MAUSAM

DOI : https:// doi.org/10.54302/mausam.v73i2.5488 Homepage: https://mausamjournal.imd.gov.in/index.php/MAUSAM

UDC No. 631 : 551.509.3 : 551.573 (540.61)

Forecasting potential evapotranspiration for Raichur district using seasonal ARIMA model

RAHUL PATIL, D. M. NAGARAJ, B. S. POLISGOWDAR* and SANTOSHA RATHOD**

*Department of Soil and Water Engineering, CAE, UAS, Raichur, India * Department of Irrigation and Drainage Engineering, CAE, UAS, Raichur, India ** Indian Institute of Rice Research, ICAR, Hyderabad – 500 030, India* (*Received 26 June 2019, Accepted 29 October 2021*)

e mail : rahul1235110@gmail.com

सार — सिंचाई प्रणालियों के विश्वसनीय प्रबंधन के लिए स्थितिज वाष्पोत्सर्जन (पीईटी) का पूर्वानुमान काफी महत्वपूर्ण कार्य है। यह लेख सामान्यतःउन मॉडलों पर आधारित है जो रायचूर जिले के लिए भविष्य में स्थितिज वाष्पोत्सर्जन की वास्तविक घटना का अनुकरण करने का प्रयास करते हैं। इस अध्ययन में थॉर्नथवेट विधि का उपयोग करके अधिकतम और न्यूनतम तापमान (°से.) डेटा की मदद से संभावित वाष्पोत्सर्जन का अनुमान लगाया गया और मौसमी स्वसमाश्रयी गतिमानऔसत विधि (SARIMA) का उपयोग करके पूर्वानुमान किया गया। मॉडल, स्व सहसंबंध फलन (ACF) और आंशिक स्व सहसंबंध फलन (PACF) के आधार पर विकसित किए गए। इसके अलावा, न्यनतम एकाइके सूचना मानदंड (एआईसी) और बायेिसयन सूचना मानदंड (बीआईसी) मान� वाले मॉडल का चयन किया गया। रायचूर , मानवी , सिंधानुरू और लिंगसुगुरु विभिन्न स्टेशनों के लिए चुने गए मॉडल क्रमशः ARIMA(2,0,2)(1,1,2)12, ARIMA(1,0,1)(2,1,0)12, ARIMA(1,0,1) (1,1,2)12 और ARIMA(1,0,1)(2,1,0)12 थे। इसके अलावा , परिणामों से पता चला कि मानवी और सिंधानुरू के लिए विकसित मॉडल अन्य दो स्टेशनों की तुलना में काफी आशाजनक पाए गए। सभी चार मॉडल एक महीने के अग्रकाल पूर्वानुमान तक बेहतर परिणाम देते पाए गए। मॉडल ने जल संसाधनों के बेहतर प्रबंधन के लिए सिंचाई योजना और कमान क्षेत्र प्रबंधन अभ्यासों में निर्णय लेने और सुधार करने में महत्वपूर्ण क्षमता प्रदान की।

ABSTRACT. The prediction of potential evapotranspiration (PET) is quite important task for reliable management of irrigation systems. This article is generally based on the models which try to mimic the actual occurrence of the Potential evapotranspiration in the future days for a Raichur district. In this study the potential evapotranspiration was estimated with the help of max and min temperature (°C) data using a Thornthwaite method and the prediction was carried out using the seasonal Autoregressive moving average method (SARIMA). The models were developed based upon autocorrelation function (ACF) and partial autocorrelation function (PACF). Furthermore, the model with the least Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) values were selected. The models selected for different stations were $ARIMA(2,0,2)(1,1,2)_{12}$, $ARIMA(1,0,1)(2,1,0)_{12}$, $ARIMA(1,0,1)(1,1,2)_{12}$ and $ARIMA(1,0,1)$ $(2,1,0)₁₂$ for Riachur, Manvi, Sindhanuru and Lingasuguru respectively. Furthermore, the results showed that the models developed for Manvi and Sindhanuru were found to be quite promising compared to the other two stations. All four models were found to be producing better results up to a lead time forecast of one month. The models provided significant potential in improving the decision making in irrigation planning and command area management practices for better management of water resources.

Key words – ACF, PACF, SARIMA, PET.

