Gaziosmanpaşa Üniversitesi Ziraat Fakültesi Dergisi
Journal of Agricultural Faculty of Gaziosmanpasa University
http://ziraatdergi.gop.edu.tr/

Araştırma Makalesi/*Research Article*

JAFAG
ISSN: 1300-2910
E-ISSN: 2147-8848
(2020) 37 (1), 17-22
doi:10.13002/jafag4680

Forecasting Reference Evapotranspiration by ARIMA Approach

Kadri YÜREKLİ[1*]
Tokat Gaziosmanpasa University, Agriculture Faculty, Department of Biosystem Engineering, Tokat
(orcid.org/0000-0003-4938-663X)
*e-mail: kadriyurekli@yahoo.com

Alındığı tarih (Received): 25.02.2020 Kabul tarihi (Accepted): 30.03.2020
Online Baskı tarihi (Printed Online): 15.04.2020 Yazılı baskı tarihi (Printed): 30.04.2019

Abstract: The effort in the study is on forecasting reference evapotranspiration (ETo) by the SARIMA models. This was accomplished by using monthly ETo dataset obtained based on some parameters from meteorology station of Konya province. Stationary of the ETo dataset was fulfilled with differencing. Three SARIMA models that provides all conditions for the stationary data set were identified. The models have very close parsimony values depending on AIC and SBC criteria.

Key Words: Reference Evapotranspiration, ARIMA approach, AIC, SBC.

Referans Evapotrasprasyonun Tahmin Edilmesinde ARIMA Yaklaşımı

Öz: Çalışmadaki amaç SARIMA modellerle referans evapotranspırasyonun (ETo) tahmin edilmesi üzerinedir. Bu Konya meteoroloji istasyonunun bazı parametrelerine bağlı olarak elde edilen aylık ETo veri seti kullanılarak gerçekleştirilmiştir. ETo veri setinin durağanlığı fark alınarak sağlanmıştır. Durağan seri için bütün koşulları yerine getiren üç SARIMA modeli tanımlanmıştır. Bu modeller AIC ve SBC kriterlerine göre çok yakın tutumluluk (parsimony) değerlerine sahip olmuşlardır.

Anahtar Kelimeler: Referans Evapotranspirasyon, ARIMA Yaklaşımı, AIC, SBC.

1. Introduction

Optimal use of fresh water resources, which are limited in our country, has gained more importance with global warming. In this sense, the sector that will attract the most attention is the agriculture. Estimating the water required in agriculture to reflect the real conditions becomes very important. This is possible with the correct estimation of reference evapotranspiration (ETo), which is one of the most important parameters in the planning and operation of water resources (Singh and Jaiswal 2006). Since it is not always possible to obtain the climatic parameters that are essential in the estimation of the ETo, in such conditions, the estimation of the future periods can be performed with simulation techniques depending on the values of the ETo calculated for the past years. Allen et al. (1998) reported that the ETo is affected only by climate parameters.

The ETo estimates were performed with the autoregressive integrated moving average (ARIMA), the most popular time series model. Some of these studies are as follows; Hamdi et al. (2008) in Jordan, Mossad and Alazba (2016) in Saudi Arabia, Trajković (1998) in Yugoslavia, Landeras et al. (2009) in Spain, Gautam and Sinha (2016) in India, Patra (2018) in India and Dwivedi and Shrivastava (2019) in India.

The aims in the study are : a) to determine the performance of the ARIMA model in the prediction of the ETo. In this context, firstly, the monthly ETo values were calculated using the climate parameters from Konya meteorology station in accordance with the FAO Penman-Monteith relationship, b) to detect the presence of the trend in the ETo data c) to apply ARIMA model to the monthly ETo data sequences, d) to perform statistical analysis of the selected model.

2. Material and Method

In this study, climate parameters measured at the meteorological station in Konya, province in the Central Anatolia Region were used as materials. Taking into account the climate parameters measured during the observation period (from 1975 to 2006), according to the FAO Penman-Monteith relationship given by Allen et al. (1998), reference plant water consumption (ETo) or in other words, reference evapotranspiration values were calculated to use as a material. The time curve of the monthly ETo is the Figure 1.

