The Andhra Agric. J 72 (1): 098-105, 2025
doi :10.61657/aaj.2025.182

Advanced time series models for rainfall prediction: A case study in Chittoor, Andhra Pradesh

A Vani, Shaik Nafeez Umar, S Hari Prasad, K Naga Vihari, and M Bhupathi Naidu
Research Scholar, Department of Statistics, S. V. University, Tirupati-517501,
Andhra Pradesh, India.

ABSTRACT

This study presents a comprehensive examination of monthly rainfall data collected in Chittoor district of Andhra Pradesh from January 1990 to December 2022. The primary objective is to develop an accurate predictive model for future rainfall patterns using time-series forecasting techniques. Traditional AutoRegressive Integrated Moving Average (ARIMA) models have been extensively applied in climate studies due to their effectiveness in capturing linear trends However, ARIMA models often fail to account for seasonal variations inherent in rainfall data. To address this limitation, the QS Test was conducted to determine the presence of seasonality, revealing strong periodic fluctuations in monthly precipitation patterns. Based on this validation, a Seasonal ARIMA (SARIMA) model SARIMA(1,0,1)(1,1,1)[12] was developed and compared against ARIMA(0,0,3)(2,0,0)[12] to assess forecasting accuracy. Model selection was guided by key statistical indicators such as AIC, RMSE, MAE, and residual diagnostics. The SARIMA model demonstrated superior residual independence ($p = 0.6631$), confirming improved white noise behavior and seasonal pattern detection. Additionally, SARIMA achieved a lower AIC (4256.71), RMSE (57.63), and MAE (36.58) compared to ARIMA, reinforcing its enhanced forecasting capabilities.

Keywords: *ARIMA, Forecasting, Rainfall Prediction, SARIMA and Time Series Analysis*

The state of Andhra Pradesh, located in southern India, is a prime example of an agricultural economy that is reliant on rainfall. Chittoor district, situated in the southern part of the state, is known for its diverse agricultural production, including crops such as groundnuts, paddy, and horticultural crops such as mangoes and bananas. To address the need for reliable rainfall prediction in Chittoor district, this study employs an autoregressive integrated moving average model to forecast monthly rainfall patterns.

The precision of rainfall forecasting plays a critical role in the efficient management of water resources, agricultural planning, and preparedness for natural disasters. Researchers have extensively applied time series models to predict precipitation patterns, with ARIMA and its derivatives proving particularly effective. In their study, Rodrigues and Deshpande (2017) examined rainfall prediction across all states in India using ARIMA and Multiple Linear Regression (MLR) models. Their findings indicated that ARIMA performed exceptionally well for short-term forecasts, while MLR effectively captured long-term trends, thereby demonstrating the poten study, Lama et al. (2021) conducted monthly rainfall forecasts for the Sub-Himalayan region using SARIMA and TDNN models. Their findings revealed that SARIMA was well-suited for capturing seasonal patterns, while TDNN exhibited superior accuracy in predicting non-linear trends. This research underscores the importance of employing flexible forecasting techniques in regions with complex topography.

The Box-Jenkins ARIMA methodology remains a popular choice for time series data modeling due to its proven reliability. In a notable application, Ogbozige (2022) utilized an ARIMA(2,1,2) model to analyze 50 years of monthly rainfall data in Calabar, Nigeria. The study yielded precise forecasts, further validating the effectiveness of the Box-Jenkins approach in hydrological research. In a comparable context, the research conducted by Sathish et al. (2017) employed SARIMA to examine monthly precipitation patterns in Gangetic West Bengal. This approach effectively captured seasonal fluctuations, contributing to improved water resource management and disaster readiness. Furthermore, investigations utilizing seasonal extensions of ARIMA models have yielded promising results across various geographical areas. For instance, Amelia et al. (2022) implemented SARIMAX for rainfall prediction in Pangkalpinang City, integrating external variables to improve forecasting precision.

Seasonal ARIMA modeling has been employed in diverse regional studies to analyze climate patterns. Ashwini et al. (2021) utilized this approach to examine monsoon rainfall in Tamil Nadu, effectively capturing the seasonal fluctuations characteristic of the area. Similarly, Dimri et al. (2020) applied Seasonal ARIMA to investigate climate variables in Uttarakhand, demonstrating the model's efficacy in handling complex seasonal trends.In addition to these applications, researchers have begun exploring sophisticated hybrid models to enhance the accuracy of climate forecasts.

