FACTA UNIVERSITATIS
Series: **Economics and Organization** Vol. 8, N° 4, 2011, pp. 421 - 431

FORECASTING ELECTRICITY CONSUMPTION BY USING HOLT-WINTERS AND SEASONAL REGRESSION MODELS*

UDC 519.246.8

Vinko Lepojević, Marija Anđelković-Pešić

Faculty of Economics, University of Niš, Serbia

Abstract. *Electricity is a key energy source in each country and an important condition for economic development. Reliable forecast of energy consumption represents a starting point in policy development and improvement of production and distribution facilities. The paper predicts the electricity consumption in the area covered by Elektroistok Ltd Nis during the period November 2011 - October 2012. For the purposes of these forecasts, Holt-Winters method and appropriate seasonal regression models were used.*

Key Words: *Time series, Holt-Winters method, seasonal regression models, electricity consummation*

INTRODUCTION

Rapid development of the theory concerning analysis of time series and their practical application began in the second half of the twentieth century. Today there are different theoretical approaches to the analysis of time series, and the impetus for their development lies in the actual problems that require resolution. The main goals of time series analysis are developing models that will describe the regularities in the behaviour of the observed time series and forecasting the future state of the occurrence, based on known state at present and the past.

Stochastic process is considered as linear if it can be expressed through mathematically linear equation in parameter. The most commonly used linear time series models are the AR, MA and ARMA models, and special cases of the general linear models, and ARIMA models used for modelling and forecasting of non-stationary time series, after the process of differentiation. The analysis of linear models is implemented in several steps,

Received December 15, 2011 / Accepted December 21, 2011
Corresponding author: Vinko Lepojević
Faculty of Economics, Trg Kralja Aleksandra 11, 18000 Niš, Serbia
Tel: +381 18 528 658 • E-mail: vinko.lepojevic@eknfak.ni.ac.rs
* The paper has been prepared under the project No. 179066, financed by Ministry of Education and Science of the Republic of Serbia

such as model selection, parameter estimation based on perceptual values of time series, as well as testing the adequacy of the model.

If the observation of time series is shown as a function of frequency, it is said that the analysis is performed in the frequency domain. For the analysis in the frequency domain, spectral density function is used. It shows the contributions of periodic components at different frequencies in the total variance of time series (Kovacic Z. 1998, p.7).

In the analysis in time domain, there are numerous methods which may be used for forecasting. If the pursuit series has clearly highlighted seasonality, it is possible to apply the equivalent Holt-Winters approach for the series with a highlighted seasonal variation, as well as multiple regression modelling analysis for modelling of the series and forecasting future values of the same. The paper demonstrates the application of these methods in order to forecast monthly electricity consumption in the footprint of Electricity "Southeast" Ltd Nis during the period November 2011 - June 2012.

Electricity is one of the most important and used forms of energy and it is widely used for different kind of needs. Nowadays electricity is essential for economic development especially for the industrial sector. Decision makers around the world widely use energy demand forecasting as one of the most important policy tools. One of the decision makers' dilemmas is how to forecast electricity demand. "The forecasting is important for electricity power system planners and demand controllers in ensuring that there would be enough supply of electricity to cope with an increasing demand. Thus, accurate load forecasting can lead to an overall reduction of cost, better budget planning, maintenance scheduling and fuel management " (Sabri, A., Humaira, Abdul L., 2011, p. 201).

According to some research results, there is a high degree of correlation between changes in GDP and changes in electricity consumption (Rakic, 2001). A positive correlation was found, i.e. increase in gross domestic product leads to an increase in electricity demand and this relationship has very high degree of correlation (between 0.95 and 0.99). This conclusion is confirmed by very high degree of elasticity between the growth rates of GDP and electricity consumption. The coefficients of elasticity usually range from 0.85 to 0.95.

1. HOLT-WINTERS FORECASTING MODEL

Unlike decomposition methods, structural models assume that each component (trend, cycle, season and random error) is unobservable stochastic process, in which these components interact with each other given observable time series (Djordjevic, Lepojevic, 2003). The main feature of these models is that during the formation of expression for each component of the "fresh" data has greater weight than the "older" data.