1. Introduction

Evapotranspiration (ET) is usually the largest component of the hydrologic cycle, given that most precipitation that falls on land is returned to the atmosphere (Asadi *et al*., 2013). Globally, ET consumes

about 60 per cent of the annual precipitation that falls over the earth's surface ET quantification is used for many purposes including crop production, management of water resources and environmental assessment (Aruna *et al*., 2017). It is a major component for water balance in the field and needs to be quantified accurately. The amount of

water supplied to meet the needs of the agricultural crops for evapotranspiration dictates the quality and quantity of production in a field. The ET data for agricultural crops has become increasingly important in irrigation as well as in water resources management. The process of ET is majorly regulated by many hydrological parameters such as temperature, relative humidity, solar radiation and wind speed etc.

The stochastic models are based on the time dependent variation and consider random effects involved in the ET process. For two main reasons, stochastic linear models are fitted to hydrological data or time series such as evapotranspiration series: enabling the integration of an on-farm system with the main system and facilitating the real-time operation of an irrigation system. It is quite important to develop a synthetic or forecast data set in order to design or plan any irrigation systems; in this context autoregressive integrated moving average (ARIMA) model is considered as one of the best model in forecasting the time series dataset. In this model the forecast of a variable is defined as a linear combination of the previous state of variable and previous forecast error. The ARIMA method is a strong time series modeling and forecasting technique which has versatility to include characteristic of time series. In past ARIMA models have been widely used to model hydrological time series (Popale and Gorantiwar, 2014). Popale and Gorantiwar (2014) used ARIMA model for prediction of rainfall of rahuri region, India. Gorantiwar and Patil (2009) did analysis of evapotranspiration of Rahuri region, India. Hamdi *et al*. (2008) developed seasonal ARIMA model for the Jordan valley. Asadi *et al.* (2014) forecasts evapotranspiration for humid and semi-humid region. salas *et al*. (1980) discussed in detail about time series modelling.

Knowledge of evapotranspiration is important for watershed management activities in meteorological and hydrological modelling, particularly water management in irrigated agriculture (Dutta *et al*., 2016). The evapotranspiration plays a major role in crop water requirement (CWR) of any crop. As CWR accounts for more than 95 % of the ET so it is quite important to understand its behaviour based on the historical data for better management of water resource, in order to understand and solve the irrigation problem the research was carried out. The objective of this study is to establish a time series model to analyse and forecast reference crop evapotranspiration for the Raichur district.

2. Materials and methodology

Raichur is an administrative district in the Indian state of karnataka. It is located in the northeastern part of the state and is bounded by Yadgir district in the north,

Bijapur and Bagalkot district in the northwest, Koppal district in the west, Bellary district in the south, Mahabubnagar district of Telangana and Kurnool district of Andhra Pradesh in the east. The district is bounded by the Krishna River on the north and the Tungabhadra River on the south. The wedge of land between the rivers is known as the Raichur Doab, after the city of Raichur. Bijapur and Yadgir districts lie to the north across the Krishna River. Bagalkot and Koppal districts lie to the west. Across the Tungabhadra lies Bellary District of Karnataka to the southwest and Mahabubnagar of Telangana to the southeast. Kurnool District of Andhra Pradesh state lies to the east and includes the lower portion of the Raichur Doab.

2.1. *Thornthwaite method (Potential evapotranspiration)*

The potential evapotranspiration is calculated by:

$$
PET = 16K \left(\frac{10T}{I}\right)^{m}
$$

where, T is monthly mean temperature (\degree C); I is heat index calculated as the sum of 12 month index values; *m* is the coefficient dependent on *I*.

$$
m = 6.75 \times 10^{-7} \cdot I^3 - 7.71 \times 10^{-7} \cdot I^2 + 1.79 \times 10^{-2} \cdot I
$$

+ 0.492

K is a correction coefficient computed as a function of the latitude and month.

2.2. *Auto correlation test (Box Ljung test)*

The null hypothesis of the Box Ljung Test, H_0 is that our model does not show lack of fit (or in simple termsthe model is just fine). The alternate hypothesis, H_a is just that the model does show a lack of fit. A significant p-value in this test rejects the null hypothesis that the time series isn't auto correlated.

2.3. *Stationary test (Dickey fuller test)*

A time series is said to be stationary (in the weak sense) if its statistical properties do not vary with time (means and variance). If the compute p values are greater than 0.05 the series is said to be non stationary. The time series need to be in stationary form in order to fit to stochastic models.

2.4. *Description of the stochastic models*

The stochastic models, also referred to as time series models, were used for the study of time series in mathematical, economic and engineering applications.

The time series modeling techniques have been shown to provide a systematic analytical tool to simulate and predict the behavior of unpredictable hydrological systems and to measure the predicted accuracy of the forecasts (Mishra and Desai, 2005).

