Journal of Agrometeorology

ISSN : 0972-1665 (print), 2583-2980 (online)
Vol. No. 27 (1) : 73-76 (March - 2025)
https://doi.org/10.54386/jam.v27i1.2838
https://journal.agrimetassociation.org/index.php/jam

Research Paper

SARIMA-based time series analysis of rainfall and temperature for the *Tarai* region of Uttarakhand

SHIVANI KOTHIYAL[1*], SHUBHIKA GOEL[1], R. K. SINGH[1], AARADHANA CHILWAL[2] and RAJEEV RANJAN[1]

[1]*Department of Agricultural Meteorology, GBPUAT, Pantnagar-263145, Uttarakhand*
[2]*ICAR-Directorate of Groundnut Research, Junagadh-362015, Gujarat*
[*]*Corresponding author email: shivani.kothiyal41418@gmail.com*

ABSTRACT

The study was conducted at the Department of Agrometeorology, GBPUAT, Pantnagar under the Gramin Krishi Mausam Seva (GKMS) Scheme. The meteorological data on temperature and rainfall for the period of 1990-2022 collected from the Norman E Borlaug Crop Research Centre, Pantnagar, were used to develop a seasonal model using MINITAB software. The software used the monthly input data for the years 1990-2017 to develop the best Seasonal Autoregressive Integrated Moving Average (SARIMA) model for forecasting the variables on monthly basis. The forecasted values using SARIMA for the time period (2018-2022) were used to validate the model against the observed data and it was observed that the SARIMA $(0,0,0)(1,0,1)_{12}$ was found suitable for rainfall and minimum temperature prediction and SARIMA $(0,0,2)(1,0,0)_{12}$ for maximum temperature prediction.

Keywords: MINITAB, ARIMA, rainfall, temperature, Pantnagar

Future trends can be easily predicted using statistical tool which considers the previous trend of data to determine the forecast. Time series analysis serves this purpose as it is useful in understanding the trends as well as forecasting the trends for the future. Various weather parameters like rainfall, monthly average temperature and relative humidity have a significant impact on agriculture and economics and its prediction aids in disaster preparedness thus, such time series analysis can have significant contribution in policy making and avoidance of such disasters (Ashraf *et al.,* 2024). The ARIMA (Auto Regressive Integrated Moving Average) model is an extension of the more sophisticated ARMA (Auto Regressive Moving Average) model and is commonly used for time series forecasting. Seasonal analysis of weather data with respect to normal for the region is a popular mean of determining the variations and vagaries in the regional weather conditions.

In order to help the farming community and the policy makers of particular region, it is very important to analyse the local weather conditions and give predictions. Rizvi *et al.,* (2024) reported that the Autoregressive Integrated Moving Average (ARIMA) model continues to be a simple and effective time series analysis technique used for weather forecasting, market condition forecasting, sales forecasting and many other. Gill *et al.,* (2023) used the SARIMA models in long term series forecasting of temperature and rainfall for different regions of Punjab with RMSE between 1.41 and 1.88. ARIMA model was used by Sharma and Singh (2024) for forecasting the monthly average rainfall in the Mandi district of Himachal Pradesh for a period of 10-year (2021–2030) and the model was found to be suitable for predicting rainfall in the region which could be applied in various hydrological and water resource studies. Kumar *et al.,* (2023) studied the rainfall forecast using ARIMA and Artificial Neural Network (ANN) and reported ARIMA to perform better for the Nadia region and this information could be effectively utilized in agricultural management, collection of rainwater and flood prediction. Kaur *et al.,* (2022) determined that the goodness of fit confirmed the adequacy of the seasonal ARIMA models and its efficiency in forecasting the long time series. In view of this, an attempt has been made to predict temperature and rainfall of Pantnagar, Udham Singh Nagar using long period data and applying SARIMA model.

Article info - DOI: *https://doi.org/10.54386/jam.v27i1.2838*
Presented in INAGMET 2024; Received: 10 December 2024; Accepted: 27 January 2025; Published online : 1 March 2025

MATERIALS AND METHODS

The study was conducted at Department of Agrometeorology, GBPUAT, Pantnagar. This is considered as the *Tarai* region of Uttarakhand. The location lies at a latitudinal and longitudinal extent of 29.1°N and 79.28°E at an altitude of 243.83 m. The daily temperature and rainfall data for the period of 1990 to 2022 were collected from the Department of Agrometeorology, GBPUAT, Pantnagar.

