Original Research Article

Time series modelling of monthly temperature and reference evapotranspiration for Navsari (Gujarat), India.

Abstract

Time series modelling has been proved its usefulness in various fields including meteorology, hydrology and agriculture. It utilizes past data and extracts useful information from them to build up a model which could simulate various processes. The prior knowledge of evapotranspiration could help in estimating the amount of water required by the crops that is useful for optimizing design of irrigation systems. In this study, the time series modelling of monthly temperature and reference evapotranspiration has been carried out utilizing past data of 35 years (1983-2017) to assist decision makers related to agriculture and meteorology. 30 years (1983-2012) of temperature and evapotranspiration data were used for training and remaining 5 years of data (2013-2017) were used for validation. The monthly evapotranspiration was estimated using Penman-Monteith FAO-56 method. Mann-Kendall test was used at 5% significant level for identifying trend component in mean temperature. The time series of temperature and evapotranspiration was made stationary for modelling the stochastic components using ARIMA (Autoregressive Integrated Moving Average) model. In order to check the normality of residuals, the Portmantaeu test was applied. The time series models for temperature and evapotranspiration which were validated for 5 years (2013-2017) and further deployed for forecasting of 5 years (2018-2022). It was found that for modelling temperature and reference evapotranspiration for Navsari, seasonal ARIMA (1,0,0)(0,1,1)12 and seasonal ARIMA (1,0,1)(1,1,2)12 were found to be appropriate models respectively. Mann Kendall test used for trend detection in monthly mean temperature revealed that October and November months had significant positive trend. Negative trend was observed only in the month of June.

Keywords: *Autoregressive integrated moving average model, Mann-Kendall test, Penman-Monteith, Portmantaeu test, Reference evapotranspiration, Temperature, Time series modelling.*

Introduction

The time series modelling aims to collect and analyse past values for developing appropriate models that describe the inherent structure and characteristics of the series [1]. Time series forecasting is the use of a developed model to forecast values based on past observed values [2]. Autoregressive Integrated Moving Average (ARIMA) model is one of the most recognized statistical models for time series forecasting owing to its simplicity. Air temperature is a variable in meteorology which indicates how hot or cold the air is. The temperature has a major influence on other meteorological variables like evaporation, relative humidity, wind speed, wind direction and precipitation patterns. Reference crop evapotranspiration is the evapotranspiration rate from a reference surface. The reference surface closely resembles an extensive area of actively growing green grass of uniform height completely shading the ground and with unlimited availability of soil moisture. The concept of reference evapotranspiration provides information of the evaporation demand of the atmosphere, independent of the type of crop, its stage of development and the management practices. Reference crop evapotranspiration is one of the most important parameters for the efficient management of available water resources and it is also a major component of the water requirement of crops and governs irrigation scheduling. The long-term values of reference evapotranspiration are required for planning and management of water resources and irrigation scheduling. Many investigators have developed equations to estimate reference evapotranspiration, however, the reference evapotranspiration estimated by Penman-Monteith method is the most common method [3]. This method includes physiological and aerodynamic parameters and it is considered to be the most reliable method to estimate reference evapotranspiration under various climatic conditions.

The ARIMA (Auto Regressive Integrated Moving Average) models was used to carry out short-term predictions of monthly maximum and minimum temperatures in the Sylhet district in north-east Bangladesh using the classical Box-Jenkins methodology. Temperature data from the year 1977 to 2011 were used for formulating the seasonal ARIMA models and the verification of the models was done for the years 2010 to 2011. For the maximum and minimum temperatures at Sylhet station, ARIMA (1,1,1) (1,1,1)12 and ARIMA (1,1,1) (0,1,1)12 respectively, were obtained as the appropriate models. Using these ARIMA-models, one-month-ahead forecasts of the temperatures for years 2010 and 2011 was carried out [4].

The monthly mean temperature was analysed in Nanjing, China, from 1951 to 2017, using SARIMA (Seasonal Autoregressive Integrated Moving Average) techniques. Data from 1951 to 2014 was used as the training set, while data from 2015 to 2017 was used as the testing set. SARIMA (1, 1, 1) × (1, 0, 1)12 showed the lowest AIC value and thus was selected as the optimal model for forecasting. The mean square errors (MSE) of the predicted values from 2015, 2016 and 2017 were 0.84, 0.89 and 0.94 respectively [5].