2.5. *ARIMA models*

Autoregressive (AR) models can be considered in conjunction with moving average (MA) models to create a specific and effective class of time series models called autoregressive integrated moving average (ARMA) models. In an ARMA model the present value of the time series is explained as a linear aggregate of p lagged values and a weighted sum of q former deviations plus a random parameter.

An ARIMA models are generally used for a time series which are stationary in nature. However these models can be used in non stationary data set by differencing the series. Box and Jenkins (1976) developed a new forecasting tool, known as the ARIMA methodology, that focus on analyzing the stochastic characteristics of time series on its own rather than constructing single or simultaneous equation models.

ARIMA models allow stating each variable by its own lagged values and stochastic error terms. The general non-seasonal ARIMA model is AR to order *p* and MA to order q and operates on dth difference of the time series z_t ; thus a model of the ARIMA family is classified by three parameters (p, d, q) that can have zero or positive integral values (Mishra and Desai, 2005)

The general non-seasonal ARIMA model may be written as :

$$
\varphi(B)\nabla^d_{Z_t}=\theta(B)a_t
$$

where, θ (*B*) are polynomials of order *p* and *q*, respectively. Non-seasonal AR operator of order *p* is written as ·

$$
\varphi(B) = \left(1 - \varphi_1 \ B - \varphi_2 \ B^2 - \cdots \varphi_p \ B^p\right)
$$

and non-seasonal MA operator of order q is written as :

$$
\varphi(B) = \left(1 - \theta_1 B - \theta_2 B^2 - \cdots \theta_q B^q\right)
$$

2.6. *Seasonal ARIMA models*

Many time series features the cyclic. Quite frequently such characteristics are on an annual period in hydrologic time series mainly due to earth's rotation

around the sun. Such type of series are cyclically nonstationary. After removing the determinist cyclic effects from a series, the ARIMA approach may be applied to obtain a linear model for the stochastic part of the series (Gorantiwar *et al*., 2011). Box *et al*. (1994) standardized the ARIMA model to address seasonality and defined a general multiplicative seasonal ARIMA model commonly referred to as SARIMA models. An inherent advantage of the SARIMA family of models is that the description of time series requires few model parameters, which exhibit non-stationarity both in season and throughout. In general the SARIMA model described as ARIMA (*p*, *d*, *q*) (P, D, Q) *s*, where (p, d, q) is the non-seasonal part of the model and (*P*, *D*, *Q*)*s* is the seasonal part of the model, which is mentioned below :

$$
\phi_p
$$
 (B) Φ_p (B^s) $\nabla^d \nabla_s^D Z_t = \theta_p(B) \phi_Q(B^s) a_t$

where, *p* is the order of non-seasonal auto regression, *d* the number of regular differencing, *q* the order of nonseasonal MA, *P* the order of seasonal auto regression, *D* the number of seasonal differencing, *Q* the order of seasonal MA, *s* is the length of season, seasonal AR parameter of order P, seasonal MA parameter of order Q.

2.6.1. *Model identification*

This step is to identify the possible ARIMA model which represents the time series behavior. The series behavior was investigated based upon the behavior of autocorrelation function (ACF) and partial autocorrelation function (PACF) (Mishra and Desai, 2005; Hsin-Fu Yeh and Hsin-Li Hsu, 2019). The ACF and PACF were used to support in determining the order of the model. The information given by ACF and PACF is useful in suggesting the type of models that may be constructed. The final model was then selected using the Akaike information criterion (AIC) and Bayesian information criterion (BIC).

These criteria help to rank models (the models with the lowest criterion value being the best). The AIC and SBC take the mathematical form as shown below.

$$
AIC = -2 \log (L) + 2k
$$

$$
SBC = -2 \log (L) + k \ln(n)
$$

where, *k* is number of parameters in the model, *L* is the likelihood function of the ARIMA model; and *n* is the number of observations.

2.6.2. *Parameter estimation*

The estimation of model parameters was achieved after identification of the appropriate model as an essential

TABLE 1

Auto correlation test for different station

Station	Chi-Square	Lag order	P-value
Raichur	233.73		< 0.001
Manvi	261.11		< 0.001
Sindanuru	237.33		< 0.001
Lingasuguru	235.86		< 0.001

TABLE 2

Stationarity test for different station

Station	Dickey fuller	Lag order	P-value
Raichur	-19.691		0.01
Manvi	-18.671		0.01
Sindanuru	-19.649		0.01
Lingasuguru	-19.841		0.01

step. The model estimate values for the AR and MA parts were calculated using Maximum likelihood. The AR and MA parameters were tested to make sure that they are statistically significant or not.