The application of two methods for forecasting electricity demand for the area Elektroistok Ltd. Nis is shown below. The first method that will be used, Holt-Winters method, belongs to the group of *ad hoc* procedures. It represents a forerunner of modern structural models that are later found to have a clear basis in statistical theory, and could be considered special education cases of the general class of structural time series models and ARIMA models.

Based on the data shown in Table 1 and Figure 1, it can be seen that electricity consumption in the area Elektroistok Ltd. Nis has highlighted seasonal character. This will determine future course of the process of analysis and forecasting of electricity consumption.

Table 1. Monthly data of electricity consumption for the territory of Elektroistok Ltd. Nis in the period January 2006 - October 2011

Year	Month	Consum. in MWh	Year	Month	Consum. in MWh	Year	Month	Consum. in MWh
2006	1	390235.8	2008	1	386108.6	2010	1	387137
	2	348096		2	364472.6		2	359273.6
	3	364650.5		3	359922.7		3	361146.9
	4	269358.7		4	324269.1		4	313071
	5	240555		5	296508.8		5	283661
	6	245675		6	277515.1		6	293101
	7	246592.1		7	302571		7	293535.2
	8	250884.3		8	290514		8	301373.5
	9	256778.6		9	307039		9	311881
	10	276384.6		10	327146		10	335094.4
	11	317551.1		11	352421		11	389505.7
	12	356103		12	384618		12	425247.2
2007	1	371028.2	2009	1	423772	2011	1	395286
	2	327965.5		2	362627		2	362661
	3	343622.3		3	368622		3	365439
	4	311044.4		4	308743		4	316152
	5	278597.9		5	284806		5	287601
	6	282964.2		6	281199		6	289552
	7	285160.7		7	294084		7	297649
	8	285683.1		8	287874		8	293240
	9	285536.2		9	285294		9	287128
	10	350492.7		10	336427		10	341803
	11	383250.2		11	348967			
	12	434766.5		12	407084			

Source: Elektroistok Ltd. Nis

Many time series, despite the presence of trends, have seasonality characteristics. Application of a simple exponential smoothing of such series is not adequate, except for very short periods. The following paragraphs will show a method of forecasting through the exponential smoothing which explicitly recognizes the trend and seasonal component in time series. Holt-Winters method is an extension of exponential smoothing method and it includes trend and seasonal fluctuations.

Holt-Winters forecasting method consists of flattening of the exponential component (μ_t), component of the trend (T_t) and seasonal component (I_t). The relationship of these three components can be represented by the following equations:

$$\mu_t = \alpha(X_t / I_{t-s}) + (1-\alpha)(\mu_{t-1} + T_{t-1})$$

$$T_t = \gamma(\mu_{t-1} - \mu_{t-1}) + (1-\gamma)T_{t-1}$$

$$I_t = \delta(X_t / \mu_t) + (1-\delta)I_{t-s}$$

$$\ddot{X}_t(h) = (\mu_t + hT_t)I_{t-s+h}$$

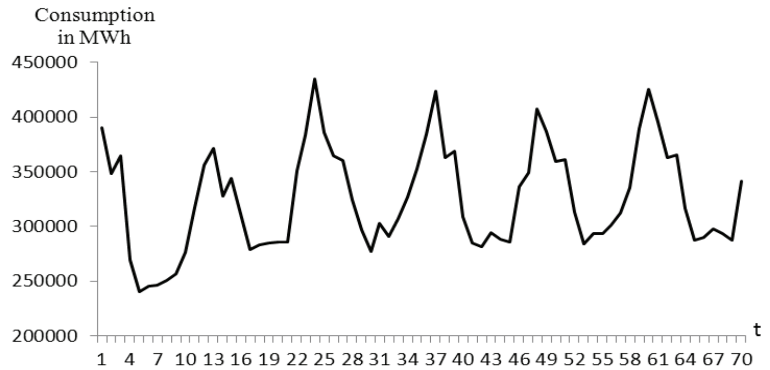

Fig. 1. Monthly electricity consumption in the territory of Elektroistok Ltd. Nis in period January 2006 - October 2011

It is notable that the three equations require constant smoothing, α, γ and δ, each of which takes values between 0 and 1. The constant α controls the smoothing of μ_t, where the choice closer to 0 emphasizes the previous values of time series, while value closer to 1 emphasizes the current values, and lessens the importance of previous ones, and the constant δ modifies the seasonal component. After selecting the constants α, γ and δ, the next step is calculating the components μ_t, T_t and I_t, based on data from time series.

