Lithuanian electricity market price forecasting model based on univariate time series analysis

Mindaugas Česnavičius

Lithuanian Energy Institute,
Breslaujos St. 3,
44403 Kaunas, Lithuania
Email: mindaugas.cesnavicius1@gmail.com

Electricity price changes can significantly affect expenses in energy intensive industries, adjust profits or losses for electricity retailers and cause problems for country's national energy strategy implementation. Forecasting models based on statistical methods and previous variable values help to predict future values and adjust strategy according to the forecast. This paper concentrates on the Lithuanian electricity market and presents the widely used ARIMA forecasting models based on the univariate time series analysis. The Lithuanian electricity market is selected due to a lack of statistical researches based on electricity market prices in Lithuania, as well as significant future electricity market liberalization projects.

Electricity price data for analysis are taken from the Nord Pool electricity market operator website. The Nord Pool represents the Northern Europe electricity market operator where Lithuania and other 14 European countries trade electricity on a daily basis. To provide a long-term electricity price outlook average monthly data from July 2012 to December 2019 are selected for analysis. Before building the ARIMA model data are tested with various statistical tests to guarantee that time series are stationary, there is no autocorrelation or structural breaks. Once the data validity is confirmed, the time series is divided into train and test sets. The train data set is used to create a fitting ARIMA model, while the test set is used to define forecasting accuracy. Created forecasts of models are compared between each other using common comparison statistics, and the most accurate models are defined. Finally, the selected model is trained on a full dataset and the electricity price forecast for the year 2020 is constructed.

The created AR (1) model had the smallest error value compared to the test dataset, while the SARIMA (1,1,1) model had the best approximation statistics. By combining both models the weighted SARIMA (1,1,1) model is constructed with the features of low forecasting error and precise actual time series approximation. The final model forecast for the year 2020 shows the monthly average electricity price decrease at the beginning of the year, a significant increase at the second half of the year and a price drop at the end of the year. Forecasting results can help companies to plan their electricity production and maintenance periods to maximize income from sold energy and minimize potential losses due to planned shutdown.

Keywords: electricity price, forecasting, univariate time series, ARIMA

INTRODUCTION

Since the end of 20th century, traditionally monopolistic and government-controlled energy sector started to become more deregulated and adopted competitive market rules. These days, many countries all around the world trade electricity under market rules according to the spot and derivative contracts [1]. The need for dynamic market appeared due to the growing installed power of renewable energy resources, introduction of smart grids, increased market competition and aging electricity grid infrastructure. The increasing electricity system complexity and requirements raised the importance of probabilistic electricity price and load forecasting that helps to ensure efficient energy system planning and operations [2]. Currently, the energy system became so deregulated that electricity demand and price forecasting emerged as one of the major research fields in electrical engineering [3]. From six months up to one-year period, electricity producers need to know the portion of electricity which can be sold via bilateral contracts and the portion which will bring more profit by being sold according to the spot prices. A reliable electricity price forecast help producers to identify potentially profitable bilateral contracts and the ones which might negatively impact the profit [4]. A number of scholars are developing various tools and algorithms to provide a reliable electricity load and price forecast [5]. The current academic literature concentrates on short-term electricity price forecasting which provides an opportunity for electricity market participants. However, there is a lack of researches considering medium to long-term electricity price forecasting models which are necessary for investment and political decisions [6]. Utility companies aiming to maximize profit under deregulated electricity market conditions have to offer the right amount of electricity, at the right time and for the right price [7]. The long-term electricity price forecast provides the necessary information for maintenance and peak production months planning to minimize losses and maximize profit.

