COMPUTATIONAL MATHEMATICS

MODEL OF SHORT-TERM FORECAST OF ELECTRICAL ENERGY CONSUMPTION OF URAL UNITED POWER SYSTEM BY SEPARATING OF A MAXIMAL SIMILARITY SAMPLE INTO THE POSITIVE AND NEGATIVE LEVELS

T. S. Demyanenko, South Ural State University, Chelyabinsk, Russian Federation, oop_flp@mail.ru

The article considers a model to forecast electrical energy consumption on the basis of the forecasting algorithm with the division of the maximal similarity sample into positive and negative levels with different approximation equations for positive and negative values. Model is tested with actual daily data of United Energy System of the Wholesale Electrical Energy and Power Market in Russia. Namely, we use Ural United Energy System data from 2009 to 2015. Based on the proposed algorithm, a forecast for the first five days of January 2016 is obtained. A forecasting error achieves 0.98% during test. We substantiate and describe step by step an algorithm to construct a model to forecast a time series. Such algorithm is more stable for stationary series. Therefore pre-time we bring a time series of electrical energy consumption to a stationary form. To this end, we find variances of the original time series of electrical energy consumption. The proposed scientific tools are recommended in the operating activity of electrical energy subjects for short-term forecast of the basic parameters of the Energy Market to reduce the penalties by improving the forecasting accuracy.

Keywords: forecasting models, main parameters, Energy Market, maximal similarity sample.

Introduction

Nowadays, the electrical energy is one of the leading sectors of Russian economy. The share of electrical energy in the country's GDP is about 10%. By electrical energy production, Russia is on the 4th place in the world after China, USA and India. In 2015, the production volume was 1.064 trillion kW o h, and the production growth achieved 0.3% [1].

The volume of electrical energy, which is sold in the framework of bilateral agreements and the market in the "day-ahead" forms the planned consumption of electrical energy. However, the actual consumption is inevitably differ from the planned consumption [2]. The trade of variances from planned production/consumption is on the balancing market in the real time. In this case the system operator (OJSC "SO UES") conducts an additional competitive selection of suppliers applications. Note that the operator takes into account the forecast consumption in the energy system, the economic efficiency of stations employment and system reliability requirements [3].

The variance between the actual consumption and the planned consumption is qualified by own or external initiatives. Own initiative arises due to actions of market participants

(customer or supplier). External initiative arises as a result of either commands of the System Operator or an accident that led to the forced change of the electrical energy production or consumption regime. On order to determine a cost of the variances for different types of initiatives we use formulas (cuttings), calculated for each hour of the day for each node of the computational model. The cuttings are defined as the maximal (minimum) values of the indicator for the balancing market (BM) and the market price for "day-ahead" (MDA). Therefore more accurate execution of the planned consumption and production of electrical energy is stimulated. On the basis of the variances cost, we define the preliminary conditions and liabilities of BM such that the difference between them forms the unbalance of the balancing market. Negative unbalance is distributed among the participants in proportion to their own initiatives. Positive unbalance is distributed between both the suppliers in proportion to the execution of external initiatives, and the consumers, which observe the planned consumption as accurately as possible. Thus, there is the following situation in BM. Participants of the market, allowing the greatest variances of actual consumption and production from planned ones by their own initiative, are "fined". The participants, which observe the planned consumption and perform the System Operator commands as accurately as possible, are "rewarded" [4].

In the paper we propose an algorithm for short-term forecasting of electrical energy consumption based on the separating of maximal similarity sample into positive and negative levels. The statistics of actual electrical energy consumption by Ural United Energy System between 2009 and 2015 is used as an initial data.