The integration of SARIMA with GARCH by Pandey et al. (2018) led to improved rainfall prediction accuracy in Agartala and Jodhpur, effectively addressing both seasonality and volatility. In a similar vein, Unnikrishnan et al. (2020) constructed a hybrid SSA-ARIMA-ANN model, which exhibited high precision in forecasting daily rainfall. Furthermore, Pham et al. (2019) employed hybrid data-intelligence algorithms across various stations in Vietnam, yielding accurate rainfall predictions and providing additional evidence for the effectiveness of hybrid methodologies.

Comparative studies have highlighted the performance of traditional and hybrid models. In a comparative study, Barman et al. (2020) evaluated the efficacy of AR, MA, and ARMA models for rainfall forecasting in Assam and Meghalaya, with results indicating the ARMA model's superior performance. The application of the SERIMA model for weekly rainfall predictions in Junagadh, Gujarat, by Damor et al. (2023) yielded accurate short-term forecasts. Additionally, Khan et al. (2023) examined the ARIMA model's effectiveness for both short- and long-term rainfall forecasting in the Klang River Basin, confirming its applicability across various temporal scales. Machine learning approaches are also gaining traction for rainfall prediction. The research conducted by K.B. et al. (2024) utilized machine learning algorithms to model precipitation patterns in Kerala, demonstrating superior predictive capabilities through the application of data-driven approaches. This paradigm shift towards machine learning methodologies represents a significant development in rainfall forecasting, emphasizing the synergistic potential of combining conventional and cutting-edge techniques. The substantial body of literature accentuates the adaptability and efficacy of ARIMA and its variants in precipitation prediction. Moreover, it demonstrates the increasing focus on hybrid and machine learning approaches, facilitating more precise and robust forecasting capabilities. This investigation builds upon these findings by delving into advanced temporal sequence models and integrated methodologies for rainfall prediction across varied geographical contexts.

MATERIAL AND METHODS

Stationary Test

Augmented Dickey–Fuller (ADF) Test

The ADF test was applied to check for stationarity in the time-series data. The null hypothesis of the ADF test states that the series has a unit root (i.e., it is nonstationary). A stationary time series is a prerequisite for ARIMA modeling. In this study, the ADF test confirmed whether the monthly rainfall data required differencing to be stationary. The following model was used for the ADF test.

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^{k} \delta_i \, \Delta Y_{t-i} + \epsilon_t$$

where ΔY_t is the difference time series, β is the trend term, and ϵ_t is the error term.

Box-Jenken's Methodology

The Box-Jenkins methodology provides a robust framework for the identification, estimation, and diagnostic checking of Autoregressive Integrated Moving Average models in time-series forecasting(Ogbozige, 2022). This study employs the Box-Jenkins approach to analyze monthly rainfall data collected in Chittoor District from January 1990 to December 2022.

The initial stage of model identification involved assessing the stationarity of the time series. If trends or seasonality are detected within the data, differencing techniques are applied to induce stationarity. The autocorrelation and partial autocorrelation function plots serve as valuable tools for identifying the appropriate autoregressive and moving average terms for the model(Davis &

Rappoport, 1974).The first and foremost step is to determine the order of differencing (d) to stationarise the series. The order of differencing (d) is selected such that it minimizes the standard deviation. This is done by fitting different ARIMA models having various orders of differencing, but a constant coefficient is selected.(Dimri, T etal ., 2020)Following the identification of the ARIMA model parameters, a Maximum Likelihood Estimation was employed to estimate the model. A comprehensive residual analysis was then conducted to ascertain whether the residuals of the fitted model exhibited the characteristics of white noise. The model adequacy was confirmed if the residuals demonstrated independence and normality. The subsequent section of this paper focuses on evaluating the performance of the selected ARIMA model using a range of statistical metrics, the ARIMA models include autoregressive (AR), moving averages (MA), and integrated processes.

