Modeling and forecasting seasonal and cyclical events using retrospective data

Anatoly Darmanyan[1] and *Aleksey* Rogachev[1,2,*]

[1]Volgograd State Agrarian University, 26, Universitetskiy Ave., Volgograd, 400002, Russian Federation
[2]All-Russian research Institute of irrigated agriculture, 9, Timiryazeva, Volgograd, 400002, Russian Federation

Abstract. Based on the data of the Ministry of energy of Russian Federation for the period of 2012-2019 years the mathematical modeling of statistics of power production were performed using time series analysis in Statistica. To describe the statistical data, the exponential smoothing model was used, its parameter was found and short-term forecasting of electricity generation in Russia for 2019 was performed. For long-term forecasting the seasonal model of autoregression and integrated moving average ARIMA - (0,1,1)(0,1,1) is chosen, its parameters are defined and its adequacy to real data is proved. With the help of the found model, the forecasting of electricity generation in the Russian Federation within 90% of the confidence interval for twelve months 2019-2020 is performed. The non-stationary time series of the studied statistics, the presence of a trend and a seasonal component in it are proved. The model parameters are defined as ARIMA(0,1,1)(0,1,1) by means of the Statistica software, and the adequacy of this model to real data is proved. Using the found mathematical model, a long – term forecast (period-12 months) of electricity generation in the Russian Federation with a 90% confidence interval was performed.

1 Introduction

To analyze statistical data in various fields of scientific activity, numerous statistical methods (more than 50) have been developed, whose computational algorithms are implemented on a PC, in particular, in the PP Statistica environment [1-3, 8, 9]. The possibilities of various statistical methods are truly enormous, especially those that use retrospective analysis, when data from past time periods are used to identify trends, as well as cyclical and seasonal components. However, as noted in [4, 10, 11], most of these methods are not used for mathematical modeling of phenomena and objects in the agro-industrial sphere. At the same time, using statistical methods, it is possible to construct adequate mathematical models that can be used not only for assessing the state and functioning of any phenomenon or process at the present time, but also for predicting the development of the phenomenon or process in the future [5].

* Corresponding author: rafr@mail.ru

This is especially important for analyzing statistical data that changes over time, in particular, having a trend, as well as cyclical and seasonal components. Among such data, the statistics of electricity generation in the Russian Federation is of interest, which reflects the state of not only this energy sector, but also the dynamics of the country's economy as a whole, since there is a positive correlation between the GDP of any country and electricity generation.

Currently, the analysis of electricity generation statistics is carried out using standard methods, when the percentage of electricity generation for certain periods of time is calculated [6] and conclusions are drawn about the dynamics of the entire process based on the magnitude of its change. This approach makes it possible to estimate the changes that have occurred over a certain period of time, but it does not allow for scientifically based forecasting of changes in the studied values for future periods of time. However, mathematical algorithms for constructing models that adequately describe, for example, non-stationary time series are well developed and implemented in the environment of the Statistica 6.0 software.

The purpose of this work was mathematical modeling of cyclical and seasonal phenomena on the example of statistics of electricity generation in the Russian Federation using time series analysis methods in the environment of PP Statistica 6.0.

2 Materials and methods

For mathematical modeling, we used retrospective statistical data of the Ministry of energy of the Russian Federation on electricity generation for the period 2012-2019 [7], which are shown graphically in figure 1.

