



International Journal of Research in Agronomy

E-ISSN: 2618-0618
P-ISSN: 2618-060X
© Agronomy
NAAS Rating (2026): 5.20
www.agronomyjournals.com
2026; SP-9(6): 92-100
Received: 25-04-2026
Accepted: 29-05-2026

Harish H Deshpande
Assistant Professor (Agronomy),
Faculty of Agriculture, Warer and
Land Management Institute,
(WALMI), Chhatrapati
Sambhajanagar, Maharashtra,
India

Harshada Deshmukh
Professor and Head (Addl),
Faculty of Agriculture, Warer and
Land Management Institute,
(WALMI), Chhatrapati
Sambhajanagar, Maharashtra,
India

Goudra Pramod Gouda
Assistant Professor, Department of
Agricultural Engineering,
Agricultural and Horticultural
Research Station, Bavikere, Keladi
Shivappa Nayaka University of
Agricultural and Horticultural
Sciences, Shivamogga, Karnataka,
India

Corresponding Author:
Harish H Deshpande
Assistant Professor (Agronomy),
Faculty of Agriculture, Warer and
Land Management Institute,
(WALMI), Chhatrapati
Sambhajanagar, Maharashtra,
India

Seasonal time series forecasting of potential evapotranspiration in northern dry zone of Karnataka

Harish H Deshpande, Harshada Deshmukh and Goudra Pramod Gouda

DOI: <https://www.doi.org/10.33545/2618060X.2026.v9.i6Sb.5696>

Abstract

Accurate estimation and forecasting of potential evapotranspiration (PET) are essential for effective irrigation management, agricultural planning and sustainable water resource utilisation, particularly in semi-arid regions. Evapotranspiration is a major component of the hydrological cycle and significantly influences crop water requirements and regional water balance. The present study aimed to develop and evaluate Seasonal Autoregressive Integrated Moving Average (SARIMA) models for forecasting PET in Ballari district, Karnataka, India, covering the major taluks of Ballari, Siruguppa, Sandur, Kurugodu and Kampli. Monthly PET values were estimated using the Thornthwaite method based on maximum and minimum temperature data obtained from the India Meteorological Department (IMD) for the period 1990 to 2025. Data from 1990 to 2020 were used for model development, while data from 2021 to 2025 were utilised for model validation. The Box-Ljung and Augmented Dickey-Fuller (ADF) tests confirmed the presence of autocorrelation and seasonal stationarity in the PET time series data. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) analyses revealed strong seasonal behaviour, leading to the adoption of SARIMA models. Based on maximum log-likelihood and minimum Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values, the best-fit models identified were SARIMA (2,0,2)(2,1,0)₁₂ for Ballari, Sandur and Kurugodu stations, and SARIMA (1,0,0)(2,1,0)₁₂ for Siruguppa and Kampli stations. Diagnostic checking confirmed that the residuals behaved like white noise, indicating the adequacy of the selected models. Forecast evaluation showed that the developed models performed satisfactorily for short-term forecasting, particularly at a 1-month lead time, with comparatively lower RMSE, MAE and MAPE values. Statistical validation using t-tests and F-tests further indicated no significant difference between observed and forecasted PET series. The study demonstrated that SARIMA models are effective tools for short-term PET forecasting and can support irrigation scheduling, water resource management and climate-resilient agricultural planning in semi-arid regions such as Ballari district.

Keywords: PET, SARIMA model, time series forecasting, box-Jenkins approach, thornthwaite method

1. Introduction

Evapotranspiration (ET) is one of the major components of the hydrological cycle and plays a critical role in agricultural water balance and watershed hydrology (Wanniarachchi & Sarukkalige, 2022) ^[22]. A substantial proportion of precipitation received over land surfaces returns to the atmosphere through the combined processes of evaporation and transpiration (Asadi *et al.*, 2013; Patil *et al.*, 2022) ^[2, 16]. Globally, nearly 60% of annual precipitation is lost through ET, making its accurate estimation essential for efficient crop production, irrigation planning, hydrological modelling and sustainable water resource management (Aruna *et al.*, 2017) ^[1]. As a dominant factor influencing field water balance, ET directly affects crop water requirement, irrigation scheduling and agricultural productivity (Hussain *et al.*, 2025) ^[12]. The quantity of water required to satisfy evapotranspiration demand largely determines both the quantity and quality of crop yield. Therefore, reliable estimation and forecasting of ET are fundamental for the optimal design and management of irrigation systems, especially in water-scarce and semi-arid regions.

The ET process is primarily governed by meteorological parameters such as air temperature, relative humidity, solar radiation and wind speed. Due to the complex interactions among these climatic variables and their temporal variability, stochastic time series modelling has emerged as an effective approach for analysing and forecasting evapotranspiration. Among various

forecasting techniques, the Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA) models are widely recognised for their robustness and efficiency in modelling hydrological and meteorological time series data. These models have been extensively applied in forecasting problems related to hydrology, agriculture, economics, engineering and environmental sciences (Khashei *et al.*, 2009; Gao *et al.*, 2021; Premkumara *et al.*, 2025) [7, 13, 14, 18]. ARIMA-based stochastic models are particularly suitable for hydrological applications because of their ability to capture seasonality, temporal dependence and random fluctuations in climatic variables. These models forecast future values as a linear function of past observations and forecast errors, thereby providing reliable short-term predictions for time-dependent hydrological processes (Torres *et al.*, 2011; Gorlapalli *et al.*, 2022) [9, 21]. In agricultural water management, ARIMA and SARIMA models are highly useful for generating synthetic datasets, improving irrigation planning and supporting real-time operation of irrigation systems.

