MAUSAM, **76**, 2 (April 2025), 365-372

MAUSAM

DOI : https://doi.org/10.54302/mausam.v76i2.5487
Homepage: https://mausamjournal.imd.gov.in/index.php/MAUSAM

UDC No. 551.577 : 551.509.3

Application of SARIMA model for precipitation modelling driven by exogenous variables

MD YEASIN, K. N. SINGH, RAMASUBRAMANIAN V., RANJIT KUMAR PAUL and ACHAL LAMA*

ICAR-Indian Agricultural Statistics Research Institute, New Delhi – 110 012, India

(*Received 7 March 2022, Accepted 23 August 2024*)

***e mail : achal.lama@icar.gov.in**

सार – मौसमी ऑटोरिग्रैसिव इंटीग्रेटेड मूविंग एवरेज (SARIMA) मॉडल ने मौसमी पूर्वानुमान लगाने की अपनी क्षमता के कारण अपनी शुरुआत से ही लोकप्रियता हासिल की है। आमतौर पर, SARIMA मॉडल मौसमीता को पकड़ता है लेकिन मौसमी प्रक्रिया में बहिर्जात चर (चरों) के प्रभाव पर विचार नहीं करता है। इसलिए, इस अध्ययन का उद्देश्य SARIMA-X मॉडल को अनुभवजन्य रूप से प्रस्तुत करना और लागू करना है जो मौसमीता के साथ-साथ प्रभावित करने वाले कारकों (X) के प्रभावों को भी ध्यान में रख सकता है। जलवायु परिवर्तन मानव अस्तित्व के सामने सबसे बड़ी वैश्विक चुनौती बन गई है और इसका प्रभाव सामाजिक, आर्थिक और पर्यावरणीय चुनौतियों के संबंध में बहुआयामी होगा। इस पांडुलिपि का उद्देश्य बैंगलोर, भारत के वर्षा समय श्रृंखला डेटा का पूर्वानुमान लगाना है। वर्षा श्रृंखला के विश्लेषण और मॉडलिंग में नियोजित कार्यप्रणाली बहिर्जात चर तापमान, सापेक्ष आर्द्रता और सतही दबाव के साथ SARIMA-X मॉडल थी। इस पांडुलिपि में, हमने SARIMA-X मॉडल के साथ-साथ इसकी अनुमान प्रक्रिया पर संक्षेप में चर्चा की है। प्रस्तावित मॉडल का निदान किया गया और परिणामों से पता चला कि मॉडल पर्याप्त और संक्षिप्त था। प्रस्तावित मॉडल की तुलना पारंपरिक SARIMA मॉडल से की गई है। इस तुलनात्मक अध्ययन से यह निष्कर्ष निकलता है कि मौसमी मॉडलिंग में बहिर्जात कारकों का उपयोग करना सर्वोत्तम है।

ABSTRACT. The Seasonal Autoregressive Integrated Moving Average (SARIMA) model has gained popularity since its inception due to its ability to forecast seasonality. Usually, the SARIMA model captures the seasonality but does not consider the effect of the exogenous variable(s) in the seasonality process. Hence, this study aims to empirically introduce and implement the SARIMA-X model which can account for seasonality as well as the effects of influencing factors (X). Climate change has become the foremost global challenge facing human existence and the effect will be multifaceted with respect to social, economic and environmental challenges. This manuscript aims to forecast the precipitation time series data of Bangalore, India. The methodology employed in the analysis and modelling of precipitation series was the SARIMA-X model with exogenous variables temperature, relative humidity and surface pressure. In this manuscript, we have briefly discussed the SARIMA-X model along with its estimation procedure. The proposed model was diagnosed and the results showed that the model was adequate and parsimonious. The proposed model has compared with the traditional SARIMA model. The supremacy of using exogenous factors in seasonality modelling is concluded by this comparative study.

Keywords – Exogenous variables, Modelling, Precipitation, SARIMA model, SARIMA-X model.