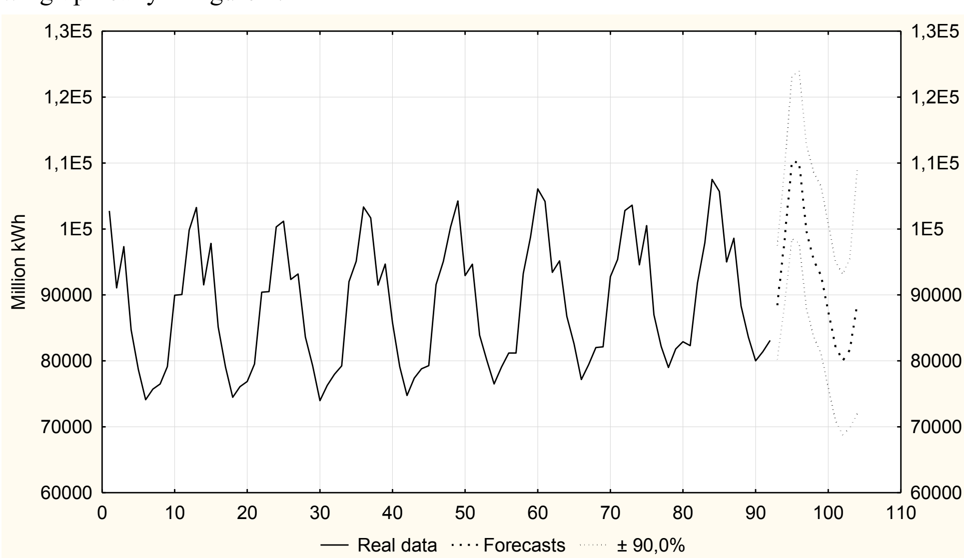

Fig. 1. Electricity generation (million kWh) in the Russian Federation by month [7]: 1 -1.01.2012; 90 – 1.08.2019 Forecast based on the ARIMA model (0,1,1)(0,1,1): 91 – 1.09.2019 G.; 103 – 1.07.2020.

As can be seen from the graph in figure 1, the statistics of electricity generation in the Russian Federation is a non-stationary time series, which has a clear seasonal component. Therefore, for mathematical modeling of such data, the tools of the Statistica software were used, in particular, the "Time Series/Forecasting" function in the non-linear Models-Time Series module [1,2]. This function provides the following tools for analyzing statistics:

Fourier analysis for identifying the seasonal component, exponential smoothing for short-term forecasting, and the ARIMA integrated moving average model for mathematical modeling and long-term forecasting.

The exponential smoothing model developed by R. Brown for short-term forecasting purposes is based on the following recurrent formula [1,3]:

$$St = \alpha * Xt + (1-\alpha) * St- \quad (1)$$

where: X = (X1, X2....... Xt) is the original time series:

α - (parameter "alpha") – smoothing coefficient.

A feature of the mathematical model (1) is that recent values given more weight than the penultimate, penultimate – even worse than the previous etc. When this formula is applied recursively, each new smoothed value is calculated as a weighted average of the current observation and the smoothed series. If $\alpha = 1$, the previous observations are completely ignored; if $\alpha = 0$, the current observations are ignored. Alpha values in the range 0 – 1 give intermediate results. The main drawbacks of this mathematical model are that it does not take into account cyclical and seasonal components and cannot be used to calculate confidence intervals when performing the short-term forecasting procedure [1,3].

A distinctive feature of the seasonal autoregression model and the integrated moving average ARIMA (hereinafter - ARIMA) is that it takes into account cyclical and seasonal components in addition to the trend, and therefore it is multiparametric. The model has the following parameters: autoregression parameter – p, seasonal autoregression parameter – P, moving average parameter – q, seasonal moving average parameter – Q. All these parameters are calculated in Statistica 6.0 in accordance with the Box-Jenkins methodology [8]. According to this methodology, the original time series is transformed by taking first-or second-order differences, resulting in a stationary time series for which the corresponding model parameters are calculated [1].

3 Results

Mathematical modeling of electricity generation statistics in the Russian Federation consisted in determining the parameters of two models: the simple exponential smoothing model and the ARIMA-arpss model, which were used for forecasting future time periods.

3.1 Exponential smoothing model

To determine the parameters of the exponential smoothing model using the "Exponential Smoothing and Forecasting" function of the "Time Series/Forecasting" module of Statistica 6.0 [1-3] for statistics of electricity generation in the Russian Federation for the period 2012-2018, a single parameter of the model (1) was found: $\alpha = 0{,}1$. Thus, the mathematical model of exponential smoothing (1) acquired the following specific form:

$$St = 0{,}1 * Xt + 0{,}9 * St-1 \quad (2)$$

According to this model, electricity generation in the Russian Federation in each month of the 2012-2018 study period is determined mainly by the value of this indicator in the previous month, i.e., mainly by the value of the term 0.9*St-1 in expression (2). At the same time, the influence of a number of values of power generation (variable Xt) is insignificant, since it contributes to the total value of St by the terms 0.1*Xt in expression (2). Using the found model (2), a short-term forecast of the values of electricity generation in the Russian Federation for 2019 was performed, which are shown in figure 2 in comparison with real data.