The time series of reference crop evapotranspiration for Bokaro district in Jharkhand state, India using data series of 102 years (1901-2002). Maximum likelihood method was used for determining the parameters of the models. The best fitted model was found to be seasonal ARIMA (0, 1, 4) (0, 1, 1)12 which was used for forecasting a period of 24 months. High coefficient of determination value 0.9821 was obtained between observed and forecasted value [6].

ARIMA model was used for forecasting reference crop evapotranspiration of Solapur region, Maharashtra state, India using 33 years (1984-2016) of daily data. One year ahead forecast (i.e. for the year 2016) of reference evapotranspiration values were obtained with the help of these selected models and compared with the values of reference evapotranspiration obtained from the weather data of 2016 by root mean square error (RMSE).The results showed that seasonal ARIMA (0,0,1) (1,0,2)52 was the best model for forecasting of weekly evapotranspiration values [7].

Study area

Navsari is located at 20.9467° N latitude and 72.9520° E longitude in the South-eastern part of Gujarat state. It receives an average annual rainfall of 1621 mm. More than 90 % of the rainfall occurs in the monsoon months namely June, July, August and September. It has an average elevation of 9m above sea level. The average maximum and minimum temperatures are 40 °C and 17 °C respectively.

Materials and methods

The mean monthly temperature and evaporation data (1983-2017) obtained from agro-meteorological cell of Navsari Agricultural University were used for modelling and trend analysis. The data were used to obtain a series of monthly reference evapotranspiration by Penman-Monteith method. The trend of the mean monthly temperature series was then analysed using Mann-Kendall test.

.

FAO Penman-Monteith method

The data of meteorological variables like humidity, maximum temperature, minimum temperature, wind, sunshine hours and radiation data were used to calculate the reference evapotranspiration using FAO Penman Monteith method. Allen *et al.,* (1998) defined and published the FAO paper no. 56 the Penman-Monteith ET_0, as the rate from a hypothetical reference crop with an assumed crop height (12 cm), a fixed surface resistance (70 sm^{-1}) and albedo (0.23), closely resembling the ET from an extensive surface of green grass cover with adequate water [3]. The following equation was used to calculate the reference evapotranspiration.

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

Where, ET_0 = reference evapotranspiration [mm day^{-1}],
R_n = net radiation at the crop surface [MJ m^{-2} day^{-1}],
G = soil heat flux density [MJ m^{-2} day^{-1}],
T = mean daily air temperature at 2 m height [°C],u_2 is wind speed at 2 m height [m s^{-1}],
e_s = saturation vapour pressure [kPa],
e_a = actualvapour pressure [kPa],
e_s –e_a = saturation vapour pressure deficit [kPa],
Δ = slope vapour pressure curve [kPa $°C^{-1}$],
γ = psychrometric constant [k\Pa $°C^{-1}$].

Auto Regressive Integrated Moving Average (ARIMA) model

In this study, the time series modelling of monthly temperature and reference evapotranspiration was carried out utilizing past data of 35 years (1983-2017) out of which 30 years (1983-2012) of data were used for training and remaining 5 years of data (2013-2017) were used for validation. The analysis of time series was carried out using ARIMA model. In the ARIMA model, the first component is the autoregressive (AR) term, the second component is the integration (I) term which accounts for stabilizing or making the data stationary and the third component is the moving average (MA) term of the forecast errors. Box-Jenkins methodology (1976) was adopted in this study which involved model identification, parameter estimation and diagnostic check (residual analysis) [8].

The identification of model consisted of specifying the appropriate structure and order of model. First, the preliminary analysis of time series data was carried out to identify periodicities and significant spikes that reflect a non-stationary process inherent in the data. Stationarity was detected from an autocorrelation plot. The differencing approach was used to achieve stationarity. Once stationarity and seasonality were addressed, the next step was to identify the order of the autoregressive and moving average terms which was accomplished by observing significant peaks in partial autocorrelogram and autocorrelogram.