1. Introduction

Climate change is one of the prime global challenges that the mother earth is facing. The effect of climate change will be multifaceted with respect to social, economic, and environmental challenges. According to an IPCC (2018) special report at 1.5 degrees Celsius warming, 6 percent of the insects, 8 percent of the plants, and 4 percent of the vertebrates will be at risk of extinction. Climate change modifies rainfall patterns across the world. Climate change affects all countries across the world in different magnitudes. In respect to impact of climate change, India is enlisted to group of vulnerable countries. According to the IMD report, India has evidence of a change in amount, frequency and intensity of rainfall in various states in the last 30 years (https://internal.imd.gov.in/). Weather forecasting plays a crucial role for crop monitoring and pest forewarning system. According to a multivariate statistical study by Puvaneswaran (1990), the key variables to examine climatic processes are a humidity factor, a temperature factor and a rainfall factor. Precipitation is highly vulnerable to climate change with major consequences for agricultural production (Tang *et al*., 2019). Precipitation is among the most significant criteria of farming management since it has a considerable effect on crop development, growth and yields, on the occurrence of pests and diseases, fertilizer requirements, etc. Precipitation is not easy to forecast, because it depends on time and space (Hashim *et al*., 2016). Numerous attempts have been made to model and forecast the frequency, intensity and magnitude of rainfall using different methodologies. Earlier time prediction approaches such as the simple quantitative approach have been used for rainfall predictions, but it has become unreliable due to evolving seasonal rainfall patterns. In the last few decades, a wide variety of techniques such as regression analysis, time series analysis, soft computing technique and hybrids methodology have been used for modelling and forecasting precipitation. Sadhuram and Murthy (2008) developed a linear multiple regression model to predict Indian summer monsoon rainfall (ISMR). Prasad *et al*. (2010) developed a multi-predictor logistic regression model for forecasting three research areas, namely India overall, Orissa on the east coast and Gujarat on the west coast, to estimate average monthly precipitation. Murthy *et al*. (2018) conducted an empirical study for modeling and forecasting South-West monsoon rainfall patterns in North-East India. The adopted Seasonal Autoregressive Integrated Moving Average (SARIMA) and Analysis of Means (ANOM) methodology to successfully capture the existing pattern of rainfall of that given area. Lama *et al.* (2021) used several parametric models (such as the SARIMA model and exponential autoregressive (EXPAR) model) and non-parametric models (Time Delay Neural Network (TDNN) model) for forecasting rainfall of the Sub-Himalayan region of India and compared their prediction ability. Harun *et al.* (2013) compared the statistical method and artificial intelligence (AI) for rainfall prediction. They selected Auto-Regressive Integrated Moving (ARIMA) and Adaptive Splines Threshold Autoregressive (ASTAR), as a statistic model, and as AI, a combination of Genetic Algorithm-Neural Network (GA-NN) was used. Dimri *et al*. (2020) investigated the monthly mean minimum and maximum temperatures and the precipitation for the state Uttarakhand, India, by using the SARIMA model. They found that the forecasted value of both parameter fits well with the estimated trend value of the data. As the SARIMA model does not include the information of exogeneous variables, it is unable to capture the effect of extraneous factors. The effect of exogenous variables plays a very crucial role in precipitation forecasting. Gutierrez-Lopez *et al*. (2019) used a number of meteorological factors such as humidity, surface temperature, atmospheric pressure, and dewpoint to predict short-term precipitation. Precipitation is influenced by different factors to a varying degree (Manandhar *et al*., 2018). According to Holley *et al*. (2014), temperature and humidity are the most important factors in forecasting precipitation. Hence, time series without exogenous variables is not sufficient for modeling and forecasting in these aspects. Exogenous variables play a very crucial role in capturing volatility in time series analysis (Yeasin *et al.*, 2021). The model of time series with explanatory variables has the potential to define the fundamental variations in data from time series and to measure the effect of environmental impacts. In recent times, the SARIMA-X model is extensively used for its greater adaptability. SARIMA-X models may characterize time series that show non-stationary behaviours both within and across periods satisfactorily and also capture the effect of exogenous variables (Raman *et al.*, 2018). In essence, the objective of this research is to model monthly precipitation data in order to have more insight into the effects of these variables on climate change and agricultural production. In order to achieve the goals, the SARIMA-X model has been used to analyze Bangalore, India, monthly Precipitation time series data with exogenous variables Temperature, Relative Humidity and Surface Pressure.