Several studies have demonstrated the applicability of ARIMA and SARIMA models in forecasting hydrological parameters. Popale and Gorantiwar (2014) [17] successfully applied ARIMA models for rainfall forecasting in the Rahuri region, while other researchers utilised these models for analysing evapotranspiration and climatic variables (Rahul *et al.*, 2020; Shrikant *et al.*, 2023) [20]. Seasonal ARIMA models have also shown promising performance in forecasting ET across diverse climatic conditions (Hamdi *et al.*, 2008; Asadi *et al.*, 2014; Damor *et al.*, 2024; Waris *et al.*, 2024; Salas *et al.*, 1980) [2, 5, 11, 19, 23].

Ballari district of Karnataka is characterised by a semi-arid climate with high temperature variability, erratic rainfall patterns and increasing pressure on available water resources due to intensive agricultural activities. The district frequently experiences water stress conditions, making efficient irrigation management and accurate estimation of crop water requirements extremely important for sustainable agricultural production. Since evapotranspiration accounts for more than 95% of crop water requirement, understanding its temporal variation and forecasting future trends are essential for effective water resource utilisation and climate-resilient agricultural planning. In this context, the present study aims to develop and evaluate Seasonal Auto-Regressive Integrated Moving Average (SARIMA) models for forecasting potential evapotranspiration

(PET) in Ballari district, Karnataka, India. PET was estimated using the Thornthwaite method based on long-term temperature data for different locations namely Ballari, Siruguppa, Sandur, Kurugodu and Kampli. The study intends to assess the suitability and predictive performance of SARIMA models for short-term PET forecasting and to support informed decision-making in irrigation scheduling, command area management and sustainable agricultural water resource planning.

2. Materials and methodology

The study was conducted in Ballari district, located in the northeastern part of Karnataka, India. Geographically, the district lies between 14.5° to 15.5° N latitude and 76.2° to 77.1° E longitude, covering an area of approximately 4,225.98 km². The study area comprises the major taluks of Siruguppa, Ballari, Sandur, Kurugodu and Kampli, as shown in Fig 1. Ballari district forms part of the Northern Dry Zone agro-climatic region of Karnataka, which is characterised by semi-arid climatic conditions, low to moderate rainfall, high temperature variability and frequent drought occurrences. Agriculture is one of the principal livelihood activities in the region and is predominantly dependent on monsoon rainfall and available irrigation sources. The major crops cultivated in the study area include paddy, cotton, groundnut, sunflower, maize, jowar, red gram and chilli. In addition to agriculture, the district also holds significant importance in Karnataka due to its mineral resources and mining-based economic activities.

The climate of the region is generally hot and dry, with high evapotranspiration demand throughout most of the year. Summer temperatures frequently rise above 40 °C, while comparatively lower temperatures are observed during the winter season. The variability in rainfall and increasing pressure on water resources make efficient irrigation management essential for sustainable agricultural production in the district. The semi-arid environmental conditions of Ballari district make it highly suitable for studying potential evapotranspiration (PET) and its temporal variability. In the present study, potential evapotranspiration (PET) was estimated using the Thornthwaite method based on long-term temperature data. For forecasting purposes, time series analysis was performed using Seasonal Autoregressive Integrated Moving Average (SARIMA) models. The adopted approach enabled the assessment of temporal trends and the development of predictive models for PET forecasting across different locations of the district.

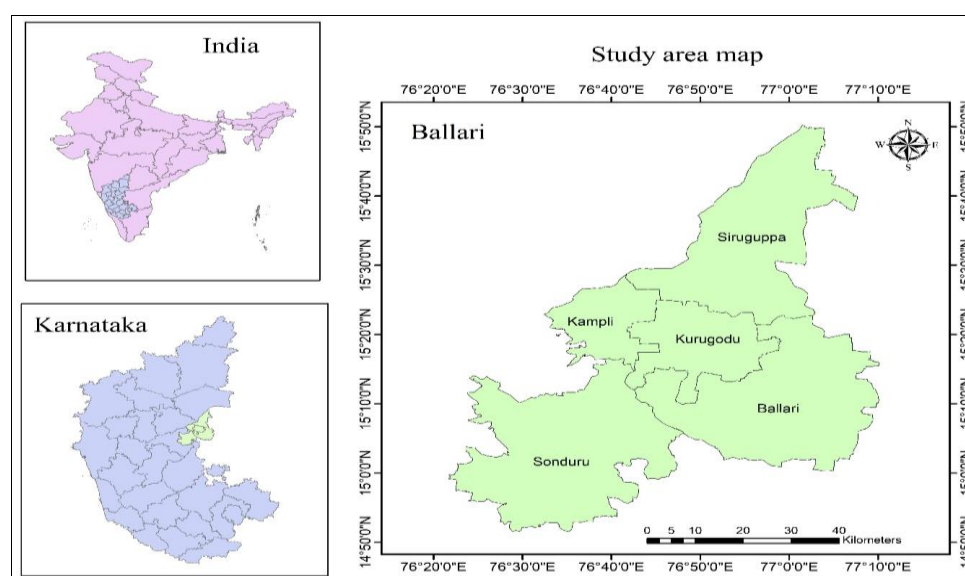


Fig 1: Study area

2.1 Estimation of potential evapotranspiration using the thornthwaite method

In the present study, potential evapotranspiration (PET) was estimated using the Thornthwaite method, which is one of the widely used temperature-based empirical approaches for estimating atmospheric water demand. The method primarily utilises mean monthly air temperature data and incorporates corrections for day length and month duration based on the latitude of the study area. Due to its simplicity and minimal data requirement, the Thornthwaite method is particularly suitable for regions where comprehensive meteorological data such as solar radiation, humidity and wind speed are not readily available.

