https://doi.org/10.56093/JISAS.V78I2.9
Available online at http://isas.org.in/isa/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 78(2) 2024 151–160

Statistical Modelling and Projection of Future Rainfall using SARIMA and Hybrid SARIMA-GARCH Models in Various Zones of Kerala

Gokul Krishnan K.B.[1], Vishal Mehta[1], P.K. Retheesh[2] and Ramakrishna S. Solanki[3]

[1]*Acharya Narendra Deva University of Agriculture and Technology, Ayodhya*
[2]*Regional Agricultural Research Station, Pilicode, Kasaragod, Kerala Agricultural University, Thrissur*
[3]*College of Agriculture, Waraseoni, Balaghat*

Received 21 March 2023; Revised 07 August 2024; Accepted 08 August 2024

SUMMARY

Water is an important natural resource considered as basic need for all living things around the world. The volume of pure water present in the Earth is regulated by the amount of rainfall received over the years. Sudden climatic changes are observed in throughout the world which led to flood, drought and uneven rainfall over the years. In this study, SARIMA and SARIMA-GARCH models are applied for forecasting rainfall in different zones of Kerala. The presence of heteroscedasticity in residuals obtained from SARIMA model was identified using ARCH-LM test and it was eliminated by applying SARIMA-GARCH model to the same. The ARCH-LM test results confirmed the presence of heteroscedasticity in residuals. The comparison of models used for predicting rainfall revealed that hybrid SARIMA-GARCH model is more efficient in projecting future values of rainfall in the northern and southern zones of Kerala whereas SARIMA model is showing more accuracy in the central zone of Kerala even in the presence of heteroscedasticity of residuals. The comparison of rainfall forecasted in different zones of Kerala clearly indicated that rainfall is higher in the northern zone whereas lower in the southern zone. In the northern and central zones, the rainfall showed a peak from June to September and almost negligible rainfall from December to February. The outperformed model in each zones of Kerala was applied for projection of future rainfall for next 5 years (2021-2025). Compare to previous years, the rainfall in the northern and central zones is expected to decrease whereas in southern zone of Kerala, rainfall will be almost same.

Keywords: ARCH-LM test; Heteroscedasticity; Rainfall; Residual; SARIMA; SARIMA-GARCH.

1. INTRODUCTION

Water is an essential natural resource considered as basic necessity for all the flora and fauna present in this world. Water plays vital role in growth and development of plants and animals. The nutrients and oxygen transportation in both plants and animals are supported by water. In plants for both photosynthesis and respiration, water act as an indispensable part. The water regulates body temperature in animals and it is mandatory for proper functioning of cells and organs. Water constituted almost 60 per cent body weight of an individual. It is an important fact that water cover nearly 71 per cent surface of our planet Earth. More than 97 per cent of water in the Earth is hold by oceans. Only 3 per cent of water present in the Earth is fresh water and rest of the amount (97 per cent) is deliberated as salt water. Rainfall is indicated as main source of water in all around the world. Rainfall is an important weather phenomenon since it plays crucial role in regulating amount of water present in the Earth. The abnormal changes can affect fresh water content, which directly impact the flora and fauna.

The prediction of rainfall with minimum error and maximum precision is very important. It helps in reducing the effects of uneven rainfall by taking necessary steps to minimize the damages. The climate change related issues lead to flood, drought and uneven distribution of rainfall all over the world. Flood and drought related problems arise in different parts of India all over the years. The flood happened in different

Corresponding author: Vishal Mehta
E-mail address: visdewas@gmail.com

parts of India include the Chennai flood occurred in 2015 and the Kerala flood take place in two consecutive years, 2018 and 2019 which lead to death of several people and damages to infrastructures. India is a tropical country which highly depends on agriculture as basic source of income. The annual rainfall of India is 1194mm which is mainly due to the south-west and north-west monsoon winds occur throughout India. The different crops cultivated in India highly get influenced by the alteration in rainfall, which can lead to crop failure, thus uneven distribution of rainfall can disturb the lives of people.

In order to reduce the effect of uneven rainfall, different models are applied for forecasting the changes and specific patterns of rainfall by different researchers over the years. Soltani *et al.* (2007) investigated about determination of rainfall pattern by employing autoregressive integrated moving average (ARIMA) in Iran such that seasonal behaviour of ARIMA was recognised by autocorrelation function (ACF) and partial autocorrelation function (PACF) plots and results indicated that there was a severe discrepancy in monthly rainfall pattern. Narayanan *et al.* (2013) examined prediction of pre-monsoon rainfall over western India using ARIMA model and forecasted values disclosed a major rise in pre-monsoon rainfall in North-West region of India. The pattern of south-west monsoon rainfall in North–East India was predicted by Murthy *et al.* (2018) by employing SARIMA process and results indicated that the predicted model was accurate in analysing and forecasting the future rainfall patterns. Exploration on fitting an ARIMA model was carried out by Swain *et al.* (2018) for predicting rainfall at Khordha district, Odisha, India in which outperformed model was selected based on smallest value for Akaike information criterion (AIC) and Bayesian information criterion (BIC) and results concluded that model gave accurate forecasting for a period of 20 years. Investigation on climate variables was undergone by Dimri *et al.* (2020) using seasonal ARIMA approach for mean rainfall, maximum and minimum temperature data collected over a period of 100 years (1901-2000) from the Bhagirathi river basin, Uttarakhand, India and concluded that the predicted data was almost equal to the actual data observed with some fluctuations for rainfall which was already expected. Krishnan *et al.* (2022) investigated a research on modelling of rainfall using SARMA, ANN and hybrid model and advocated that farmers must take safety measures to overcome the negative impacts of fluctuations in rainfall