2. Methodology

SARIMA-X is a rational extension of the SARIMA model that allows integrating independent variables that put on some explanatory value to the process. If the SARIMA model is not adequate to provide appropriate efficiency, it is very normal to search for other driving phenomena whose effect over time is not adequately rooted in the past values of the dependent time series. The time series model building approach of SARIMA-X has two phases. In the first phase, we begin with a statistically and conceptually sound regression model. And in the second phase, the residuals from the regression are modeled with SARIMA to remove the seasonality and serial correlation that is present in the residual series. The final SARIMA-X model comprises the effect of exogenous variables along with non-seasonal AR and/or MA terms and seasonal AR and/or MA terms to maximize the explanatory power while eliminating the significant autocorrelation exhibited by the residuals. To ensure that the established SARIMA-X model is statistically accurate, there are statistical assumptions need to fullfilled (Andrews *et al*., 2013). These assumptions are: the series must be stationary, there should be no significant serial correlation in the residuals and explanatory variables must have significant, non-zero coefficients and logical signs, with high correlation to the dependent variable.

The external variables can be modeled by multiple linear regression equation and can be expressed as:

$$Y_t = \beta_1 x_{1,t} + \beta_2 x_{2,t} + ... + \beta_k x_{k,t} + \omega_t \quad (1)$$

where, $x_{1,t}$, $x_{2,t}$,…,$x_{k,t}$ are observations of k number of external variables corresponding to the dependent variable Y_t; β_1,…,β_k are regression coefficients of external variables; ω_t is a stochastic error, *i.e.*, the error series that is independent of the input series.

The residual series ω_t can be represented in the form of the SARIMA model as follows (Peter & Silvia, 2012):

$$\varphi(B)\phi(B^s)(1-B)^d(1-B^s)^D \omega_t = \theta(B)\Theta(B^s)\varepsilon_t \quad (2)$$

where,

s = length of periodicity (seasonality);

$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - ... - \varphi_p B^p$

= the non-seasonal autoregressive operator of order p;

$\varphi_1, \varphi_2, …, \varphi_p$ = the corresponding non-seasonal autoregressive parameters;

$\Phi(B_s)$ = 1-$\Phi_1 B^s$-$\Phi_2 B^{2s}$-…-$\Phi_p B^{ps}$

= the seasonal autoregressive operator of order p;

Φ_1, Φ_2,…,Φ_p = the equivalent seasonal autoregressive parameters;

$\theta(B)$= 1-$\theta_1 B$-$\theta_2 B_2$-…-$\theta_q B^q$

= the non-seasonal moving average operator of order q;

$\theta_1, \theta_2, …, \theta_q$ = the associated non-seasonal moving average parameters;

$\Theta(B^s)$= 1-$\Theta_1 B^s$ - $\Theta_2 B^{2s}$-…-$\Theta_p B^{qs}$

= the seasonal moving average operate of order q;

$\theta_1, \theta_2, …, \theta_q$ = the corresponding seasonal moving average parameters;

$(1-B)^d$ = the non-seasonal differencing operator of order d to produce non-seasonal stationarity of the d[th] differenced data (usually d = 0, 1, or 2);

$(1-B^s)^D$ = the seasonal differencing operator of order D to produce non-seasonal stationarity of the d[th] differenced data (usually d = 0, 1, or 2).

ε_t = error of the model

The general SARIMA-X model equation can be obtained by substituting ω_t in regression equation (Cools *et al*., 2009; Aburto & Weber, 2007). Formally, the SARIMA-X model can be represented by the following equation

$$Y_t = \beta_1 x_{1,t} + \beta_2 x_{2,t} + ... + \beta_k x_{k,t} + \\ + \frac{\theta(B)\Theta(B^s)\varepsilon_t}{\varphi(B)\phi(B^s)(1-B)^d(1-B^s)^D} \quad (3)$$

The basic steps of SARIMA-X methodology are as follows :

Step 1 : *Stationarity and Seasonality*

As the SARIMA modeling process has been discussed earlier, we tested the dependent time series for stationarity using the ADF test and examined seasonality by using the WO test. Based on the test results, we apply appropriate seasonal and non-seasonal differencing schemes for the dependent variable. For consistency and reliability, the differentiation scheme applied to the dependent variable may be applied to all explanatory variables to stationarize them as well. The associations between them are further robust over time since both the dependent and independent variables become stationary.

Step 2 : *Examine the correlation between variables*

After ensuring both seasonal and non-seasonal stationarity, exogenous variables are screened based on the correlation. If any exogenous variables do not show significant evidence of linear association with the study variable, then those variables are eliminated from the exogenous variable list as discussed above in Assumption 4.