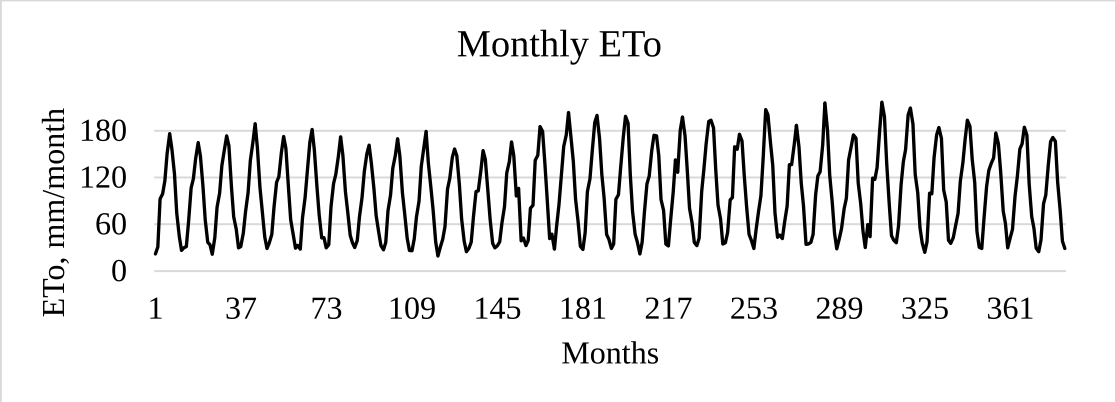

Figure 1. The plot of Monthly ETo values
Şekil 1. Aylık ETo değerlerinin değişim grafiği

Konya is located between 36º 41 'and 39º16' north latitudes and 31º41 'and 34º26' east longitudes. The average annual precipitation is 321 mm in Konya, which has an average height of 1016 meters from the sea. For this reason, dry agriculture is dominant in the region. The typical steppe climate, which is hot and dry in summers and snowy in winters, is predominant in the study area (Tapur 2008).

2.1. Time Series Analysis of Monthly Data

Many hydrologic time series are with a non-stationarity characteristic owing to feature cyclic and stochastic in hydrologic phenomena (Tao and Delleur 1976). The mathematical relationship given below is used, while applying the SARIMA model of order (p,d,q)(P,D,Q), also called Box-Jenkins, which has a linear characteristic, to the monthly hydro-meteorological data sequences (Janacek and Swift 1993).

$$\emptyset(B)\Phi(B^s)(w_i - \mu) = \theta(B)\Theta(B^s)a_i \quad (1)$$

$$w_i = (1\text{-}B)^d\,(1\text{-}B^s)^D\,x_i \quad (2)$$

In Equation 1, w_i should be taken as z_i if the series is stationary. Where $\emptyset(B), \theta(B)$, $\Phi(B)$ and $\Theta(B)$for the regular and seasonal series are the autoregressive and moving-average operators having polynomials of order p, q, P and Q, respectively. Besides, “d and D” in the equation 2 above represent the order of the non-seasonal and seasonal differencing components. The a_t is called as white noise time series at time t, “s” is referred to the length of season.

Whether seasonal or not, the ARIMA model has three stages, which are identification, estimation and diagnostic checking (Box and Jenkins 1976). For the purpose, each stage related to the ARIMA is carried out meticulously to form the model that best fits a given data set.

In the first stage, a disquisition is performed on the stationarity of the series or necessity of differencing. Besides, the order of AR and MA operators is determined. The existence of seasonality in the series can be revealed by plotting seasonal series. On the other hand, variability (trend) in mean and variance of the data set should be detected to achieve stationary series. This can be done with different statistical methods. In this study, the Spearman’s Rho test being a non-parametric approach was preferred to reveal the presence of trend. All details about this test are available in Chiew and Siriwardena (2005).

Hipel et al. (1977) also stated that autocorrelation (ACF) and partial autocorrelation (PACF) functions are a good tool to have knowledge of AR and MA operators. The ACF evaluates serial dependence among the observations. With the visual appearance of ACF and PACF, the definition of ARIMA model can be carried out tentatively. For statistically analyzing of the ACF belonging to a seasonal series, the autocorrelation coefficient (r_k) with the order of k should be calculated for a maximum lag up to 5s (5s < N/4). The PACF is mostly obtained from 20 to 40 lags (40 < n/4). Statistical analysis of whether autocorrelation coefficients differ from zero significantly can be performed with the Q (r) statistic (Ljung and Box 1978). Its relationship is given in the Equation 3. However, autocorrelation $r_k(a)$ values of residual (error) terms have been taken into consideration in this relationship. Instead of $r_k(a)$, "r_k" should be preferred in the analysis of the autocorrelation coefficients of real data.