Wind speed prediction was carried out by Lv and Yue (2011) employing ARIMA, ARIMA-ARCH, Wavelet-ARIMA and Wavelet-ARIMA-ARCH methods. The result revealed that Wavelet ARIMA-ARCH model showed more accuracy compared to other models used for forecasting. Relationship between rainfall–runoff was investigated by Modarres and Ouarda (2013a) using multivariate GARCH model. The result specified that conditional variance of stream-flow showed a stronger short-run endurance whereas conditional variance of rainfall consists of higher persistency. Javari (2017) investigated about determination of dynamic linear and non-linearity of monthly and annual rainfall variations using ARIMA and GARCH models in Iran. The ARIMA-GARCH hybrid model was designated as finest model for predicting rainfall and it also confirmed presence of random and non-random variability. Mishra *et al.* (2021) undertook a study about modelling and forecasting of metrological aspects using ARCH process beneath specification of various error distribution specifications. The results revealed that ARIMA-GARCH model indicated more precision in future projection of weather parameters.

The researchers all around the world developed different methods for fitting and projecting future values of weather parameters. The recent approaches regarding forecasting of rainfall include non parametric machine language approach (Praveen *et al.*, 2020), non parametyric time delay neural network (Lama *et al.*, 2022), coupled neural network-wavelet technique (Singh *et al.*, 2024), long short term memory recurrent neural network (Dotse, 2024), exponential smoothing neural network (Krishnan and Mehta, 2024) and various other hybrid methods (Krishnan *et al.,* 2022). The present study mainly engaged in predicting the rainfall with high precision using SARIMA and SARIMA-GARCH methods such that effect of heteroscedasticity is also determined.

2. MATERIALS AND METHODS

2.1 Study Area and Location

In order to undergo the research work, monthly rainfall data was collected from different zones of Kerala. The data for northern and central zones was collected from regional agricultural research station (RARS) Pilicode and Pattambi respectively, for a period of 39 years (1982-2020) whereas data was collected for 36 years (1985-2020) from RARS, Vellayani for the southern zone of Kerala. The RARS, Pilicode lies in Kasaragod district of Kerala. It is having an elevation of 15m from sea level and it is located in between $12.1997^o N$ latitude and $75.1633^o E$ longitude. It was noted that average annual rainfall at Pilicode was 3379mm. The RARS, Pattambi situated in Palakkad, district of Kerala. It is positioned in among $10.8057^o N$ latitude and $76.1957^o E$ longitude with of 63m elevation. The average annual rainfall observed at Pattambi was 1838mm.The RARS, Vellayani is located in Trivandrum, district of Kerala. The RARS, Vellayani is placed at $8.4316^o N$ latitude and $76.986^o E$ longitude with height of 8m from sea level. The average annual rainfall at Vellayani is 1704mm. The three different locations are marked in Fig.1.

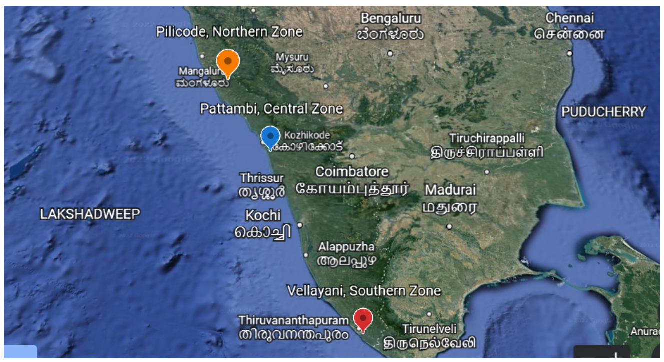

Fig. 1. Study area map for the northern, central and southern zones of Kerala (Source: Google Earth)

2.2 2.2. Methodology

2.2.1 Seasonal autoregressive integrated moving average (SARIMA)

The Seasonal ARIMA model is considered as the extension of Box-Jenkins ARIMA model. The ARIMA (p,d,q) model was developed by Box and Jenkins (1976) for modelling and predicting future values of time series data, where p denotes autoregressive part, d indicates degree of differencing and q represented the moving average part (Sultana and Hasan, 2015). The Seasonal ARIMA model is expressed as the consolidation of both seasonal and non-seasonal parameters. Non-seasonal part of Seasonal ARIMA model is same as ARIMA (p,d,q) model whereas seasonal part of Seasonal ARIMA model is represented as $(P,D,Q)_s$ where, S indicates periodicity of the data.

SARIMA model is mathematically represented by Wang *et al.* (2005a) as:

$$Y_t = (1-B)^d \left(1-B^S\right)^D e_t \tag{1}$$

where, Y_t is original value, e_t is the residual term, B represents back shift operator, $\left(1-B^d\right)$ is the differencing of order d of non-seasonal part and $\left(1-B^S\right)^D$ represent the differencing of order D of seasonal part of SARIMA model (Rahman and Lateh, 2017).

2.2.1.1 Test for checking stationary using augmented ducky fuller (ADF)

Prior to undergo fitting of SARIMA model, stationary of time series data is ensured by employing ADF test (Banerjee *et al.* 1993; Gerretsadikan and Sharma, 2011). The time series data with trend and seasonality are non stationary. ADF test indicate whether the data is having constant mean and variance such that it never get altered with time (Said and Dickey, 1984). The hypothesis for ADF test is expressed as the following, if H_o : presence of unit root or data is not stationary, if H_1 : absence of unit root or data is stationary. If data is stationary, separate test for seasonality is not needed (Franses, 1991; Kwiatkowski *et al.*, 1992).

