American Journal of Engineering Research (AJER)
e-ISSN : 2320-0847 p-ISSN : 2320-0936
Volume-02, Issue-12, pp-296-302
www.ajer.org

Research Paper Open Access

The forecasting of Potential Evapotranspiration using time series analysis in humid and semi humid regions

Arash Asadi, Seyed Farnood Vahdat, Amirpooya sarraf

Islamic Azad University, Dehdasht branch, Iran
Ph.D of hydrology and water resources engineering, Tehran, Iran
Assistant professor in department of soil sciences,roudehen branch, Islamic azad university, roudehen,iran

Abstract: - Stochastic models have been proposed as one technique to generate scenarios of future climate change. The goal of this study is the simulation and modeling of monthly Potential Evapotranspiration (PET) using stochastic methods. Also, in this research for calculation of PET was used the Thornthwaite method.The 28-year data PET at yasuj Synoptic Station in southwest of Iran have been used in this study and based on ARIMA model, the autocorrelation and partial autocorrelation methods, assessment of parameters and types of model, the suitable models to forecast monthly PET was obtained. After models validation and evaluation, the forecasting was made for the crop years 2012-13 and 2013-14. In view of the forecasting made, the findings of the forecasting show an increase in PET along with a narrowing of the range of variations.

Keywords: - *Potential Evapotranspiration(PET), Time Series Analysis , ARIMA*

I. INTRODUCTION

The study of meteorological parameters is highly important in hydrology problems, since the same parameters generally form the climate of a region and is due to variations caused by water, wind, rain, etc. that issues problems, such as drought. Therefore, accuracy in data collection of such parameters is of particular importance. The study of long statistical term of the behavior and fluctuations in climatic parameters and analysis of the results obtained as well as the study of the behavior of a phenomenon in the past can analyze its probable trend in the future, too. Therefore, one can study the climatic variations using forecasting and estimation of parameters, such as precipitation and temperature and studying their behaviorin the past. In order to modeling and forecasting, stochastic and time series methods can be used. Statistical methodsinclude two objectives: 1- understanding of random processes, 2- Forecasting of series (Anderson, 1971). Time series analysis has rapidly developed in theory and practice since 1970s to forecast and control. This type of analysis is generally related to data which are not independent and are consecutively dependent to each another. In a study, the mean of monthly temperature of Tabriz Station in Iran was investigated based on Box & Jenkins ARIMA (Autoregressive Integrated Moving Average) model, In this study the monthly temperature of Tabriz for a 40-year statistical period (1959-98) was examined based on autocorrelation and partial autocorrelation methods as well as controlling the normality of residues using above mentioned models. Based on the obtained models, the variations of the mean of temperature of Tabriz Station are forecasted up to the year 2010 (Jahanbakhsh and Babapour Basser, 2003). A study was conducted to analyze the climate of Birjand Synoptic Station in Iran and recognize climatic fluctuations, especially drought and wetness to provide a suitable model to forecast the climatic fluctuations and the best model using statistical methods and Box-Jenkins models of time series of precipitation and temperature. Among the necessities to conduct this study are climatic forecasting to be used in the state planning at large concerning natural disasters, thus, the precipitation and temperature of Birjand Station have been studied to identify the climatic fluctuations and their possible forecasting (Bani Vaheb and Alijani, 2005). Bouhaddou et al. (1997) used the AutoRegressive Moving Average model (ARMA) model for simulation of weather parameters such as ambient temperature, humidity and clearness index. Frausto et al. (2003) implied that autoregressive (AR) and ARMA could be used to describe the inside air temperature of an unheated. Kurunc et al. (2005) applied the ARIMA approach to water quality constituents and streamflows of the Yesilirmak River in Turkey. Yurekli and Kurunc (2006) performed prediction of drought periods based on water consumption of the selected critical crops by using ARIMA approach. Yurekli et al. (2005) used the ARIMA model to simulate monthly stream flow of Kelkit Stream in Turkey. Yurekli and Ozturk (2003) showed whether the daily extreme stream flow sequences concerning with Kelkit Stream could be generated by stochastic models. In a study, modeling of drought in Fars Province in Iran was made using Box-Jenkins method and ARIMA model and the model to forecast drought in any region was obtained after zoning of different regions (Shamsnia et al., 2009). Shahidi et al., (2010) used ITSM software for Modeling and Forecasting Groundwater Level Fluctuations of Shiraz Plain in Iran. The autoregressive (order 24) fitted to the series with AIC=165.117. Coefficient of the fitted model was finalized by the residual tests. In a another study, the monthly maximum of the 24-h average time-series data of ambient air quality—sulphur dioxide (SO2), nitrogen dioxide (NO2) and suspended particulate matter (SPM) concentration monitored at the six National Ambient Air Quality Monitoring (NAAQM) stations in Delhi, was analysed using Box–Jenkins modelling approach. The model evaluation statistics suggest that considerably satisfactory real-time forecasts of pollution concentrations can be generated using the Box–Jenkins approach. The developed models can be used to provide short-term, real-time forecasts of extreme air pollution concentrations for the Air Quality Control Region (AQCR) of Delhi City, India (Sharma et al., 2009).

