background data include not only preceding measurement results, but also the enterprise load, which sufficiently affects the electric energy consumed by the enterprise;

- If the calculation error decreases, expenses for the purchase (supply) of missing (surplus) electric energy also decrease, so do the retail markups. When the error decreases by 11.78%, the electric energy expenses reduce by 8.48%, which amounts to approx. 160,000 USD per year. That is, the decrease in the error by approx. 1% allows the enterprise to save approx. 13,500 USD;

- Predictions shall be made with account for the electric energy price and an extra charge for a deviation from the planned value (during some hours, the extra charge is zero, which can be helpful in a slight reduction in losses).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Liu, Z. (2015) Chapter 1. Global Energy Development: The Reality and Challenges. In: *Global Energy Interconnection*, Academic Press, Cambridge, 1-64.
https://www.sciencedirect.com/science/article/pii/B9780128044056000014
https://doi.org/10.1016/B978-0-12-804405-6.00001-4

[2] Kim, T.-Y. and Cho, S.-B. (2019) Predicting Residential Energy Consumption Using CNN-LSTM Neural Networks. *Energy*, **182**, 72-81.
https://www.sciencedirect.com/science/article/pii/S0360544219311223
https://doi.org/10.1016/j.energy.2019.05.230

[3] Stingl, C., Hopf, K. and Staake, T. (2018) Explaining and Predicting Annual Electricity Demand of Enterprises—A Case Study from Switzerland. *Energy Informatics*, **1**, Article No.: 50.
https://www.researchgate.net/publication/328191457_Explaining_and_predicting_annual_electricity_demand_of_enterprises_-_a_case_study_from_Switzerland
https://doi.org/10.1186/s42162-018-0028-0

[4] Abdurafikov, R. (2009) Russian Electricity Market, 26, 27, 40-43
https://www.vttresearch.com/sites/default/files/pdf/workingpapers/2009/W121.pdf

[5] Kennedy, D. (2003) Liberalisation of the Russian Power Sector. *Energy Policy*, **31**, 745-758. https://www.sciencedirect.com/science/article/pii/S0301421502001258
https://doi.org/10.1016/S0301-4215(02)00125-8

[6] Romanova, V. (2019) Problems and Tasks of Legal Regulation of Exchange Trade in Energy Resources in the Russian Federation and in the Eurasian Economic Union. 67, 68. https://www.elibrary.ru/item.asp?id=37348185
https://doi.org/10.18572/2410-4396-2019-1-65-71

[7]   Rules for the Application of Prices and Tariffs on Wholesale Power Market.
      https://www.dvec.ru/electric_market/rr_prices_rules/

[8]   StatSoft Inc. ARIMA.
      http://statsoft.ru/home/textbook/modules/sttimser.html#arima

[9]   Azadeh, A., Ghaderi, S.F. and Sohrabkhani, S. (2016) Forecasting Electrical Con-
      sumption by Integration of Neural Network, Time Series and ANOVA. *Applied
      Mathematics and Computation*, **186**, 94-96.
      https://www.sciencedirect.com/science/article/abs/pii/S0096300306011106

[10] Taylor, J.W. (2012) Short-Term Load Forecasting with Exponentially Weighted.
      IEEE Transactions on Power Systems, **27**, 458-464.
      https://ieeexplore.ieee.org/abstract/document/5978238
      https://doi.org/10.1109/TPWRS.2011.2161780

[11] Chen, J.F. (1995) Analysis of an Adaptive Time-Series Autoregressive Moving-Average
      (ARMA) Model for Short-Term Load Forecasting. *Electric Power System Research*,
      34, 187-196. https://www.sciencedirect.com/science/article/pii/0378779695009771
      https://doi.org/10.1016/0378-7796(95)00977-1

[12] Mohamad, A. (2012) Finding the Best ARIMA Model of Forecast Daily Peak Elec-
      tricity: Applied Statistics education and Research Collaboration (ASEARC), 11.

[13] Barak, S. and Sadegh, S.S. (2016) Forecasting Energy Consumption Using Ensemble
      ARIMA-ANFIS Hybrid Algorithm. *International Journal of Electrical Power & Ener-
      gy Systems*, **82**, 92-104.
      https://www.sciencedirect.com/science/article/pii/S0142061516303702
      https://doi.org/10.1016/j.ijepes.2016.03.012

[14] Tso, G.K.F. and Yau, K.K.W. (2007) Predicting Electricity Energy Consumption: A
      Comparison of Regression Analysis, Decision Tree and Neural Networks. *Energy*,
      32, 1761-1768.
      https://www.sciencedirect.com/science/article/pii/S0360544206003288
      https://doi.org/10.1016/j.energy.2006.11.010

[15] Haykin, S. (2007) Neural Networks: A Comprehensive Foundation. 3rd Edition, 10-
      24, 164-175.

[16] Bramer, M. (2007) Principles of Data Mining. Springer, 5-6, 23-24, 41-43.

[17] Tikhonov, E.E. (2006) Forecasting Models for Electricity Market. 206-208

[18] Gilliland, M. (2010) Forecasting FAQs. 196.
      https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119199885.app1

[19] Surkov, A. (2019) Improving the Accuracy of Economic Forecasting by Combining
      Economic Forecasts. 270-271.
      https://cyberleninka.ru/article/n/povyshenie-tochnosti-ekonomicheskogo-prognozi
      rovaniya-na-osnove-obedineniya-ekonomicheskih-prognozov

[20] Tomskenergosbyt Inc. (2016) Tariffs and Prices. https://ensb.tomsk.ru/

[21] Babinovich, D. and Shutov, E. (2012) The Role of Forecasting in Energy Efficiency
      of Enterprise. 2-5. https://www.elibrary.ru/item.asp?id=18789210