The ADF test statistic is expressed as the following:

$$t = \frac{\hat{\rho}-1}{\hat{\sigma}_p} \tag{2}$$

where, $\hat{\rho}$ indicate maximum likelihood estimate of ordinary least square estimator and $\hat{\sigma}_{\hat{p}}$ is the standard error for ordinary least square estimator. If the value of $\hat{\rho}$ =1, the data is not stationary, otherwise data is considered as stationary (Wang *et al.*, 2005b).

2.2.1.2 Selection criteria of model

The best SARIMA model suitable for the time series data is based AIC, BIC and Akaike information corrected criterion (AICC) (Chaudhuri and Dutta, 2014). The model showing least value for above mentioned criteria is considered as best performing model. At present conditions, statistical software directly select suitable model.

$$AIC = -2 \times \ln(L) + 2 \times k \tag{3}$$

$$BIC = -2 \times \ln(L) + 2 \times \ln(N) \times k \quad (4)$$

$$AICC = -2 \times \ln(L) + 2 \times k + \frac{2k(k+1)}{N-k-1} \quad (5)$$

where, L is maximum likelihood estimate, N is number of noted observations and k is number of parameters which are estimated.

2.2.1.3 Diagnostic checking using Ljung-Box test

The residuals obtained from SARIMA model is tested for detecting the presence of autocorrelation using Ljung-Box (1978) test. If the autocorrelation is absent in residuals, fitted model is accurate and otherwise not. The hypothesis for Ljung-Box test is described as, if H_o: absence of autocorrelation in residuals whereas, if H_1: presence of autocorrelation in residuals (Krishan *et al*., 2023). The test statistics of Ljung-Box test is expressed as the following:

$$Q = n(n+2)\sum_{k=1}^{m}\frac{T_k^2}{n-k} \quad (6)$$

where, n is total number of measurements, T_k is the autocorrelation among residuals at lag k and m is tested number of lags (Wang *et al.*, 2005a).

2.2.2 Generalised autoregressive conditional heteroscedasticity (GARCH)

GARCH model is fitted to the residuals obtained from SARIMA model after conducting the test for identifying the presence of time dependent variance for the same. It is an important fact that even a well fitted SARIMA model may show variance which changes with time. ARCH model developed by Engle (1982) and it was improved by Bollerslev (1986) to establish the GARCH model through addition of lagged conditional variance for better smoothening of series. GARCH (A, B) is mathematically expressed by Modarres and Ouarda (2013b) as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{A}\alpha_A\varepsilon_{t-A}^2 + \sum_{j=1}^{B}\beta_B\sigma_{t-B}^2 \quad (7)$$

Where parameter of GARCH (A, B) is indicated as $\alpha_1 \ldots \alpha_A$ and $\beta_1 \ldots \beta_B$ respectively and w is a constant.

2.2.2.1 Tests for heteroscedasticity using ARCH-LM test

ARCH-LM test is employed for detecting presence of ARCH effect in the residuals obtained after fitting SARIMA model for time series data. In ARCH-LM test, Lagrange Multiplier was employed to determine the ARCH effect. The hypothesis for ARCH-LM test is given as, if H_o: presence of homogeneity in residuals whereas, if H_1: absence of homogeneity in the residuals. The test statistics of ARCH-LM test is described as the following:

$$LM = nR^2 \quad (8)$$

where, R^2 indicate coefficient of determination and n denote total number of observations (Oktaviani and Setiawan, 2021).

2.2.3 SARIMA-GARCH hybrid model

The hybrid model was obtained after merging both SARIMA and GARCH models which was proposed by Pandey *et al.* (2019). The hybrid SARIMA- GARCH model performed in such a manner that conditional mean was explained by SARIMA model whereas conditional variance was defined by GARCH model. SARIMA-GARCH hybrid model is expressed as:

$$\begin{cases} Y_t = (1-B)^d (1-B^S)^D e_t \\ \sigma_t^2 = \omega + \sum_{i=1}^{A}\alpha_A\varepsilon_{t-A}^2 + \sum_{j=1}^{B}\beta_B\sigma_{t-B}^2 \end{cases} \quad (9)$$

2.3 Criteria for model evaluation

Evaluation of fitted models is assessed using various error values comprising mean square error (MSE), root mean square error (RMSE) and mean absolute error (MAE). DM test was also carried out to determine whether forecasted values using SARIMA and SARIMA-GARCH are significantly different or not.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (10)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (11)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (12)$$

Where y_i indicate actual value, $\hat{y}_i$ denote projected value and n designate the total number of observations.

2.3.1 Diebold Mariano (DM) test

Diebold and Mariano (2002) test was carried to check whether two different models used for prediction are significantly different or not.

h_o: The forecasted values using two different models are not significantly different

h_1: The forecasted values using two different models are significantly different

$$\text{DM} = \frac{\bar{d}}{\sqrt{\left[\gamma_0 + 2\sum_{k=1}^{h-1}\gamma_k\right]}} \qquad (13)$$

Where $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ and $\gamma_k = \frac{1}{n}\sum_{i=k+1}^{n}\left(d_i - \bar{d}\right)\left(d_{i-k} - \bar{d}\right)$, d_i is the deviation among square of the difference between error values (Actual-Forecasted) using two different models, h is the forecast horizon and k is the lag.