Therefore, considering the importance of climatic parameters and the importance they have in determining the roles of other climatic elements, their modeling and recasting using advanced statistical methods is a necessity and could be a basic pillar in agricultural and water resource managements. Also, in this research for calculation of PET was used the Thornthwaite method. The goal of the present study is the simulation and providing a model to forecast Potential Evapotranspiration (PET) under study using the statistical models of time series analysis in The Synoptic Station of the yasuj City.

II. RESEARCH METHODOLOGY

In this study, the monthly data on the precipitation and the mean temperature of yasuj Synoptic Station were used for calculation of Potential Evapotranspiration (PET) and the required information was collected from the tables and the databases available. Yasuj city located in kohgilouye and boyerahmad Province in southwest part of Iran, is at 51 35 E longitude and 30 42 N latitude with the area of 26416 square kilometers. The mean annual precipitation 860 mm and mean annual temperature for the study area about 15 °C (I.R. of Iran Meteorological Org.). The geographical location of the study region is shown in Figure 1. The statistical period under study is the crop years 1983-84 through 2011-12. Initially, the homogeneity of data was confirmed using the run test statistical method. Essentially homogenous test before statistical analysis on data should be taken to ensure the stochastic data. Homogeneous data was done using SPSS software.

Then, based on the results obtained and studying the sequence of observations and the past behavior of the phenomenon the appropriate model was devised to forecast using time series analysis and stochastic methods. In order to model the data, they were fixed after preparing the time series of observations of Potential Evapotranspiration.

For fitting ARIMA model to the time series of the new data sequence, the basis of the approach consists of three phases: model identification, parameter estimation and diagnostic testing (Yurekli and Ozturk, 2003). Identification stage is proposed to determine the differencing required to produce stationary and also the order of AR and MA operators for a given series. Stationary is a necessary condition in building an ARIMA model that is useful for forecasting. A stationary time series has the property that its statistical characteristics such as the mean and the autocorrelation structure are constant over time. When the observed time series presents trend and heteroscedasticity, differencing and power transformation are often applied to the data to remove the trend and stabilize variance before an ARIMA model can be fitted. Estimation stage consists of using the data to estimate and to make inferences about values of the parameters conditional on the tentatively identified model. The parameters are estimated such that an overall measure of residuals is minimized. This can be done with a nonlinear optimization procedure.

The diagnostic checking of model adequacy is the last stage of model building. This stage determines whether residuals are independent, homoscedastic and normally distributed. Several diagnostic statistics and plots of the residuals can be used to examine the goodness of fit; the tentative model should be identified, which is again followed by the stage of parameter estimation and model verification. Diagnostic information may help to suggest alternative model(s). This three-step model building process is typically repeated several times until a satisfactory model is finally selected. The final selected model can then be used for prediction purpose. By plotting original series trends in the mean and variance may be revealed (Box and Jenkins, 1976).The ARIMA model is essentially an approach to forecasting time series data. However, the ARIMA model requires the use of stationary time series data (Dickey and Fuller, 1981).