MINITAB Software

MINITAB (https://www.minitab.com/en-us/company/). is a commonly used tool by the statisticians for solving the real-world problems. It has an easy ARIMA model forecasting component which provides various statistical results with easy interpretations and varietal applications. The software uses ARIMA to determine a time series model which contains autoregressive, differencing, and moving average components and this model can be used in generating forecasts. ARIMA model use correlational techniques to study the model patterns that are not visible in the plotted data. MINITAB software version 21.1.0 was used to determine the best SARIMA model for forecasting in the region. ARIMA tool in the software helps in time series analysis of the long-term data based on different statistical parameters. SARIMA model is built on the basis of the historical data collected for the region.

The monthly values for the observed temperature and rainfall data for a 27-year period (1990-2017) was used as an input in determining the SARIMA model in the MINITAB software. The calculation of monthly values was performed using Weathercock software developed by ICAR-CRIDA, Hyderabad. The MINITAB software used several iterations with the ARIMA tool in the software to determine the best model based on the ACF, PACF, AIC, BIC and various comparative statistical parameters between the observed and predicted datasets. The SARIMA model is represented as SARIMA (p, d, q), (P, D, Q) $_M$ where P, D, Q are the AR, Differencing and MA terms, respectively while M is the seasonal part of the model (Kumar *et al.,* 2022).

The SARIMA model forecast was statistically analysed using root mean square error (RMSE) mean absolute error (MAE), mean absolute percentage error (MAPE) and coefficient of determination (R^2) as given below:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(P_i - O_i)^2}{n}} \; ; \qquad MAE = \frac{1}{n}\sum_{i=1}^{n}|P_i - O_i|$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|P_i - O_i|}{O_i} \times 100 \; ; \qquad R^2 = 1 - \frac{\sum(P_i - O_i)^2}{\sum(P_i - \overline{O_i})^2}$$

P_i = Predicted value at ith time, O_i= Observed value at ith time, n = number of observations

RESULTS AND DISCUSSION

Rainfall

The Augmented Dickey-Fuller (ADF) Test was done to check the stationarity of the time series data. The observed p and test statistic value, respectively was 0.95 and -0.12 i.e. greater than the significance level of 0.05 and critical value of -2.87, thus indicating the consideration of differencing to make the data stationary. Differencing method was used to make the data stationary and the autocorrelation (ACF) and partial autocorrelation (PACF) results that followed clearly defined the stationary mean of the raw data. As the data was seasonal, a seasonal ARIMA forecast was used to determine the best ARIMA model. The corrected Akaike's Information Criterion (AICc) and Bayesian Information Criterion (BIC), respectively was observed to be lower (Table 1) for the seasonal ARIMA combination of $(0,0,0)(1,0,1)_{12}$. The P-value for the parameters seasonal autoregressive and moving average is less than 0.05 which indicates that these terms are statistically significant and should be included in the model. The final estimates of the model for seasonal autoregressive and moving average, respectively was observed as 1.00 and 0.96 (coefficient), 0.002 and 0.02 (standard error of coefficient), 469.76 and 39.16 (T-value), 0 and 0 (P-value). The model was tested for independent residuals by modified Box-Pierce (Ljung-Box) which showed P-value < 0.05 representing independent residuals without any autocorrelation at different lags. The ACF and PACF results for rainfall residuals showed that they were mostly in significant limit. The forecasts obtained for the selected ARIMA model when compared with the observed data for the year 2018-22 showed the values to be nearly similar (Fig. 1). The ARIMA model results showed moderate performance for rainfall parameter with RMSE of 114.3; MAE of 72.76; MAPE of 1.83 and R^2 of 0.45 (Table 2). This performance of the SARIMA model for rainfall could be attributed to the varying rainfall pattern and amount in the region. The rainfall variability could be the reason for the moderate performance of the model. Similar results were obtained by Dimri *et al.,* (2020) for the Bhagirathi basin (Tehri and Uttarkashi) where the 20-year forecast was used for the time series analysis and the 10 years data (2001-2010) was validated with the observed data at respective stations. The comparison showed that precipitation forecasts fit well with the observed data. However, the RMSE value was high for the two stations (Tehri = 59.27 and Uttarkashi=57.43) which was due to over-prediction of heavy rainfall events while on normal rainfall days the observed and forecasted rainfall events matched well. Nkeshita *et al.,* (2022) observed two seasonal ARIMA models to perform well with high R^2, low RMSE and BIC value. Monthly weather forecasts for Odeda region were developed using Seasonal ARIMA and model $(3,1,3)$ $(2,1,1)_{12}$ was selected having maximum value of R^2, minimum RMSE and BIC. Thus, SARIMA could be an effective tool in rainfall forecast within the area.