Step 3 : *Built regression model and check the residuals*

In this step, we start building a regression model by the stepwise regression procedure in which includes significant variables in the model and eliminates variables from the model that is insignificant simultaneously. After fitting the regression model, the residuals from the model are collected and analyzed. We tested the presence of serial correlation. If autocorrelation is present in the residuals series that indicate that AR and/or MA terms may be inserted in the model.

Step 4 : *Identification of SARIMAX model parameters*

Seasonal AR and/or MA and non-seasonal AR and/or MA terms in the regression model are driven by the significance lags of the ACF and PACF. The process of identification of SARIMA parameters have already been discussed in the SARIMA model. Identification of SARIMAX model parameters are similar to the identification of SARIMA model parameters. AIC and SBC of the favorable models are examined and select the most suitable model from the list of the favorable model has the lowest AIC and SBC value.

Step 5 : *Parameters estimation*

Generally, the SARIMA-X model can be represented by the following equation

$$Y_t = \beta_1 x_{1,t} + \beta_2 x_{2,t} + \ldots + \beta_k x_{k,t} + \frac{\theta(B)\Theta(B^s)\varepsilon_t}{\varphi(B)\phi(B^s)(1-B)^d(1-B^s)^D} \qquad (4)$$

The parameters of the SARIMA-X models are commonly estimated using the maximum likelihood estimation technique.

On assuming $\varepsilon_t \sim NID(0, \sigma^2_e)$, then the MLE will be asymptotically equivalent to the minimization of $\sum \varepsilon_t^2$, which can be obtained by using the non-linear technique. The vector of parameters of order (k+p+q+P+Q+1) that need to be estimated are Ω and σ^2_e.

$$\Omega = (\beta_1, \beta_2, \ldots \beta_k, \varphi_1, \varphi_2, \ldots, \varphi_p, \phi_1, \phi_2, \ldots,' \phi_p, \theta_1, \theta_2, \ldots, \theta_q, \Theta_1, \Theta_2, \ldots, \Theta_q) \qquad (5)$$

The maximum likelihood function can be expressed as :

$$L(\Omega, \sigma_e^2) = \frac{\exp\left[-\frac{1}{2}(Y-XB)^T \Gamma_N^{-1}(Y-XB)\right]}{\sqrt{(2\pi)^N \det(\Gamma_N)}} \qquad (6)$$

where Γ_N is the auto-covariance matrix of ε. Estimate Ω and σ^2_e that maximizes L is equivalent to maximize the logarithm if *L*.

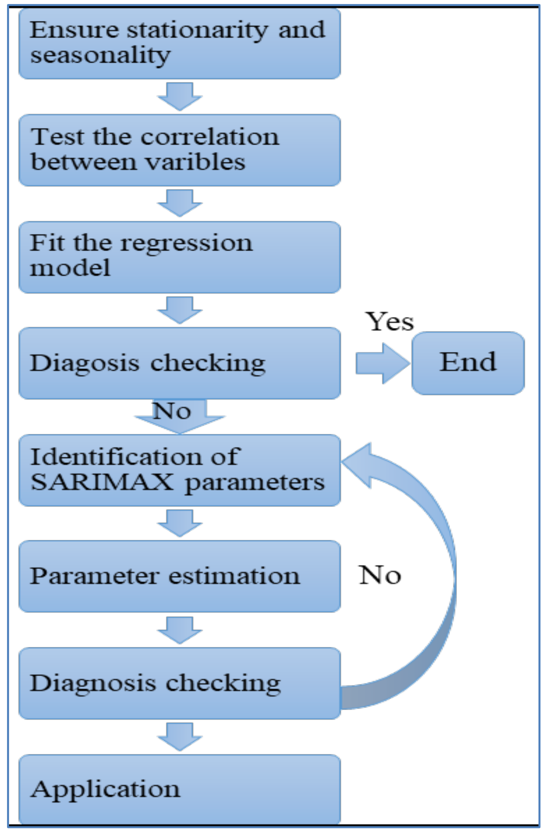

Fig. 1. Flowchart of SARIMA-X model

$$LL = \log L = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log\det(\Gamma_N) - \frac{1}{2}(Y-XB)^T \Gamma_N^{-1}(Y-XB)$$

$$= \frac{N}{2}\log(2\pi) - \frac{N}{2}\log\det(\Gamma_N) - \frac{1}{2}SSE \qquad (7)$$