III. THE MODELING PROCEDURES

Modeling is made using time series analysis by several methods, one of which is the ARIMA or Box-Jenkins method, being called the (p,d,q) model, too (Box and Jenkins, 1976). In the (p,d,q) model, p denotes the number of autoregressive values, q denotes the number of moving average values and d is the order of differencing, representing the number of times required to bring the series to a kind of statistical equilibrium. In an ARIMA model, (p,d,q) is called the non-seasonal part of the model, p denotes the order of connection of the time series with its past and q denotes the connection of the series with factors effective in its construction. The mathematical formulation of ARIMA models shown by equation (1). Analysis of a time series is made in several stages. At the first stage, the primary values of p, d and q are determined using the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). A careful study of the autocorrelation and partial autocorrelation diagrams and their elements, will provide a general view on the existence of the time series, its trend and characteristics. This general view is usually a basis for selection of the suitable model. Also, the diagrams are used to confirm the degree of fitness and accuracy of selection of the model. At the second stage, it is examined whether p and q (representing the autoregressive and moving average values, respectively) could remain in the model or must exit it. At the third stage, it is evaluated whether the residue (the residue error) values are stochastic with normal distribution or not. It is then, that one can say the model has a good fitness and is appropriate. If the time series is of seasonal type, then the modeling has a two-dimensional state, and in principle, a part of the time series variations belongs to variations in any season and another part of it belongs to variations between different seasons. A special type of seasonal models that shows deniable results in practice and coin sides with the general structure of ARIMA models is devised by Box and Jenkins (1976), which is called multiplicative seasonal model. It is in the form of ARIMA (pdq) (PDQ). Then, for the model being ideal, the schemes must be used to test the model and for the comparison purpose, so as the best model is chosen for forecasting.

$$X_{(t)} = X_{(t-1)} \pm X_{(t-2)} \pm X_{(t-3)} \pm X_{(t-n)} \pm Z_{(t)} \quad (1)$$

IV. MODEL SELECTION CRITERIA

Several appropriate models may be used to select a model to analyze time series or generally data analysis to present a given set of data. Sometimes, selection is easy, whereas, it may be much difficult in other times. Therefore, numerous criteria are introduced to compare models which are different from methods for model recognition. Some of these models are based on statistics summarized from residues (that are computed from a fitted scheme) and others are determined based on the forecasting error (that is computed from forecasting outside the sample). For the first method, one can point to AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and SBC (Schwartz-Bayesian Criterion) and for the scheme based on the forecasting error, one can point to the Mean Percent Error (MPE) method, the Root Mean Square Error (RMSE), the Mean Absolute Value Error (MAE), and the Mean Absolute Value Percent Error (MAPE). The model, in which the above statistics are the lowest, will be selected as the appropriate model. Akaike (1974) suggests a mathematical formulation of the parsimony criterion of model building as Akaike Information Criterion (AIC) for the purpose of selecting an optimal model fits to a given data. Mathematical formulation of AIC is defined as:

$$AIC\ (M) = n.Ln(\sigma_a^2) + 2M \quad (2)$$

Where "M" is the number of AR and MA parameters to estimate, "σ_a^2 " is Residual variance and " n "is the number of observation. The model that gives the minimum AIC is selected as a parsimonious model. Akaike (Akaike, 1974) has shown that the AIC criterion trends to overestimate the order of the autoregression. But, Akaike (Akaike, 1978; Akaike, 1979) has developed a Bayesian extension of minimum AIC procedure, called as BIC. The another index for model evaluating is efficiency factor. The model efficiency (EF), which indicates the robustness of the model (Raes et al., 2006). EF ranges from $-\infty$ to 1with higher values indicating a better agreement. If EF is negative, the model prediction is worse than the mean observation:

$$EF = \frac{\sum_{i=1}^{n}(O_i - \bar{O})^2 - \sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}(} \quad (3)$$

Where Oi and Pi are respectively the observed and predicted (simulated) values for each of the n study cases and O the mean observed value.