Table 1: Selected ARIMA model and its performance for rainfall and temperature

ARIMA Model	Parameter	AICc	BIC	Interpretation
SARIMA (0,0,0) $(1,0,1)_{12}$	Rainfall	4195.65	4207.03	Moderate
SARIMA (0,0,2) $(1,0,0)_{12}$	Max. temp.	1381.2	1400.1	Good
SARIMA (0,0,0) $(1,0,1)_{12}$	Mini. temp.	766.6	778.0	Good

Table 2: Statistical analysis of monthly ARIMA model forecast against the observed weather parameters

Statistical parameter	Rainfall	Maximum temperature	Minimum temperature
RMSE	114.3	1.86	1.22
MAE	72.8	1.34	0.90
MAPE	1.8	0.06	0.08
R^2	0.4	0.89	0.97

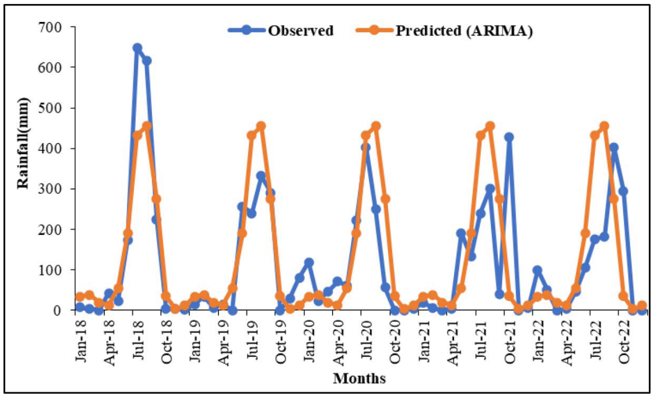

Fig. 1: Monthly comparison of observed data and ARIMA model forecast of rainfall for the period (2018-2022)

Maximum temperature

The ADF Test was done to check the stationarity of the time series. The observed p and test statistic value, respectively was 0.02 and -3.25 i.e. smaller than the significance level of 0.05 and greater than the critical T-value of -2.87, thus indicating the data to be stationary and not requiring differencing. However, the mean monthly plot of the original data series showed the autocorrelation and partial autocorrelation function indicating slight non-stationarity of the data. So, differencing method was used to make the data stationary and the ACF and PACF results that followed clearly defined the stationary mean of the raw data. As the data was seasonal, a seasonal ARIMA forecast was used to determine the best ARIMA model. The corrected AICc and BIC, respectively was observed to be lower for the seasonal ARIMA combination of (0,0,2) $(1,0,0)_{12}$ (Table 1). The P-value for the parameters seasonal autoregressive and moving average is less than 0.05 which indicates that these terms are statistically significant and should be included in the model. The final estimates of the model for seasonal autoregressive and moving average, respectively was observed as 0.95 and-0.27 (Coefficient), 0.002 and 0.05 (Standard error of coefficient), 43.18 and -4.94 (T-value), 0 and 0 (P-value). The model was tested for independent residuals by modified Box-Pierce (Ljung-Box) which showed P-value < 0.05 representing independent residuals without any autocorrelation at different lags. The ACF and PACF results for maximum temperature residuals showed that they were mostly in significant limit. The forecasts obtained for the selected ARIMA model when compared with the observed data for the year 2018-22 showed the values to be nearly similar (Fig. 2). The ARIMA model forecasts performed good for maximum temperature parameter with RMSE of 1.86; MAE of 1.34; MAPE of 0.06 and R^2 of 0.89 (Table 2). The maximum temperature follows a synonymous trend in the region with very low variability due to which the model showed good statistical results for the parameter. Similarly, the study by