The general form of the ARIMA (p,d,q) given below

Autoregressive Model(AR)

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t \quad \text{----------(1)}$$

Moving Averages(MA) (q);

$$Y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \quad \text{---------- (2)}$$

The general form of ARIMA model(p d q) is

$$\phi(B)(1-B)^d Y_t = \theta(B)\epsilon_t \quad \text{---------- (3)}$$

Where ‘p’ is order of autoregressive process, ‘q’ is order of moving average process, and ‘d ’ is order of differencing the series to make it stationary.

Seasonal ARIMA (SARIMA) Model

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model extends the ARIMA model by incorporating seasonal components to better model time series data that exhibit periodic behavior. A SARIMA model is typically denoted as: SARIMA (p, d, q) (P, D, Q)s

The general SARIMA model equation is expressed as

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D Y_t = \Theta_Q(B^s)\theta_q(B)\epsilon_t \quad \text{------------ (4)}$$

Model Performance Evolutiona. Akaike Information Criterion (AIC)

The Akaike Information Criterion (AIC) serves as a robust tool for comparing model performance by incorporating both the goodness-of-fit and model complexity. A lower AIC value indicates a better trade-off between accuracy and simplicity. In the present study, SARIMA models outperformed traditional ARIMA models, with the SARIMA(1,0,1)(1,1,1)[12] model achieving the lowest AIC of 4256.71, followed closely by SARIMA(1,0,2)(0,1,1)[12] at 4256.84 and SARIMA(0,0,2)(2,1,1)[12] at 4258.61. In contrast, the best-performing ARIMA model, ARIMA(0,0,3)(2,0,0)[12], recorded a significantly higher AIC of 4473.77. These results underscore the importance of incorporating seasonal components in modeling, as SARIMA models provided more parsimonious fits to the rainfall data in Chittoor.

$$AIC = 2k - 2ln(\hat{L})$$

Bayesian Information Criterion (BIC)

The Bayesian Information Criterion (BIC), like AIC, rewards model accuracy while penalizing complexity more strongly, making it especially useful when model simplicity is prioritized. In this study, the SARIMA(1,0,1)(1,1,1)[12] model also achieved the lowest BIC value of 4276.46, further validating its efficiency and appropriateness for seasonal rainfall prediction. Compared to this, the BIC values for alternative SARIMA models such as SARIMA(0,0,2) and SARIMA(1,0,2) were marginally higher (4282.31 and 4276.60 respectively), and notably higher in ARIMA modelssuch as 4501.64 for ARIMA(0,0,3) and 4537.71 for ARIMA(1,0,0). This further strengthens the case for SARIMA as the optimal modeling approach in seasonal rainfall forecasting.

$$BIC = k\,.\ \ln(n) - 2\,.\ln \hat{L}$$

Mean Absolute Error (MAE):

Mean Absolute Error (MAE) provides an intuitive measure of average prediction error without penalizing large deviations as heavily as RMSE. It is especially valuable for understanding the general forecasting accuracy of a model. In this analysis, the SARIMA(0,0,2)(2,1,1)[12] model achieved the lowest MAE of 36.35, followed closely by SARIMA(1,0,1)(1,1,1)[12] at 36.58 and SARIMA(2,0,2)(1,1,1)[12] at 36.67. Conversely, the ARIMA models displayed substantially higher MAEs, with ARIMA(0,0,3)(2,0,0)[12] registering 47.25. This again highlights the superiority of SARIMA models in handling seasonally influenced time series like rainfall, offering more precise and consistent forecasting performance. Represents the average of absolute differences between forecasted and actual values.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is a widely used metric that quantifies the standard deviation of prediction errors, emphasizing larger errors due to squaring. A lower RMSE indicates better model accuracy. Among all models tested, SARIMA(0,0,2)(2,1,1)[12] exhibited the lowest RMSE of 57.52, closely followed by SARIMA(1,0,1)(1,1,1)[12] with 57.63. In contrast, the best ARIMA model, ARIMA(0,0,3)(2,0,0)[12], produced a significantly higher RMSE of 68.16. These findings clearly demonstrate that SARIMA models—particularly those with well-specified seasonal components—are more adept at minimizing forecast errors and improving predictive reliability.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

RESULTS AND DISCUSSIONS

Analysis of monthly rainfall data (measured in millimeters) for Chittoor District from January 1990 to December 2022 revealed distinct seasonal patterns. The most substantial rainfall occurred during October and November, coinciding with the northeast monsoon season. Boxplot analysis illustrated considerable rainfall variability during these months, as evidenced by higher medians and wider interquartile ranges(Fig1). Conversely, rainfall was significantly lower from January to May, with negligible amounts recorded in certain years, underscoring the strong seasonal dependence of rainfall in this region.