ARIMA model can be written as

$$\nabla^d X_t = (1-B)^d X_t \qquad (2)$$

Where B is the backshift operator: $BX_t = X_{t-1}$

Seasonal ARIMA model can be written as

$$\varphi(B)(1-B)^d X_t = \theta(B)\omega_t \qquad (3)$$

Let, $E(\nabla^d X_t) = \mu$

$$\varphi(B)(1-B)^d X_t = \alpha + \theta(B)\omega_t \qquad (4)$$

Where $\alpha = \mu(1-\varphi-\ldots-\varphi_p)$

The ARIMA model is also written as

$$W_t = \mu + \frac{\theta(B)}{\emptyset(B)} a_t \qquad (5)$$

where,

W_t= the response series $Y_t = (1-B)^d Y_t$

μ = the mean time of weekly parameter.

B = the backshift operator, that is, $BX_t = X_{t-1}$

$\emptyset(B) = 1 - \emptyset_1 B^1 - \ldots\ldots\ldots\ldots - \emptyset_p B^p$ = the autoregressive operator.

$\Theta(B) = 1 - \Theta_1 B^1 - \ldots\ldots\ldots\ldots - \Theta_q B^q$ = the moving average operator.

α_t = the independent disturbance (random error).

The seasonal ARIMA model is given as follows:

$$\Phi_P\,(B^s)\varphi_p(B)\nabla_s^D\,\nabla^d z_t = \theta_q\,(B)\Theta_Q(B_s)a_t \qquad (6)$$

$\Phi_P\,(B^s) = seasonal\ autoregressive\ operator\ of\ order\ P$

$\varphi_p = regular\ autoregressive\ operator\ of\ orer\ p$

$\nabla_s^D = seasonal\ differences$

$\nabla^d = regular\ differences$

$\Theta_Q(B_s) = seasonal\ moving\ average\ operator\ of\ order\ P$

$\theta_q\,(B) = regular\ moving\ average\ operator\ of\ order\ p$

$a_t = white\ noise\ process$

Table 1: Determination of model by using ACF and PACF patterns

MODEL	**ACF**	**PACF**
AR (*p*)	Dies down	Cut off after lag q
MA (*q*)	Cut off after lag p	Dies down
ARMA (*p, q*)	Dies down	Dies down

Parameters were estimated by Marquardt's algorithm available in statistical toolbox of Matlab software and the model was selected based on mean square error criteria. A portmanteau test, also called the Ljung-Box test was used for testing for autocorrelation in the residuals of a model. If the residuals were found to be significant then the model was rejected, and the test was conducted on other candidate models. The absence of any significant spikes in the residual ACF and PACF plots demonstrated proper fitting. For Ljung Box test [9], the null hypothesis was that the set of autocorrelations for residuals is white noise.

It was computed as:

$$X_m^2 = n(n+2)\sum_{k=1}^{m}\frac{r_k^2}{n-k} \qquad (7)$$

where n is the size of sample, r_k is the sample autocorrelation at lag k, and the m is the number of lags being tested .

Mann Kendall test

It is a non-parametric test used for the purpose of statistically assessing if there is a monotonic upward or downward trend of the variable of interest over time [10,11]. According to this test, the null hypothesis H_0 assumes that there is no trend (the data is independent and randomly ordered) and this is tested against the alternative hypothesis H_1, which assumes that there is a trend [12].

The Mann-Kendall statistic S is calculated as

$$S = \sum_{i=1}^{n-1}\sum_{j=j+1}^{n} sgn(x_j - x_i) \qquad (8)$$

x_i is ranked from i = 1,2,.......n-1

x_j, is ranked from j = i+1,2,......n.

$$sgn(x_j - x_i) = \begin{cases} 1\ if(x_j - x_i) > 0 \\ 0\ if\ (x_j - x_i) = 0 \\ -1\ if\ (x_j - x_i) < 0 \end{cases} \qquad (9)$$

For n ≥ 10, the statistic S is approximately normally distributed with the mean E(S)=0 and variance as follows:

$$Var\ (S) = \frac{n\ (n-1)(2n+5) - \sum_{i=1}^{m} t_i(i)(i-1)(2i+5)}{18} \qquad (10)$$

Where t_i is considered as the number of ties up to sample i.