The goal of Holt-Winters smoothing of series is to use the same for forecasting future time series values. The idea is to make forecasts by combining the most recent - the latest, direct assessment μ_t, with an assessment of the expected increase (or decrease), which corresponds to the trend T_t, all adjusted to the seasonal component I_t.

The trend component of the series is determined using the calculated average of the most recent changes in the level and trend assessment from the previous period. By choosing the value of γ closer to 0, a greater emphasis is on the trend of the previous assessment (marked as T_{t-1}), while the choice of γ closer to 1 gives more importance to the existing change of level (marked as $\mu_t - \mu_{t-1}$).

Since monthly data is available, it is necessary to use the first year or the first twelve months, to form the initial assessment of seasonal component. The average monthly consumption of electricity in the first year is 296,905.4MWh. The consumption of each month in the first year is divided by this amount. In this way the seasonal indices are obtained, which are then used to evaluate the average values in January 2007.

The assessment of initial values of trend is provided through assesses of linear trend in the period January 2006 - October 2011. As the assessed value of the linear trend is

$$\ddot{X}_t = 325087,4 + 376.156t ,$$

for the assessment of the initial value of trend in the 13th period (January 2007) coefficient with time equals $T_{13} = 376,156$. A further step is to determine the seasonal component score in the 13th period, so that starting from the 14th period the recursive calculation of the level, trend and season commences. Starting from the seasonal component in the 13 period

$$I_{13} = \delta \frac{X_{13}}{\mu_{13}} + (1-\delta) I_1 ,$$

while for the 14th period assessment of trend and season can be provided as it follows:

$$\mu_{14} = \alpha \frac{X_{14}}{I_2} + (1-\alpha)(\mu_{13} + T_{13}) ,$$

$$T_{14} = \gamma(\mu_{14} - \mu_{13}) + (1-\gamma) T_{13} ,$$

$$I_{14} = \delta \frac{X_{14}}{\mu_{14}} + (1-\delta) I_2 .$$

Table 2. Forecasts of electricity consumption based on Holt-Winters method

Year	Month	HW smoot. values	Year	Month	HW smoot. values	Year	Month	HW smoot. values
2006	1	357328.3		5	296008.1		9	294515.5
	2	323683.5		6	296813.5		10	333079.0
	3	331195.4		7	302066.3		11	363747.0
	4	290700.8		8	303273.4		12	410024.7
	5	260602.4		9	308858.5	2011	1	406113.5
	6	257527.6		10	347419.9		2	366497.5
	7	263929.0		11	379899.6		3	370695.0
	8	262043.0		12	421083.4		4	321542.3
	9	267667.4	2009	1	409355.6		5	291915.9
	10	303202.8		2	374135.1		6	289647.0
	11	333537.2		3	378330.4		7	296651.2
	12	374785.4		4	330343.7		8	295436.7
2007	1	366469.4		5	298907.0		9	300988.1
	2	328308.3		6	294869.4		10	335029.7
	3	332994.1		7	300922.6		**11**	**368340.2**
	4	283892.4		8	298583.5		**12**	**410865.9**
	5	257852.2		9	303760.7	2012	**1**	**402707.9**
	6	259094.4		10	337232.3		**2**	**364532.0**
	7	268431.9		11	369441.1		**3**	**368752.0**
	8	270706.6		12	409017.3		**4**	**319932.0**
	9	279428.2	2010	1	404168.1		**5**	**290814.5**
	10	315859.0		2	362001.1		**6**	**289273.0**
	11	354557.3		3	365855.8		**7**	**296302.2**
	12	401721.5		4	315968.1		**8**	**294607.6**
2008	1	402877.9		5	285933.0		**9**	**299258.1**
	2	364137.6		6	282091.7		**10**	**336610.5**
	3	371880.0		7	290449.2			
	4	323782.5		8	288075.0			