Electricity price time series exhibit a nonstationary component due to seasonal price swings caused by different seasonal electricity consumption patterns. During the mild summer consumers will need much less energy compared to the cold winter when electricity is one of the heating options. The academic literature proposes different stationary models of electricity price forecasting. For example, dynamic regression models, linear transfer function models or ARIMA models can be employed [8]. The ARIMA models are a class of stochastic processes used to analyse time series based on the Box and Jenkins method [9]. The method allows one to transform nonstationary time series into stationary ones, detect seasonality and identify autocorrelation. The ARIMA models are widely used to forecast various commodities prices, such as oil and natural gas [10]. Commodities and electricity price share common seasonality and interdependence features which suggests that similar modelling tools could be used for forecasting the future values.

The Lithuanian electricity market especially lacks for researches of statistical electricity price forecasting. Up to date, only one paper related with short-term electricity price forecasting models can be found. In the paper, the authors construct a short-term electricity price forecast by using average, seasonal naive and exponential smoothing methods [11]. The long-term Lithuanian electricity market price research based on univariate ARIMA models could have a meaningful academic contribution and provide a guidance for the industrial companies, electricity retailers and power plants. The long-term electricity price forecast would suggest times when maintenance could happen with the least costs, provide the average price benchmark for electricity retailers and identify the periods when energy could be sold for the highest price.

METHODOLOGY

Methodology of this research consists of three parts: data collection, construction of different forecasting models and forecasts evaluation. In the first part, a reliable electricity price data source is selected and a primary inspection with general observations about time series are made. The second part explores the features of ARIMA models and provides the forecasting guidance. Finally, the third part concentrates on statistics selection which would evaluate the forecasting accuracy of different models and identify the most precise one.

The research is based on the Lithuanian average monthly electricity market price. The selected time series dataset contains 90 points and represents the price changes from July 2012 to December 2019. The data are gathered from the Nord Pool, Northern Europe electricity exchange market operator where Lithuania is one of the 14 most active power exchange participants. Figure 1 represents the plotted time series, which are used in the research.

The graph shows constant average electricity price movements from the lower to the higher value. The highest electricity price in the observed period is 64.14 €/MWh, while the lowest value is 29.65 €/MWh, more than double price difference can be noticed. Seasonality can be identified due to a significantly lower price at the end of each year. The selected time series represent the process, which is similar to the stationary one with signs of seasonality. These reasons suggest ARIMA models for the possible analysis and forecasting.

The ARIMA models are constructed using past time series values and error terms. Previous researches show that these models perform poorly when forecasting commodity prices in the long-term horizon because of lack of non-linear relationships [13]. To achieve the long-term forecast using ARIMA models, the average monthly electricity price was selected. By doing so, prediction for one year can be made by forecasting 12 values. From the ARIMA model point of view it can be considered as a mid-term forecast while in reality it produces one-year forecast results.

Four different ARIMA models were selected during the research. The first model was the AR (1) process with the mathematical expression:

$$y_t = c + \alpha y_{t-1} + \varepsilon_t. \tag{1}$$

The AR (1) model depends on the constant term, one lagged value and the error term. This is the simplest ARIMA model, which can also be described as ARIMA (1,0,0). The second model was ARIMA (1,1,0) with the mathematical expression:

$$y_t = c + \alpha(y_{t-1} - y_{t-2}) + \varepsilon_t. \tag{2}$$

The ARIMA (1,1,0) model is very similar to AR (1). The difference is that instead of the lagged value the difference between the first lag and second lag is taken as a variable in the equation. Using the difference between the lagged values can help to create stationary time series, which are necessary for an accurate forecast. The third model was ARIMA (1,1,1) with the equation expression:

$$y_t = c + \alpha(y_{t-1} - y_{t-2}) + \beta\varepsilon_{t-1} + \varepsilon_t. \tag{3}$$