The Thornthwaite method estimates monthly PET as a function of mean air temperature and heat index, while accounting for seasonal variations in daylight hours. The method assumes that air temperature is the dominant climatic factor influencing evapotranspiration. The monthly PET values were computed using the standard Thornthwaite equation, adjusted for the geographical latitude of the study area. The mathematical expression used for PET estimation is given as

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

Where, (T) represents the mean monthly air temperature (°C), I denotes the annual heat index computed as the summation of the monthly heat index values for all twelve months and m is an empirical exponent that varies as a function of the annual heat index. The parameter (K) is a correction coefficient introduced to account for variations in day length and the number of days in a month, which are dependent on the geographical latitude of the study area.

The annual heat index reflects the thermal characteristics of the region and plays an important role in determining the evapotranspiration demand. The exponent (m) was estimated using an empirical polynomial relationship proposed by Thornthwaite, which is expressed as follows

$$m = 6.75 \times 10^{-7} \cdot I^3 - 7.71 \times 10^{-7} \cdot I^2 + 1.79 \times 10^{-2} \cdot I + 0.492$$

The correction coefficient (K) accounts for seasonal variations in day length and the number of days in each month, which vary according to the geographical latitude of the study area. This adjustment improves the accuracy of PET estimation by incorporating the influence of solar exposure and photoperiod on evapotranspiration processes throughout the year. The Thornthwaite method provides a simplified temperature-based estimation of potential evapotranspiration and is widely adopted in climatological, hydrological and agricultural water management studies. Owing to its minimal data requirements and computational simplicity, the method is particularly suitable for regions where comprehensive meteorological observations such as solar radiation, humidity and wind speed are limited or unavailable. Therefore, the Thornthwaite approach serves as an effective and practical tool for analysing long-term evapotranspiration trends and supporting water resource planning in data-scarce regions.

2.2 Autocorrelation test (Box-Ljung Test)

The Box-Ljung test was employed in the present study to examine the presence of autocorrelation in the residuals of the developed ARIMA/SARIMA models and to evaluate the adequacy of model fitting. The test assesses whether the residuals are independently distributed or exhibit significant

serial correlation over different lag periods. The null hypothesis (H_0) states that the residuals are independently distributed and that the fitted model adequately represents the observed time series data, indicating the absence of autocorrelation. In contrast, the alternative hypothesis (H_a) suggests the presence of autocorrelation in the residuals, implying that the model fails to capture all the underlying information in the series. The Box-Ljung test statistic is computed using the autocorrelation coefficients of residuals at various lag intervals. A non-significant p-value ($p > 0.05$) indicates acceptance of the null hypothesis, confirming that the residuals behave like white noise and that the fitted model is statistically adequate. Conversely, a significant p-value ($p < 0.05$) indicates rejection of the null hypothesis and suggests the presence of residual autocorrelation, thereby signifying inadequacy of the developed model. Hence, the Box-Ljung test serves as an important diagnostic tool for validating the reliability and forecasting performance of ARIMA/SARIMA time series models.

2.3 Stationarity test (Augmented Dickey-Fuller Test)

The stationarity of the time series data was examined using the Augmented Dickey-Fuller (ADF) test prior to model development. Stationarity is an essential requirement in ARIMA/SARIMA modelling because non-stationary data can lead to unreliable parameter estimation and inaccurate forecasting results. A stationary time series is characterised by constant statistical properties such as mean, variance and autocorrelation over time. The null hypothesis (H_0) of the Augmented Dickey-Fuller test states that the time series contains a unit root and is therefore non-stationary. The alternative hypothesis (H_1) indicates that the series is stationary and does not contain a unit root. The test was applied to evaluate whether differencing was required to transform the series into a stationary form suitable for time series modelling. A statistically significant p-value ($p < 0.05$) leads to the rejection of the null hypothesis, indicating that the time series is stationary. On the other hand, a non-significant p-value ($p > 0.05$) suggests that the series is non-stationary and requires differencing or transformation before fitting the ARIMA/SARIMA model. The ADF test therefore plays a crucial role in identifying the order of differencing necessary for developing an appropriate and reliable forecasting model.

2.4 Description of the stochastic models

Stochastic models, also known as time series models, are widely used for analysing and forecasting temporal data in fields such as hydrology, agriculture, economics and engineering. These models help in understanding the underlying patterns, trends and seasonal variations present in time-dependent datasets. In hydrological studies, stochastic models are particularly useful for analysing complex and variable climatic parameters such as rainfall and evapotranspiration. By utilising historical observations, these models can generate reliable forecasts and evaluate prediction accuracy. Among the different stochastic approaches, ARIMA and SARIMA models are extensively used due to their capability to capture temporal dependence, seasonality and random fluctuations in the data series, thereby supporting effective water resource planning and irrigation management (Mishra and Desai, 2005) ^[15].