3. RESULTS AND DISCUSSION

In current study, the main focus was to obtain future values of rainfall in the northern, central and southern zones of Kerala with maximum accuracy. The methods employed in the study are SARIMA and SARIMA-GARCH hybrid models. The study also consider the heteroscedasticity present in residuals obtained after fitting SARIMA model which may lead to decrease the efficiency of fitted model. Prior to undergo fitting of SARIMA model, the time series plots of rainfall in different zones are depicted in Fig. 2.

The time series plot presented in Fig. 2 clearly indicated that the northern zone is receiving the maximum rainfall over the years and least amount of rainfall is indicated in the Southern zone of Kerala. The plots also suggested that rainfall is increasing and decreasing over the years which concluded the absence of specific pattern. After depicting the time series plots for rainfall, next step is ensuring whether the time series data is stationary using ADF test, which is mandatory for fitting SARIMA model which is presented in Table 1.

The results described in Table 1 declared that rainfall data is stationary collected from different zones of Kerala, which indicated that separate test for seasonality is not required. Next step was fitting

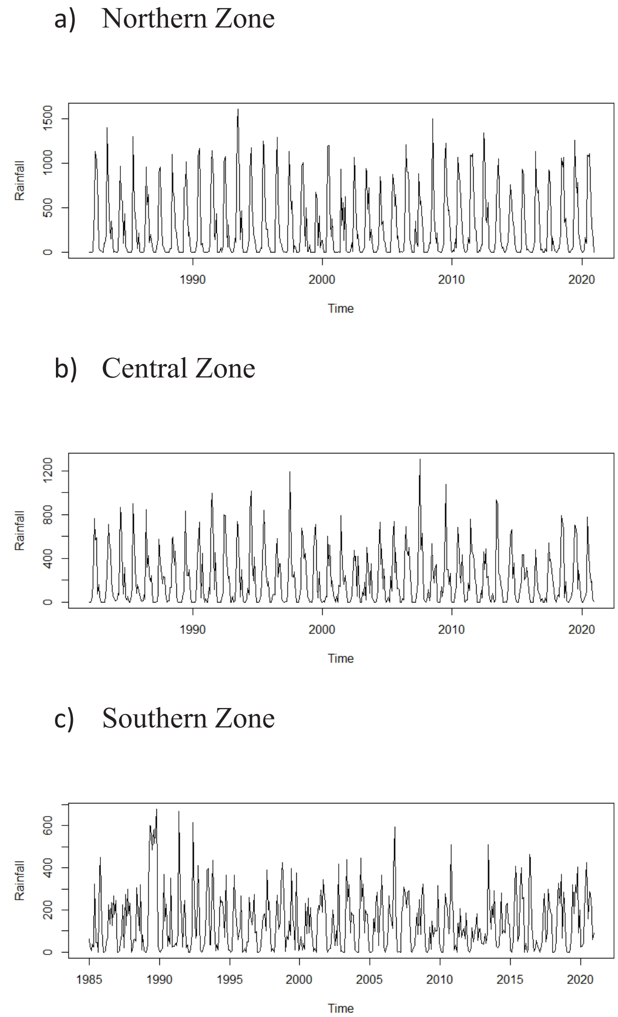

Fig. 2. Time series plot for rainfall indifferent zones of Kerala

Table 1. ADF test results for rainfall in different zones of Kerala

Weather Parameter	Zone	Test Statistic	P value
Rainfall	Northern	-15.66	0.01
	Central	-14.49	0.01
	Southern	-8.69	0.01

of SARIMA model for rainfall data collected from different zones of Kerala. Before undergoing fitting of SARIMA model, the time series data was split in two parts, training and testing sets. Training set for the northern and central zones consists of data from 1982 to 2015 (34 years) and for the southern zone it was from 1985 to 2015 (31 years). The data from 2016 to 2020 was considered as the testing set for the northern, central and southern zones of Kerala. R software was used in this study for fitting SARIMA (Trapletti and Hornik, 2013; Hyndman and Khandakar, 2008) and Gretl (Mixon Jr, 2009; Adkins, 2010; Cottrell and Lucchetti, 2012) was used for obtaining SARIMA-GARCH hybrid model and also for undergoing forecasting for next 5 years (2021-2025). The R software identified outperforming model using different criteria, which is described in Table 2.

Table 2. Model selection criteria for different zones of Kerala

Weather Parameter	Zone	AIC	AICC	BIC
Rainfall	Northern	5263.49	5263.55	5275.43
	Central	5048.43	5048.46	5056.39
	Southern	4585.21	4585.44	4608.72

The best performing model for rainfall in different zones are selected on least value for AIC, BIC and AICC which clearly represented in Table 2. The SARIMA model selected for rainfall in the northern, central and southern zones of Kerala with its parameters are illustrated in Table 3.

Table 3. SARIMA model with its parameters for rainfall in different zones of Kerala

Weather Parameter	Zone	Model	Parameters
Rainfall	Northern	SARIMA $(0,0,0)\,(2,1,0)_{12}$	sar1-0.7509 sar2 -0.3121
	Central	SARIMA $(0,0,0)\,(1,1,0)_{12}$	sar1 -0.5949
	Southern	SARIMA $(1,0,1)\,(2,0,0)_{12}$	ar1 0.5995 ma1 -0.3053 sar1 0.3154 sar2 0.2782

Model showing maximum performance in fitting rainfall is showed in Table 3 with its parameters. In order to check the autocorrelation between residuals gained from SARIMA model, ACF plots are obtained which is represented in Fig. 3.