Visual analysis of monthly rainfall data from January 1990 to December 2022, encompassing approximately 400 months (33 years), revealed several

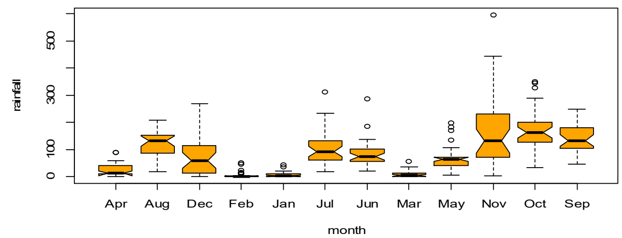

Fig.1. Monthly wise rainfall in Chittoor District in Andhra Pradesh

key characteristics. The time series exhibited significant fluctuations with periodic spikes indicating months of heavy rainfall. These cyclical patterns suggest a distinct seasonal component that likely corresponds to the monsoon season. Notably, a few extreme peaks were observed, representing outliers characterized by exceptionally high rainfall compared to the overall data distribution. These observations highlight the inherent variability and periodicity within rainfall data, which are crucial factors to consider when developing forecasting models. The cyclical nature of the time series, as depicted in (Fig 2), supports the use of a time series model, such as ARIMA, which can effectively capture both trend and seasonality for improved prediction accuracy.

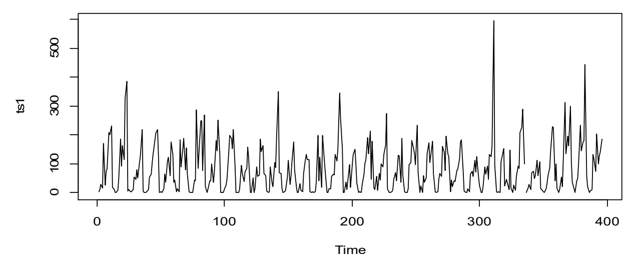

Fig.2. Rainfall trend in Chittoor District of Andhra Pradesh

The Autocorrelation Function plot illustrates the correlation between rainfall values and their past values at different time lags. The presence of significant spikes exceeding the confidence intervals at early lags (e.g., 1, 4, 6, and 10) indicates substantial autocorrelation within the data. The periodic nature of these spikes suggests a potential seasonal pattern, possibly linked to the monsoon cycles. This observation aligns with the expectation of seasonality in the rainfall data. The gradual decay of autocorrelation over time further supports this hypothesis.

Conversely, the Partial Autocorrelation Function plot reveals the correlation between the time series and its lags, after accounting for the influence of intermediate lags. The significant spike at the first lag, followed by scattered significant spikes at lag six and beyond, suggests that the rainfall data may be effectively modeled using autoregressive terms. PACF is particularly useful in determining the appropriate number of AR terms for an ARIMA model. The gradual decay observed in the Autocorrelation Function plot suggests the presence of potential autoregressive terms, indicating that past rainfall values have a persistent influence on the current values. Significant spikes at lower lags in the Partial Autocorrelation Function plot further support this finding, suggesting that incorporating a few AR terms could enhance the accuracy of the model. Given these observations, a combination of AR and Moving Average terms, likely in the form of an ARMA or ARIMA model, appears suitable for modeling the rainfall data. This model should also account for seasonality, as suggested by the ACF Plot (Fig 3). These plots are crucial for determining the optimal parameters (p, d, q) for the ARIMA models, which are essential for generating accurate rainfall predictions based on historical patterns.