2.5 ARIMA models

Autoregressive Integrated Moving Average (ARIMA) models are among the most widely used stochastic approaches for time series analysis and forecasting. These models combine the Autoregressive (AR) and Moving Average (MA) components

and are particularly effective for modelling temporal data exhibiting autocorrelation. ARIMA models are generally applied to stationary time series data; therefore, non-stationary series are transformed into stationary form through differencing before model development. In an ARIMA model, the present value of a time series is expressed as a linear combination of its previous observations, past forecast errors and a random error component. The general ARIMA model is represented as ARIMA (p,d,q), where p denotes the order of the autoregressive component, (d) represents the degree of differencing required to achieve stationarity and (q) indicates the order of the moving average component. The ARIMA modelling framework, originally proposed by Box and Jenkins (1976), is widely used for forecasting hydrological and climatological variables because it relies on the intrinsic stochastic characteristics of the observed time series without requiring external explanatory variables. The general form of a non-seasonal ARIMA model can be expressed as:

$$\phi(B)\nabla_z^d = \theta(B)a_t$$

Where z_t represents the original time series, ∇^d denotes the differencing operator applied (d) times to achieve stationarity and a_t represents the random white noise error term. The terms $\phi(B)$ and $\theta(B)$ are polynomials of the backshift operator (B), corresponding to the autoregressive (AR) and moving average (MA) components of the model, respectively. The autoregressive component describes the dependence of the present observation on its previous lagged values, whereas the moving average component accounts for the influence of past random errors. The autoregressive operator of order (p) is expressed as

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$$

The moving average operator of order (q) is expressed as a polynomial function of the backshift operator and represents the dependence of the current observation on the previous random error terms. It accounts for the short-term random fluctuations present in the time series and improves the forecasting capability of the model by incorporating past error structures. The moving average operator of order (q) is expressed as

$$\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$$

ARIMA models provide a flexible and efficient approach for modelling and forecasting time-dependent data by describing the current value of a variable as a function of its past observations and stochastic error terms. These models are capable of capturing temporal dependence, trends and random variations present in the time series, making them highly suitable for hydrological and climatological applications. Due to their strong forecasting capability and adaptability, ARIMA models are widely used for analysing complex environmental datasets and supporting effective decision-making in water resource and agricultural management (Mishra and Desai, 2005) ^[15].

2.6 Seasonal ARIMA models

Many hydrological and climatological time series exhibit cyclic or seasonal behaviour, often associated with annual variations caused by the Earth's rotation around the sun. Such series are generally cyclically non-stationary and require specialised modelling approaches to adequately capture both seasonal and non-seasonal patterns. After removing deterministic seasonal effects from the data, the ARIMA approach can be effectively

applied to model the stochastic component of the series (Gorantiwar *et al.*, 2011; Mallikarjun *et al.*, 2023) ^[8, 14]. To address seasonality in time series data, Box *et al.* (1994) extended the ARIMA framework and developed the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. SARIMA models are particularly advantageous because they can efficiently represent both seasonal and non-seasonal variations using a limited number of parameters. The general SARIMA model is denoted as ARIMA (p,d,q) (P,D,Q)_s, where (p,d,q) represents the non-seasonal autoregressive, differencing and moving average components, while (P,D,Q)_s represents the corresponding seasonal components with seasonal period (s). These models are widely used in hydrological forecasting due to their capability to capture temporal dependence, seasonality and stochastic variations present in climatic and environmental time series data.

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

where (p) is the order of the non-seasonal autoregressive component, (d) is the degree of regular differencing, (q) is the order of the non-seasonal moving average component, (P) is the order of the seasonal autoregressive component, (D) is the degree of seasonal differencing, (Q) is the order of the seasonal moving average component and (s) represents the seasonal period of the time series.

2.6.1 Model identification

The first step in ARIMA/SARIMA modelling involves identifying an appropriate model structure that best represents the behaviour of the time series data. This was accomplished by analysing the patterns of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF), which are widely used to determine the suitable orders of the autoregressive (p) and moving average (q) components (Mishra and Desai, 2005; Yeh and Hsu, 2019) ^[15, 24]. The patterns, spikes and cut-off behaviour observed in the ACF and PACF plots provide important diagnostic information regarding the potential model structure. Based on these analyses, several tentative ARIMA/SARIMA models were developed and subsequently evaluated using statistical selection criteria such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), also known as the Schwarz Bayesian Criterion (SBC). The model with the lowest AIC and BIC values was considered the most appropriate, as these criteria provide an optimal balance between model accuracy and complexity.

The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are statistical measures used for selecting the most appropriate ARIMA/SARIMA model. Lower AIC and BIC values indicate a better balance between model accuracy and complexity. The AIC and BIC are calculated using the following expressions

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

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

Where (k) represents the number of estimated parameters in the model, (L) denotes the likelihood function of the fitted ARIMA/SARIMA model and (n) is the total number of observations in the time series dataset.

2.6.2 Parameter estimation

After identifying the appropriate ARIMA/SARIMA model structure, the next step involved estimating the model

parameters. The autoregressive (AR) and moving average (MA) parameters were estimated using the Maximum Likelihood Estimation (MLE) method, which provides efficient and reliable parameter estimates for time series models. The estimated parameters were subsequently evaluated for their statistical significance to ensure the suitability and adequacy of the fitted model for forecasting purposes.

2.6.3 Diagnostic checking

Diagnostic checking is the final and essential step in ARIMA/SARIMA model development, performed to evaluate the adequacy and reliability of the fitted model. This process involves examining whether the residuals behave like white noise, meaning that they are independently distributed, normally distributed and free from significant autocorrelation. In the present study, the Residual Autocorrelation Function (RACF) was analysed to assess the randomness and independence of the residuals. The absence of significant autocorrelation in the residual series indicated that the model successfully captured the underlying structure and seasonal behaviour of the time series data. Thus, residuals exhibiting white noise characteristics confirmed the suitability and adequacy of the developed ARIMA/SARIMA models for forecasting potential evapotranspiration.

