UNIVERSITY OF NIŠ
The scientific journal **FACTA UNIVERSITATIS**
Series: **Architecture and Civil Engineering** Vol.1, N° 5, 1998 pp. 617 - 625
Editors of series: Dragan Veličković, Dušan Ilić, e-mail: facta@ni.ac.yu
Address: Univerzitetski trg 2, 18000 Niš, YU
Tel: +381 18 547-095, Fax: +381 18 547-950

COMPARISON OF PREDICTION MODELS OF REFERENCE CROP EVAPOTRANSPIRATION

UDC 626.85(045)

Slaviša Trajković

Faculty of Civil Engineering and Architecture, University of Niš
Beogradska 14, 18000 Niš, Yugoslavia

Abstract. *The prediction of evapotranspiration is necessary for a reliable management of irrigation systems. This paper is based on several models used for the prediction of reference crop evapotranspiration in the area of Niš, Yugoslavia. Two simple mathematical models (Yearly Differencing model and Monthly AVerage model), in which prediction is based on appropriate previous values of reference crop evapotranspiration, are present here. A seasonal ARIMA (AutoRegressive Integrated Moving Average) model is identified and compared to the previous models. Based on the results of model comparisons it is concluded that seasonal ARIMA can provide a reasonably accurate prediction of reference crop evapotranspiration. The obtained results show that the seasonal ARIMA model is a very effective and reliable prediction model.*

Key words: *Prediction, Evapotranspiration, ARIMA model*

1. Introduction

An important factor of each water resource study is a successful prediction of climatological variables which have the influence upon the behavior of the water system. This is particularly important for the agricultural areas where the evapotranspiration prediction guarantees a reliable project planning, design and operating of an irrigation systems.

A crop's water demands are directly connected with crop evapotranspiration, ET, and they vary depending on the crop growth. The prediction of reference crop evapotranspiration (ET_o) is usually done first and after that the value of crop evapotranspiration (ET) is obtained from the following relation:

$$ET = k_c \cdot ET_o \tag{1}$$

Received November 15, 1999

where k_c = the crop coefficient. Reference crop evapotranspiration is the rate of evapotranspiration from an extended surface of 8 to 15 cm tall, green cover of uniform height, actively growing, completely shading the ground and not short of water [4]. Doorenbos and Pruitt (1977) suggested four methods for estimating reference crop evapotranspiration using the connections between the physical parameters which are involved in the process of crop evapotranspiration. They are: Blaney-Criddle method, Penman method, pan method and radiation method. The researches which have been made in Yugoslavia show that the best results were obtained by Blaney-Criddle method [12].

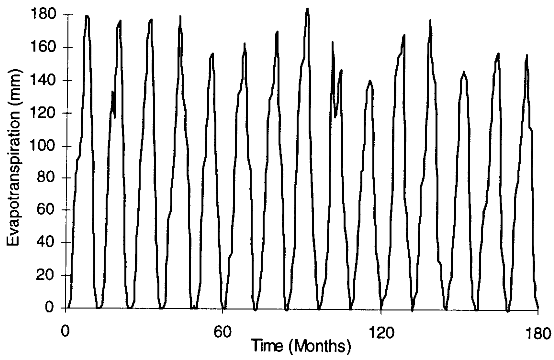

Fig. 1. Reference crop evapotranspiration in Niš from 1961 to 1975

The values of meteorological parameters such as air temperature, relative humidity, wind speed and sunshine were obtained from the HydroMeteorological Station (HMS) in Niš from 1961 to 1975. The values of reference crop evapotranspiration for this period were produced by Blaney-Criddle method (Fig. 1). The procedure which was described in the paper [9] could not be used because the HMS Niš does not have either lysimeter or the Class A pan.

The data (the monthly values of reference crop evapotranspiration) were divided into two groups. The data from 1961 to 1970 were used for calibrating the prediction models which were suggested. The data from 1971 to 1975 were used for verifying the prediction results obtained by suggested models.