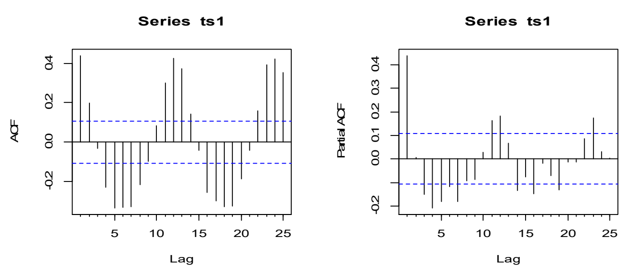

Fig.3. ACF and PACF for rainfall data

The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots provide essential insights into the residual behavior of the ARIMA model, helping to assess whether the model has effectively accounted for time-dependent structures within the rainfall data.

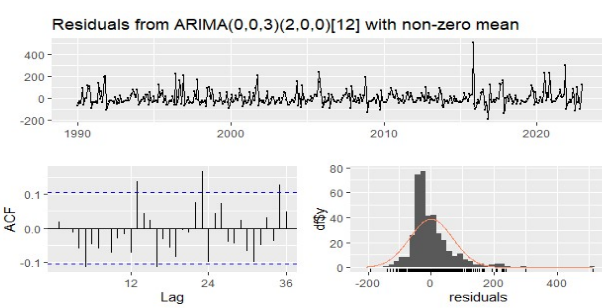

Fig.4. ARIMA Residual analysis for rainfall data

The residual time series plot shows random distribution around zero, indicating the model has captured overall trends and dependencies. However, spikes suggest it may struggle with extreme rainfall fluctuations and monsoonal effects. The ACF plot shows correlation between residuals at different lags. Most spikes fall within confidence intervals, indicating removed temporal dependencies, though some exceed bounds at higher lags, suggesting unaccounted seasonal effects. The PACF plot helps determine autoregressive terms needed. Significant spikes at early lags confirm the need for autoregressive components, while sharp cutoff after few lags indicates effective short-term dependency modeling. The residuals histogram evaluates error distribution, ideally showing normal distribution around zero. Residuals form a bell-shaped curve, showing good alignment with rainfall patterns, though a slight rightward skew indicates under-prediction of heavy rainfall events, suggesting possible need for additional seasonal components.

Table 1: ARIMA Model Performance Metrics, different ARIMA (p d q) models

Model	Order (p,d,q)	AIC	BIC	RMSE	MAE
ARIMA (0,0,3)	(0,0,3)	4473.77	4501.64	68.16	47.25
ARIMA (1,0,0)	(1,0,0)	4525.77	4537.71	73.83	54.78
ARIMA (2,0,1)	(2,0,1)	4528.19	4548.09	73.68	54.59
ARIMA (4,0,1)	(4,0,1)	4484.86	4512.73	69.35	49.01

To evaluate various ARIMA configurations for forecasting monthly rainfall, multiple models with different autoregressive, differencing, and moving average parameters were compared. Table 1 summarizes performance metrics including AIC, BIC, RMSE, and MAE for seasonal and non-seasonal ARIMA models. The seasonal ARIMA(0,0,3)(2,0,0)[12] outperformed non-seasonal counterparts, recording the lowest AIC (4473.77) and BIC (4501.64), indicating better model fit. This model achieved the lowest RMSE (68.16) and MAE (47.25), suggesting more precise predictions. Non-seasonal models, including ARIMA(1,0,0), ARIMA(2,0,1), and ARIMA(4,0,1), showed higher error rates, with even ARIMA(4,0,1)'s competitive AIC (4484.86) unable to surpass the seasonal model's performance. These results confirm seasonal patterns in rainfall data and justify adopting SARIMA for enhanced prediction.

Table 2: ARIMA Forecasted Rainfall for 2023

Month	Point Forecast (mm)	95% Confidence Interval (Lower–Upper)
Jan-23	64	[-70.6, 198.6]
Feb-23	52.3	[-84.8, 189.4]
Mar-23	49	[-89.3, 187.3]
Apr-23	51.6	[-87.0, 190.2]
May-23	89.7	[-48.9, 228.3]
Jun-23	97.4	[-41.2, 236.0]
Jul-23	115.2	[-23.4, 253.8]
Aug-23	132.9	[-5.7, 271.5]
Sep-23	109.4	[-29.2, 248.0]
Oct-23	118.6	[-20.0, 257.2]
Nov-23	184.3	[45.6, 322.9]
Dec-23	107.2	[-31.4, 245.8]

The QS (Quasi-Seasonal) Test is used to statistically determine the presence of seasonal patterns in time series data. It is particularly useful before fitting seasonal models like SARIMA.