The ACF plots given in Fig. 3 suggested that autocorrelation is absent in all the three cases. For northern zone, all lags lies within the limit but for central zone the lags 12 and 24 are passing the boundary whereas for southern zone, lags 4 and 24 cross the limit. Even though some lags are crossing, it may not indicate the presence of autocorrelation in residuals. However, for confirming the presence or absence of autocorrelation between residuals of SARIMA model, diagnostic checking was carried out using Ljung-Box test. Results of Ljung-Box test is clearly conveyed in Table 4.

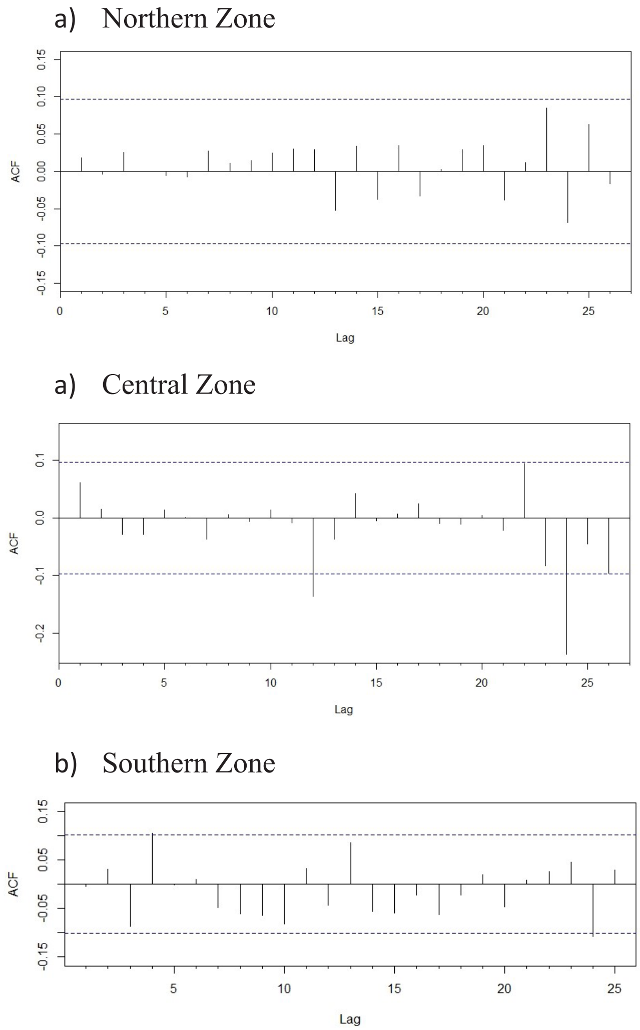

Fig. 3. ACF plots for residuals of SARIMA model for rainfall in different zones

Table 4. Results of Ljung-Box test statistic for different zones of Kerala

Weather Parameter	Zone	Q Statistic	P value
Rainfall	Northern	11.87	0.96
	Central	29.32	0.89
	Southern	30.27	0.30

The Ljung-Box results described in Table 4 clearly suggested that autocorrelation is absent between the residuals obtained after fitting the models for the rainfall in the northern, central and southern zones of Kerala, which clearly indicated that SARIMA model fitted is performing with good precision. Next step was detecting the presence of heteroscedasticity in residuals using ARCH-LM test. The results of ARCH-LM test for rainfall in different zones described in Table 5.

Table 5. Results of ARCH-LM test statistic for different zones of Kerala

Weather Parameter	Zone	Test Statistic	P value
Rainfall	Northern	24.441	0.01***
	Central	56.230	0.01***
	Southern	36.950	0.01***

*** significance at 1% level

The results presented in Table 5 clearly showed the incidence of ARCH effect in all three cases. After detecting the presence of ARCH effect, next step was fitting GARCH model for the residuals. Gretl open software was used for identifying suitable GARCH model for the residuals. The suitable SARIMA-GARCH model and its parameters were presented in Table 6.

Table 6. SARIMA model with its parameters for different zones of Kerala

Weather Parameter	Zone	Model	Parameters	Standard error	Z value
Rainfall	Northern	SARIMA $(0,0,0)$ $(2,1,0)_{12}$ - GARCH $(1,1)$	μ=20876.3 ω= 2.465 α = 0.489 β =3.087e-012	2538.24 7.478 0.159 0.063	8.225 0.329 3.071 4.87e-11
	Central	SARIMA $(0,0,0)$ $(1,1,0)_{12}$ - GARCH $(1,1)$	μ= 11209.2 ω= 2.779 α = 0.358 β =0.111	1797.40 6.056 0.106 0.084	6.236 0.458 3.337 1.331
	Southern	SARIMA $(1,0,1)$ $(2,0,0)_{12}$ - GARCH $(1,1)$	μ=8956.52 ω= -4.097 α = 0.301 β =0.002	1622.41 5.398 0.087 0.117	5.52 −0.759 3.466 0.018

Results presented in Table 6 indicated that GARCH (1, 1) is the best model selected for the residuals obtained after fitting SARIMA model for rainfall in different zones of Kerala. Even though the value of GARCH coefficient value in north zone is low, ignoring it affected the accuracy of rainfall forecasting (The error value are comparatively higher for ARCH (1) than GARCH (1, 1)). Succeeding step was comparison and evaluation of performance of SARIMA and hybrid SARIMA-GARCH model in such a way that effects of removal of heteroscedasticity was determined. The rainfall for next 5 years (2016-2020) was predicted using SARIMA and hybrid SARIMA-GARCH models and it was compared with the testing set. The evaluation criterion was based on MSE, RMSE and MAE values and also based on DM test results which are illustrated in Table 7.