1. Forecasting Algorithm with the Separation of Maximal Similarity Sample into Positive and Negative Levels with Different Approximation Equations

Note that a consumption volume is a numerical indicator ordered in time, which describes the level of consumed electrical energy at successive discrete times. Therefore, such consumption data can be considered as the time series. For example, we consider the data of Ural United Energy System from 2009 to 2015. For the analysis we use the actual values of electrical energy consumption. The data are taken from the official website of the System Operator of Russia electrical energy. A time series is called stationary, if the following two conditions hold. First, levels of the series are approximately uniform and have the form of continuous random fluctuations around a mean value. Second, neither the average amplitude, nor the nature of such fluctuations do not show significant changes over time. Note that the original time series is not stationary. This fact is evidenced by Dickey-Fuller test. Therefore, it is necessary to obtain a series of variances $\tilde{Y}_t = \{(y_{i+1} - y_i, t_i), i = 1, 2, ..., n-1\}$. To this end, we consider the difference between the level of series y_{i+1} and the previous level of series: $\tilde{Y}(t) = y_{t+1} - y_t$.

The forecast is based on a number of consumption deviations in the first five days of January 2016.

Let us formally describe each step of the algorithm.

Step 1. Define a sample prior to the forecast.
In time series $\tilde{Y}_t$, $t \in \{1, 2, ..., n-1\}$, the sample $\tilde{Y}_t^M = \tilde{Y}_{t_{n-1}}, ..., \tilde{Y}_{t_{n-1}-M}$ of length M, where $M \in \{1, 2, ..., t_{n-1}\}$, precedes the forecast value [5].

Step 2. Decompose a sample prior to the forecast.

The sample $\tilde{Y}_t^M$, prior to the forecast, is decomposed into two sub-samples. They are $\tilde{Y}_t^{M+}$ of length $v \in \{n-1-M; n-1\}$ for positive values and $\tilde{Y}_t^{M-}$ of length $M - v \in \{n-1-M; n-1\}$ for negative values.

Step 3. Define an approximation sample.

The modelling of the time series using samples is based on the assumption that the time series is a sequence of samples. Based on this assumption, let us assume that in the studied time series $\tilde{Y}_t$ there is a sample $\tilde{Y}_{t-k}^M = \tilde{Y}_{t_{n-1}-k}, ..., \tilde{Y}_{t_{n-1}-k+M}$ of length M and such that its origin is at point $k \in \{1, 2, ..., t_{n-1} - M\}$.

To define the sample $\tilde{Y}_{t-k}^M$, we use the following algorithm:

1. Sequentially compare the values of sample $\tilde{Y}_t^M$ and the values of time series $\tilde{Y}_t$ with a lag of one.

2. For each shift, calculate a value of the correlation coefficient:

$$r_k^M = \frac{\left| M \sum_{i=1}^M (\tilde{Y}_{t-k+i}^M \cdot \tilde{Y}_{t+i}^M) - \sum_{i=1}^M (\tilde{Y}_{t-k+i}^M) \cdot \sum_{i=1}^M \tilde{Y}_{t+i}^M \right|}{\sqrt{\left[M \sum_{i=1}^M (\tilde{Y}_{t-k+i}^M)^2 - (\sum_{i=1}^M \tilde{Y}_{t-k+i}^M)^2 \right] \cdot \left[M \sum_{i=1}^M (\tilde{Y}_{t+i}^M)^2 - (\sum_{i=1}^M \tilde{Y}_{t+i}^M)^2 \right]}} \in [0, 1].$$

3. Select maximum of the obtained correlation coefficients. Determine the value k of lag, which corresponds to the maximum.

Step 4. Determine a prediction sample.

Suppose it is necessary to forecast p levels of the time series. Then, the prediction sample is an interval of the time series, which follows after $\tilde{Y}_{t-k}^M$ and is of length p, i.e. $\tilde{Y}_t^p = \tilde{Y}_{t-k+M+1}, ..., \tilde{Y}_{t-k+m+1+p}$.

Step 5. Decompose an approximation sample.

Note that values of the time series can be both positive and negative. Therefore, we can not calculate the approximation by the classical linear model. So, the resulting sample $\tilde{Y}_{t-k}^M$ is divided into two parts: $\tilde{Y}_{t-k}^{M+} \geq 0$ for positive sample values and $\tilde{Y}_{t-k}^{M-} < 0$ – for negative sample values. A length of sample $\tilde{Y}_{t-k}^{M+}$ is $q \in \{n-1-M; n-1\}$, a length of sample $\tilde{Y}_{t-k}^{M-}$ is $(M - q) \in \{n-1-M; n-1\}$.