The estimate of the parameters $\hat{\Omega}$ and σ^2_e are obtained by minimizing

$$SSE = \sum_{t=1}^{N} \varepsilon_t^2 \qquad (8)$$

Step 6 : *Diagnostic checking*

Different models can be obtained with various combinations of seasonal AR, non-seasonal AR, seasonal MA, and non-seasonal MA individually and collectively. The best model is obtained with the help of AIC and SBC values. A new model should be identified by repeating the above steps if the model is not optimal. Practically, this move is to verify whether the model assumptions about

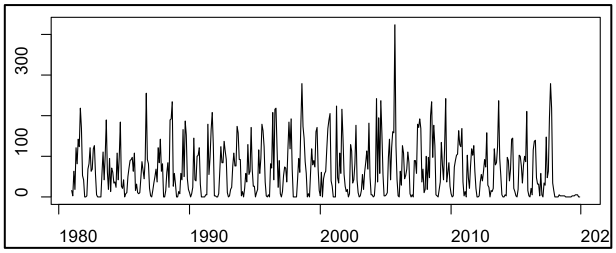

Fig. 2. Time plot of precipitation series

TABLE 1

Summary statistic

	Precipitation	Temperature	Relative Humidity	Surface Pressure
Mean	66.64	24.09	64.59	92.59
Median	50.01	23.76	69.13	92.56
Mode	0.00	23.67	74.69	92.40
Standard Deviation	65.87	2.70	15.40	0.22
Kurtosis	1.72	-0.44	-0.85	-1.14
Skewness	1.21	0.38	-0.56	0.23
Minimum	0.00	18.48	23.73	92.20
Maximum	425.00	31.15	88.02	93.10

the errors are fulfilled. This can be accomplished by carrying out the Ljung-Box test.

Step 7 : *Application*

Finally, the forecasting ability of the method is analyzed by utilizing different criteria. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are the most widely used accuracy metrics. Now the model is ready for various practical applications (Fig. 1).

3. Results and discussion

Data description

In this study, precipitation as the dependent variable and temperature, relative humidity, and surface pressure as independent variables have been considered. The monthly data on precipitation, temperature, relative humidity, and surface pressure for 39 years (from 1981 to 2019) period of Banglaore city has been collected from NASA Prediction of Worldwide Energy Resources (https://power.larc.nasa. gov/).

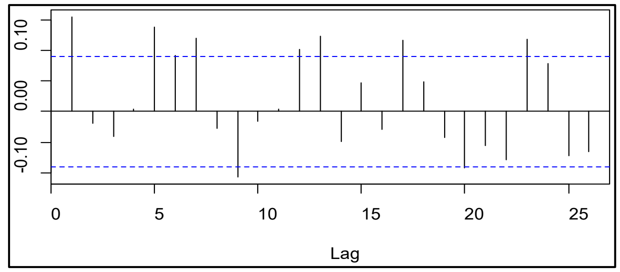

Fig. 3. ACF plot for residuals of the regression model

TABLE 2

Correlation of precipitation with other variables

Variables	Correlation	p-value
Temperature	-0.09	0.06
Relative Humidity	0.61	<0.01
Surface Pressure	-0.96	<0.01

TABLE 3

Parameters estimate of regression analysis

Parameters	Estimate	Std. Error	p-value	Significance
Intercept	2960.70	1611.63	0.07	No
Temperature	9.76	1.73	< 0.01	Yes
Relative Humidity	3.59	0.27	< 0.01	Yes
Surface Pressure	-36.30	16.89	0.03	Yes

To get a clear idea about datasets, a summary of datasets and the time plots of precipitation series are given in the following Table 1 and Fig. 2 respectively:

The summary Table 1 provides insights into the distribution and variability of precipitation, temperature, relative humidity and surface pressure. Precipitation has a high mean (66.64) and standard deviation (65.87), indicating significant variability and skewness (1.21). Temperature shows a mean of 24.09 with low variability and slight positive skewness (0.38), suggesting a relatively normal distribution. Relative humidity has a mean of 64.59, a higher median (69.13) and negative skewness (-0.56), indicating a left-leaning distribution. Surface pressure shows minimal variability (mean 92.59, standard deviation 0.22) and slight positive skewness (0.23), with values closely clustering around the mean.

Correlation analysis for exogenous variables

The compatibility of the exogenous variables is screened using the test significance of correlation. The following table gives details about correlation analysis.