Based on this score in the 14th period the level of the observed phenomena is forecast for one period ahead:

$$\ddot{X}_{14}(1) = (\mu_{14} + T_{14}) I_3 .$$

In forming the score of the three equations, the following values of constants smoothing were used: α = 0.100, γ = 0.100 and δ = 0100. The chosen values of smoothing constants in the previously shown way minimize the mean square error values of forecasts. The choice of constants in Holt-Winters model requires special attention. In the case of a priori choices the values from 0.01 to 0.3 are suggested for all three components (level, trend and season). However, this method of determining the constants of smoothing does not take care of specificities of time series which smoothes out. Therefore, it is recommended to assess all three constants on the basis of time series, using one of statistical criteria. As a criterion, mean square error of forecasts is used mostly. For seasonal models it should be borne in mind that the conditions for stable solutions are often quite complex, especially when it is about monthly time series. Therefore, it is inconceivable to find adequate solutions without the use of statistical software. For the purposes of forecasting electricity consumption, optimal value smoothing constants were determined by searching the space of all possible values of the constants α, γ and δ in such a way as to minimize the selected trio mean squared error forecast.

In order to determine the adaptation of model to empirical data, it is desirable to use the same graph for evaluated and empirical values, as it is done in Figure 2.

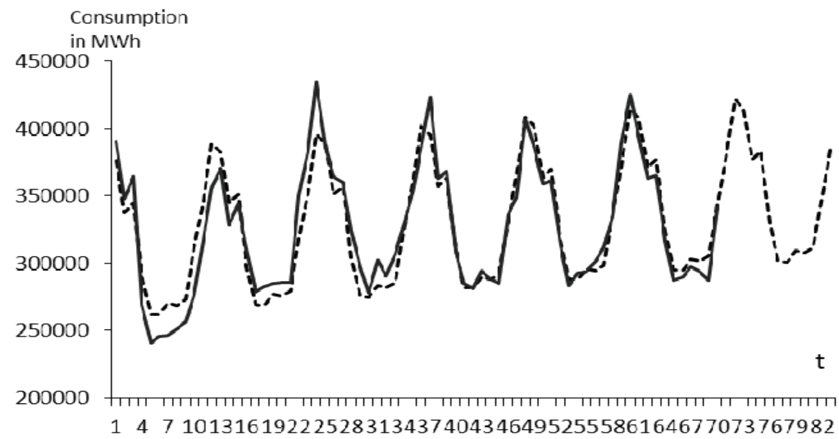

Fig. 2. Evaluated and empirical energy consumption during period January 2006 - October 2012

2. Sesonal Regression Models

Multiple regression models can be used for forecasting future values of time series with a highlighted seasonal component. In order to realize that the mean time series $E(X_t)$ is given in mathematical form that describes the secular trend and seasonal components of time series. Although seasonal models use a wide variety of mathematical forms, the use of dummy variables to display the seasonal difference is a commonly used approach.

The model that combines the expected growth and the seasonal component is given in the following form:

$$E(X_t) = \beta_0 + \beta_1 t + \beta_2 M_1 + \beta_3 M_2 + \ldots + \beta_{12} M_{11}$$

where t represents Time period, ranging from t=1 for January 2006. to t=70 for October 2011.

$$M_1 = \begin{cases} 1 & \text{if month 1} \\ 0 & \text{if month 2,3,...,12} \end{cases}$$

$$M_2 = \begin{cases} 1 & \text{if month 2} \\ 0 & \text{if month 1,3,...,12} \end{cases}$$

$$\ldots$$

$$M_{11} = \begin{cases} 1 & \text{if month 11} \\ 0 & \text{if month 1,2,...,10,12} \end{cases}.$$

Based on empirical data on electricity consumption in the observed area in the period 2006-2011, the appropriate regression model is formulated. Its parameters are given in Table 3. It is evident that the coefficient of multiple determination R^2 equals 0.89, while the adjusted value of this coefficient is 0.87, which indicates that the evaluated model correctly approximates changes in level of given phenomenon, i.e., the model explained 87% of variability of the monthly electricity consumption in the observed five-year period. Another confirmation of the validity and adequacy of the model for forecasting the future consumption of electricity in this area is the value of statistics F = 39.74, and p-value of 0.00.