Step 6. Addition decompose approximation samples zeros

So, we divide samples prior both the forecast and the approximation into positive and negative values. In practice, a length of sample $\tilde{Y}_t^{M+}$ may not be correspond to a length of sample $\tilde{Y}_{t-k}^{M+}$. Typically, such situation occurs when the levels with different signs are located at the same position in the samples $\tilde{Y}_t^{M+}$ and $\tilde{Y}_{t-k}^{M+}$. For example, for index 100, there are a positive value in sample $\tilde{Y}_t^{M+}$ and negative value in sample $\tilde{Y}_{t-k}^{M+}$. To resolve this situation, both samples are constructed of length M, level values, which are suitable for the sample parameters, remain unchanged and occupy their respective positions, and all the gaps are filled with zeros.

Samples $\tilde{Y}_{t-k}^{M-}$ and $\tilde{Y}_t^{M-}$ are complemented in the same way.

Step 7. Decompose the forecasting sample.

Forecasting sample $\tilde{Y}_t^p$ is divided into two sub-samples. A length of sub-sample $\tilde{Y}_t^{p+}$ (for positive values) is $w \in \{t - K + m + 1; t - k + m + 1 + p\}$. A length of sub-sample $\tilde{Y}_t^{p-}$ (for negative values) is $(p - w) \in \{t - K + m + 1; t - k + m + 1 + p\}$.

Step 8. Construct the approximation equations for positive and negative values of the approximation sample.

For obtained approximation samples $\tilde{Y}_{t-k}^{M+}$ and $\tilde{Y}_{t-k}^{M-}$ we construct the approximation equations

$$\hat{\tilde{Y}}_t^{M+} = \alpha_1^+ \tilde{Y}_{t-k}^{M+} + \alpha_0^+ I^{M+},$$

$$\hat{\tilde{Y}}_t^{M-} = \alpha_1^- \tilde{Y}_{t-k}^{M-} + \alpha_0^- I^{M-},$$

where I^{M+} and I^{M-} are unit vectors, α_0^+, α_0^-, α_1^+, α_1^- are approximation coefficients. The values of the coefficients should be such that

$$\sigma^2 = \sum_{t=0}^{M-1} (\tilde{Y}_{t+i}^{M+} - \hat{\tilde{Y}}_{t+i}^{M+})^2 \to \min,$$

$$\sigma^2 = \sum_{t=0}^{M-1} (\tilde{Y}_{t+i}^{M-} - \hat{\tilde{Y}}_{t+i}^{M-})^2 \to \min,$$

that is, a square of the deviations between the actual and the model values should be minimal.

The coefficients of equations are found by the least squares method:

$$Z_X^+ \cdot A^+ = Z_Y^+,$$

where

$$A^+ = \begin{pmatrix} \alpha_1^+ \\ \alpha_0^+ \end{pmatrix},$$

$$Z_X^+ = \begin{pmatrix} \sum_{i=0}^{M-1} (\tilde{Y}_{t-k+i}^{M+})^2 & \sum_{i=0}^{M-1} (\tilde{Y}_{t-k+i}^{M+}) \\ \sum_{i=0}^{M-1} (\tilde{Y}_{t-k+i}^{M+}) & M \end{pmatrix}, \qquad Z_Y^+ = \begin{pmatrix} \sum_{i=0}^{M-1} (\tilde{Y}_{t-k+i}^{M+}) \cdot (\tilde{Y}_{t+i}^{M+}) \\ \sum_{i=0}^{M-1} (\tilde{Y}_{t+i}^{M+}) \end{pmatrix},$$

$$Z_X^- \cdot A^- = Z_Y^-,$$

where

$$A^- = \begin{pmatrix} \alpha_1^- \\ \alpha_0^- \end{pmatrix},$$

$$Z_X^- = \begin{pmatrix} \sum_{i=0}^{M-1} (\tilde{Y}_{t-k+i}^{M-})^2 & \sum_{i=0}^{M-1} (\tilde{Y}_{t-k+i}^{M-}) \\ \sum_{i=0}^{M-1} (\tilde{Y}_{t-k+i}^{M-}) & M \end{pmatrix}, \qquad Z_Y^- = \begin{pmatrix} \sum_{i=0}^{M-1} (\tilde{Y}_{t-k+i}^{M-}) \cdot (\tilde{Y}_{t+i}^{M-}) \\ \sum_{i=0}^{M-1} (\tilde{Y}_{t+i}^{M-}) \end{pmatrix}.$$

Step 9. Obtain the forecast.