In the present study, ARIMA model, ITSM software, AIC, RMSE and EF criterion were used for modeling and forecasting the precipitation and temperature. The ITSM software determine the best model with minimum AIC and BIC. Also the best model validated using model efficiency.

Time series of monthly Potential Evapotranspiration in yasuj Station were showed in Fig (2). Trend and seasonal components recognized by ACF/ PACF diagrams (Figure 3), shows the peaks in 12 and 24 lag times. These deterministic parameters removed by difference operator. Residual testing was used for validation. ACF/PACF of residuals shows all covered by 95% confidence interval (Figure 4). The RACFs drawn for the best models indicated that the residuals were not significantly different from a white noise series at 5% significance level. Inspection of the RACFs and the residuals integrated periodogram confirmed a strong model fit.

V. DISCUSSION

5.1 Modeling of monthly Potential Evapotranspiration (PET)

To model using ACF and PACF methods, assessment of values related to auto regression and moving average were made and eventually, an appropriate model for estimation of Potential Evapotranspiration values for yasuj Station was found as ARIMA (2 0 0) (0 1 1)12. To prevent excessive fitting errors, AIC and EF criterion was used. In comparison between schemes, regarding the lowest AIC and EF value, the final model with the best fitting of data, obtained using the method of maximum likelihood and ITSM software.the evaluation criterions are shown in table 1.

Figure 5. shows the correlation between observed and predicted data from ARIMA models in crop years 2009-10 through 2011-12. Therefore, because of the strong correlation of data, the selected model is suitable for simulating the monthly Potential Evapotranspiration. According to the results, Predicted data for the agriculture years 2012-13 and 2013-14 are shown in Figure 6.

VI. CONCLUSION

Recent droughts in kohgilouye and boyerahmad Province with yasuj, as its center, have led to much damage. To prevent such huge damage, knowledge of the fluctuations during the statistical period and forecasting of them in planning is necessary. The findings of the study of climatic parameter of monthly Potential Evapotranspiration and evaluation of diagram showed that, variations of Potential Evapotranspiration in yasuj region denote the existence of severe and, in some instances, long-term droughts. The Box-Jenkins model was used to forecast the studied parameters and the final model was tested using AIC and EF criterion and the results showed that it can be used to forecast the monthly variations in Potential Evapotranspiration in the city of Shiraz regarding its high accuracy. For model validation, EF value calculated 0.9 for Potential Evapotranspiration. Also R^2 for climate variables obtained .99. Consequently, the models can be used for forecasting of studied variables. As regards the mean monthly Potential Evapotranspiration, the trend of increasing, especially in recent years, has continued and the findings of the forecasting show an increase in Potential Evapotranspiration along with a narrowing of the range of variations.

VII. REFERENCES

[1] Akaike H. A look at the statistical model identification. (1974). IEEE Transactions on Automatic Control AC. 19(6):716–23.

[2] Akaike H. A. (1978). Bayesian analysis of the minimum AIC procedure. Annals of the Institute of Statistics and Mathematics. 30(A):9–14.

[3] Akaike H. (1979). Bayesian extension of the minimum AIC procedure of autoregressive model fitting. Biometrika. 66:237–42.

[4] Anderson, T.W. (1971). The statistical Analysis of time series. John Wiley & Sons, NewYork.

[5] Bani Waheb, E. and Alijani, B. (2005). On the drought and wetness years and forecasting of climatic variations of Birjand region using statistical models. Journal of Geographic Researches. 37(52): 33-46. [In Persian].

[6] Bouhaddou, H. Hassani, M.M. Zeroual, A. and Wilkonson, A.J. (1997). Stochastic simulation of weather data using higher statistics. Renewable Energy.12 (1):21–37.

[7] Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis, Forecasting and Control. Holden-Day Pulication.

[8] Brockwell, P. J. and Davis, R. A. (2002). Introduction to time series and forecasting. New York: Springer-Verlag.

[9] Dickey, D.A. and Fuller, W.A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica. 49:1057–72.

[10] Frausto, H.U. Pieters, J.G. and Deltour, J.M. (2003). Modelling greenhouse temperature by means of Auto Regressive Models. Biosystems Engineering. 84(29):147–57.