The test statistics Z_s is computed as

$$Z_s = \begin{cases} \dfrac{S-1}{\sqrt{var(S)}}\ for\ S > 0 \\ 0\ for\ S = 0 \\ \dfrac{S+1}{\sqrt{var(S)}}\ for\ S < 0 \end{cases} \qquad (11)$$

The test statistic Zs is used a measure of significance of trend. In fact, this test statistic is used to test the null hypothesis, H_0. If $|Zs|$ is greater than $Z_{\alpha/2}$, where α represents the chosen significance level, then the null hypothesis is invalid implying that the trend is significant.

Results and Discussion

The time series of mean temperature and reference evapotranspiration was analysed for periodicities using autocorrelation plot. The presence of significant periodicities was evident from the autocorrelation plot shown in Fig. 1, therefore, the series was differenced to make it stationary. The autocorrelation plot of the differenced series is shown in Fig. 2 in which it could be seen that the number of significant periodicities was reduced.

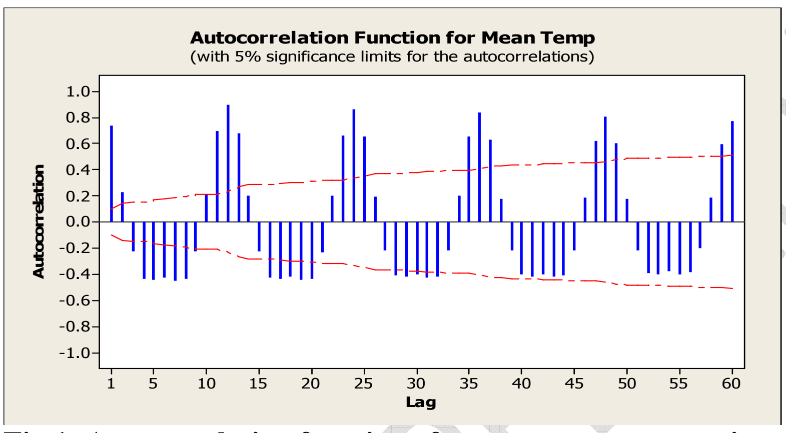

Fig 1: Autocorrelation function of mean temperature series

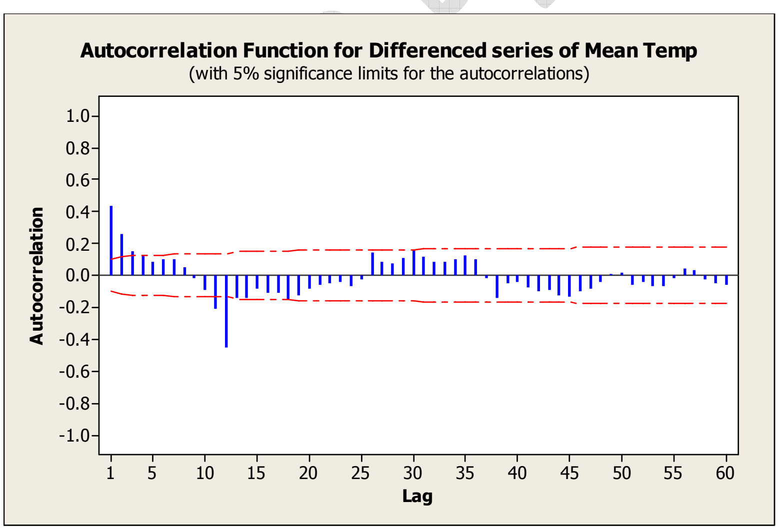

Fig 2: Autocorrelation function of the differenced series of mean temperature

Numerous ARIMA models were used to for modelling mean temperature and their residuals were also tested. Finally, it was found that seasonal ARIMA (1,0,0)(0,1,1)12 was found to be the appropriate model for mean temperature. The parameter estimates of the model are given in Table 2. One autoregressive term and one seasonal moving average term has been used in the model. The model was used for simulating the mean temperature series. The coefficient of determination values for training and testing period were respectively found to be 0.94 and 0.92 as shown in Fig. 3 and Fig. 4 respectively. The model was then tested for their residuals. Ljung box test revealed that the autocorrelations and partial autocorrelations were not found to be significant as shown in Fig. 7 Also, the histogram of the residual series and normal probability plot exhibited normal distribution as shown in Fig. 5 and Fig. 6 respectively. The seasonal ARIMA (1,0,0)(0,1,1)12 was then used to predict mean temperature for five years from the year 2018 to 2022 which is given in Appendix.