The ARIMA (1,1,1) model has all the components as the previous model with an additional lagged error term variable. The additional variable represents stochastic time series movement and increases forecasting accuracy by the model. The fourth model was SARIMA (1,1,1) with the mathematical expression:
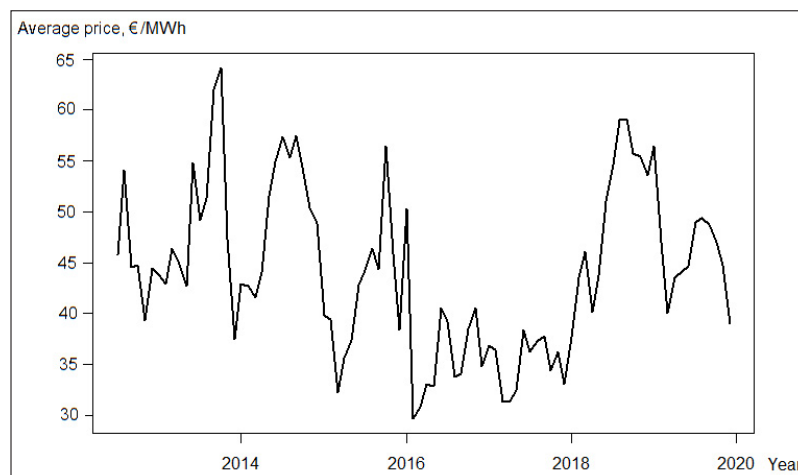

Fig. 1. Historical average monthly electricity market price in Lithuania [12]

$$y_t = c + \alpha_1(y_{t-1} - y_{t-2}) + \beta_1\varepsilon_{t-1} +$$

$$\alpha_2(y_s - y_{s-1}) + \beta_2\varepsilon_s + \varepsilon_t. \qquad (4)$$

The SARIMA (1,1,1) model has all the variables as the previous model but additional seasonal components are added. Seasonal variables follow the same logic as the corresponding lagged values, but they are constructed according to the previous seasonal values. The seasonal variable of the current month would be made according to the same month's value of the last year.

To be able to define the forecast accuracy, the dataset was divided into training and testing datasets. The training dataset represents the period from July 2012 to December 2018, while the testing dataset contains one-year period from January 2019 to December 2019. The training dataset allows model creation while the testing dataset can be used to compare the forecast with the real values and define forecasting accuracy.

To select the best model, common comparison statistics were evaluated. Models' fitted values approximation of the actual time series was defined calculating the AIC and BIC statistics. The lower statistic value shows that better fitted values match with the actual time series data. The AIC and BIC criteria can only be used to compare models between each other but do not have a stand-alone meaning. The forecasting accuracy was compared using RMSE, MAE, MPE and MAPE forecasting error statistics. These statistics represent the forecasting error in a percentage value and the statistic with the highest value was selected as a representative one.

RESULTS

Four different ARIMA models were constructed to define the best approximation for the given average monthly electricity price training dataset. Figure 2 represents the plots showing how the fitted values of each model approximate the actual time series and compare one-year forecast with the actual data.