The results given in Table 7 declared that SARIMA-GARCH model is showing better performance in forecasting rainfall in northern and southern zones of Kerala whereas in the central zone of Kerala SARIMA model is expected to predict future values with more efficiency. However, the DM statistic concluded that forecasted values of rainfall in the northern zone using SARIMA and SARIMA-GARCH model is significantly different but for the central and southern zones of Kerala it was not significantly different.

The best selected model was used to provide future values of rainfall from 2021 to 2025 in different zones of Kerala which is portrayed in Fig. 4.

The comparison of rainfall forecasted in the northern, central and southern zones of Kerala depicted in Fig. 4 clearly indicated that rainfall in northern zone is comparatively higher. The southern zone indicated with the least amount of rainfall for next 5 years with a steady rainfall over the year. In the northern and central zones, the rainfall showed a peak from June

Table 7. Comparison and Evaluation of SARIMA and hybrid SARIMA-GARCH model for rainfall in different zones of Kerala

Weather Parameter	Zone	Model	MSE	RMSE	MAE	DM Test Statistic	P Value
Rainfall	Northern	SARIMA	29697.804	172.330	102.312	3.155	0.0025
		SARIMA-GARCH	**29375.830**	**171.393**	**101.903**		
	Central	**SARIMA**	**18612.328**	**136.427**	**94.495**	-0.255	0.6825
		SARIMA-GARCH	18799.645	137.111	95.556		
	Southern	SARIMA	14783.950	121.589	105.466	0.471	0.6392
		SARIMA-GARCH	**14723.010**	**121.338**	**104.783**		

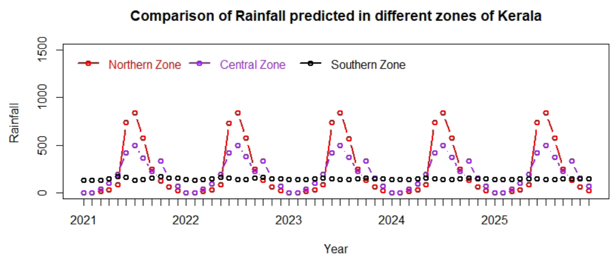

Fig. 4. Predicted rainfall from 2021 to 2025 in different zones of Kerala

to September and almost negligible rainfall from December to February. Compare to previous years, the rainfall in the northern and central zones is expected to decrease whereas in southern zone of Kerala, rainfall will be almost same for next 5 years (2021-2025).

4. CONCLUSION

Water is an important natural resource considered as basic need for all living things around the world. The volume of pure water present in the Earth which is consumed by the human beings and plants are regulated by the amount of rainfall received over the years. Sudden climatic changes are observed in throughout the world which led to flood, drought and uneven rainfall over the years. Thus developing new strategies for forecasting the rainfall with maximum efficiency is mandatory for taking precautionary measures. In this study, SARIMA and SARIMA-GARCH models are applied for forecasting rainfall in different zones of Kerala. The presence of heteroscedasticity in residuals obtained from SARIMA model was identified using ARCH-LM test and it was eliminated by applying SARIMA-GARCH model to the same. R software was used in the study for fitting SARIMA and gretl software was to fit GARCH model for residuals. The outperformed SARIMA model for rainfall in the northern, central and southern zones are SARIMA $(0,0,0)\ (2,1,0)_{12}$, SARIMA$(0,0,0)\ (1,1,0)_{12}$ and SARIMA$(1,0,1)\ (2,0,0)_{12}$ respectively. The ARCH-LM test results confirmed the presence of heteroscedasticity in all the three cases. The most efficient hybrid model selected for predicting rainfall in the northern, central and southern zones of Kerala are SARIMA$(0,0,0)\ (2,1,0)_{12}$-GARCH$(1,1)$, SARIMA $(0,0,0)\ (1,1,0)_{12}$-GARCH$(1,1)$ and SARIMA$(1,0,1)$ $(2,0,0)_{12}$-GARCH$(1,1)$ respectively. The comparison of models used for predicting rainfall revealed that hybrid SARIMA-GARCH model is more efficient in projecting future values of rainfall in the northern and southern zones of Kerala whereas SARIMA model is showing more accuracy in the central zone of Kerala even in the presence of heteroscedasticity of residuals. However, the DM statistic concluded that forecasted values of rainfall in the northern zone using SARIMA and SARIMA-GARCH model is significantly different but for the central and southern zones of Kerala it was not significantly different. The most efficient model in forecasting rainfall based on error values in each zones of Kerala is applied for projecting future values of rainfall for next 5 years. The future values predicted indicated that the northern and central zones will have fall in rainfall and the southern zone will have almost same rainfall for next 5 years (2021-2025).

ACKNOWLEDGEMENTS

The authors are expressing our gratitude towards the reviewer for the valid suggestions which enhanced the quality of the manuscript. The authors are thankful towards Associate Director of Research (ADR), Regional Agricultural Research Station (RARS), Pilicode (Northern Zone), Pattambi (Central Zone) and Vellayani (Southern Zone) under Kerala Agricultural University (KAU) Thrissur for providing data and support while undertaking research. The authors are also showing our gracefulness towards the PG and PhD research scholars, teaching and non-teaching staffs of Department of Agricultural Statistics, College of Agriculture, Kumarganj, Acharya Narendra Deva University of Agriculture and Technology, Ayodhya, Uttar Pradesh.