Table 2: Parameter estimates of seasonal ARIMA (1,0,0)(0,1,1)12

Terms	Estimate	Std. Error	T- value	P -value
AR - 1	0.4889	0.046	10.63	0.000
SMA - 12	0.9479	0.0238	39.9	0.000
Constant	0.009391	0.003916	2.4	0.017

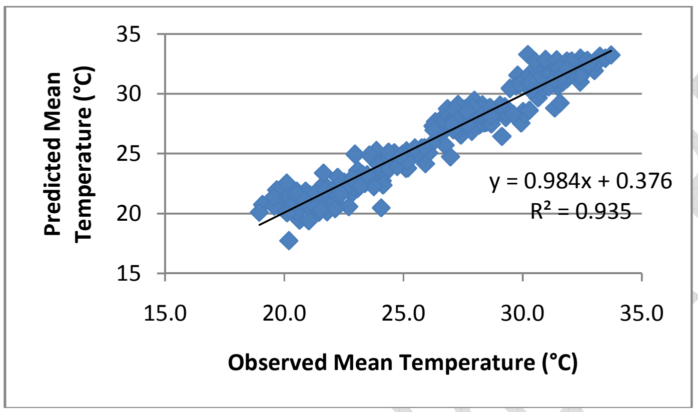

Fig. 3 : Observed vs Predicted Mean temperature (°C) for testing period (1983-2012)

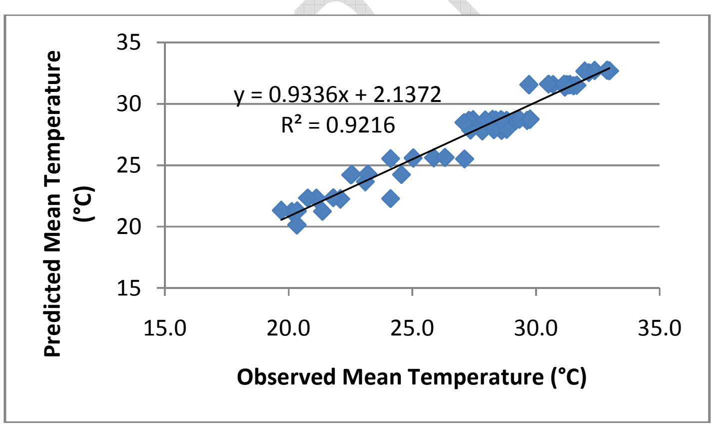

Fig. 4: Observed vs. Predicted Mean temperature (°C) for validation period (2013-2017)