2.6.4 PET forecasting

Potential evapotranspiration (PET) was forecasted for lead periods ranging from 1 to 6 months using the best-fit ARIMA/SARIMA models developed from historical time series data. For each lead time, the forecasted PET values were compared with the corresponding observed values to evaluate the consistency and predictive capability of the developed models. For instance, a one-month lead forecast generated for January 2017 corresponds to the predicted PET value for February 2017. The performance of the forecasting models was assessed using statistical evaluation measures such as the correlation coefficient (R), root mean square error (RMSE) and mean absolute error (MAE) between the observed and predicted PET values across all lead periods. These statistical indicators were used to determine the accuracy, reliability and forecasting efficiency of the developed models.

2.7 Input Dataset and software

The time series dataset comprising maximum and minimum temperature data for the period 1990 to 2025 was obtained from the India Meteorological Department (IMD) gridded meteorological database. Data from 1990 to 2020 were utilised for model development and calibration, while the period from 2021 to 2025 was used for model validation and forecasting performance assessment. Potential evapotranspiration (PET) was estimated using the Thornthwaite method in Microsoft Excel, whereas the development, analysis and implementation of SARIMA models were carried out using the statistical software R Studio.

3. Results and discussion

The model development process commenced with preliminary statistical tests to examine the autocorrelation and stationarity characteristics of the PET time series data. Autocorrelation in the datasets was assessed using the Box-Ljung test, and the corresponding Chi-square statistics along with p-values are presented in Table 1. The calculated Chi-square values were 183.07, 172.19, 188.12, 188.60 and 181.19 for Ballari, Siruguppa, Sandur, Kurugodu and Kampli stations, respectively,

with p-values less than 0.001 for all locations. These highly significant results confirmed the presence of autocorrelation in the PET series, indicating the suitability of stochastic time series modelling approaches.

Further, the stationarity of the time series was examined using the Augmented Dickey-Fuller (ADF) test, and the results are presented in Table 2. The Dickey-Fuller test statistics obtained were -17.754, -18.721, -15.306, -16.598 and -16.721 for Ballari, Siruguppa, Sandur, Kurugodu and Kampli stations, respectively, with corresponding p-values of 0.01 for all locations. The significant p-values indicated that the differenced PET series attained stationarity and were suitable for ARIMA/SARIMA model development. The results also revealed the presence of seasonal behaviour in the datasets, necessitating seasonal differencing prior to model fitting.

The principal step in the Box-Jenkins ARIMA modelling approach involves identifying an appropriate model structure that best represents the characteristics of the time series data. Different combinations of autoregressive (AR) and moving average (MA) parameters were evaluated, and the models yielding the highest log-likelihood values along with the lowest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values were selected as the most suitable models. The results of model identification and selection for Ballari, Siruguppa, Sandur, Kurugodu and Kampli stations are presented in Tables 3 and 4. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots presented in Figs. 2 and 3 exhibited significant seasonal patterns and serial dependence in the PET time series data. The ACF plot showed gradually decaying seasonal spikes, while the PACF plot displayed significant spikes at initial and seasonal lags, indicating the presence of seasonal autocorrelation in the datasets. These characteristics confirmed the necessity of applying seasonal differencing and adopting Seasonal ARIMA (SARIMA) models for PET forecasting.

Based on the model evaluation criteria, the best-fitting models identified were SARIMA (2,0,2)(2,1,0)₁₂ for Ballari, Sandur and Kurugodu stations, and SARIMA (1,0,0)(2,1,0)₁₂ for Siruguppa and Kampli stations. These models produced maximum log-likelihood values of -1688.64, -1635.32, -1640.64, -1593.37 and -1635.41 for Ballari, Siruguppa, Sandur, Kurugodu and Kampli stations, respectively. The corresponding AIC values were 3411.82, 3299.43, 3315.48, 3220.93 and 3299.43, while the BIC values were 3441.39, 3317.46, 3345.06, 3250.51 and 3317.46, respectively. The estimated model parameters and their statistical significance are presented in Table 4. Most of the autoregressive, moving average and seasonal parameters were found to be statistically significant, indicating the adequacy of the selected models in capturing the temporal and seasonal variability of PET across the study locations.

Further, diagnostic checking of the developed SARIMA models was carried out by analysing the residual series using the Box-Ljung test, and the results are presented in Table 5. The calculated Chi-square values were 2.20, 0.05, 0.01, 2.24 and 2.14 for Ballari, Siruguppa, Sandur, Kurugodu and Kampli stations, respectively, with corresponding p-values greater than 0.05 for all stations. The non-significant p-values indicated the absence of autocorrelation in the residuals, confirming that the residual series behaved like white noise. This demonstrated that the selected SARIMA models adequately captured the structure and seasonality of the PET time series data and were therefore suitable for reliable forecasting applications.

Following the development of SARIMA models for the five stations, forecasting of potential evapotranspiration (PET) was carried out for lead times ranging from 1 to 6 months. The

forecasting performance of the developed models is presented in Table 6. The results indicated that the 1-month lead forecasts exhibited comparatively better performance across all stations, with lower values of root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). The corresponding correlation coefficients ((R)) also indicated strong agreement between the observed and forecasted PET values during shorter lead periods. For Ballari station, the SARIMA model produced RMSE, MAE and MAPE values of 17.69, 13.90 and 23.57%, respectively, for the 1-month lead forecast. Similarly, Siruguppa station showed comparatively better forecasting accuracy with RMSE, MAE and MAPE values of 10.97, 9.25 and 9.53%, respectively. Sandur, Kurugodu and Kampli stations also demonstrated satisfactory forecasting performance at shorter lead times. However, the forecasting errors increased progressively with increasing lead time, particularly beyond 3 months, indicating a reduction in forecasting accuracy for longer prediction horizons. This behaviour suggests that the developed SARIMA models are more suitable for short-term PET forecasting, especially for 1-month lead predictions.