REFERENCES

Adkins, L.C. (2010). Using gretl for Principles of Econometrics, Version 1.3131. https://learneconometrics.com/gretl/ebook.pdf

Banerjee, A., Dolado, J.J., Galbraith, J.W. and Hendry, D. (1993). *Co-integration, error correction, and the econometric analysis of non-stationary data*. Oxford university press. https://global.oup.com/academic/product/co-integration-error-correction-and-the-econometric-analysis-of-non-stationary-data-9780198288107?cc=pl&lang=en&

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, **31**, 307-327. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf.

Box, G.E.P. and Jenkins, G.M. (1976). Time series analysis: forecasting and control. Holden-Day, Boca Raton. https://books.google.co.in/books/about/Time_Series_Analysis.html?id=5BVfnXaq03oC&redir_esc=y

Chaudhuri, S. and Dutta, D. (2014). Mann–Kendall trend of pollutants, temperature and humidity over an urban station of

India with forecast verification using different ARIMA models. *Environmental monitoring and assessment*, **186**(8), 4719-4742. https://www.researchgate.net/profile/Sutapa-Chaudhuri/publication/261363864.pdf.

Cottrell, A. and R, Lucchetti. 2012. Gretl User's Guide. http://sourceforge.net/projects/gretl/files/manual/gretl-guide.pdf/download.

Diebold, F.X. and Mariano, R.S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, **20**(1), 134-144. https://www.tandfonline.com/doi/abs/10.1198/073500102753410444

Dimri, T., Ahmad, S. and Sharif, M. (2020). Time series analysis of climate variables using seasonal ARIMA approach. *Journal of Earth System Science*, **129**(1), 1-16. https://www.ias.ac.in/article/fulltext/jess/129/00/0149.pdf.

Dotse, S.Q. (2024). Deep learning–based long short-term memory recurrent neural networks for monthly rainfall forecasting in Ghana, West Africa. *Theoretical and Applied Climatology*, **155**(4), 3033-3045. https://ui.adsabs.harvard.edu/abs/2023ThApC.155.3033D/abstract

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**(4), 987-1008. http://www.econ.uiuc.edu/~econ536/Papers/engle82.pdf

Franses, P.H. (1991). Seasonality, non-stationarity and the forecasting of monthly time series. *International Journal of forecasting*, **7**(2), 199-208. https://doi.org/10.1016/0169-2070(91)90054-Y

Gerretsadikan, A. and Sharma, M.K. (2011). Modeling and forecasting of rainfall data of mekele for Tigray region (Ethiopia). *Statistics and Applications*, **9**(1-2), 31-53. https://www.ssca.org.in/media/3MKSharma.pdf. Google Earth. https://www.google.com/intl/en_in/earth/

Hyndman, R.J. and Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, **26**(3). https://www.jstatsoft.org/article/view/v027i03

Javari, M. (2017). Assessment of dynamic linear and non-linear models on rainfall variations predicting of Iran. *Agricultural Engineering International: CIGR Journal*, **19**(2), 224-240. https://cigrjournal.org/index.php/Ejounral/article/view/4193

Krishnan, G.K.B., Mehta, V. and Yadav, R.S. (2022). Assesment of Future Pattern of Rainfall in Different Zones of Kerala Using Incorporation of SARIMA, ANN and Hybrid SARIMA-ANN Models. *Economic Affairs*, **67**(05), 823-832. http://ndpublisher.in/admin/issues/EAv67n5q.pdf.

Krishnan G.K.B., Mehta V. and Rai, V.N. (2023). Stochastic modelling and forecasting of relative humidity and wind speed for different zones of Kerala. Mausam 74 (04):1053-1064. https://mausamjournal.imd.gov.in/index.php/MAUSAM/article/view/5603

Krishnan, G.K.B. and Mehta, V. (2024). Comparison Study on Modelling and Prediction of Weather Parameters Combining Exponential Smoothing and Artificial Neural Network Models in Different Zones of Kerala. *Environment and Ecology*, **42**(3), 1094-1103. https://environmentandecology.com/wp-content/uploads/2024/07/MS25-Comparison-Study-on-Modelling-and-Prediction-of.pdf

Kwiatkowski, D., Phillips, P.C., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. *Journal of econometrics*, **54**(1-3), 159-178. https://doi.org/10.1016/0304-4076(92)90104-Y

Lama, A., Singh, K.N., Singh, H., Shekhawat, R., Mishra, P. and Gurung, B. (2022). Forecasting monthly rainfall of Sub-Himalayan region of India using parametric and non-parametric modelling approaches. *Modeling Earth Systems and Environment*, **8**, 837-845. https://www.semanticscholar.org/paper/Forecasting-monthly-rainfall-of-Sub-Himalayan-of-Lama-Singh/77297644698b184293276a51ad31c6fb0120129e

Ljung, G.M. and Box G.E.P. (1978). On a measure of a lack of fit in time series models. *Biometrika*, **65**, 297-303. https://apps.dtic.mil/sti/pdfs/ADA049397.pdf

Lv, P. and Yue, L. (2011). Short-term wind speed forecasting based on non-stationary time series analysis and ARCH model. In 2011 International Conference on Multimedia Technology, 2549-2553. https://www.researchgate.net/publication/241189798_Shortterm_wind_speed_forecasting_based_on_nonstationary_time_series_analysis_and_ARCH_model

Mishra, P., Fatih, C., Vani, G., Lavrod, J.M., Jain, V., Dubey, A. and Choudhary, A.K. (2021). Modeling and forecasting of metrological factors using ARCH process under different errors distribution specification. *MAUSAM*, **72**(2), 301-312. https://mausamjournal.imd.gov.in/index.php/MAUSAM/article/view/618/535

Mixon Jr, J. 2009. GRETL: an econometrics package for teaching and research. *Managerial Finance*, **36**(1), 71-81. https://doi.org/10.1108/03074351011006856

Modarres, R. and Ouarda, T.B.M.J. (2013b). Generalized autoregressive conditional heteroscedasticity modelling of hydrologic time series. *Hydrological Processes*, **27**(22), 3174–3191. https://rezamodarres.iut.ac.ir/sites/rezamodarres.iut.ac.ir/files/u129/pdf.