In order to obtain the forecast, we approximate each value of sub-sample $\tilde{Y}_t^{p+}$ using equation $\hat{\tilde{Y}}_t^{p+} = \alpha_1^+ \tilde{Y}_t^{p+} + \alpha_0^+ I^{M+}$, and each value of sub-sample $\tilde{Y}_t^{p-}$ using equation $\hat{\tilde{Y}}_t^{p-} = \alpha_1^- \tilde{Y}_t^{p-} + \alpha_0^- I^{M+}$. Obtained forecasting values are arranged in the same order as the corresponding values in the sample $\tilde{Y}_t^p$.

2. Results

Graphical analysis of time series of electrical energy consumption (Figure 1) shows that there exist the seasonal fluctuations whose period is 1 year. Therefore, it is advisable to define a length of sample $\tilde{Y}_t^M = \tilde{Y}_{t_{n-1}}, ..., \tilde{Y}_{t_{n-1}-M}$ as the one-half period, where $M = 183$.
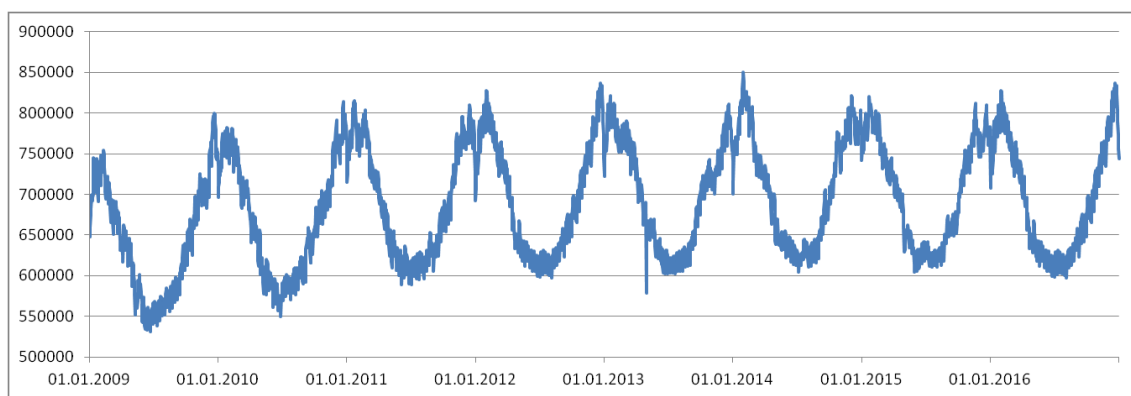

Fig. 1. The consumption of electrical energy of Ural United Power System is presented by daily data 2009–2015

The correlation coefficient between the sample in prior the forecast and the approximation sample is 0.899. Graphical representation of the samples is shown in Figure 2.
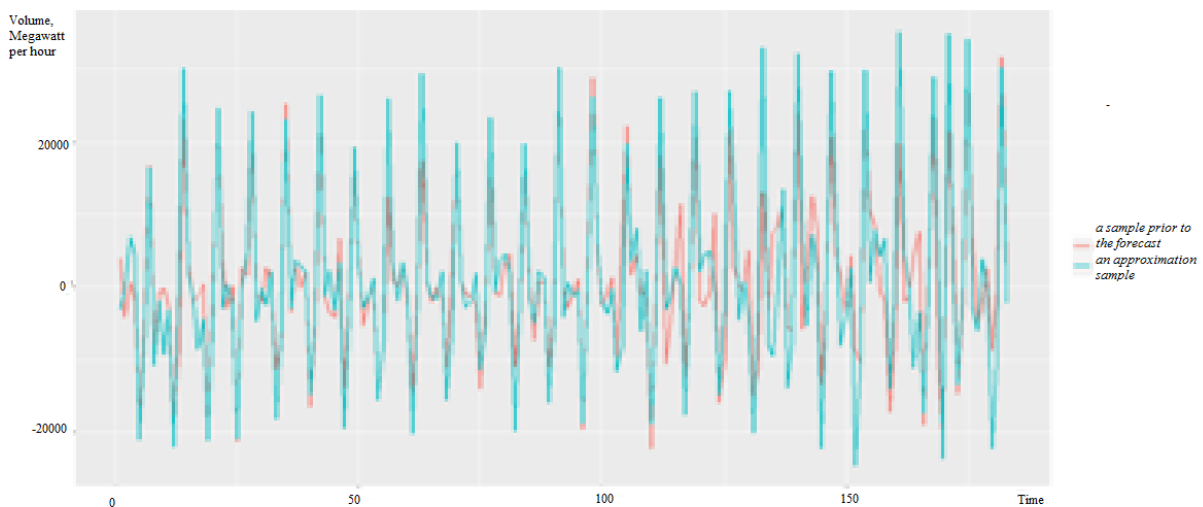