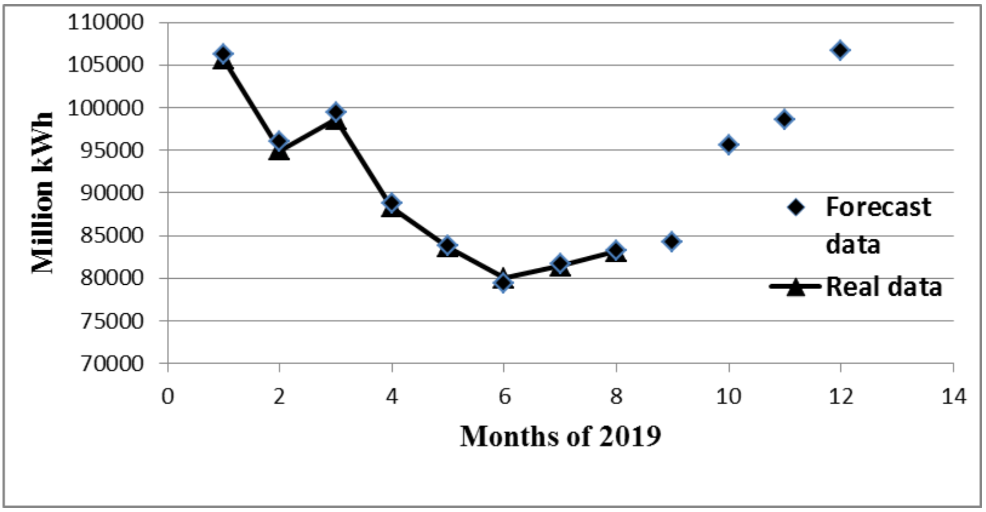

Fig. 2. Real data on electricity generation in Russia in 2019. in comparison with the predicted values for the exponential smoothing model (2).

As can be seen from the graph in figure 2, between the real values of electricity generation in the Russian Federation in 2019. there is a good correspondence between the values calculated using the exponential smoothing model.

However, it should be noted that the short-term forecast of electricity generation values in the Russian Federation for September-December 2019. it was performed without determining the confidence interval, since this is not provided for by the calculation algorithm using this model, which was already noted above when describing this model. Long-term forecasting with the calculation of the confidence interval is possible using the ARIMA model, finding the parameters of which is a step-by-step research procedure, which is used later.

At the first step of mathematical modeling of electricity generation statistics in the Russian Federation, it is necessary to answer the question whether this statistics is a non-stationary time series. Therefore, first of all, a study was conducted to determine whether the time series under study has cyclical or seasonal components. To do this, using the "Spectral (Fourier) analysis" function of the "Time Series/Forecasting" module of the Statistica 6.0 software, a Fourier spectral analysis of all statistical data for the period 2012-2018 was performed, resulting in a periodogram with a narrow single peak at the value of the number of months t=12. This peak indicates the presence of a seasonal cycle with a period of 12 months, and there are no other irregular cycles in the studied statistics. Thus, it is revealed that electricity generation in the Russian Federation for the period 2012-2018. it had a stable seasonal cycle with a period of 12 months, i.e. with a period of 1 year. Thus, the studied statistics have one of the signs of non-stationarity of the time series [1-3].

The next step in mathematical modeling is to determine the trend in the time series under study. Since it is difficult to do this visually using the graph in figure 1, a sample autocorrelation function was constructed using the "ARIMA and autocorrelation functions" function of the "Time Series/Forecasting" module of the Statistica 6.0 software to identify the presence (or absence) of a trend in the studied statistical data, which is shown in figure 2. As can be seen from this graph, the sample autocorrelation function does not tend to fade, but has peaks of a stable seasonal cycle with a period of 12 months. All this indicates that the time series is non-stationary, i.e. there is a weak trend.

3.2 The ARIMA model

The model (ARIMA) with three parameters (p,q,d,) was chosen as a mathematical model for describing the studied statistics. To determine the value of the parameter d, first-order differences were calculated, resulting in a time series that was checked for stationarity using a partial autocorrelation function. It turned out that the time series of the first differences became stationary, which allowed us to choose the value of the model parameter d=1. For other parameters of the ARIMA-arpss model, the following values are selected: p=0 and q=1.