The other stage is associated with parameter estimation and making inference about the tentatively defined model. The final stage is on the fulfillment of some statistical features of the residuals. It is assumed in ARIMA model that the residuals from the model have no serial dependence, homoscedastic and normally distributed. By obtaining the residual autocorrelation function (RACF), based on the Q(r) statistics (Equation 3), the serial dependence among the residues is tested. The normality condition associated with the residuals was fulfilled according to Smirnov-Kolmogorov method. Fisher test specifying with an exhaustive explanation in Dahmen and Hall (1990) was used for homoscedasticity (stability in variance) of the residuals. The stability in variance and normality condition of the residuals may be performed by a Box-Cox transformation Yurekli et al (2005) or other normalization methods in Jayalakshmi and Santhakumaran (2011), before the relevant statistical tests. Besides, the Fisher transform in Ehlers (2002) can be used for this purpose.

$$Q(r) = n(n+2)\sum_{k=1}^{m}(n-k)^{-1} r_k(a)^2 \qquad (3)$$

Box and Jenkins (1976) suggested the parsimonious model, that is, use of the model with as few parameters as possible. Some criteria (Akaike Information Criterion, AIC; Schwarz Bayesian Criterion, SBC) are presented in the literature for the analysis of this purpose (Akaike 1974; Schwarz 1978). The model with minimum AIC and SBC among alternative SARIMA models carrying out all conditions is accepted as a parsimonious with respect to these criteria.

3. **Results and Discussion**

In the study, firstly, the Spearman's Rho test (SR) was used to determine whether the monthly reference plant water consumption (ETo) values estimated from Konya meteorology station data are stationary. The null hypothesis (H_o) associated with the SR test statistically endorses the series which is independent and identically distributed. Alternative hypothesis (H_1) is related to non-stationary of the series. The detection of this is done by comparing the statistical measure (t_{cal}) from the SR test to the critical table value (± 1.96) from standard normal distribution (Chiew and Siriwardena, 2005). The result of the SR test belonging to the monthly ETo for Konya province is that the data set has non-stationary, because of (t_{cal} = 2.140>1.96). The presence of the monotonic trend, that shows, the non-stationary behaviour of the relevant series, highlights the need for the non-seasonal differencing in the series.

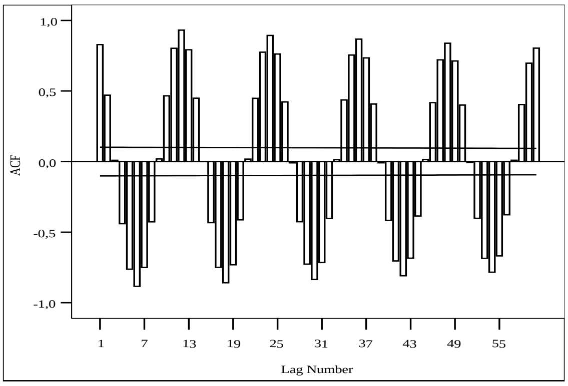

Figure 2. The ACF graph related to actual ETo dataset of Konya.

Şekil 1. Konya ilinin gerçek ETo verisetiyle ilişkili ACF grafiği.

The seasonal behavior of the ETo dataset can be visually explored with the ACF (autocorrelation function) plot. The ACF plot of the ETo dataset was obtained according to 5s < N/4 rule (up to r_k = 60 for the study) in Figure 2. This figure puts emphasis on seasonal behavior.

Therefore, the application of seasonal differencing to the series is mandatory to eliminate the seasonal effect. In Figure 2, the autocorrelation coefficients seem to exceed the upper and lower confidence levels shown by the continuous line in the plot. This result highlights statistically significant serial dependence among the ETo data values. The autocorrelation coefficient was obtained as 0.804 in lag 60. Based on this value, the Ljung-Box Q(r) statistic value was calculated as 9114.8. When the Q (r) statistic was compared with the table value of the chi square distribution, it was determined that the r_k in lag 60 was statistically significantly different from zero.

After the first differencing non-seasonal and seasonal to the data set to reach a time series not having monotonic trend and seasonality, The ACF and PACF (partial autocorrelation function) plots were obtained for r_k = 60 (Figure 3). Alternative SARIMA model structures were estimated from these plots.