Null Hypothesis (H_0): No seasonal pattern exists. Alternative Hypothesis (H_1): Seasonality is present. If the p-value < 0.05, we reject the null hypothesis, indicating significant seasonality in the data. This test aids in justifying the use of SARIMA over non-seasonal ARIMA.

Table 3: QS Test for Seasonality

Test Used	Test Statistic	P-value
QS Test	12.71	0.00174

To determine seasonality in monthly rainfall data from Chittoor district (1990–2022), the QS (Quasi-Seasonal) test was employed. The test yielded a statistic of 12.71 with p-value 0.00174. Since the p-value is below 0.05, we reject the null hypothesis of no seasonality, providing strong evidence of seasonality in the rainfall series. This detection validates using seasonal SARIMA over non-seasonal ARIMA models.

Table 4: Different SARIMA (p d q) models, Model Performance Metrics

Model	Order (p,d,q)	Seasonal (P,D,Q)[s]	AIC	BIC	RMSE	MAE
SARIMA (1,0,1)	(1,0,1)	(1,1,1)[12]	4256.7	4276.5	57.63	36.58
SARIMA (2,0,1)	(2,0,1)	(1,1,0)[12]	4386.9	4406.7	71.83	45.1
SARIMA (0,0,2)	(0,0,2)	(2,1,1)[12]	4258.6	4282.3	57.52	36.35
SARIMA (1,0,2)	(1,0,2)	(0,1,1)[12]	4256.8	4276.6	57.69	36.72
SARIMA (2,0,2)	(2,0,2)	(1,1,1)[12]	4260.4	4288	57.61	36.67

SARIMA Model Summary

Since QS Test confirmed seasonality, SARIMA models were tested to improve forecast accuracy.

To capture seasonal dynamics in monthly rainfall patterns, several Seasonal ARIMA (SARIMA) models were tested. Table 3 presents model performance using key indicators: AIC, BIC, RMSE, and MAE. The SARIMA(1,0,1)(1,1,1)[12] model emerged as the most robust, recording the lowest AIC (4256.71) and BIC (4276.46) values, indicating optimal model fit. This model achieved low RMSE (57.63) and MAE (36.58) scores, reflecting high predictive accuracy. SARIMA(0,0,2)(2,1,1)[12] and SARIMA(1,0,2)(0,1,1)[12] showed similar RMSE and MAE values but higher AIC and BIC, while SARIMA(2,0,1)(1,1,0)[12] and SARIMA(2,0,2)(1,1,1)[12] displayed higher error metrics. The performance metrics favor SARIMA(1,0,1)(1,1,1)[12] as the optimal model for seasonal rainfall forecasting in Chittoor, Andhra

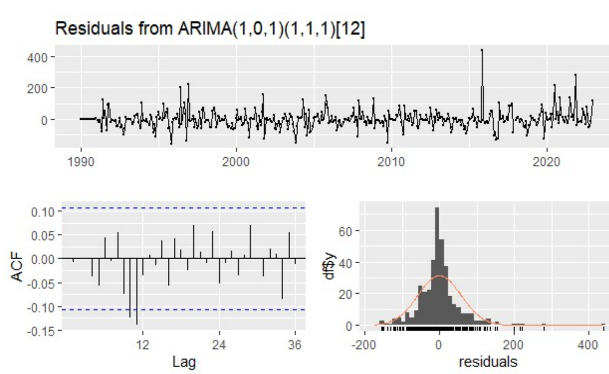

Fig.5.SARIMA Residual analysis for rainfall data

Pradesh, making it excellent for reliable monthly rainfall predictions crucial for agricultural and water management decisions.