2.6.3. *Diagnostic checking*

Diagnosis of the ARIMA model is a crucial part of model development and the last step. It involves checking the appropriateness of the model chosen. Several diagnostic statistics and residual plots are examined to see whether or not the residuals are correlated to white noise. In this study we obtained the residual ACF function (RACF) to determine whether residuals are white noise.

2.6.4. *Drought forecasting*

The prediction of Potential evapotranspiration was done for 1-6 month lead time using the best fit models from historical data. Basic statistical properties of the observed and predicted data for 1-6 month lead time was computed and tested whether the predicted data preserve the basic statistical properties of the observed PET series. The predictions are calculated for different lead time. For instance, a 1-month lead time prediction means that during January 2017, the prediction for February 2017 is computed. The correlation coefficients (R), RMSE and MAE were observed between the observed and predicted data for 1 to 6 month lead times.

Fig. 1. Autocorrelation function plot of PET time series for Raichur Station

Fig. 2. Partial autocorrelation function plot of PET time series for Raichur Station

2.7. *Input dataset and software*

The time series of temperature data set (Max and Min) was taken from the Main Agriculture Research Station (MARS) Raichur. The data set were from 1984- 2018, out of which 1984-2016 was used for the development of the model and the 2017-2018 was used for the validation purpose. The estimation of Potential evapotranspiration was estimated using MS Excel and SARIMA models were developed in the R studio.

3. Results and discussion

Development of model was done with prerequisite tests namely Stationary and autocorrelation test. The autocorrelation test was carried out using box test and corresponding probability levels are presented in Table 1. The results revealed that the test statistic for box.test with a Chi square and *P* values were 233.73 (0.01), 261.11 (0.01), 237.33 (0.01), 235.86 (0.01) and 235.86 (0.01) for Raichur, Manvi, Sindhanuru and Lingasugur respectively,

TABLE 3

Log likelihood AIC and BIC values of ARIMA model for different station

TABLE 4

Parameter estimation of SARIMA by maximum likelihood method for different station

were observed to be significant at 5% level of significance reflecting autocorrelation in data. On the other hand adf.test was carried out to check whether the data is stationary or not. The data was observed to have a seasonality thereby seasonal differencing was done to the data sets Table 2.

The principal step in Box-Jenkins ARIMA model building is identification of the model. Different orders of Autoregressive (AR) and Moving Average (MA) parameters *p* and *q* are considered and combination of the order which yields maximum log-likelihood and lowest values of Akaike Information Criteria (AIC) and Bayesian

TABLE 5

Auto correlation check for residuals of Seasonal ARIMA model at different station

TABLE 6

Performance measure of seasonal ARIMA models at different stations

TABLE 7

Comparison of statistic properties of the observed and predicted data

Information Criteria (BIC) are considered as best model. The results pertaining to Raichur, Manvi, Sindhanuru and Lingasuguru stations regarding model development are presented in Tables 3&4. The ACF and PACF were plotted (Figs. 1 and 2) to determine the model, the data were observed have a seasonality thereby seasonal ARIMA models were selected with a seasonal differencing as shown in Table 4. The best selected models for different stations were ARIMA $(2,0,2)(1,1,2)_{12}$, ARIMA $(1,0,1)(2,1,0)_{12}$, ARIMA $(1,0,1)(1,1,2)_{12}$ and ARIMA $(1,0,1)(2,1,0)_{12}$ with an maximum likelihood values of -1380.54, -1205.10, -1338.08 and -1356.04 respectively for Raichur, Manvi, Sindhanuru and Lingasuguru. The parameters estimated for different stations are presented in Table 4. In addition, the residuals were obtained by differencing original series with the fitted series and residuals were found to be white noise as presented in Table 5.

After the development of models for 4 taluks the forecasting part was carried out at different lead time (1-6 months) and the results Table 6 reveal that initially for all stations the forecast was observed to be good at 1 lead time with a correlation coefficient of 0.90, 0.93, 0.87 and 0.90 for Raichur, Manvi, Sindhanur and Lingasugur respectively. The RMSE and MAE were observed to be least at 1 leads and increases as the lead time increase, these stochastic models were found to suitable to forecast up to 1 lead time. A view at the Table 6 can be easily noticed that as the lead time increases the error rate showed increase tremendously. It can be easily concluded that Seasonal ARIMA models suited well for forecasting at 1-month lead time for Potential evapotranspiration forecasting under Raichur Region. Basic statistical properties are compared between observed and forecasted data for 1-month lead time, using t-test for the means and F-test for standard deviation (Haan 1977), shown in Table 7. Since t_{cal} values related to means were between *t*critical table values (± 1.71) for two tailed at a 5% significance level), the data shows that there is no significant difference between the mean values of observed and predicted data. Similarly, the F_{cal} values of standard deviation were smaller than the F-critical values at a 5% significance level. Thus, the results show that predicted data preserves the basic statistical properties of the observed series.