[11] Jahanbakhsh, S. and Babapour Basseri, E.A. (2003). Studying and forecasting of the mean monthly temperature of Tabriz, using ARIMA model. Journal of Geographic Researches. 15(3): 34-46. [In Persian].

[12] Kurunc, A. Yurekli, K. C- evik, O. (2005). Performance of two stochastic approaches for forecasting water quality and streamflow data from Yes-ılırmak River, Turkey. Environmental Modelling & Software. 20:1195–200.

[13] Raes, D. Greets, S. Kipkorir, E. Wellens, J. and Sahli, A. (2006). Simulation of yield decline as a result of water tress with a robust soil water balance model. Agricultural Water Management. 81:335-357.

[14] Shahidi, N. Rahnemaei, M. Sharifan, R.A. and Nematollahi, A.R. (2010). Modeling and forecasting groundwater level fluctuations of Shiraz Plain using advanced statistical models. International Conference on Environmental Engineering and Applications (ICEEA 2010). 10-12 Sep. Singapore.

[15] Shamsnia, S.A., Amiri. S.N. and Pirmoradian, N. (2009) Drought simulation in Fars province using standardized precipitation index and time series analysis (ARIMA model). International Journal of Applied Mathematics. 22(6):869-878. [Available on site http://www.diogenes.bg/ijam/].

[16] Sharma, P., Chandra, A. and Kaushik, S.C. (2009) Forecasts using Box–Jenkins models for the ambient air quality data of Delhi City. Environmental Monitoring and Assessment. 157(1-4):105-112.

[17] Yurekli, K. and Kurunc, A. (2006) Simulating agricultural drought periods based on daily rainfall and crop water consumption. Journal of Arid Environments. 67: 629-640.

[18] Yurekli, K. and Ozturk, F. (2003) Stochastic modeling of annual maximum and minimum streamflow of Kelkit Stream. International Water Resources Association. 28(4):433–41.

[19] Yurekli K, Kurunc- A, Ozturk F. (2005) Application of linear stochastic models to monthly flow data of Kelkit Stream. Ecological Modelling. 183:67–75.