Visually, the AR (1) model's fitted time series constantly miss the real values by one forecasting period and do not catch the peaks. The AR (1) forecast presents exponential electricity price decrease, which could have a reasonable average price meaning but does not show price movement dynamics. The ARIMA (1,1,0) model's fitted values still miss the actual values by one period but have a precise estimation of the peak values. The ARIMA (1,1,0) forecast has a constant value and lack of meaningful
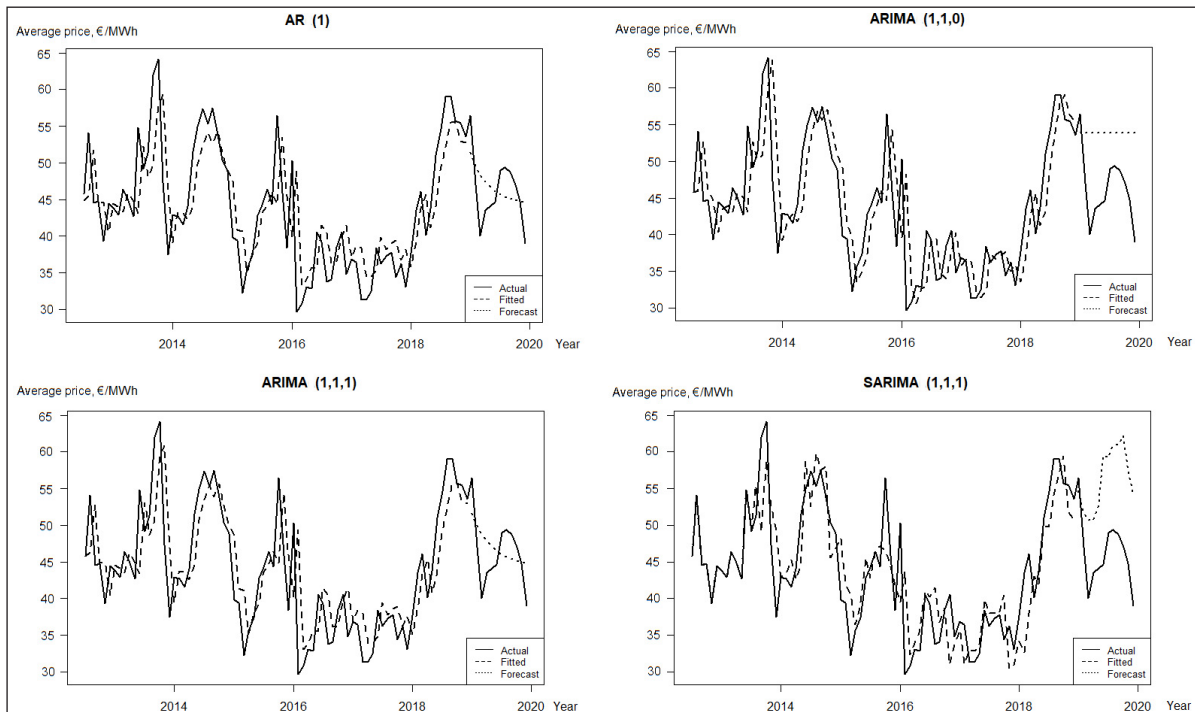

Fig. 2. Comparison of models' fitted and forecasted values with the actual data

economic interpretation. Visually, the ARIMA (1,1,1) model continues to miss the actual time series values but has a better estimation of peak prices and exponential average price forecast, similar to the AR (1) model prediction. SARIMA (1,1,1) fairly approximates the actual time series and catches the peak values. The forecast of the fourth model has a precise estimation of the actual time series dynamics but significantly misses the real values by forecasting a much higher price than it was. The visual inspection of models concludes that the AR (1) and ARIMA (1,1,1) models have the most precise forecasts while SARIMA (1,1,1) has the most precise time series estimation and the best dynamical results.

To statistically find the model with the best approximation and the smallest forecasting error, several model selection criteria are calculated. Table 1 summarizes the AIC/BIC criteria of each model for approximation and the RMSE/MAE/MPE/MAPE criteria for forecasting accuracy.

The statistics shows that the AR (1) model has the worst values for approximation criteria but holds the lowest forecasting error, contrary to the SARIMA (1,1,1) model, which has the best approximation statistics and the highest forecasting error. The results of statistical models confirm assumptions, which were made from the visual inspection of models.

None of the four constructed models offer a precise approximation and a reliable forecast. The additional fifth weighted SARIMA (1,1,1) model was created by combining the results of the AR (1) and SARIMA (1,1,1) models. The combined model uses the dynamic SARIMA forecast result, but adjusting it according to the percentage value defined by the exponential AR forecast. The forecast error for the joint model is two times lower than that for the SARIMA's model, approximately reaching 12%. Figure 3 summarizes the forecast results obtained by all models in one plot.