2. Prediction Models

2.1. Yearly Differencing model

Yearly Differencing (YD) model can be described mathematically as:

$$ET_{o,t} = ET_{o,t-12} + e_t \tag{2}$$

where $ET_{o,t}$ = reference crop evapotranspiration during the month t; $ET_{o,t-12}$ = reference crop evapotranspiration during the month $t-12$; and e_t = prediction error in month t. This simple model predicts the reference crop evapotranspiration in such a way that it is equated with the reference crop evapotranspiration during the same month of the previous year. The application of this model is even more important when there is a limited number of information about the previous values of reference crop evapotranspiration. YD model does not need either calibration or parameters estimation, and it is very simple to use. The successful application of this model is obtained through the domination of seasonal component of evapotranspiration time series.

2.2. Monthly Average model

Monthly Average (MAV) model is described mathematically as :

$$ET_{o,t} = \frac{1}{N}\sum_{i=1}^{N} ET_{o,t-12\cdot i} + e_t \tag{3}$$

where N = the number of years of observation. Monthly values of reference crop evapotranspiration is obtained as an average value of reference crop evapotranspiration during the previous years in the corresponding month. This method decreases the influence that extreme years have on the prediction by averaging them in with the "usual" years. MAV model needs a small calibration, and its parameters are easily calculated from the data set.

2.3. AutoRegressive Integrated Moving Average (ARIMA) model

The Box-Jenkins method is one of the most widely used time series prediction methods in practice [2], [5], [7]. The method uses a systematic procedure to select an appropriate model from a rich family of models (ARIMA models).

AutoRegressive (AR) models estimate values for the dependent variable X_t as a regression function of previous values $X_{t-1}, ..., X_{t-p}$ plus some random error et. Moving Average (MA) models give a series value X_t as a linear combination of some finite past random errors, $e_{t-1}, ..., e_{t-q}$. p and q are referred as orders of the models. AR(p) and MA(q) models can be combined to form an ARMA(p,q) model. This model can provide additional flexibility in describing of the time series.

However, a large number of time series is nonstationary and for modelling such time series, simple AR, MA or ARMA models are not appropriate. Box and Jenkins (1976) [1] suggested that a nonstationary series can be transformed into a stationary one by differencing. The ARMA models applied to the differenced series are called integrated models, denoted by ARIMA (AutoRegressive Integrated Moving Average) models.

A time series involving seasonal data will have relations at a specific lag s which depends on the nature of the data, e.g. for monthly data s = 12. Such series can be successfuly modeled only if the model includes the connections with the seasonal lag as well.

The general multiplacitive seasonal ARIMA (p,d,q)(P,D,Q)s model has the following form:

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D x_t = c + \theta_q(B)\Theta_Q(B^s)e_t \tag{4}$$

where C = constant; B = a backshift operator; d = order of nonseasonal difference operator; D = order of the seasonal difference operator; p = order of nonseasonal AR operator; P = order of seasonal AR operator; q = order of nonseasonal MA operator; and Q =order of seasonal MA operator.

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p \tag{5}$$

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - ... - \phi_P B^{Ps} \tag{6}$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q \tag{7}$$

$$\Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - ... - \Theta_Q B^{Qs} \tag{8}$$

The conditions of stationarity and invertibility are met only if all the roots of the characteristics equation $\phi_p(B)=0, \Phi_P(B^s)=0, \theta_q(B)=0, \Theta_Q(B)=0$ lie outside the unit circle.

The Box-Jenkins method performs prediction through the following process:

1. Model Identification: The orders of the model are determined.
2. Model Estimation: The linear model coefficients are estimated.
3. Model validation: Certain diagnostic methods are used to test the suitability of the estimated model.
4. Forecasting: The best model chosen is used for forecasting.

One of the basic conditions for applying the ARIMA model on a particular time series is its stationarity. A time series with seasonal variation may be considered stationary if the theoretical autocorerelation function (ρ_k) and theoretical partial autocorrelation function (ρ_{kk}) are zero after a lag $k = 2s + 2$. It is considered that ρ_k and ρ_{kk} equal zero if [8]:

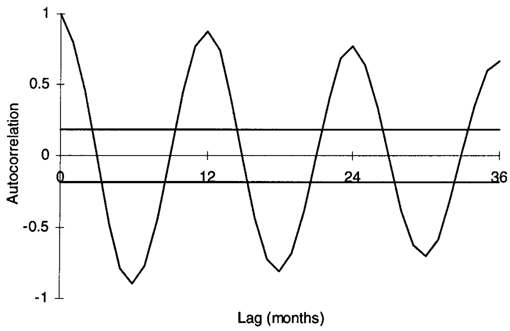

Fig. 2. ACF for evapotranspiration in Niš

$$\rho_k = 0 \text{ ako} |r_k| \le 2/(T)^{0.5} \quad (9)$$

$$\rho_{kk} = 0 \text{ ako} |r_{kk}| \le 2/(T)^{0.5} \quad (10)$$

where r_k = sample autocorrealtion at lag k; r_{kk} = sample partial autocorrelation at lag k; and T = number of observations.

The sample autocorrelation function (ACF) of analysed series do not meet the above condition already mentioned decreasing extremly slowly in a sinusoidal fashion (Fig. 2).

That is why the time series is being transformed into a stationary one using differencing ($d = 0$, $D = 1$, $s = 12$) according to the following expression:

$$y_t = (1-B)^d (1-B^s)^D x_t = (1-B^{12}) ET_{o,t} \quad (11)$$

Procedure was repeated for the new time series y_t. The ACF of the transformed time series y_t goes beyond the limits included in the expression (9) and (10) only with the lag of 12 and the PACF (sample Partial AutoCorrelation Function) with the lag of 12 and 24 months (Fig. 3). The time series y_t is stationary because the ACF and PACF cut off at lags less than 26 months ($2s + 2 = 26$).

On basis of information obtained from the ACF and PACF, several forms of the ARIMA model were identified tentatively. The parameters of model were calculated by maximum likelihood estimation. The data from 1961 to 1970 were used for estimating the unknown parameters.

Once a model has been selected and parameters calculated, the adequacy of the model has to be checked. This process is also called diagnostic checking. Box-Pierce method is used for this purpose, as well as the Portmanteau lack-of-fit test and t-statistics.

The Box-Pierce method is based on the calculation of ACF residuals. If the model is adequate at describing the behavior of the time series, the residuals are not correlated, i.e. all ACF values lie within the the limits included in the expression (9) and (10). The Portmanteau lack-of-fit test investigates the first m ACF values of the residuals using

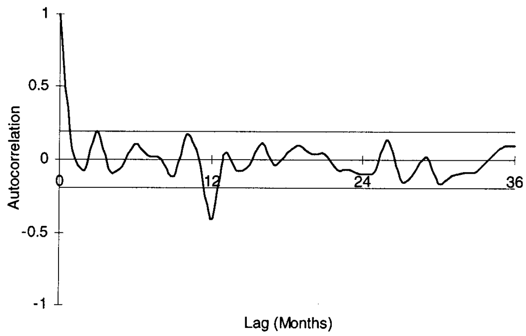

Fig. 3(A). ACF for transformed series y_t

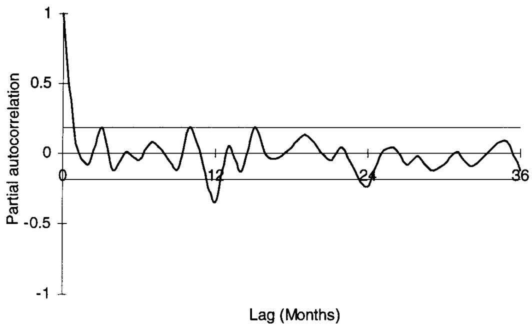

Fig. 3(B). PACF for transformed time series y_t

Box-Pierce chi-square statistics which is given in the following expression :

$$Q = (T - d)\sum_{j=1}^{m} r_j^2 \qquad (12)$$

where $m =$ the number of residual autocorrelation used in the estimation of Q ($m = 3s$ [8]); r_j = autocorrelation at lag j. ARIMA model is adequate if $Q < \chi^2_{0.5}$ $(m - n_p)$ where $n_p =$ the number of model parameters. The third test of model adequacy is the examination of standard errors of the model parameters. A high standard error in comparison with the parameter values points out a higher uncertainty in parameter estimation which questions the stability of the model. The model is adequate if it meets the following condition:

$$t = cv / se > 2 \qquad (13)$$

where cv = parameter value; and se = standard error.