Modarres, R. and Ouarda, T.B. (2013a). Modeling rainfall–runoff relationship using multivariate GARCH model. *Journal of Hydrology*, **499**, 1-18. https://rezamodarres.iut.ac.ir/sites/rezamodarres.iut.ac.ir/files/file_pubwdet/pdf.

Murthy, K.N., Saravana, R. and Kumar, K.V. (2018). Modeling and forecasting rainfall patterns of southwest monsoons in North–East India as a SARIMA process. *Meteorology Atmospheric Physics*, **130**(1), 99-106. https://d1wqtxts1xzle7.cloudfront.net/87826879/s00703-017-0504-220220621-1-ocm8nt-libre.pdf

Narayanan, P., Basistha, A., Sarkar, S. and Kamna, S. (2013). Trend analysis and ARIMA modelling of pre-monsoon rainfall data for western India. *ComptesRendus Geoscience*. **345**(1), 22-27. https://pdf.sciencedirectassets.com/272261/1-s2.0-S1631071313X00021/1-s2.0-S1631071312002416/main.pdf.

Oktaviani, F. and Setiawan, I. (2021). Forecasting sea surface temperature anomalies using the SARIMA ARCH/GARCH model. In *Journal of Physics: Conference Series*, **1882**(1), 012-020. https://iopscience.iop.org/article/10.1088/1742-6596/1882/1/012020/pdf.

Pandey, P.K., Tripura, H. and Pandey, V. (2019). Improving prediction accuracy of rainfall time series By Hybrid SARIMA–GARCH modeling. *Natural Resources Research*, **28**(3), 1125-1138. https://www.researchgate.net/publication/329627377_Improving_Prediction_Accuracy_of_Rainfall_Time_Series_By_Hybrid_SARIMA-GARCH_Modeling

Praveen, B., Talukdar, S., Shahfahad, Mahato, S., Mondal, J., Sharma, P., Islam, A.R.M.T. and Rahman, A. (2020). Analyzing trend and forecasting of rainfall changes in India using non-parametrical and machine learning approaches. *Scientific reports*, **10**(1), 10342. https://pubmed.ncbi.nlm.nih.gov/32587299/

Rahman, M.R. and Lateh, H. (2017). Climate change in Bangladesh: a spatio-temporal analysis and simulation of recent temperature and rainfall data using GIS and time series analysis model. *Theoretical and Applied Climatology*, **128**(1-2), 27-41. https://www.researchgate.net/profile/Habibah-Lateh/publication/285589104.pdf.

Said, S.E. and Dickey, D.A. (1984). Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika*, **71**, 599–607. http://www.larrylisblog.net/WebContents/Financial%20Models/ADFTest.pdf.

Singh, S., Kumar, D., Vishwakarma, D.K., Kumar, R. and Kushwaha, N.L. (2024). Seasonal rainfall pattern using coupled neural network-wavelet technique of southern Uttarakhand, India. *Theoretical and Applied Climatology*, **155**, 5185-5201. https://www.researchgate.net/publication/362206924_Seasonal_rainfall_pattern_using_coupled_neural_network-wavelet_technique_of_of_southern_Uttarakhand_India

Soltani, S., Modarres, R. and Eslamian, S.S. (2007). The use of time series modeling for the determination of rainfall climates of Iran. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, **27**(6), 819-829. https://rezamodarres.iut.ac.ir/sites/rezamodarres.iut.ac.ir/files/file_pubwdet/pdf.

Sultana, N. and Hasan, M.M. (2015). Forecasting Temperature in the Coastal Area of Bay of Bengal-An Application of Box-Jenkins Seasonal ARIMA Model. *Civil and Environmental Research*, **7**(8), 256-272. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf.

Swain, S., Nandi, S., Patel, P. (2018). Development of an ARIMA Model for Monthly Rainfall Forecasting over Khordha District, Odisha, India. In: Sa, P., Bakshi, S., Hatzilygeroudis, I., Sahoo, M. (eds) Recent Findings in Intelligent Computing Techniques . Advances in Intelligent Systems and Computing, vol 708. Springer, Singapore. https://www.springerprofessional.de/development-of-an-arima-model-for-monthly-rainfall-forecasting-o/16251136

Trapletti, A. and Hornik, K. (2013). tseries: Time Series Analysis and Computational Finance, R package version 0.10-32. *https://cran.r-project.org/package=tseries*

Wang, W., Van Gelder, P.H.A., Vrijling, J.K. and Ma, J. (2005 (a)). Testing and modelling autoregressive conditional heteroscedasticity of streamflow processes. *Nonlinear Processes in Geophysics*, **12**, 55-6. https://npg.copernicus.org/articles/12/55/2005/npg-12-55-2005.pdf.

Wang, W., Van Gelder, P.H.A.J. M. and Vrijling, J.K. (2005 (b)). Trend and stationarity analysis for streamflow processes of rivers in western Europe in the 20th century. In Proceedings: IWA International Conference on Water Economics, Statistics, and Finance Rethymno, Greece (Vol. 810). London: IWA. https://www.researchgate.net/profile/Phajm-Gelder/publication/228636735.pdf.