Table 1. **A summary of models approximation and accuracy criteria**

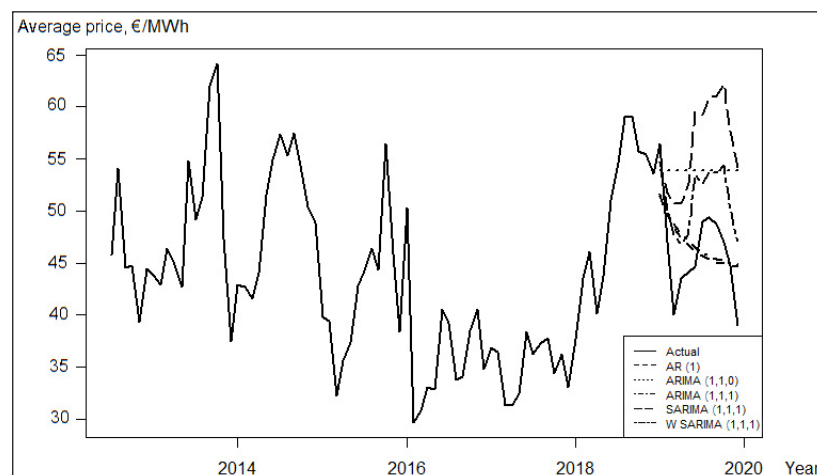| | **Models** | | | | |
	AR (1)	ARIMA (1,1,0)	ARIMA (1,1,1)	SARIMA (1,1,1)	Weighted SARIMA (1,1,1)
	Approximation criteria				
AIC	492.64	491.27	488.91	429.75	–
BIC	499.71	495.95	495.94	440.62	–
	Accuracy criteria				
RMSE	4.13	8.95	4.22	11.21	5.64
MAE	3.58	8.20	3.65	10.44	5.19
MPE	−1.97	−17.91	−2.67	−22.74	−11.13
MAPE	7.94	18,68	8.16	23.27	11.65

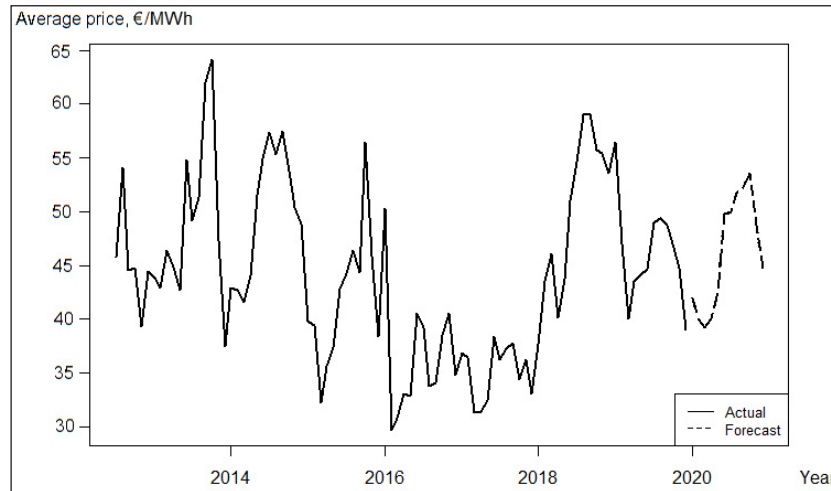**Fig. 3.** Comparison of forecast accuracy

Fig. 4. Weighted SARIMA (1,1,1) forecast for 2020

The plotted comparison of forecasts confirms that the weighted SARIMA (1,1,1) forecast has the closest shape to the actual time series graph. The defined forecast clearly has a smaller error and mimics the real values dynamics.

Due to the best approximation and the lowest forecasting error the weighted SARIMA (1,1,1) model is selected to forecast the Lithuanian average monthly electricity market price for the year 2020. Figure 4 presents the actual electricity price graph from July 2012 to December 2019 and the forecasted values from January 2020 to December 2020.