If several tentative models pass the diagnostic checking, AIC (Akaike Information Criteria) or BIC (Bayes Information Criteria) is applied to select the best model. Two ARIMA models were identified for detailed evaluation and the results of model estimation and verification are presented in Table 1.

Table 1. Results of Model Estimation and Verification

	Model 1 (0,0,0)(0,1,1)12	Model 2 (0,0,0)(1,1,1)12	
	SMA(1)	SAR(1)	SMA(1)
Parametar value	0.6785	-0.2143	0.6169
Std Error	0.0785	0.1333	0.1144
t-ratio	8.6401	-1.6022	5.3940
Q-value	27.5	31.8	
AIC	6.529	6.562	

The results of model verification suggest that model 1 is more appropriate and parsimonious than model 2. Model 1 has less AIC and Q statistics. Also, t statistics of seasonal AR component of model 2 don't meet the necessary condition ($t > 2$). As can be seen in Table 2, model 1 provides slightly better forecasts than model 2 from period from 1970 to 1975 (out of sample data set). In this table MSE is the mean square error of the verifying period, MXE denotes the maximum absolute error (the difference between the observed and the predicted values), MAE is the mean absolute error, and MPEVII denotes the mean percent error for July.

Table 2. Out of sample error statistics for ARIMA models

ARIMA	MSE (mm^2)	MXE (mm)	MPEVII (%)	MAE (mm)
Model 1	219.2	36.2	0.3	11.5
Model 2	222.6	40.3	0.3	11.3

Model 1 is given in this form :

$$ET_{o,t} = ET_{o,t-12} - 0.6785 e_{t-12} + e_t \tag{14}$$

VAR = 219.24; AIC = 6.53; t = 8.64

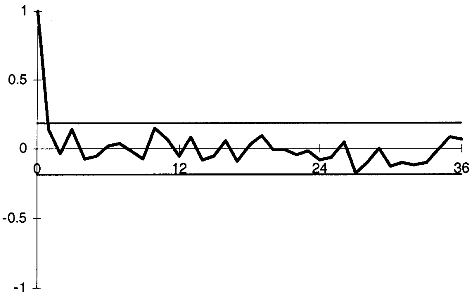

Fig. 4. ACF of residuals

where VAR = residual variance. The ARIMA model which was obtained in this paper, with the help of the ASTSA computer package, has the similar structure as the model used in the [6], with a slightly smaller value of the AIC and VAR statistics.

The ACF values of residuals are given in Fig. 4. All the values can be considered to be negligible. The Q statistics at lag 36 also show that the residuals are not correlated ($Q = 27.5 < 49.5 = \chi^2_{0.5}$), and it is concluded that the time series et is the white noise. Value of t-statistics ($t = 8.64$) show that the model is stable.

3. Comparison of Models

The comparison of the observed and predicted values of the reference crop evapotranspiration is shown in Fig. 5 and Table 3. The predicted values were made by YD, MAV and ARIMA models.

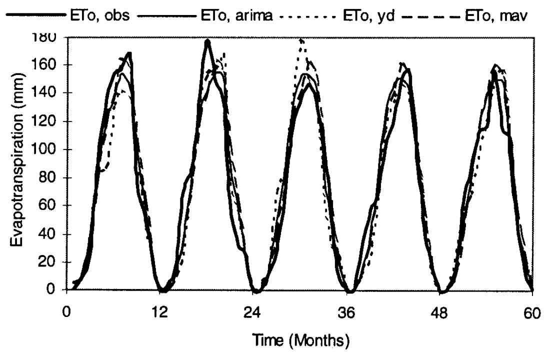

Fig. 5. Comparison of predicted to actual evapotranspiration from 1971 to 1975

Table 3. Observed and predicted ET (mm)