The residual time series plot shows randomly distributed residuals around zero, indicating the SARIMA(1,0,1)(1,1,1)[12] model effectively captures trends and seasonal patterns in rainfall data. Unlike ARIMA, SARIMA accounts for periodic dependencies from monsoonal effects. Some small residual spikes suggest under-prediction during high-rainfall months. The ACF plot examines residual correlation across lags to ensure past errors don't influence future predictions. Most residual spikes lie within confidence intervals, confirming SARIMA removes major temporal dependencies, though minor spikes at higher lags suggest residual autocorrelation. The SARIMA model shows improved seasonal corrections versus ARIMA, supported by the Ljung-Box test (p = 0.6631). The PACF plot's early-lag spikes confirm the necessity of autoregressive components, with SARIMA's AR(1) term effectively modeling rainfall dependencies. A sharp cutoff after few lags indicates proper handling of seasonal dependencies. The residual histogram shows a bell-shaped curve centered at zero, confirming SARIMA's alignment with rainfall trends. Minimal skewness compared to ARIMA suggests fewer forecasting errors, though slight rightward skew indicates occasional under-prediction during peak rainfall months. This improved residual behavior demonstrates SARIMA's stronger forecasting capability.

Table 5: SARIMA Forecasted Rainfall for 2023

Month	Point Forecast (mm)	95% Confidence Interval (Lower–Upper)
Jan-23	15.1	[-101.9, 132.2]
Feb-23	11.8	[-105.7, 129.3]
Mar-23	14.4	[-103.3, 132.2]
Apr-23	27.9	[-89.9, 145.8]
May-23	69.6	[-48.3, 187.5]
Jun-23	83.6	[-34.3, 201.5]
Jul-23	104.3	[-13.7, 222.2]
Aug-23	121.2	[3.3, 239.2]
Sep-23	142.2	[24.2, 260.1]
Oct-23	173.8	[55.8, 291.7]
Nov-23	167.1	[49.1, 285.0]
Dec-23	68	[-50.0, 186.0]

Fig.6. Forecast value of rainfall data

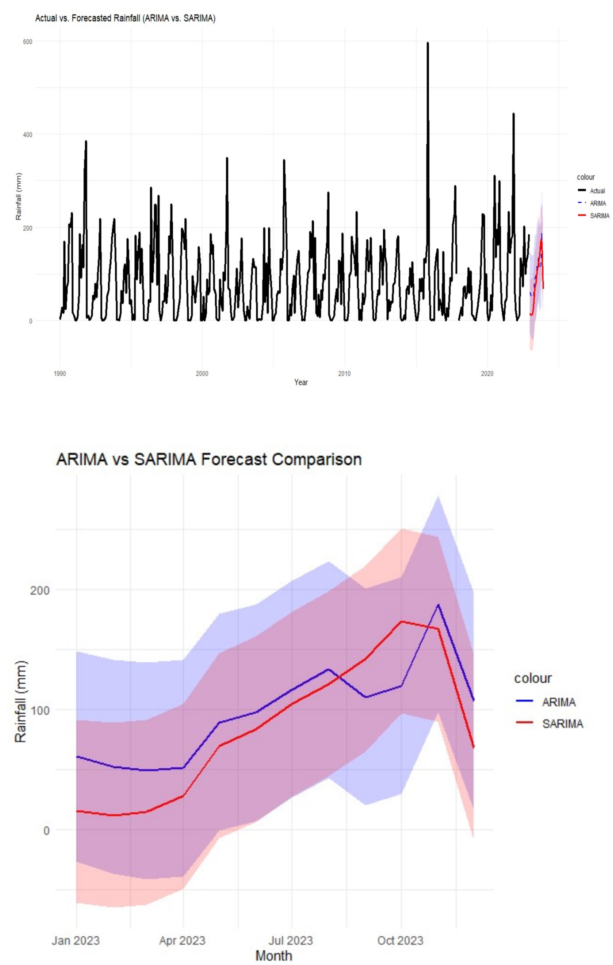

Table 6: Residual Diagnostics

Model	Ljung-Box Q Statistic	p-value
ARIMA(0,0,3)(2,0,0)[12]	46.85	0.00038
SARIMA(1,0,1)(1,1,1)[12]	16.84	0.6631

Residual diagnostics assessed the fitted models by testing for autocorrelation in forecast errors. The Ljung-Box Q statistic and p-values indicate whether residuals behave like white noise—essential for reliable time series forecasting. The ARIMA(0,0,3)(2,0,0)[12] model's Q statistic of 46.85 with p-value 0.000375 showed significant residual autocorrelation, indicating incomplete pattern capture. The SARIMA(1,0,1)(1,1,1)[12] model's Q statistic of 16.84 with p-value 0.6631 indicated statistically independent residuals, meeting white noise assumptions and successfully accounting for time-dependent structures in rainfall data.