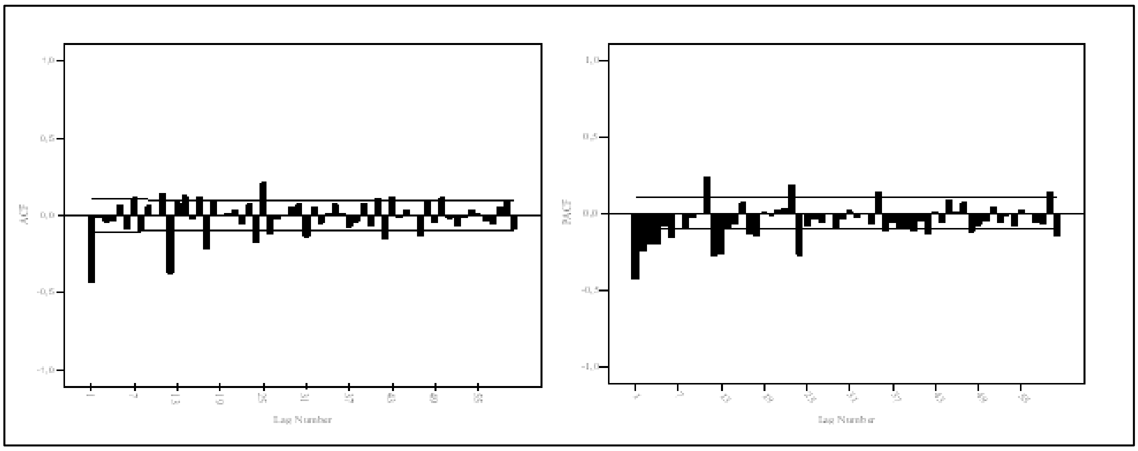

Figure 3. The ACF and PACF graphs after the first differencing to actual ETo dataset.
Şekil 3. **G**erçek ETo verisinin birinci farkı alındıktan sonraki verinin ACFve PACF grafikleri.

The tentatively SARIMA models from the ACF and PACF charts were detected and analyzed whether the required conditions were carried out. Three SARIMA models providing conditions in each stage were selected. These models some of their features were given in the Table1. As can be seen from the table, all parameters were statistically significant for the t test, except the MA1 parameter of SARIMA model (3,1,2)(0,1,1). The diagnostic checking results of these models are in Table 2. For this reason, the normality, serial dependence and homoscedasticity tests belonging to the residuals from the models were implemented. Besides, The AIC and SBS criteria used to reveal the parsimonious model are also given in this table.

Table 1. The selected SARIMA models for ETo dataset
Çizelge 1. ETo veriseti için seçilen SARIMA modelleri

Parameter	(2,1,1)(0,1,1)	t	P	(3,1,3)(0,1,1)	t	P	(3,1,2)(0,1,1)	t	P
AR1	0.227	3.90	0.000	-1.276	-18.71	0.000	-0.731	-12.39	0.000
AR2	0.116	2.07	0.039	-0.569	-5.58	0.000	0.331	4.16	0.000
AR3	---	---	---	0.247	4.08	0.000	0.156	2.76	0.000
MA1	0.919	32.40	0.000	-0.611	-12.88	0.000	-0.070	-1.70	0.090
MA2	---	---	---	0.386	7.63	0.000	0.919	24.01	0.000
MA3	---	---	---	0.856	20.77	0.000	---	---	---
SMA1	0.938	24.80	0.000	0.944	23.77	0.000	0.927	25.38	0.000

It is evident in the Table 2 that the residuals from the SARIMA models accomplished normality, serial dependence and homoscedasticity conditions. In the context of parsimony, all three models gave close results according to AIC and SBC criteria. Therefore, it is concluded that these three models can be used for ETo estimates in Konya province.

Table 2. The diagnostic checking results associated with SARIMA models
Çizelge 1. SARIMA modelleriyle ilişkili teşhis control sonuçları

ARIMA Model	SK		Q(r)-RACF		F-Test		AIC	SBC
	z	p	Q(r)	P	F_{cal}	F_{Table}		
(2,1,1)(0,1,1)	0.725	0.669	71.20	0.153	0.635	1.183	2863	2879
(3,1,3)(0,1,1)	0.694	0.721	71.73	0.143	0.638		2865	2893
(3,1,2)(0,1,1)	0.646	0.798	65.09	0.304	0.633		2857	2880

SK, Smirnov-Kolmogorov normality test
Q(r)-RACF, Ljung-Box Q statistic for residuals
F-Test, Homoscedasticity test

4. Conclusion

In conditions where climate change affects our globe very heavily, the significance and management of fresh water has become more and more important. Considering that a significant part of clean water is consumed in agriculture, accurate estimation of reference evapotranspiration is very important for reliable planning and management of water resources. Simulation techniques can be used to estimate the reference evapotranspiration where the required climate parameters are not sufficient or where there are lack of its measurement. In the study, there SARIMA models were selected for forecasting ETo dataset from Konya meteorology station.