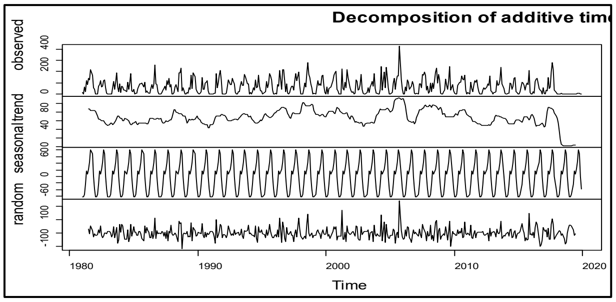

Fig. 4. Decomposition plot of precipitation series

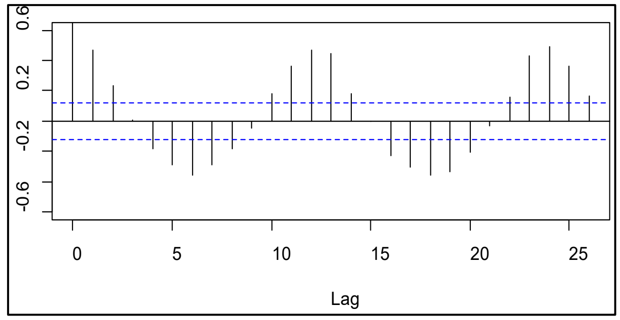

Fig. 5. ACF of precipitation series

TABLE 4

ADF test

Variables	Test statistic	p-value
Precipitation	-10.34	0.01
Temperature	-17.91	0.01
Relative Humidity	-23.60	0.01
Surface Pressure	-9.60	0.01

TABLE 5

Parameters estimate of SARIMA model

Parameters	Estimate	Std. Error	p-value	Significance
AR1	0.52	0.13	< 0.01	Yes
MA1	-0.33	0.15	0.02	Yes
SAR1	0.27	0.04	< 0.01	Yes
SAR2	0.32	0.05	< 0.01	Yes
Intercept	63.78	7.92	< 0.01	Yes

The next step is regression model building. The regression model is fitted with precipitation as the dependent variable and temperature, relative humidity and surface pressure as independent variables. The final results of the regression process are displayed in the following Tables 2 & 3.

For validation of this model, the residuals must be checked for auto-correlation. The residuals of the regression are checked with the help of the Box-Ljung test and ACF plot (Fig. 3) the Box-Ljung test is significant. That means the regression model shows a lack of fit and the residuals are serially correlated. ACF plots also show significant lags. So improvement is needed for these models. To remove serial correlation from the residuals AR and/or MA terms must be added.

Test for stationary and seasonality

After that, SARIMA model has been implanted. In time series analysis, the most important assumption is the stationary of the data sets. To test the stationary we used the ADF test. The null hypothesis of the ADF test is "The dataset is not stationary" (Table 4).

Before the SARIMA model building, we should check the seasonality of the precipitation series. For this, we apply Webel-Ollech (WO) test for seasonality and check the plot of the decomposed series (Fig. 4).

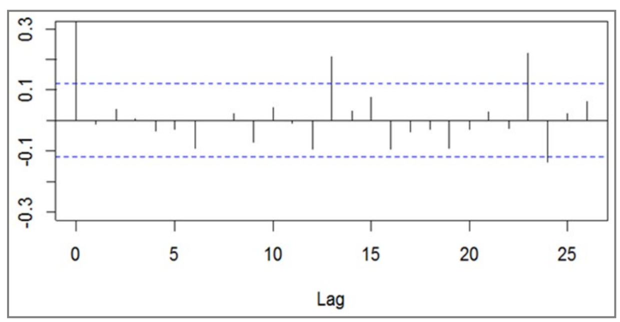

Fig. 6. ACF of residuals of the SARIMA model

The null hypothesis of the WO test is the series is non-seasonal. The WO test is significant, *i.e.*, the series has seasonality. Also, from the above plot, it is clear that the series exhibits seasonality.

Fitting SARIMA model

Now the next step is to fit the SARIMA model for the selected datasets. We selected the order of the SARIMA model in such a way that the AIC and SBC values are minimum. ACF and PACF helps to find the order of the SARIMA model. ACF and PACF plots are presented in Fig. 5. The best fitted model for study data is SARIMA (1,0,1) $(2,0,0)_{12}$ (Table 5).