To further validate the forecasting capability of the developed models, the statistical characteristics of the observed and forecasted PET series at the 1-month lead time were compared using the t-test for mean values and the F-test for variance, following the procedure suggested by Haan (1977) [10]. The results are presented in Table 7. The calculated t-values for Ballari, Siruguppa, Sandur, Kurugodu and Kampli stations were 1.02, 0.16, 0.008, -0.076 and -0.22, respectively, which were all within the critical limit of ± 1.71 at the 5% significance level. This indicated that there was no significant difference between the observed and forecasted mean PET values. Similarly, the calculated F-values for variance comparison were lower than the critical value of 4.05 for all stations, confirming that the variability of the forecasted PET series closely matched that of the observed data. These results demonstrated that the developed SARIMA models effectively preserved the statistical characteristics of the original PET series and provided reliable short-term forecasts for the study area.

4. Conclusion: The present study demonstrated the applicability and effectiveness of Seasonal Autoregressive Integrated Moving

Average (SARIMA) models for forecasting potential evapotranspiration (PET) in Ballari district, Karnataka. PET was estimated using the Thornthwaite method based on long-term maximum and minimum temperature data obtained from the India Meteorological Department (IMD) for the period 1990-2025. The PET time series exhibited significant autocorrelation and seasonal behaviour, as confirmed through the Box-Ljung and Augmented Dickey-Fuller (ADF) tests, indicating the suitability of stochastic seasonal time series modelling approaches. The analysis of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots revealed strong seasonal patterns in the PET datasets across all stations. Based on model selection criteria such as maximum log-likelihood and minimum AIC and BIC values, the most suitable SARIMA models identified were SARIMA (2,0,2)(2,1,0)₁₂ for Ballari, Sandur and Kurugodu stations, and SARIMA (1,0,0)(2,1,0)₁₂ for Siruguppa and Kampli stations. The estimated model parameters were found to be statistically significant, and diagnostic checking confirmed that the residuals behaved like white noise, indicating the adequacy and reliability of the fitted models.

Forecasting performance evaluation showed that the developed SARIMA models provided satisfactory results for short-term PET forecasting. Among the different lead times analysed, the 1-month lead forecasts demonstrated the highest accuracy with comparatively lower RMSE, MAE and MAPE values and strong agreement between observed and forecasted PET values. However, forecasting errors increased progressively with increasing lead times, indicating a reduction in model performance for long-term predictions. Statistical validation using t-tests and F-tests further confirmed that the forecasted PET series preserved the mean and variability characteristics of the observed data. Overall, the study highlights the usefulness of SARIMA models as an effective tool for short-term PET forecasting in semi-arid regions such as Ballari district. The developed models can support irrigation scheduling, agricultural water management and climate-resilient decision-making by providing reliable PET forecasts. The findings of the study may also assist planners, researchers and water resource managers in improving water use efficiency and sustainable agricultural management under changing climatic conditions.

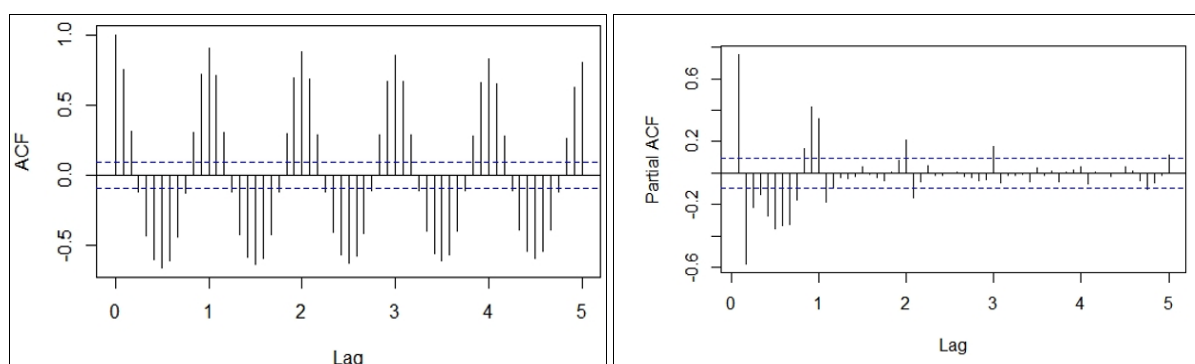


Fig. 2: Autocorrelation function plot of PET time series for Ballari Station, **Fig. 3:** Partial autocorrelation function plot of PET time series for Ballari Station

Table 1: Auto correlation test for different station

Station	Chi-Square	Lag order	P-value
Ballari	183.07	1	<0.001
Siruguppa	172.19	1	<0.001
Sandur	188.12	1	<0.001
Kurugodu	188.6	1	<0.001
Kampali	181.19	1	<0.001

Table 2: Stationarity test for different station

Station	Dickey fuller	Lag order	P-value
Ballari	-17.754	7	0.01
Siruguppa	-18.721	7	0.01
Sandur	-15.306	7	0.01
Kurugodu	-16.598	7	0.01
Kampli	-16.721	7	0.01