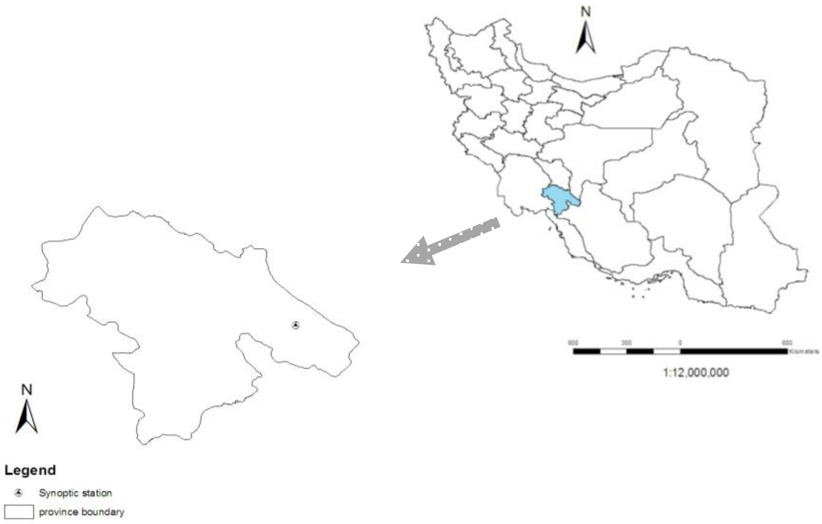

Fig 1. Regional map of Iran , location of study area and synoptic station

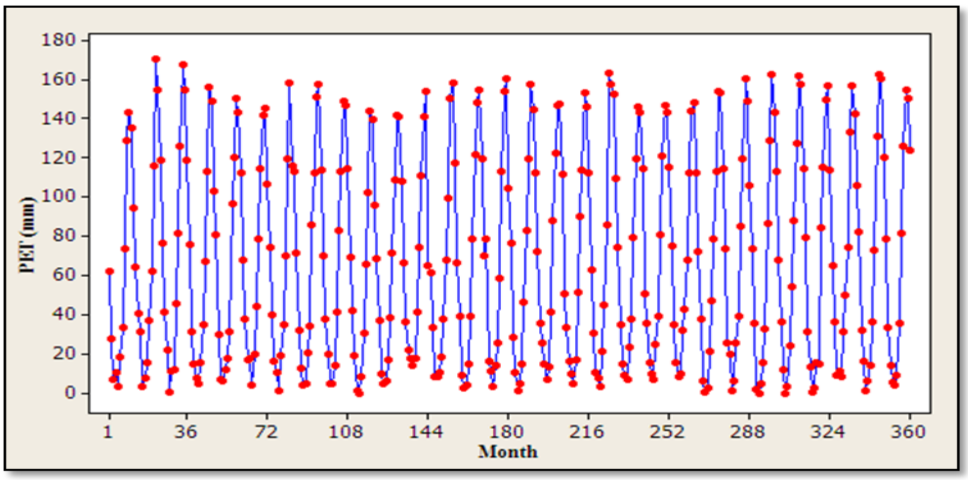

Figure 2. Time series of monthly PET in yasuj Station

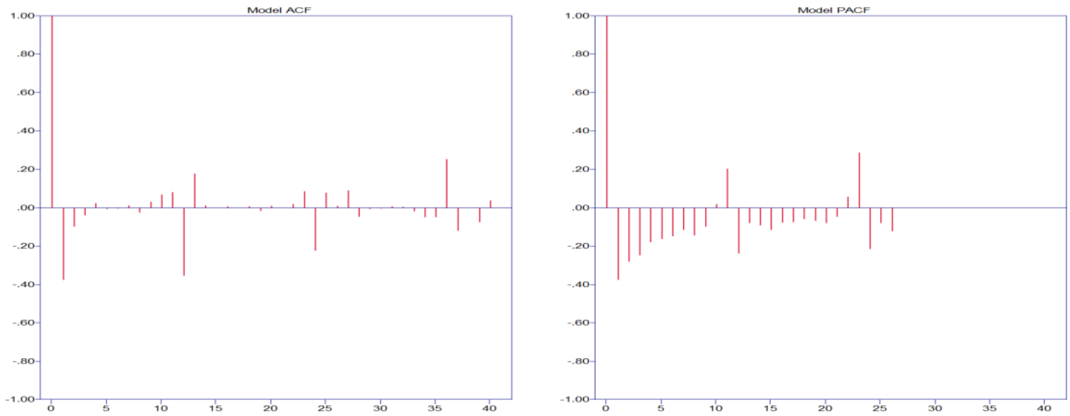

Figure 3. ACF/PACF of monthly PET in yasuj Station

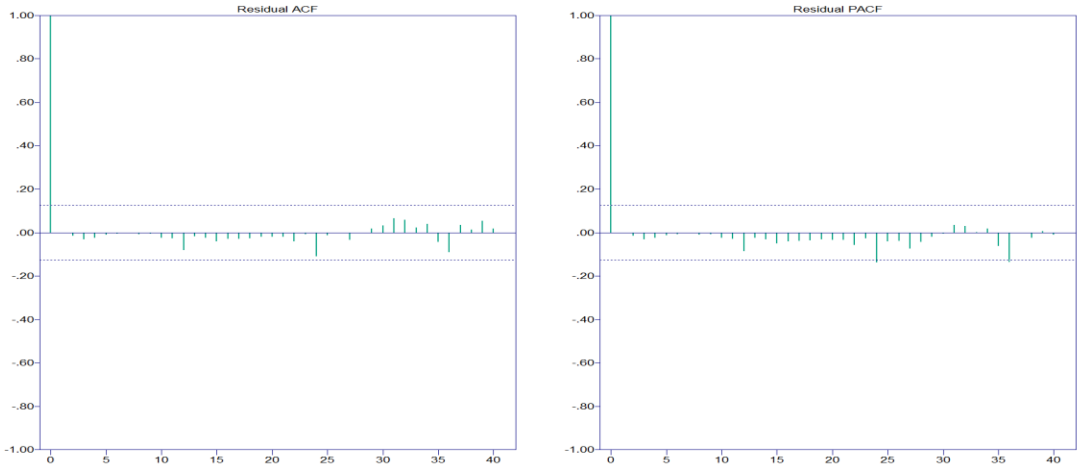

Figure 4. ACF/PACF of residual for monthly PET

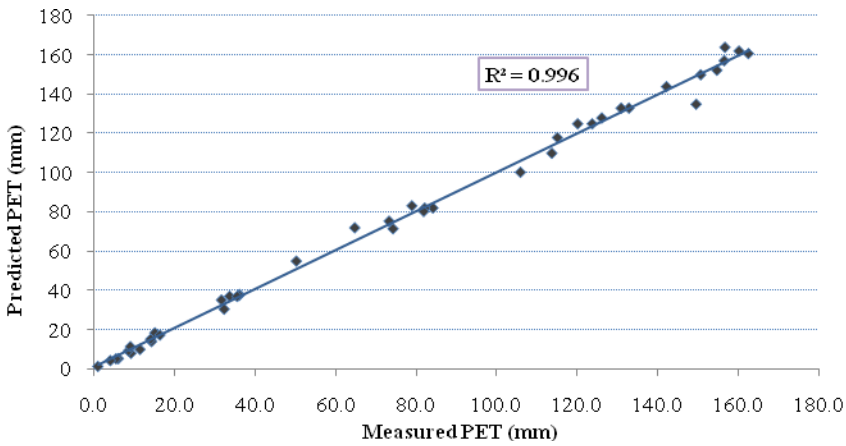