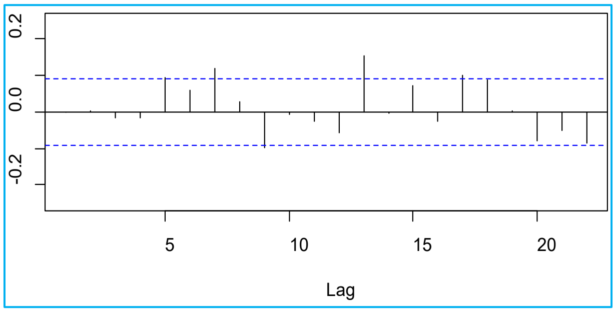

Fig. 7. ACF of residuals of SARIMAX model

TABLE 6

Parameters estimate of SARIMAX model

Parameters	Estimate	Std. Error	p-value	Significance
AR1	0.17	0.35	0.64	No
MA1	-0.05	0.35	0.90	No
SAR1	0.15	0.05	< 0.01	Yes
SAR2	0.16	0.05	< 0.01	Yes
Intercept	4997.36	1955.15	0.01	Yes
Temperature	7.28	2.18	< 0.01	Yes
RH	3.47	0.33	< 0.01	Yes
Surface Pressure	-57.58	20.55	<0.01	Yes

TABLE 7

Comparison among the models

Model	RMSE	MAE
Regression model	45.00	33.11
SARIMA model	52.57	39.60
SARIMAX model	43.50	31.81

The residuals of the SARIMA model are analyzed with the help of the Box-Ljung test and ACF plot. The Box-Ljung test is significant and the ACF plot is given in Fig. 6.

Fitting SARIMA-X model

Now we incorporate the exogenous variables into the SARIMA model and quantify its effect on the forecasting accuracy. Parameters estimate of the SARIMA-X model are given in Table 6.

For study data, the Box-Ljung test is not significant, *i.e.*, models do not evidence lack of fit and the residuals are not serially correlated (Fig. 7).

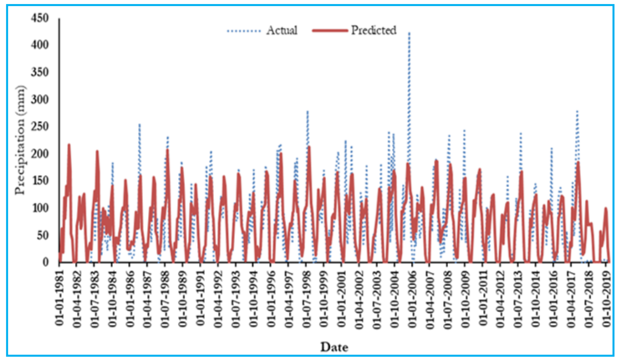

Fig. 8. The actual predicted plot

The following table gives the comparison forecasting accuracy of the above discussed models (Table 7).

From the above comparison, we find that the SARIMA-X model has the lowest RMSE and MAE values. We use the Diebold-Mariano test to determine whether the prediction of SARIMA and SARIMAX is significantly different or not, based on residuals generated by the models. The null hypothesis is that the two forecasts have the same accuracy. The alternative hypothesis is taken as the first forecast is less accurate than the second forecast. The test is significant. The result shows that the SARIMA-X model produces a better forecast than the SARIMA model. The actual predicted plot has been presented in Fig. 8.

4. Conclusion

This manuscript examined the seasonality behavior of precipitation data series and the potential impact of exogenous variables, *viz*., temperature, relative humidity, and surface pressure on the study variable. The inclusion of relevant external variables into the SARIMA model is examined to be capable to enhance the forecasting accuracy. We empirically compare the SARIMA model and types of SARIMA-X models. From this empirical study, we can infer that the SARIMA-X established its supremacy over the SARIMA model. The findings of this study has provided direct support for the potential use of accurate forecasts in decision-making for the farmers, Agri-industry as well as the government of India for policymaking. Further research can be done on the optimum criteria to select numbers and types of exogenous variables to include in the SARIMA-X model.

Author declarations

Conflicts of interest : Authors state no known conflict of interest.

Availability of data and material : The monthly data on precipitation, temperature, relative humidity, and surface pressure for 39 years (from 1981 to 2019) period of Banglaore cities has been collected from NASA Prediction of Worldwide Energy Resources (https://power.larc.nasa. gov/).

Code *availability* : Code will be made available as per requirements.

Author's *contribution* : All the authors contributed equally.