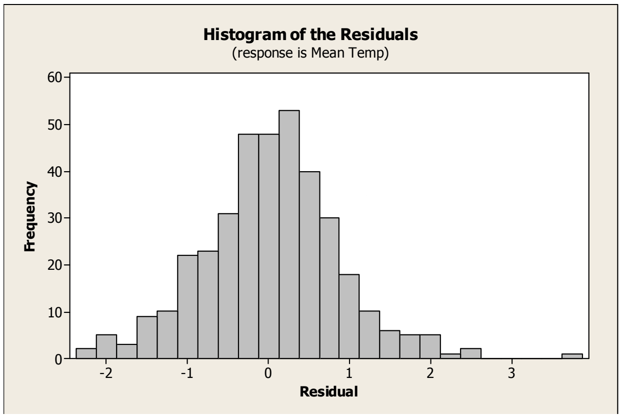

Fig. 5: Histogram of residuals

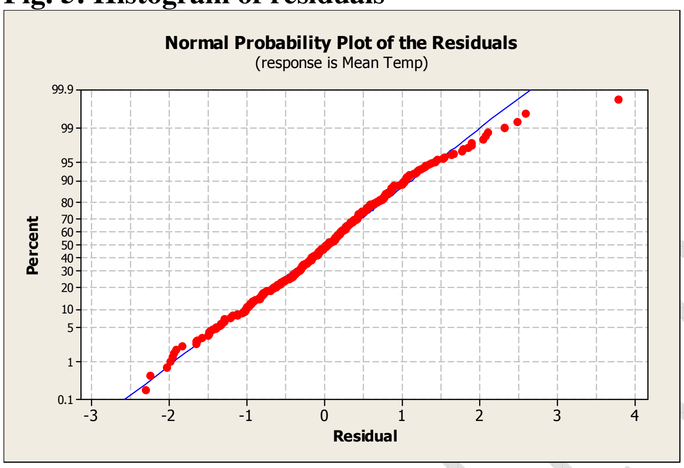

Fig. 6: Normal probability plot of the residuals

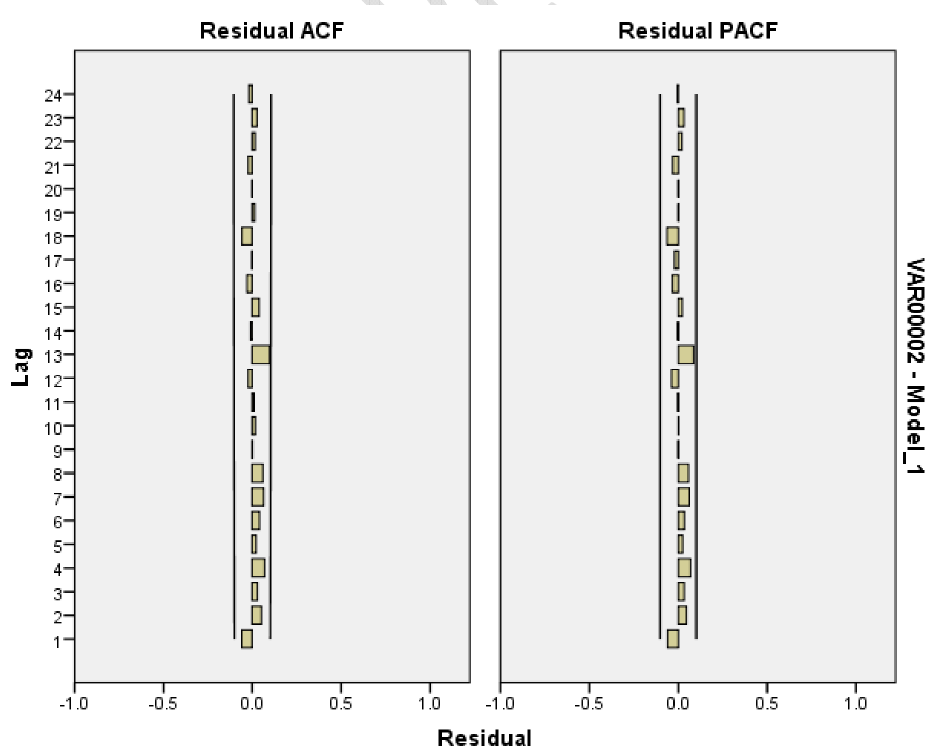

Fig.7: ACF and PACF plot of residuals

The modelling approach for reference evapotranspiration was similar to that of mean temperature; therefore, only the parameter estimates of the selected model for reference evapotranspiration is given in this study. The approach of differencing to make the series stationary was also carried out on the series of reference evapotranspiration. The seasonal ARIMA model (1,0,1)(1,1,2)12 was found to be the appropriate model for forecasting reference evapotranspiration. The coefficient of determination was found to be higher than 0.91 for training period and 0.86 for testing period. The model also passed the residual tests which included Ljung Box test and normality test. The model was then used to predict evapotranspiration for the next five years from the year 2018 to 2022. The parameter estimates are given for the seasonal ARIMA (1,0,1)(1,1,2)12 are given in Table 3.