The forecast shows that during the year 2020 the average monthly electricity price in Lithuania will drop from January to March. From April the price will start to grow significantly and will reach the peak in October. During November and December, the average monthly electricity price will drop again. The lowest price level is predicted in March – 39.2 €/MWh, while the highest price level is forecasted in October – 53.61 €/MWh.

CONCLUSIONS

The literature analysis identified the growing importance of reliable electricity price forecasting due to generation from renewable energy sources and electricity market linearization. Up to date, the academic literature mostly concentrates on short-term forecasting models while there is a scarcity of long-term researches. The ARIMA models are widely used to forecast commodities prices and can be used to construct long-term electricity price forecasting models.

The univariate long-term Lithuanian electricity market price forecast can be made using the average monthly electricity price. Monthly periods between the time series allow to make a meaningful one-year forecast without forecast repetition. This would not be possible using hourly, daily or weekly time periods.

The analysis based on the ARIMA models shows that the AR (1) model had the smallest forecast error, while the SARIMA (1,1,1) model had the best actual data approximation statistics. Using both models the weighted SARIMA (1,1,1) model was constructed. The defined model can mimic actual time series dynamics and have a significantly lower forecast error compared to the SARIMA (1,1,1) model.

The weighted SARIMA (1,1,1) model forecast for 2020 showed a drop of the average monthly electricity price in Lithuania during the beginning of the year, a significant growth from April to October and another price drop at the end of the year. Forecasted electricity price dynamics can be used by electricity generators while planning maintenance and peak production time during the year. National electricity transmission and distribution operators can plan electricity grid infrastructure projects. Independent electricity retailers can predict their quarterly and yearly income based on the average forecasted electricity price for the quarter and during the whole year.

The weighted SARIMA (1,1,1) model had an approximate forecasting error of 12%. The error value is acceptable to define the abstract electricity price movement dynamics but is too high to confidently conclude the average price for every month. Univariate time series forecasting depends only on the past values and does not include any external factors, which could greatly increase forecasting accuracy. Further researches of a long-term electricity price forecasting model should be conducted considering models, which include economically justified external factors.

References

1. Weron R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*. 2014. Vol. 30. No. 4. P. 1030–1081.

2. Nowotarski J., Weron R. Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renewable and Sustainable Energy Reviews*. 2017. Vol. 81(P1). P. 1548–1568.

3. Bunn D. W. Forecasting loads and prices in competitive power markets. *Proceedings IEEE*. 2000. Vol. 88. No. 2. P. 163–169.

4. Contreras J., Espinola R., Nogales F. J., Conejo A. J. Arima models to predict next-day electricity prices. *IEEE Transactions on Power System*. 2003. Vol. 18. No. 3. P. 1014–1020.

5. Aggarwal S. K., Saini L. M., Kumar A. Electricity price forecasting in deregulated markets: A review and evaluation. *International Journal of Electrical Power & Energy Systems*. 2009. Vol. 31. No. 1. P. 13–22.

6. Ziel F., Steinert R. Probabilistic mid- and long-term electricity price forecasting. *Renewable and Sustainable Energy Reviews*. 2018. Vol. 94(C). P. 251–266.

7. Yan X., Chowdhury N. A. Mid-term electricity market clearing price forecasting: A hybrid LSSVM and ARMAX approach. *International Journal of Electrical Power & Energy Systems*. 2013. Vol. 53. P. 20–26.

8. Barquin J., Julian G.-G., Duenas P. A hybrid approach for modelling electricity price series in the medium term. *Proceedings of 16th Power Systems Computation Conference – PSCC, Glasgow, Scotland, UK, 14–18 July 2008*.

9. Box G. E. P., Jenkins G. M., Reinsel G. C. *Time Series Analysis Forecasting and Control*. 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 1994.

10. Weiss E. Forecasting commodity prices using ARIMA. *Technical Analysis of Stocks & Commodities*. 2000. Vol. 18. No. 1. P. 18–19.