4. Conclusions

The Seasonal ARIMA models revealed that the models have an ability to forecast up to 1-month lead month with a higher accuracy over all the stations. Of the all stations, Seasonal ARIMA model provided excellent results at Sindhanuru station with an MAE, RMSE and MAPE values of 7.70, 10.52 and 5.28 respectively. The seasonal ARIMA models for different station are found to

forecasting potential evapotranspiration accurately up to one ahead with a least error. Similarly for the basic statistical analysis the difference between the observed and forecasted mean were found to be non-significant, which in turn reveal that the models are found to be quite promising in forecasting the potential evapotranspiration over the study period. The prediction of evapotranspiration guarantees reliable project planning, design and operating of irrigation systems.

Acknowledgements

The authors would like to thank Department of Soil and Water Engineering, College of Agricultural Engineering, University of Agricultural Sciences, Raichur for providing necessary facilities.

Disclaimer : The contents and views expressed in this research paper/article are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

References

- Aruna, K. T., Satish Kumar, U., Ayyanagowdar, M. S., Srinivasa Reddy, G.V. and Shanwad, U. K., 2017, "Estimation of crop coefficient (Kc) values for groundnut crop with evaluated crop evapotranspiration at different moisture deficit levels under agro climatic condition of Raichur, Karnataka", *Green Farming*, **8**, 5, 1161-1164.
- Asadi, A., Vahdat, S. F. and Sarraf, A., 2013, "The forecasting of potential evapotranspiration using time series analysis in humid and semi humid regions", *American Journal of Engineering Research*, **2**, 1, 296-302.
- Box, G. E. P. and Jenkins, G. M., 1976, "Time series analysis forecasting and control", Holden-Day, San Francisco.
- Box, G. E. P., Jenkins, G. M. and Reinsel, G. C., 1994, "Time Series Analysis, Forecasting and Control", (3rd Edition), Englewood Cliffs, N.J. Prentice Hall.
- Dutta, B., Smith, W. N., Grant, B. B., Pattey, E., Desjardins, R. L. and Li, C., 2016, "Model development in DNDC for the prediction of evapotranspiration and water use in temperate field cropping systems". *Environmental Modelling & Software*, **80,** 1, 9-25.
- Gorantiwar, S. D. and Patil, P. D., 2009, "Stochastic modelling for crop evapotranspiration for Rahuri Region", *International Journal of Agricultural Engineering*, **2**, 1, 140-145.
- Gorantiwar, S. D., Meshram, D. T. and Mittal, H. K., 2011, "Seasonal ARIMA model for generation and forecasting evapotranspirtion of Solapur district of Maharashtra" *Journal of Agrometeorology*, **12,** 2, 119-122.
- Haan, C. T., 1977, "Statistical methods in hydrology", Iowa State Press, Iowa.
- Hamdi, M. R., Bdour, A. N. and Tarawneh, A. N., 2008, "Developing reference crop evapotranspiration time series simulation model using class a pan : A case study for the Jordan Valley, Jordan", *Journal of Earth and Environmental Sciences*, **1**, 1, 33-44.
- Hsin-Fu Yeh and Hsin, Li Hsu., 2019, "Stochastic model for drought forecasting in the Southern Taiwan basin". *Water*, **11**, 10, 1-15.
- Mishra, A. K. and Desai, V. R., 2005, "Drought forecasting using stochastic models", *J. Stoch. Environ. Res. Risk Assess*., **19**, 1, 326-339.
- Mohan, S. and Arumugam, N., 1995, "Forecasting weekly reference crop evapotranspiration series". *Hydrological Sciences Journal*, **40**, 6, 689-702.
- Popale, P. G. and Gorantiwar, S. D., 2014, "Stochastic generation and forecasting of weekly rainfall for Rahuri Region", *International Journal of Innovative Research in Science, Engineering and Technology*, **3**, 4, 185-196.
- Salas, J. D., Delleure, J. W., Yevjevich, V. and Lane, W. L., 1980, "Applied Modeling of Hydrologic Time Series", *Water Res. Publ.*, Colorado, USA