Table 3. Multiple Regression Results

Dependent: Demand		
Multiple R = .94510458	R^2= .89322266	R^2= .87074322 (adjusted)
F = 39.73510 p = 0.000000	df = 12,57	No. of cases: 70
Standard error of estimate:16957.748894		
Intercept: 382133.85106	Std.Error: 8413.399	t(57) = 45.420 p = 0.0000
t beta=540	M1 beta=-6604	M2 beta=-45222
M3 beta=-39377	M4 beta=-93378	M5 beta=-122402
M6 beta=-123229	M7 beta=-115505	M8 beta=- 117715
M9 beta=-114240	M10 beta=-75831	M11 beta=-42685

In order to forecast consumption for the last two months of this year and first ten months of 2012, it is necessary to determine the assessed value of $\widehat{X}$ for t = 71-82 substituting the appropriate value of the dummy variable for each month. These values are calculated as follows:

$$\widehat{X}_{\text{XI mesec 2011}} = \widehat{\beta}_0 + \widehat{\beta}_1(71) + \widehat{\beta}_{12}$$

$$\widehat{X}_{\text{XII mesec 2011}} = \widehat{\beta}_0 + \widehat{\beta}_1(72)$$

$$\ldots$$

$$\widehat{X}_{\text{X mesec 2012}} = \widehat{\beta}_0 + \widehat{\beta}_1(82) + \widehat{\beta}_{11}.$$

Table 4. Predicted values of electricity consumption based on the seasonal regression

Year	Month	Predicted values	Year	Month	Predicted values	Year	Month	Predicted values
2006	1	376070		5	275392		9	298674
	2	337992		6	275105		10	337623
	3	344377		7	283369		11	371309
	4	290916		8	281699		12	415074
	5	262432		9	285714	2011	1	408470
	6	262145		10	324663		2	370392
	7	270409		11	358349		3	376777
	8	268739		12	402114		4	323316
	9	272754	2009	1	395510		5	294832
	10	311703		2	357432		6	294545
	11	345389		3	363817		7	302809
	12	389154		4	310356		8	301139
2007	1	382550		5	281872		9	305154
	2	344472		6	281585		10	344103
	3	350857		7	289849		**11**	**377789**
	4	297396		8	288179		**12**	**421554**
	5	268912		9	292194	2012	**1**	**414950**
	6	268625		10	331143		**2**	**376872**
	7	276889		11	364829		**3**	**383257**
	8	275219		12	408594		**4**	**329796**
	9	279234	2010	1	401990		**5**	**301312**
	10	318183		2	363912		**6**	**301025**
	11	351869		3	370297		**7**	**309289**
	12	395634		4	316836		**8**	**307619**
2008	1	389030		5	288352		**9**	**311634**
	2	350952		6	288065		**10**	**350583**
	3	357337		7	296329			
	4	303876		8	294659			

Seasonal model used for forecasting electricity production is an additive model, because the trend component ($\beta_1 t$) is added to the seasonal component ($\beta_2 M_1 + \beta_3 M_2 + \ldots + \beta_{12} M_{11}$) in order to formulate the model. Multiplicative model might have the same expression, except that the dependent variable becomes the logarithm of the real consumption i.e.:

$$\ln X_t = \beta_0 + \beta_1 t + \beta_2 M_1 + \beta_3 M_2 + \ldots + \beta_{12} M_{11} + \varepsilon_t$$

For both sides of the equation antilogarithm has to be calculated in order to provide multiplicative expression of this model as it is shown below:

$$\begin{aligned} X_t &= \exp\{\beta_0 + \beta_1 t + \beta_2 M_1 + \beta_3 M_2 + \ldots + \beta_{12} M_{11} + \varepsilon_t\} \\ &= \exp\{\beta_0\}\exp\{\beta_1 t\}\exp\{\beta_2 M_1 + \beta_3 M_2 + \ldots + \beta_{12} M_{11}\}\exp\{\varepsilon_t\} \end{aligned}$$