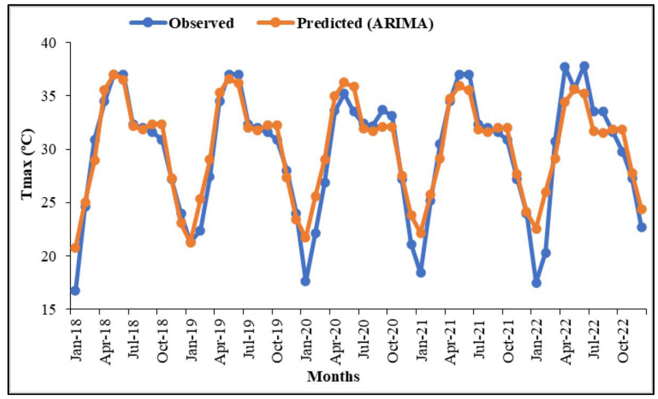

Fig. 2: Monthly comparison of observed data and ARIMA model forecast of maximum temperature for the period (2018-2022)

Dimri *et al.,* (2020) observed that the temperature forecast for maximum and minimum temperature fitted well with the observed data with a RMSE=1.5 for maximum and RMSE=1.4 for minimum temperature at both the stations, the SARIMA model over predicted the temperature data at the locations. Thus, the SARIMA model could be effectively used to predict maximum temperature in the region.

Minimum temperature

The ADF Test was done to check the stationarity of the time series. The observed p and test statistic value, respectively was 0.95 and -0.12 i.e. greater than the significance level of 0.05 and critical value of -2.87, thus indicating the consideration of differencing to make the data stationary. Differencing method was used to make the data stationary and the ACF and PACF results that followed clearly defined the stationary mean of the raw data. As the data was seasonal, a seasonal ARIMA forecast was used to determine the best ARIMA model. The corrected AICc and BIC, respectively was observed to be lower for the seasonal ARIMA combination of (0,0,0) $(1,0,1)_{12}$ (Table 1). The P-value for the parameters seasonal autoregressive and moving average is less than 0.05 which indicates that these terms are statistically significant and should be included in the model. The final estimates of the model for seasonal autoregressive and moving average, respectively was observed as 1.00 and 0.89 (Coefficient), 0.00 and 0.03 (Standard error of coefficient), 2748.32 and 27.99 (T-value), 0 and 0 (P-value). The model was tested for independent residuals by modified Box-Pierce (Ljung-Box) which showed P-value < 0.05 representing independent residuals without any autocorrelation at different lags. The ACF and PACF results for minimum temperature residuals showed that they were mostly in significant limit. The final estimates showed the selected SARIMA model to be the best one with low AICc. The forecasts obtained for the selected ARIMA model when compared with the observed data yielded nearly same results (Fig. 3). The ARIMA model forecasts showed very good performance for minimum temperature parameter (Table 2) with RMSE of 1.22; MAE of 0.90; MAPE of 0.08 and R^2 of 0.97. The low variability in the minimum temperature for the region could be attributed to the very good performance of the model for the region. The ARIMA model technique performed in the Karachi station of Pakistan from the available monthly temperature data from January 1989 to December 2018 showed that the residual series was normally distributed with the negligent deviation in the RMSE and MAE of the validation set from the observed data thus,

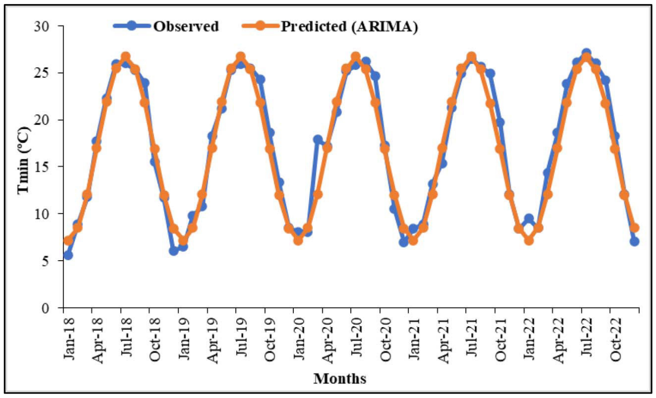

Fig. 3: Monthly comparison of observed data and ARIMA model forecast of minimum temperature for the period (2018-2022)

considering them to be in an acceptable range (Amjad *et al.,* 2023). For minimum temperature, MAE ranged from 0.74-0.49 (Ludhiana and Bathinda), with lower values indicating higher model accuracy. Thus, the SARIMA model could be effectively used to predict minimum temperature in the region.