Since the time series under study, as shown above, has a pronounced seasonal component with a period of 12 months, it is necessary to make a seasonal adjustment to the ARIMA model [1]. Therefore, seasonal parameters with the following values were introduced into this model: the seasonal autoregression parameter Ps=0, the seasonal difference Ds=1, and the seasonal moving average parameter Qs=1. Thus, the following ARIMA model was formed (0,1,1)(0,1,1).

The parameters of this model, calculated using the Statistica 6.0 software, were highly significant at a level below 0.05, which proves their statistical significance. In addition, to check the adequacy of the found ARIMA model (0,1,1)(0,1,1) to real data, a number of residuals were analyzed for compliance with its normal distribution. This analysis showed that, despite some outliers, a number of residues in General correspond to the normal distribution, which indicates the adequacy of the found ARIMA model (0,1,1)(0,1,1) to real data.

Using the found model ARIMA (0,1,1)(0,1,1), the forecast of electricity generation in the Russian Federation for the future period of 12 months is made: from 1.09.2019 to 1.07.2020. The results of calculations together with the 90% confidence interval are shown in figure 1. As can be seen from the comparison of the graphs in figure 1, the dynamics of the predicted values of electricity generation is in good agreement with the dynamics of the entire time series for the period 2012-2018.

5 Conclusions

The research methodology was developed and mathematical modeling of electricity generation statistics in the Russian Federation was performed based on the analysis of retrospective data of the Ministry of energy of the Russian Federation for the period 2012-2019. It is shown that the studied statistics represent a non-stationary time series, in which there is a trend and a seasonal component. For mathematical modeling of such a time series, an exponential smoothing model is used, for which its parameter $\alpha = 0{,}1$ is determined by means of the Statistica 6.0 software. Using the found model, a short-term forecast of electricity generation in the Russian Federation for 2019 was made. and the calculated values were compared with real data, which showed a good match.

The seasonal model of autoregression and integrated moving average ARIMA(p,q,d,)(Ps,Ds,Qs) was chosen as a mathematical model for describing statistics of electricity generation in the Russian Federation and conducting long-term forecasting. Using statistical data on electricity generation in the Russian Federation for the period 2012-2019, the non-stationary time series of the studied statistics, the presence of a trend and a seasonal component in it are proved. The model parameters are defined as ARIMA(0,1,1)(0,1,1) by means of the Statistica 6.0 software, and the adequacy of this model to real data is proved. Using the found mathematical model, a long – term forecast (period-12 months) of electricity generation in the Russian Federation with a 90% confidence interval was performed.

The article was prepared with the financial support of the RFBR and the Administration of the Volgograd region for the project № 19-416-340014.

References

1. A.A. Khalafyan, Statistica 6, Static data analysis Textbook (Moscow: Binom-Press, 512, 2007)
2. V.P. Borovikov, V.P. Ivchenko, Forecasting in the STATISTICA system in the WINDOWS environment: fundamentals of theory and intensive practice on the computer (Textbook. Moscow: Finance and statistics, 368, 2006)
3. Yu.P. Lukashin, Adaptive methods for short-term time series forecasting (Moscow: Finance and statistics, 368, 2003)
4. A.P. Darmanian, N.M. Veselova, *Scientific significance of statistical methods for analysing experimental data in the system of postgraduate agricultural education,* Current issues of professional education **3(12)**, 33-37 (2018)
5. Forecast of world and Russian energy development (Moscow: INEI RAN – Analytical center under the government of the Russian Federation, 2016)
6. S.A. Ayvazyan, Applied statistics, Fundamentals of econometrics 2001 **2,** 432 (2001)
7. Statistics of the Ministry of energy Retrieved from: https://minenergo.gov.ru/activity/statistic/
8. A.F. Rogachev, *Energy assessment and optimization of programmed agricultural production using retrospective data,* E3S Web of Conferences. The conference proceedings Innovative Technologies in Environmental Science and Education. Don State Technical Universyty (2019)
9. K.E. Tokarev et al. *IOP Conf. Ser.:* Earth Environ. Sci. **488**, 012047 (2020)
10. A.F. Rogachev, Mathematical modeling of economic dynamics in agricultural production: Monograph (Volgograd: Publishing house of Volgograd State University 172, 2014)
11. A.F. Rogachev et al. *A set of data on retrospective grain yield for neural network modeling* IOP Conf. Ser.: Earth Environ. Sci. **577**, 012006 (2020)