References

Akaike H. (1974). A look at the statistical model identification. IEEE Transactions on Automatic Control, AC-19(6): 716-723.

Allen RG, Pereira LS, Raes D and Smith M (1998) .Crop Evapotranspiration (guidelines for computing crop water requirements). FAO Irrigation and Drainage Paper No. 56.

Box GEP and Jenkins GM (1976). Time Series Analysis Forecasting and Control. 575p., Holden-Day: San Francisco.

Chiew F and Sriwardena L (2005). Trend (trend/change detection software), User Guide. CRC for Catchement Hydrology, 23p., Australia.

Dahmen ER and Hall MJ (1990). Screening of Hydrological Data: Tests for Stationarity and Relative Consistency. International Institute for Land Reclamation and Improvement, Publication 49, 58p., Wageningen, The Netherlands.

Dwivedi DK and Shrivastava P K (2019). Time Series Modelling of Monthly Temperature and Reference Evapotranspiration for Navsari (Gujarat), India. Current Journal of Applied Science and Technology, 35:1-13.

Ehlers JF (2002). Using the Fisher Transform. Stocks & Commodities, V(20:11): 40-42.

Gautam R and Sinha AK (2016). Time series analysis of reference crop evapotranspiration for Bokaro District, Jharkhand, India. Journal of Water and Land Development, 30: 51-56.

Hamdi MR, Bdour ANand Tarawneh ZS (2008). Developing Reference Crop Evapotranspiration Time Series Simulation Model Using Class a Pan: A Case Study for the Jordan Valley /Jordan. Jordan Journal of Earth and Environmental Sciences, 1: 33-44.

Hipel KW, McLeod AI and Lennox WC. (1977). Advances in Box-Jenkins modeling: I. model construction. Water Resources Research 13(3): 567-575.

Janacek G and Swift L. (1993). Time Series Forecasting, Simulation, Application. Ellis Horwood, 333p., New York.

Jayalakshmi T and Santhakumaran A (2011). Statistical Normalization and Back Propagation for Classification. International Journal of Computer Theory and Engineering, 3(1): 1793-8201.

Landeras G, Barredo AO and Lopez JJ (2009). Forecasting Weekly Evapotranspiration with ARIMA and Artificial Neural Network Models. Journal of Irrigation and Drainage Engineering, 135: 323-334.

Ljung GM and Box GEP (1978). On a measure of lack of fit in time series models. Biometrika, 65(2): 297-303.

Mossad A and Alazba AA (2016). Simulation of temporal variation for reference evapotranspiration under arid climate. Arabian Journal Geosciences, 9: 1-9.

Patra SR (2018). Time Series Analysis of Reference Crop Evapotranspiration Using Machine Learning Techniques For Ganjam District, Odisha, India. Proceeding of the 2nd International Conference on Compute and Data Analysis, March 2018, DeKalb IL USA.

Schwartz G (1978). Estimating the dimension of a model. The Annals Statistics, 6: 461-464.

Sing S and Jaiswal CS (2006).Numerical Solution of 2D free surface to ditch drains in presence of transient recharge and depth-dependent ET in sloping aquifer. Water Resources Management, 20:779-793.

Tao PC and Delleur JW (1976). Seasonal and Nonseasonal ARMA Models in Hydrology. Journal of The Hydraulics Division, HY10: 1541-1559,

Tapur T (2008) Kazımkarabekir İlçesi’nde (Karaman) Tarım ve Hayvancılık. Selçuk Üniversitesi, Sosyal Bilimler Enstitüsü Dergisi, Sayı: 20: 603-620.

Trajković S (1998). Comparıson of predıctıon models of reference crop evapotranspıratıon. The scientific journal Facta Universitatis, 1: 617-625.

Yurekli K, Kurunc A and Ozturk F (2005). Application of Linear Stochastic Models to Monthly Flow Data of Kelkit Stream. Ecological Modeling, 183(1): 67-75.