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
						1971						
ETobs	5.6	9.0	25.4	90.3	136.1	154.2	157.5	167.4	76.8	51.8	22.8	-2.5
ETarima	-5.3	13.3	39.7	83.4	122.3	135.3	153.9	144.0	99.8	57.8	24.5	-2.8
ETyd	-0.6	12.3	45.3	83.4	88.0	131.1	141.1	135.2	97.2	45.0	23.1	0.6
ETmav	-5.8	10.7	38.5	80.5	121.1	140.0	164.4	153.5	105.1	63.5	23.4	-3.3
						1972						
ETobs	4.3	16.8	70.4	88.2	138.3	177.6	147.3	134.9	64.2	32.6	25.5	2.5
ETarima	-1.8	11.9	35.1	85.6	126.7	141.4	155.0	151.5	92.4	55.8	23.9	-2.7
ETyd	5.6	9.0	25.4	90.3	136.1	154.2	157.5	167.4	76.8	51.8	22.8	-2.5
ETmav	-4.8	10.6	37.3	81.4	122.5	141.3	163.8	154.8	102.5	62.4	23.4	-3.2
						1973						
ETobs	-0.6	11.5	24.5	68.1	114.7	137.4	146.3	136.7	100.2	56.4	12.3	-0.9
ETarima	0.2	13.5	46.5	86.5	130.4	153.0	152.5	146.2	83.3	48.4	24.4	-1.0
ETyd	4.3	16.8	70.4	88.2	138.3	177.6	147.3	134.9	64.2	32.6	25.5	2.5
ETmav	-4.0	11.1	40.1	82.0	123.8	144.3	162.4	153.1	99.3	59.9	23.6	-2.8
						1974						
ETobs	0.0	27.7	56.7	69.0	100.4	122.7	149.7	155.9	100.2	46.5	21.9	-3.1
ETarima	-0.1	12.8	39.4	80.6	125.4	148.0	150.5	143.1	88.8	50.9	20.5	-1.0
ETyd	-0.6	11.5	24.5	68.1	114.7	137.4	146.3	136.7	100.2	56.4	12.3	-0.9
ETmav	-3.7	11.1	38.9	80.9	123.1	143.8	161.1	151.8	99.4	59.6	22.7	-2.6
						1975						
ETobs	0.3	12.0	62.3	85.8	113.2	117.0	156.6	113.8	108.0	52.7	12.3	1.6
ETarima	-0.1	17.6	45.0	76.8	117.4	139.9	150.3	147.2	92.4	49.5	21.0	-1.7
ETyd	0.0	27.7	56.7	69.0	100.4	122.7	149.7	155.9	100.2	46.5	21.9	-3.1
ETmav	-3.5	12.3	40.2	80.0	121.5	142.3	160.3	152.1	99.4	58.7	22.6	-2.7

The differences between the observed and the predicted values from 1971 to 1975 (verifying period) are given in Fig. 6.

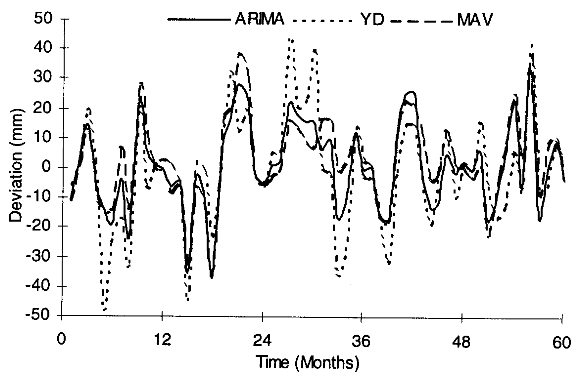

Fig. 6. Differences between observed and predicted ET from 1971 to 1975

Table 4 summarizes the predicting error statistics for each of the three models.

Table 4. Error statistics for verifying period

model	MSE (mm2)	MXE (mm)	MPEVII (%)	MAE (mm)
YD	355.6	48.1	0.5	13.8
MAV	235.6	38.4	0.7	11.8
ARIMA	219.2	36.2	0.3	11.5

The best results were obtained using the ARIMA model. The MSE for this model is 219.2 mm^2, the maximum absolute error is 36.2 mm and the mean absolute error is 11.5 mm. The mean percent error for July is 0.3%. The smallest statistics values are obtained using ARIMA model, which emphasizes the best agreement between the observed and the predicted evapotranspiration values. The MAV model provides better forecasts as compared to the YD model.