Figure 5. Correlation between observed and predicted data from ARIMA models in crop years 2009-10 through 2011-12.

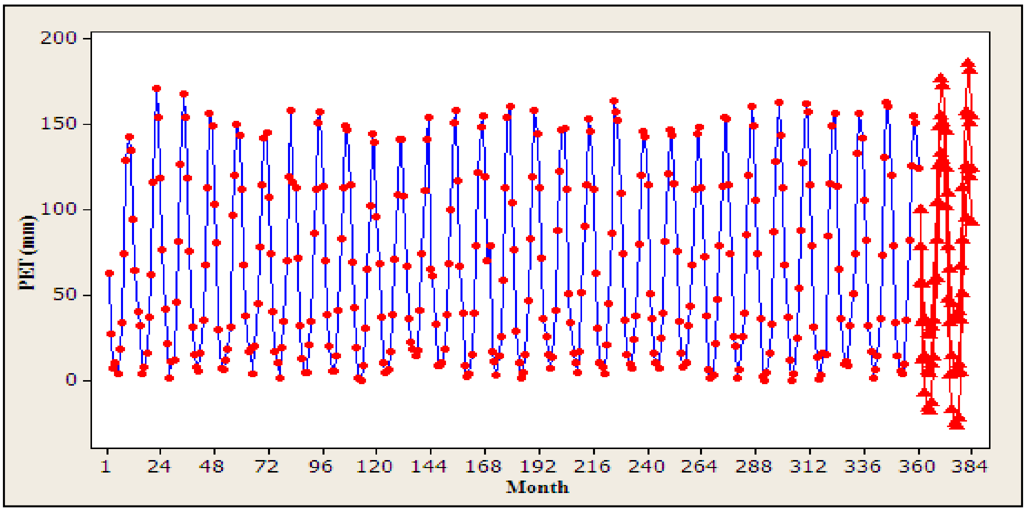

Figure 6. Predicted data of the mean of monthly PET for 2012-13 and 2013-14

Table 1. The ARIMA models selected for PET variable

ARIMA Model	AIC	RMSE	EF
(2 0 0) (0 1 1)12	4.18283	8.02173	0.91
(1,0,1) (0,1,1)12	4.18443	8.02816	0.85
(1,0,0) (0,1,1)12	4.18482	8.05446	0.83
(2,0,2) (0,1,1)12	4.18508	7.98135	0.76
(0,0,2) (0,1,1)12	4.18863	8.04502	0.72