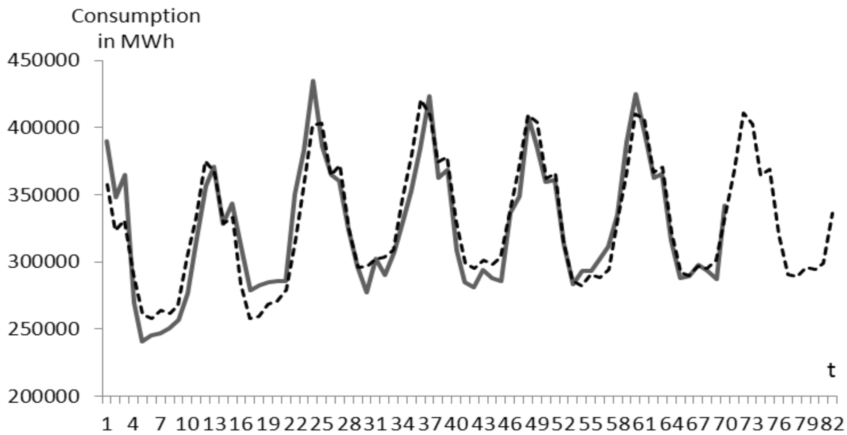

Fig. 3. Realized and assessed energy consumption during January 2006 - October 2012

This kind of model - multiplicative model, often provides better predictions when the time series change over time by increasing rate. Therefore, they are suggested when it is about forecasting the demand at the growing markets.

3. Comparative Analysis

Predicted values of monthly electricity consumption in the period November 2011 - October 2012 the area covered by Elektroistok Ltd. Nis, differ with respect to the forecasting method used. In order to provide that forecasts are as close as possible to actual values, the choice of the method should be based on charts, which show the empirical values of the observed phenomena and the corresponding flattening values. According to Figure 2 and Figure 3, there is no significant difference between the assessed value of electricity consumption and the corresponding empirical values, which confirms good adaptation of selected models to empirical data as well as validity of performed predictions. Therefore it can be said that it is hard to choose among the approaches based on the charts. For this reason, the other criteria for selection of appropriate methods of adjustment and prognosis are used.

Hence, it is possible to use the smoothing error for the sake of evaluating the accuracy of forecasts, but also as criteria in the selection of forecasts methods to be used. There are two popular measures for determining the accuracy of forecasts - the mean absolute deviation (MAD) and the root mean squared error (RMSE). Their formulas are given as:

$$MAD = \frac{\sum_{t=1}^{N} | \ddot{X}_t - X_t |}{N}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N} (\widehat{X}_t - X_t)^2}{N}}$$

Criteria for the analysis of the accuracy of forecasts, such as MAD and RMSE, require special attention in the interpretation. Number of time periods included in the evaluation is critical regarding the decision which model is better. The choice also depends on how long the period for which the predictions should be made is. Once the decision is made, the criteria can be used to measure the accuracy of forecasts.

For comparison of the two models discussed above, mean absolute deviation is calculated first:

$$MAD(HW) = 13288{,}46$$
$$MAD(SR) = 12219{,}44$$

The calculation of the root mean squared errors between smoothed and empirical data is calculated in the next step:

$$RMSE(HW) = 16542{,}54$$
$$RMSE(SR) = 15303{,}01$$

It is notable that both criteria lead to the same conclusion. Forecasts which are made based on seasonal regression models show lower absolute values of the average deviation between the empirical values smoothed and lower mid-squared value of deviations from the predictions which were obtained using the Holt-Winters methods. Consequently, seasonal forecasts obtained through regression model are better, and therefore the same is recommended for forecasting electricity consumption in a given area.

Finally, by using these relatively simple methods for forecasting electricity consumption in the observed area, satisfactory prognosis can be provided very fast. Used measures of dispersion, MAD and RMSE, are a good criterion for making decision about the model which will be used for prediction.