Fig. 2. The sample in prior the forecast and the approximation sample

A diagram of both forecasting values for the first 5 days of January 2016 and the actual values is shown in Figure 3.
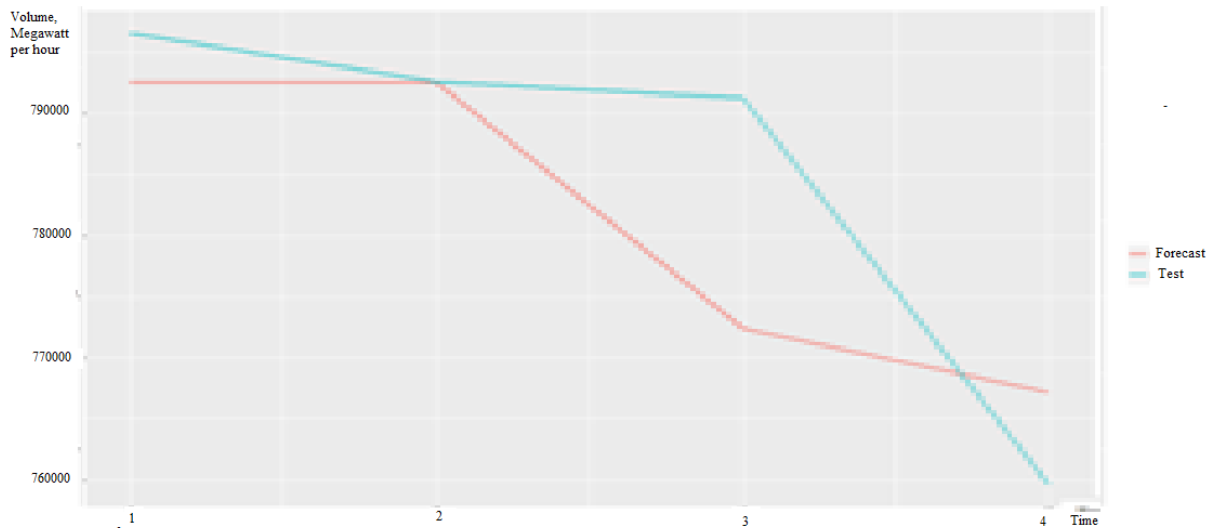

Fig. 3. Forecasting and actual values of the electrical energy consumption

An accuracy of the time series forecast is evaluated using the mean absolute percentage error (MAPE):

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{\left| y(i) - \widehat{y}(i) \right|}{y(i)} \cdot 100\%,$$

where $y(i)$ is an actual value, $\widehat{y}(i)$ is a forecasting value. For the current forecast, $MAPE = 0.98\%$.

Conclusion

The article considers a model of short-term electrical energy market forecast using a forecasting algorithm with the division of the maximum similarity sample on the positive and negative levels with different approximation equations for positive and negative values. The algorithm is more stable for stationary series. Therefore, pre-time the time series of electrical energy consumption is reduced to a stationary form. To this end, we find the deviations of the original time series of electrical energy consumption. The average error of approximation is 0.98%. Therefore, we can say that the model is suitable for short-term planning of the market of both electrical energy and power.