Table 3: Parameter estimates of seasonal ARIMA (1,0,1)(1,1,2)12

Terms	Estimate	Std. Error	t-value	p-value
Constant	0.068	0.021	3.218	0.002
AR	0.907	0.046	19.895	0
MA	0.549	0.086	6.391	0
AR, Seasonal (1)	-0.725	0.295	-2.455	0.015
MA, Seasonal (1)	0.376	1.594	0.236	0.814
MA, Seasonal (2)	0.615	1.144	0.538	0.591

The performance measures of seasonal ARIMA model in terms of root mean square error (RMSE) and coefficient of determination (R^2) are for temperature as well as reference evapotranspiration are given in Table 4.

Table 4: Performance measures of seasonal ARIMA model for mean temperature and reference evapotranspiration

	Time series	**Mean Temperature**	**Reference Evapotranspiration**
Performance criteria	Period	Seasonal ARIMA (1,0,0)(0,1,1)12	Seasonal ARIMA (1,0,1)(1,1,2)12
Root Mean square error (RMSE)	Training	0.96	0.41
	Testing	0.98	0.53
Coefficient of determination (R^2)	Training	0.93	0.91
	Testing	0.92	0.86

The RMSE value and coefficient of determination values suggest that the ARIMA models selected in this study are reliable and they can be used for forecasting mean temperature and reference evapotranspiration. Gautam *et al.* (2016) obtained high coefficient of determination value of 0.9821 between observed and forecasted value for the testing period of 24 months [6].

Mann Kendall test was used for trend analysis of monthly mean temperature of Navsari and it was found that there was a significant positive trend for October and November month as indicated by the p-value. Mishra *et al.* (2004) utilized 101 years of temperature data for trend analysis of mean temperature by Mann Kendall test in Upper Ganga Canal Command and concluded that the annual mean temperatures increased by 0.60ºC in 101 years [13]. Chinchorkar *et al.* (2016) used Mann Kendall test for trend detection of mean monthly maximum temperature for Junagadh and it was found in 32 years, the highest increase in mean monthly maximum temperature occurred in November [14]. In this study, negative trend was found in the month of June while positive trend was found in the remaining months. The Mann Kendall test results for trend analysis of mean temperature are given in Table 5.

Table 5: Mann-Kendall test results for trend analysis of mean temperature

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
M-K Test Value (S)	46	99	82	79	37	-14	21	94	24	125	118	22
Critical Value (0.05)	1.65	1.65	1.65	1.65	1.65	-1.65	1.65	1.65	1.65	1.65	1.65	1.65
Standard Deviation of S	61.56	61.60	61.51	61.47	61.51	61.50	61.32	61.54	61.28	61.57	61.56	61.56
Standardized Value of S	0.73	1.59	1.32	1.27	0.59	-0.21	0.33	1.51	0.38	2.01	1.90	0.34
p-value	0.23	0.06	0.09	0.10	0.28	0.42	0.37	0.07	0.35	0.02	0.03	0.37

Conclusions

It was concluded that for modelling temperature and reference evapotranspiration for Navsari, seasonal ARIMA (1,0,0)(0,1,1)12 and seasonal ARIMA (1,0,1)(1,1,2)12 were found to be appropriate models respectively. The coefficient of determination for the testing period were found to be 0.92 and 0.86 respectively for mean monthly temperature and monthly reference evapotranspiration when the observed and predicted values were compared. The predicted temperature and reference evapotranspiration by ARIMA models can thus be used for irrigation planning and management. Mann Kendall test used for trend detection in monthly mean temperature revealed that October and November months had significant positive trend. Negative trend was observed only in the month of June.

References

1. Adhikari, R., & Agrawal, R. K. (2013). An introductory study on time series modeling and forecasting.

2. Raicharoen, T., Lursinsap, C., & Sanguanbhokai, P. (2003). Application of critical support vector machine to time series prediction. International Symposium on Circuits and Systems (Vol.5, pp.V-741-V-744 vol.5). IEEE.

3. Allen, R. G., Pereira, L. S., Raes, D., Smith, M. (1998). Crop Evapotranspiration, Guideline for Computing Crop Water Requirements. FAO Irrigation and Drainage Paper 56. FAO Rome, Italy. p. 300.

4. Nury, A. H., Koch, M., & Alam, M. J. B. (2013). Time Series Analysis and Forecasting of Temperatures in the Sylhet Division of Bangladesh. In *4th International Conference on Environmental Aspects of Bangladesh (ICEAB), August* (pp. 24-26).