4. CONCLUSION

The prediction of evapotranspiration guarantees reliable project planning, design and operating of irrigation systems. This paper presents the use of YD, MAV and ARIMA models for predicting reference crop evapotranspiration in the area of Niš, Yugoslavia. The YD and MAV models base their prediction on the appropriate previous values of reference crop evapotranspiration. The application of these models is successful because of the domination of seasonal component in the evapotranspiration time series. Beside these models, a seasonal ARIMA model was also presented. According to the comparison results it can be said that seasonal ARIMA models guarantee the most reliable prediction of reference crop evapotranspiration, and that they are superior when compared to the YD and MAV models.

The author of this paper plans to develop a new model for predicting ETo, based on the artificial neural networks, which have already been successfully used for predicting climatological time series [3], [10], [11].

Acknowledgments. *Acknowledgment is given to Miguel A. Marino, Professor at University of California, Davis, for pioneering work in this area and for stimulating the work in the present study. The writer is grateful to Robert H. Shumway, Professor at University of California, Davis, for allowing him to use the ASTSA computer package.*

REFERENCES

1. Box, G. E. P. and Jenkins, G. M., *Time Series Analysis, Forecasting and Control*, Holden-Day Inc., San Francisko, California, USA, 553p., 1976.
2. Chang, J. T., Stochastic Forecast of Water Losses, J.Irrig. and Drain.Engrg., ASCE, 114(3), pp. 547-558, 1988.
3. Cook, D. F. and Wolfe, M. L., *A Back-Propagation Neural Network to Predict Average Air Temperature*, AI Applications 5(1), pp. 40-46, 1991.
4. Doorenbos, J., and Pruitt, W. O., *Crop water requirments*, FAO irrigation and drainage paper No. 24, 2nd Ed., Food and Agricultural Organization of UN, Rome, Italy, 1977.

5. Elango, L., Ramachandran, S., and Sreenivasan, M., *Time series analysis of ground-water quality parameters*, Proc., Conf. Hydroinformatics'96, Zurich, pp. 613-620, 1996.
6. Hameed, T. , Marino, M.A., and Shumway, R.H., *Evapotranspiration Transfer-Function-Noise Modeling*, J.Irrig. and Drain.Engrg., ASCE, 121(2), pp. 159-169, 1995.
7. Irvine, K. N. and Eberhardt, A.J., *Multiplicative, Seasonal ARIMA Models for Lake Erie and Lake Ontario Water Levels*, Water Resources Bulletin, AWRA, 28(2), pp. 385-396, 1992.
8. Maier, H. R., and Dandy, G. C., *Comparison of the Box-Jenkins procedure with artificial neural network methods for univariate time series modelling*, Resarch Report No. R 127, Department of Civil and Environmental Engineering, University of Adelaide, Australia, 102p., 1995.
9. Marino, M.A., Tracy, J.C., and Taghavi, S. A., *Forecasting of reference crop evapotranspi ration*, Agricultural Water Management, 24, pp.163-187, 1993.
10. Miyano, T. and Girosi, F., *Forecasting Global Temperature Variations by Neural Networks*, A.I. Memo No. 1447 & C.B.C.L. No. 101, Massachusetts Institute of Technology, 10 p. , 1994.
11. Schizas, C. N., Pattichis, C. S., and Michelides, S. C., *Forecasting Minimum Temperature with Short Time-Length Data Using Artificial Neural Networks*, Neural Network World 4(2), pp. 219- 230, 1994.
12. Skoric, M. and Beric, M, Selection of appropriate method for determing of reference evapotranspiration at location "Rimski Sancevi" (in Serbian), Water Resource 26(147-149), pp. 71-76, 1994.

UPOREDJIVANJE MODELA PREDIKCIJE REFERENTNE EVAPOTRANSPIRACIJE

Slaviša Trajković

Pouzdano upravljanje irigacionim sistemima se ne može zamisliti bez predikcije evapotranspiracije. U radu se vrši uporedjivanje nekoliko modela predikcije referentne evapotranspiracije. Dva jednostavna matematička modela kod kojih se predikcija zasniva na odgovarajućim prethodnim vrednostima su prikazana u ovom radu. Sezonski ARIMA model je identifikovan i uporedjen sa prethodnim modelima. Na osnovu rezultata uporedjivanja može se zaključiti da sezonski ARIMA model može obezbediti najpouzdaniju predikciju referentne evapotranspiracije.

Ključne reči: *Predikcija, Evapotranspiracija, ARIMA model*