CONCLUSION

The findings of this study confirm the significance of incorporating seasonal effects in rainfall forecasting through the QS test, which identified strong periodic patterns in the data. Comparisons between ARIMA and SARIMA models demonstrated that SARIMA(1,0,1)(1,1,1)[12] outperformed ARIMA(0,0,3)(2,0,0)[12], achieving lower AIC (4256.71), RMSE (57.63), and MAE (36.58), thereby enhancing predictive reliability. Furthermore, Ljung-Box test results (p = 0.6631 for SARIMA) indicate residual independence, proving that SARIMA effectively captures rainfall trends without significant autocorrelation. Forecasted rainfall values highlight SARIMA's superior ability to model seasonal fluctuations, providing narrower confidence intervals and more stable predictions compared to ARIMA. These results reinforce SARIMA's suitability for monthly rainfall forecasting in Chittoor, supporting more informed environmental and agricultural decision-making. Future research could explore hybrid approaches, such as integrating machine learning models with SARIMA, to further optimize predictive performance in climate forecasting applications.

LITERATURE CITED

Amelia R, Kustiawan E, Sulistiana I and Dalimunthe D Y 2022. Forecasting Rainfall Pangkalpinang City Using Seasonal Autoregressive Integrated Moving Average With Exogenous (Sarimax). 16(1):137-146.

Ashwini U, Kalaivani K, Ulagapriya K and Saritha A 2021. Time Series Analysis based Tamilnadu Monsoon Rainfall Prediction using Seasonal ARIMA.

Barman U, Choudhury R D, Hussain A E, Dahal M J, Barman p and Hazarika M 2020. Comparative Assessment of AR , MA and ARMA for Time Series Forecasting of Assam and Meghalaya Rainfall division.

Damor P A, Ram B and Kun Apara A N 2023. Stochastic Time Series Analysis, Modeling, and Forecasting of Weekly Rainfall Using Serima Model. Science domain International, 13(12): 773-782.

Dimri T, Ahmad S and Sharif M 2020. Time series analysis of climate variables using seasonal ARIMA approach. Springer Science + Business Media, 129(1).

K B, G K, Mehta V and Rai V 2024. Machine Language approach for Modeling and Predicting Rainfall in Different Zones of Kerala., 78(2): 89-96.

Khan m, Mustafa M R U, Hossain M S and Julius A D 2023. Short Term and Long Term Rainfall forecasting using ARIMA Model, 14(5): 292-298.

Lama A, Singh K N, Singh H, Shekhawat R, Mishra p and Gurung B 2021. Forecasting monthly rainfall of Sub-Himalayan region of india using parametric and non-parametric modelling approach

Ogbozige F J 2022. Modeling monthly rainfall of Calabar, Nigeria Using Box-Jenkins (ARIMA) method., 41(2): 131-140.

Pandey P K, Tripura H and Pandey V K (2018, December 13). Improving Prediction Accuracy of rainfall Time Series By Hybrid SARIMA-GARCH modeling springer Science + Business media, 28(3): 1125-1138.

Pham Q B, Abba S I, Usman A G, Linh N T T, Gupta V, Malik, Costache R, Vo N D and Tri D Q 2019. Potential of Hybrid Data-Intelligence Algorithms for multi-Station Modelling of rainfall. Springer Science + Business Media, 33(15): 5067-5087.

Rodrigues J and Deshpande A 2017. Prediction of Rainfall for all the States of India Using Auto-Regressive Integrated Moving Average Model and Multiple Linear Regression.

Sathish G, Narasimha L, Babu P M, Laha S and Kumar N B 2017. Time Series Analysis of monthly Rainfall for Gangetic West Bengal Using Box Jenkins SARIMA Modeling. Excellent Publishers, 6(7): 2603-2610.

Unnikrishnan P and Jothi Prakash V 2020. Hybrid SSA-Arima-ANN Model for Forecasting Daily Rainfall. Springer Science + Business Media, 34(11): 3609-3623.

Received on 10.01.2025 and Accepted on 15.02.2025