5. Chen, P., Niu, A., Liu, D., Jiang, W., & Ma, B. (2018, July). Time Series Forecasting of Temperatures using SARIMA: An Example from Nanjing. In *IOP Conference Series: Materials Science and Engineering* (Vol. 394, No. 5, p. 052024). IOP Publishing.

6. Gautam, R., & Sinha, A. K. (2016). Time series analysis of reference crop evapotranspiration for Bokaro District, Jharkhand, India. *Journal of Water and Land Development*, *30*(1), 51-56.

7. Meshram, D. T., Gorantiwar, S. D., Sangale, U. R., Nagraj, B. K., & Pal, R. K. (2017). ARIMA model for forecasting reference crop evapotranspiration of Solapur region, Maharashtra, India, 7(2), 287-295.

8. Box, G. E., & Jenkins, G. M. (1976). Time Series Analysis: forecasting and control, rev. edn. *Holder-Day, San Francisco, CA, USA*.

9. Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*(2), 297-303.

10. Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society*, 245-259.

11. Kendall MG (1975) Rank correlation methods. Griffin, London12. Onoz, B., Bayazit, M., 2012. The Power of Statistical Tests for Trend Detection. Turkish Journal of Engineering & Environmental Sciences 27 (2003), 247 – 251.

13. Mishra, N., Khare, D., Shukla, R., & Kumar, K. (2014). Trend Analysis of Air Temperature Time Series by Mann Kendall Test-A Case Study of Upper Ganga Canal Command (1901-2002). *British Journal of Applied Science & Technology*, *4*(28), 4066.

14. Chinchorkar, S. S., Ram, D. M., & Trivedi, M. M. (2016). A case study of temperature and rainfall trends using Mann-Kendall test in Saurashtra Region (Junagadh) of Gujarat, India. *International Journal of Environmental Sciences*, *6*(6), 954-969.

Appendix

Predicted Mean Temperature and Reference Evapotranspiration (2018-2022)

	Mean Temperature (˚C)	Reference evapotranspiration (mm)
Jan-18	20.3	4.92
Feb-18	23.1	7.16
Mar-18	27.9	7.75
Apr-18	31.2	8.60
May-18	32.1	9.14
Jun-18	31.5	6.55
Jul-18	28.4	5.07
Aug-18	27.8	4.45
Sep-18	27.5	6.14
Oct-18	29.2	5.25
Nov-18	27.1	6.91
Dec-18	22.1	5.22
Jan-19	20.1	4.82
Feb-19	22.5	6.80
Mar-19	27.3	7.49
Apr-19	31.3	8.71
May-19	32.0	8.64
Jun-19	31.7	6.58
Jul-19	29.6	5.41
Aug-19	28.6	4.22
Sep-19	28.2	6.17
Oct-19	28.8	5.49
Nov-19	24.1	6.07
Dec-19	24.1	5.37
Jan-20	21.4	4.82
Feb-20	24.6	6.75
Mar-20	28.8	8.17
Apr-20	31.4	8.36
May-20	33.0	8.48
Jun-20	29.7	6.14
Jul-20	27.5	5.44
Aug-20	27.4	4.22
Sep-20	28.6	5.92
Oct-20	28.4	4.91
Nov-20	25.0	6.77
Dec-20	21.1	4.84
Jan-21	20.4	4.56

Feb-21	22.6	7.42
Mar-21	27.1	7.89
Apr-21	30.7	7.92
May-21	32.4	9.05
Jun-21	31.2	6.38
Jul-21	28.2	5.03
Aug-21	28.3	4.21
Sep-21	27.3	6.56
Oct-21	28.0	4.83
Nov-21	25.9	6.31
Dec-21	20.8	5.05
Jan-22	19.7	5.19
Feb-22	23.2	8.41
Mar-22	27.1	9.38
Apr-22	30.5	9.77
May-22	32.9	9.20
Jun-22	31.1	8.04
Jul-22	29.3	6.39
Aug-22	28.8	4.86
Sep-22	28.6	7.33
Oct-22	29.8	6.14
Nov-22	26.3	7.08
Dec-22	21.8	5.54