11. Beigaitė R., Krilavičius T. Electricity price forecasting for Nord Pool data. *Proceedings of the International Conference on Platform Technology and Service*. 2017. Vol. 1856. P. 37–42.

12. *Day-ahead Prices, LT, Monthly.* Nord Pool, 2020. https://www.nordpoolgroup.com/Market-data1/Dayahead/Area-Prices/LT/Monthly/?view=table

13. Marcos R. A., Reneses J., Bello A. Long-term Spanish electricity market price forecasting with cointegration and VEC models. *Proceedings of the International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), Beijing, 2016*. P. 1–7.

Mindaugas Česnavičius

LIETUVOS ELEKTROS RINKOS KAINOS PROGNOZAVIMO MODELIS, PAREMTAS VIENMATE LAIKO EILUČIŲ ANALIZE

Santrauka

Elektros energijos kainos pokyčiai gali daryti įtaką energijai imlių pramonės šakų išlaidoms, paveikti nepriklausomų energijos tiekėjų pelną ir sukelti problemų įgyvendinat energetikos strategiją nacionaliniu mastu. Statistiniai prognozavimo metodai, paremti buvusiomis kintamojo reikšmėmis, leidžia nuspėti ateities vertes ir atitinkamai pritaikyti strategiją. Šiame straipsnyje pristatoma vienmačiu ARIMA modeliu paremta Lietuvos elektros energijos kainos prognozė. Lietuvos elektros rinka pasirinkta dėl nedidelio esamų statistinių tyrimų kiekio ir numatomų reikšmingų elektros rinkos liberalizavimo projektų.

Tyrime buvo naudojama istorinė elektros energijos kaina iš „Nord Pool" elektros rinkos operatoriaus duomenų bazės. „Nord Pool" yra Šiaurės Europos šalių elektros rinkos operatorius. Jo paslaugomis naudojasi Lietuva ir dar 14 Europos šalių, kiekvieną dieną vykdydamos prekybą elektros energija. Siekiant sudaryti ilgalaikę elektros kainos prognozę, tyrime buvo analizuojama vidutinė mėnesinė elektros energijos kaina nuo 2012 m. liepos iki 2019 m. gruodžio mėn. Prieš sudarant ARIMA modelį duomenys buvo apdoroti statistiniais testais. Siekta įsitikinti, kad laiko eilutė yra stacionari, nėra autokoreliacijos ir struktūrinių trūkių. Nustačius, kad duomenys tinkami tolimesnei analizei, laiko eilutė buvo padalyta į modelio apmokymo ir testavimo duomenų rinkinius. Apmokymo duomenų rinkinys naudotas sukurti įvairių parametrų ARIMA modelius, o testavimo duomenų rinkinys padėjo nustatyti skirtingų modelių prognozavimo tikslumą. Modelių prognozės buvo lyginamos tarpusavyje naudojant standartinius statistinius palyginimo kriterijus ir taip nustatant tiksliausią modelį. Galiausiai, nustatytas modelis buvo apmokomas naudojant visą duomenų rinkinį ir sudaryta 2020 m. elektros energijos kainos prognozė.

Sukurtas AR(1) modelis turėjo mažiausią prognozavimo paklaidą, o SARIMA (1,1,1) modelis geriausiai atkartojo laiko eilutės sezoniškumą. Sujungus abu šiuos modelius į koeficientu pasvertą SARIMA (1,1,1) modelį, gaunama prognozė turi mažą paklaidą ir sezoniškumą. Galutinė prognozė rodo, kad vidutinė mėnesinė elektros energijos kaina 2020 m. pradžioje sumažės, vėliau reikšmingai padidės antroje metų pusėje ir vėl sumažės metų gale. Sudaryta prognozė padeda elektros rinkos dalyviams planuoti gamybos ir remonto laikotarpius siekiant maksimizuoti pelną ir minimizuoti nuostolius.

Raktažodžiai: elektros kaina, prognozavimas, vienmatės laiko eilutės, ARIMA