Table 3: Log likelihood AIC and BIC values of ARIMA model for different station

Stations	Model	Log-Likelihood	AIC	BIC
Ballari	SARIMA(2,0,2)(2,1,0) ₁₂	-1688.64	3411.82	3441.39
Siruguppa	SARIMA(1,0,0)(2,1,0) ₁₂	-1635.32	3299.43	3317.46
Sandur	SARIMA(2,0,2)(2,1,0) ₁₂	-1640.64	3315.48	3345.06
Kurugodu	SARIMA(2,0,2)(2,1,0) ₁₂	-1593.37	3220.93	3250.51
Kampli	SARIMA(1,0,0)(2,1,0) ₁₂	-1635.41	3299.43	3317.46

Table 4: Parameter estimation of SARIMA by maximum likelihood method for different station

Station	Model	Parameters	Estimate	S.E.	Z value	P-value
Ballari	SARIMA (2,0,2)(1,1,2) ₁₂	AR1	-0.242	0.156	-1.800	0.069
		AR2	0.625	0.129	5.176	< 0.001
		MA1	0.58	0.181	3.335	< 0.001
		MA2	-0.369	0.169	-2.466	0.013
		SAR1	0.276	0.335	0.802	0.429
		SMA1	-1.248	0.322	-4.085	< 0.001
Siruguppa	SARIMA (1,0,0)(2,1,0) ₁₂	SMA2	0.380	0.297	1.275	0.206
		AR1	0.656	0.123	5.747	< 0.001
		MA1	-0.259	0.155	-1.938	0.05
		SAR1	-0.727	0.059	-15.776	< 0.001
Sandur	SARIMA (2,0,2)(2,1,0) ₁₂	SAR2	-0.353	0.060	-7.782	< 0.001
		AR1	0.689	0.126	5.869	< 0.001
		MA1	-0.35	0.158	-2.522	0.01
		SAR1	0.121	0.320	0.338	0.74
		SMA1	-1.145	0.312	-3.876	< 0.001
Kurugodu	SARIMA (2,0,2)(2,1,0) ₁₂	SMA2	0.284	0.297	0.874	0.38
		AR1	0.644	0.134	5.120	< 0.001
		MA1	-0.296	0.164	-2.072	0.03
		SAR1	0.142	0.321	0.407	0.69
		SMA1	-1.157	0.311	-3.924	< 0.001
Kampli	SARIMA (1,0,0)(2,1,0) ₁₂	SMA2	0.277	0.297	0.914	0.36
		AR1	0.345	0.062	6.394	< 0.001
		SAR1	-0.738	0.066	-14.026	< 0.001
		SAR2	-0.272	0.065	-5.427	< 0.001

Table 5: Auto correlation check for residuals of Seasonal ARIMA model at different station

Station	Chi-Square	Lag order	P-value
Ballari	2.20	1	0.13
Siruguppa	0.05	1	0.72
Sandur	0.01	1	0.81
Kurugodu	2.24	1	0.12
Kampli	2.14	1	0.15

Table 6: Performance measure of Seasonal ARIMA models at different stations

Station	Model	Performance measures	Lead time					
Ballari	ARIMA (2,0,2)(1,1,2) ₁₂		1	2	3	4	5	6
		RMSE	17.69	25.03	25.52	25.05	23.25	22.37
		MAPE	23.57	14.79	14.48	13.96	15.25	14.15
		MAE	13.9	21.33	20.78	20.02	20.21	17.57
Siruguppa	ARIMA (1,0,1)(2,1,0) ₁₂	R	1.16	1.06	1.08	1.04	0.91	1.02
		RMSE	10.97	18.22	19.85	18.7	17.64	17.62
		MAPE	9.53	16.79	18.33	17.87	17.42	17.08
		MAE	9.25	16.15	16.91	16.16	15.24	14.08
Sandur	ARIMA (1,0,1)(1,1,2) ₁₂	R	1.19	1.06	1.07	1.05	0.92	1.02
		RMSE	10.78	20.45	39.77	52.36	59.71	61.68
		MAPE	5.54	10.78	22.73	29.42	33.6	38.07
		MAE	7.96	16	33.04	43.36	49.04	51.95

		R	1.13	1	0.93	0.92	0.76	0.76
Kurugodu	ARIMA (1,0,1)(2,1,0) ₁₂	RMSE	21.09	23.28	53.95	62.19	64.69	66.93
		MAPE	10.77	40.57	30.9	36.11	41.13	45.23
		MAE	15.85	33.77	44.82	51.5	54.99	56.54
		R	1.16	1.04	1.05	1.02	0.9	0.99
Kampli	ARIMA (1,0,0)(2,1,0) ₁₂	RMSE	21.09	23.28	53.95	62.19	64.69	66.93
		MAPE	10.77	40.57	30.9	36.11	41.13	45.23
		MAE	15.85	33.77	44.82	51.5	54.99	56.54
		R	1.16	1.04	1.05	1.02	0.9	0.99

Table 7: Comparison of statistic properties of the observed and predicted data

Stations	Mean observed	Mean forecasted	Decision (t<1.71)	Observed variance	Forecast variance	Decision (f < 4.05)
Ballari	140.22	135.42	1.02	1396.05	995.44	0.13
Siruguppa	100.11	98.56	0.16	643.79	548.864	0.0006
Sandur	139.1	137.77	0.008	1091.48	1399.05	0.0002
Kurugodu	139	138.19	-0.076	1260.84	1547.64	0.0003
Kampli	129.86	130.53	-0.22	1244.67	1358.44	0.0077

Disclaimer (Artificial Intelligence)

The authors hereby declare that generative Artificial Intelligence (AI) tools were used only for language refinement and grammatical improvement during the preparation of this manuscript. The scientific content, data analysis, interpretation of results and conclusions were entirely developed and verified by the authors. No AI tools were used for data generation, analysis or manipulation.