As values assessed through seasonal regression model are better adapted to empirical data, measured by value of MAD and RMSE, the forecast is considered more reliable. So, in the next twelve months a similar pattern in electricity consumption in the observed area is expected, which was one of the assumptions in formulating the model. In addition to that, stability of spending is expected to be at a higher level than in the previous five years. One should bear in mind a well-known fact that is valid for forecasting: as the period of forecasting is more distant from the baseline period, by which forecast is made, the reliability of forecasts declines.

CONCLUSION

Quality and adequacy of prognosis depend on the quality of data analysis. Based on these forecasts (and other knowledge) companies manage the production, trade and other activities, while economic policy makers take certain activities. In the last two decades there has been significant progress in the mathematical and statistical analysis of stochastic processes. Their usage in the analysis of time series (especially in the area of macroeconomic and financial series) proceeded from this fact. These models are trying to connect knowledge from the fields of mathematics and statistics with economic theory. They open up new opportunities in modeling and forecasting of economic time series.

There are many approaches to modeling and forecasting economic time series. The old methods are sophisticated, but the introduction of new methods, is expected to provide more reliable forecasts. The success of the applied forecasting method depends partly on the properties of time series. With proper choice of the method, taking into account the assumptions of its application, it is possible to provide high-quality forecasts.

Predicted values of electricity consumption for the next twelve months indicate that there is a positive trend in demand for this phenomenon. The forecasts are significant since they provide information for entities engaged in the production and distribution of electricity, so that they can adapt production and distribution capacities, especially in the months when consumption reaches a maximum value.

REFERENCES

1. Brockwell, D.(2002) *Introduction to Time Series and Forecasting*, Springer-Verlag, New York
2. Djordjevic V., Lepojevic V., (2003) *Henderson's Approach to Variance Components Estimation for Unbalanced Data*, Facta Universitatis, Vol.2, No 1, 2003 pp. 59-64.
3. Kirchgassner, G., Wolters, J. (2007) *Introduction to Modern Time Series Analysis*, Berlin, Springer
4. Kovačić, Z. (1998) *Time Serias Analysis*, Faculty of Economics, Belgrade
5. McClave, J. (1988) *Statistics for Business and Economics*, San Francisco, Dellen Publishing Company
6. Mladenović, Z., Nojković, A. (2008) *Time Serias Analysis: Examples from Serbian Practice*, Faculty of Economics, Belgrade
7. Rakic B., (2001) *Restructuring of the Power Sector in Serbia*, Facta Universitatis, Vol.1, No 9, 2001 pp. 65-72.
8. Sabri, A., Humaira' Abdul L. (2011) Competencies Analysis Of Box-Jenkins Method In Forecasting Electricity Demand, *Umtas 2011, 201-2004*
9. Sarah G., Roland F., Christophe C. (2000) Robust Forecasting with Exponential and Holt-Winters Smoothing, *Business*, 1-21
10. Taylor, J. (2003) Short-Term Electricity Demand Forecasting Using Double Seasonal Exponential Smoothing. *Journal of Operational Research Society*, 54, 799-805.
11. Taylor, J. (2003) Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of Operational Research Society*, 54:799–805
12. Yaffee, R., McGee, M. (2000) Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS. New York: Academic Press.

PROGNOZIRANJE POTROŠNJE ELEKTRIČNE ENERGIJE KORIŠĆENJEM HOLT-WINTERS I SEZONSKOG REGRESIONOG MODELA

Vinko Lepojević, Marija Anđelković-Pešić

Električna energija predstavlja ključni energent svake zemlje i bitan uslov privrednog razvoja. Pouzdane prognoze potrošnje ovog energenta su polazna osnova u kreiranju politike razvoja i unapredjenja proizvodnih i distributivnih kapaciteta. U radu se prognozira potrošnje električne energije na području koje pokriva Elektroistok Niš d.o.o. u periodu novembar 2011 - oktobar 2012. Za potrebe ove prognoze korišćen je Holt-Wintersov metod izravnanja i odgovarajući sezonski regresioni modeli.

Ključne reči: *vremenske serije, Holt-Wintersov metod, sezonski regresioni modeli, potrošnja električne energije*