Ethics *declarations* : The authors did not receive financial support from any organization for the submitted work.

The authors declare they have no financial interests.

Consent to participate : Not applicable

Consent for publication : Not applicable

Disclaimer : The contents and views expressed in this study are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

References

Aburto, L. and Weber, R., 2007, "Improved supply chain management based on hybrid demand forecasts", *Applied Soft Computing*, **7**, 1, 136-144.

Andrews, B. H., Dean, M. D., Swain, R. and Cole, C., 2013, "Building ARIMA and ARIMAX models for predicting long-term disability benefit application rates in the public/private sectors", *Society of Actuaries*, 1-54.

Cools, M., Elke, M., and Geert, Wets., 2009, "Investigating the variability in daily traffic counts through use of ARIMAX and SARIMAX models: Assessing the effect of holidays on two site locations", *Transportation research record,* 2136, 1 : 57-66.

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.

Dubey, A. D., 2015, "Artificial neural network models for rainfall prediction in Pondicherry", *International Journal of Computer Applications*, **120**, 3, 30-35.

Gutierrez-Lopez, A., Cruz-Paz, I. and Muñoz Mandujano, M., 2019, "Algorithm to Predict the Rainfall Starting Point as a Function of Atmospheric Pressure, Humidity, and Dewpoint", *Climate*, **7**, 11, 131.

Harun, N., Pallu, M. S. and Achmad, A., 2013, "Statistic approach versus artificial intelligence for rainfall prediction based on data series", *International Journal of Engineering and Technology*, **5**, 2, 1962-1969

Hashim, R., Roy, C., Motamedi, S., Shamshirband, S., Petković, D., Gocic, M. and Lee, S. C., 2016, "Selection of meteorological parameters affecting rainfall estimation using neuro-fuzzy computing methodology", *Atmospheric Research*, **171**, 21-30.

Holley, D. M., Dorling, S. R., Steele, C. J. and Earl, N., 2014, "A climatology of convective available potential energy in Great Britain", *International Journal of Climatology*, **34**, 14, 3811-3824.

Intergovernmental Panel on Climate Change (IPCC), 2018, *Global Warming of 1.5°C.* https://www.ipcc.ch/sr15/.

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 approaches", *Modeling Earth Systems and Environment*, **17**, 1-9.

Manandhar, S., Dev, S., Lee, Y. H., Winkler, S. and Meng, Y. S., 2018, "Systematic study of weather variables for rainfall detection", *IEEE International Geoscience and Remote Sensing Symposium*, 3027-3030.

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 and Atmospheric Physics*, **130**, 1, 99-106.

Peter, Ď., & Silvia, P., 2012, "ARIMA vs. ARIMAX-which approach is better to analyze and forecast macroeconomic time series", In *Proceedings of 30th international conference mathematical methods in economics*, 2, 136-140.

Prasad, K., Dash, S. K. and Mohanty, U. C., 2010, "A logistic regression approach for monthly rainfall forecasts in meteorological subdivisions of India based on DEMETER retrospective forecasts", *International journal of climatology*, **30**, 10, 1577-1588.

Puvaneswaran, M., 1990, "Climatic classification for queensland using multivariate statistical techniques", *International Journal of Climatology*, **10**, 591-608.

Raman, R. K., Mohanty, S. K., Bhatta, K. S., Karna, S. K., Sahoo, A. K., Mohanty,. P. and Das, B. K., 2018, "Time series forecasting model for fisheries in Chilika lagoon (a Ramsar site, 1981) Odisha, India: a case study", *Wetlands Ecology and Management*, **26**, 4, 677-687.

Sadhuram, Y. and Murthy, T. R., 2008, "Simple multiple regression model for long range forecasting of Indian summer monsoon rainfall", *Meteorology and Atmospheric Physics*, **99**, 1-2, 17-24.

Tang, Y., Tang, Q. H., Wang, Z. G., Chiew, F. H. S., Zhang, X. J. and Xiao, H., 2019, "Different Precipitation Elasticity of Runoff for Precipitation Increase and Decrease at Watershed Scale", *Journal of Geophysical Research: Atmospheres*, **124**, 1-12.

Yeasin, M., Singh, K. N., Lama, A. and Paul, R. K., 2021, "Modelling Volatility Influenced by Exogenous Factors using an Improved GARCH-X Model", *Journal of the Indian Society of Agricultural Statistics*, **74**, 3, 209-216.