Competing interests

The authors have declared that no competing interests exist.

References

1. Aruna KT, Satish Kumar U, Ayyanagowdar MS, Srinivasa Reddy GV, Shanwad UK. 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*. 2017;8(5):1161-1164.
2. Asadi A, Vahdat SF, Sarraf A. The forecasting of potential evapotranspiration using time series analysis in humid and semi humid regions. *American Journal of Engineering Research*. 2014;2(12):296-302.
3. Box GEP, Jenkins GM. Time series analysis forecasting and control. San Francisco: Holden-Day; 1976.
4. Box GEP, Jenkins GM, Reinsel GC. Time series analysis, forecasting and control. 3rd ed. Englewood Cliffs (NJ): Prentice Hall; 1994.
5. Damor PA, Mod AA, Ram B, Parmar HV. Time series analysis and development of simulation model for monthly rainfall using ARIMA model. *International Journal of Agricultural Sciences*. 2024;20(1):226-235.
6. Dutta B, Smith VN, Grant BB, Pattey E, Desjardins RL, Li C. Model development in DNDC for the prediction of evapotranspiration and water use in temperate field cropping systems. *Environmental Modelling & Software*. 2016;80:9-25. DOI: 10.1016/j.envsoft.2016.02.004
7. Gao Y, Mosalam KM, Chen Y, Wang W, Chen Y. Auto-regressive integrated moving-average machine learning for damage identification of steel frames. *Applied Sciences*. 2021;11(13):6084. DOI: 10.3390/app11136084
8. Gorantiwar SD, Meshram DT, Mittal HK. Seasonal ARIMA model for generation and forecasting evapotranspiration of Solapur district of Maharashtra. *Journal of Agrometeorology*. 2011;13(2):119-122.
9. Gorlapalli A, Kallakuri S, Sreekanth PD, Patil R, Bandumula N, Ondrasek G, Rathod S. Characterization and prediction of water stress using time series and artificial intelligence sustainability models. *Sustainability*. 2022;14(11):6690. DOI: 10.3390/su14116690
10. Haan CT. Statistical methods in hydrology. Ames (IA): Iowa State Press; 1977.
11. Hamdi MR, Bdour AN, Tarawneh ZS. 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*. 2008;1(1):33-44.
12. Hussain S, Niyazi B, Elfeki A, Masoud M. Forecasting 21st century hydrological challenges in arid regions: Climate-induced shifts in runoff and evapotranspiration. *Water Resources Management*. 2025;1-31. DOI: 10.1007/s11269-025-030273
13. Khashei M, Bijari M, Ardali GAR. Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks (ANNs). *Neurocomputing*. 2009;72(4-6):956-967. DOI: 10.1016/j.neucom.2008.07.017
14. Mallikarjun Reddy B, Basavareddy R, Patil R, Guguloth P, Premkumara S. Forecasting potential evapotranspiration for Yadgir district, Karnataka, India using seasonal ARIMA model. *Ecology, Environment & Conservation*. 2023;29(November Suppl. Issue):S154-S160.
15. Mishra AK, Desai VR. Drought forecasting using stochastic models. *Stochastic Environmental Research and Risk Assessment*. 2005;19:326-339.
16. Patil R, Nagaraj DM, Polisgowdar BS, Rathod S. Forecasting potential evapotranspiration for Raichur district using seasonal ARIMA model. *Mausam*. 2022;73(2):433-440.
17. Popale PG, Gorantiwar SD. Stochastic generation and forecasting of weekly rainfall for Rahuri Region. *International Journal of Innovative Research in Science, Engineering and Technology*. 2014;3(4):185-196.
18. Premkumara, Praveen P, Pragna G, Reddy M, Anjanakumara G. Accuracy Assessment of Box-Jenkins SARIMA Model for Forecasting Potential Evapotranspiration in Bidar District, Karnataka, India. *Journal of Geography, Environment and Earth Science International*. 2025;29(6):62-71.
19. Salas JD, Delleure JW, Yevjevich V, Lane WL. Applied

- modeling of hydrologic time series. Littleton (CO): Water Resources Publications; 1980.
20. Shrikant RS, Manjunath KM, Patil R, Kulkarni PS. Predicting potential evapotranspiration for Kalaburagi district using a seasonal ARIMA model. *International Journal of Environment and Climate Change*. 2023;13(11):2073-2082.
 21. Torres AF, Walker WR, McKee M. Forecasting daily potential evapotranspiration using machine learning and limited climatic data. *Agricultural Water Management*. 98(4):553-562.
 22. Wanniarachchi S, Sarukkalige R. A review on evapotranspiration estimation in agricultural water management: Past, present, and future. *Hydrology*. 2022;9(7):123. DOI: 10.3390/hydrology9070123
 23. Waris U, Sarif S, Batool SA. Exploring association and forecasting of evapotranspiration based on meteorological factors over megacity Lahore (Pakistan) and central place of Indo-Gangetic Basin. *Environment, Development and Sustainability*. 2024;26(8):20255-20277.
 24. Yeh HF, Hsu HL. Stochastic model for drought forecasting in the Southern Taiwan Basin. *Water*. 2019;11(10):2041. DOI: 10.3390/w11102041