References

1. Mokhov V.G., Demyanenko T.S. Modelling of the Time Series Digressions by the Example of the Ups of the Ural. *Bulletin of the South Ural State University. Series: Mathematical modelling, programming and computer software*, 2015, vol. 8, no. 4, pp. 106–110. doi: 10.14529/mmp150412.

2. Drucker P.F. *Management Challenges for the 21st Century*. New York, Harper Business, 1999.

3. Salvatore D.*Microeconomics: Theory and Applications*. USA, Oxford University Press, 2003.

4. Mokhov V.G., Demyanenko T.S. Prognosing the Consumption of Electric Energy on the Wholesale Electric Power Market. *Bulletin of the South Ural State University. Series: Economics and Management*, 2014, vol. 8, no. 2, pp. 86–92. (in Russian).

5. Chuchueva I.A. The Time Series Extrapolation Model Based on Maximum Likeness Set. *Information Technologies*, 2010, no. 12, pp. 43–47. (in Russian).

Tatyana S. Demyanenko, Senior Tutor, Department of Mathematical and Computer Modeling, South Ural State University (Chelyabinsk, Russian Federation), oop_flp@mail.ru.

МОДЕЛЬ КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ ОБЪЕМА ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ ОЭС УРАЛА С ПОМОЩЬЮ РАЗДЕЛЕНИЯ ВЫБОРКИ МАКСИМАЛЬНОГО ПОДОБИЯ НА ПОЛОЖИТЕЛЬНЫЕ И ОТРИЦАТЕЛЬНЫЕ УРОВНИ

Т.С. Демьяненко

В статье рассмотрена модель прогнозирования объемов потребления электроэнергии на основе алгоритма прогнозирования с разделением выборки максимального подобия на положительные и отрицательные уровни с разными уравнениями аппроксимации для положительных и отрицательных значений. Модель протестирована на фактических посуточных данных Объединенной энергосистемы Оптового рынка электроэнергии и мощности России, а именно на данных ОЭС Урала с 2009 по 2015 гг. На основании предложенного алгоритма построен прогноз на первые 5 дней января 2016 г. При тестировании была достигнута ошибка прогноза 0,98%. Обоснован и описан пошаговый алгоритм построения прогнозирования модели временного ряда. Данный алгоритм более устойчив для стационарных рядов, поэтому предварительно временной ряд объема потребления электроэнергии был приведен к стационарному виду путем нахождения отклонений исходного временного ряда объема потребления электроэнергии. Разработанный научный инструментарий рекомендуется в операционной деятельности субъектов электроэнергетики при краткосрочном прогнозировании основных параметров энергетического рынка для снижения штрафных санкций за счет повышения точности прогнозов.

Ключевые слова: модели прогнозирования, основные параметры, рынок электроэнергетики, выборка максимального подобия

Литература

1. Мохов, В.Г. Моделирование основных параметров рынка на сутки вперед и индекса балансирующего рынка / В.Г. Мохов, Т.С. Демьяненко // Вестник ЮУрГУ. Серия: Математическое моделирование и программирование. – 2015. – Т. 8, № 4. – С. 127–130.

2. Drucker, P.F. Management Challenges for the 21st Century / P.F. Drucker. – New York: Harper Business, 1999.

3. Salvatore, D. Microeconomics: theory and applications / D. Salvatore. – USA: Oxford University Press, 2003.

4. Мохов, В.Г. Прогнозирование потребления электрической энергии на оптовом рынке энергии и мощности / В.Г. Мохов, Т.С. Демьяненко// Вестник ЮУрГУ. Серия: Экономика и менеджмент. – 2014. – Т. 8, № 2. – С. 86–92.

5. Чучуева, И.А. Модель экстраполяции временных рядов по выборке максимального подобия / И.А. Чучуева // Информационные технологии. – 2010. – № 12. – С. 43–47.

Демьяненко Татьяна Сергеевна, старший преподаватель, кафедра математического и компьютерного моделирования, Южно-Уральский государственный университет (г. Челябинск, Российская Федерация), oop_flp@mail.ru