CONCLUSION

Temperature and rainfall parameters were used to build a SARIMA model for the region suitable for further applications to understand rainfall and temperature trends. The selected SARIMA model was $(0,0,0)(1,0,1)_{12}$ for rainfall and minimum temperature and $(0,0,2)(1,0,0)_{12}$ for maximum temperature. The SARIMA model forecast was compared with the observed data values graphically and statistically and the results were observed to be good. The SARIMA model for rainfall was found to give moderate statistical results due to high variability in the rainfall pattern and amount in the region while model accuracy was good for temperature due to their low variability and high synonymity in the region. Thus, SARIMA models are utilizable tool in forecasting the weather patterns in the region which can help in understanding the trends. These models use a large amount of data to build the most accurate model which can be used as a tool in forecasting future weather conditions and policy making. The time series forecasting using SARIMA still remains a crucial step in deciding the future agricultural plan of work.

ACKNOWLEDGEMENT

The data maintenance and support provided by India Meteorological Department regularly under the Gramin Krishi Mausam Sewa (GKMS) scheme to Agromet Field Unit (AMFU), Pantnagar is highly acknowledged.

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Conflict of interest: The authors declare that there is no conflict of interest related to this article.

Data availability: Data shall be provided on request from the author.

Author's contribution: **S K** and **R K S**: Conceptualization, investigation, writing the original draft and methodology; **R K S** and **R R**: Supervision and editing; **A C, S K** and **S G**: Formal analysis and data curation.

Disclaimer: The contents, opinions and views expressed in the research article published in the Journal of Agrometeorology are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

Publisher's Note: The periodical remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

Amjad, M., Khan, A., Fatima, K., Ajaz, O., Ali, S. and Main, K. (2023). Analysis of temperature variability, trends and prediction in the Karachi region of Pakistan using ARIMA models. *Atmosphere*, 14(1):88. https://doi.org/10.3390/atmos14010088

Ashraf, S., Ali, N., Dr. Muhammad, I. K., Niaz, A. and Masood (2024). Weather Forecasting: Using Time Series Analysis under different Stations of the UK. *Migration Letters,* 21: S9, pp. 782-794

Dimri, T., Ahmad, S. and Sharif, M. (2020). Time series analysis of climate variables using seasonal ARIMA approach. *Earth Syst. Sci.,* 129:149. https://doi.org/10.1007/s12040-020-01408-x

Gill, K. K.., Bhatt, K., Kaur, B. and Sandhu, S. S. (2023). ARIMA approach for temperature and rainfall time series prediction in Punjab. *J. Agrometeorol.,* 25 (4): 571-576. https://doi.org/10.54386/jam.v25i4.2250

Kaur, B., Kaur N., Gill K.K., Singh J., Bhan, S.C. and Saha, S. (2022). Forecasting mean monthly maximum and minimum air temperature of Jalandhar district of Punjab, India using seasonal ARIMA model. *J. Agrometeorol.,* 24 (1): 42-49. https://doi.org/10.54386/jam.v24i1.777

Kumar, Utkarsh, Dhirendra Kumar Singh, Suresh Chandra Panday, Jaideep Kumar Bisht and Lakshmi Kant (2022). Development and evaluation of seasonal rainfall forecasting (SARIMA) model for Kumaon region of Uttarakhand. *Indian J. Soil Cons.,* 50(3):190-198

Kumar, K. Sathees, Gowthaman, T. and Bhattacharyya, B. (2023) Comparison of Arima and Ann for Forecasting the Annual Rainfall of Nadia District, West Bengal, India. *Eco. Env. Cons.,* 29 (August Suppl. Issue): pp. (S24-S29). http://doi.org/10.53550/EEC.2023.v29i04s.005

Nkeshita, F., Adekunle, A., Owosho, S. and Adewale, A. (2022). Rainfall forecasting for Odeda, south-western Nigeria using a seasonal ARIMA time-series model. *J. Eng. Res.,* 27(1): 1-10. http://jer.unilag.edu.ng/article/view/1589

Rizvi Mohd, F., Sahu, S. and Rana, S. (2024). ARIMA Model Time Series Forecasting. *Intern. J. Res. Applied Sci. Engg. Technol., (IJRASET).* 12(5): 3782-85

Sharma, A. and Singh, K. (2024). ARIMA-Based forecasting of monthly rainfall in Mandi district, Himachal Pradesh. *Water Supply*, 24 (9): 3226-3237. https://doi.org/10.2166/ws.2024.200