Vol. 50, No. 3, pp 190-198, 2022
Indian Journal of Soil Conservation

Development and evaluation of seasonal rainfall forecasting (SARIMA) model for Kumaon region of Uttarakhand

Utkarsh Kumar[1,2,*], Dhirendra Kumar Singh[2], Suresh Chandra Panday[1], Jaideep Kumar Bisht[1] and Lakshmi Kant[1]

[1]ICAR-Vivekananda Parvatiya Krishi Anusandhan Sansthan, Almora, Uttarakhand; [2]Division of Agricultural Engineering, ICAR-Indian Agricultural Research Institute, New Delhi, Delhi.

**Corresponding author:*
E-mail: utkarsh.kumar@icar.gov.in (Utkarsh Kumar)

ARTICLE INFO

Article history:
Received : March, 2022
Revised : November, 2022
Accepted : November, 2022

Key words:
AIC
Himalaya
Precipitation
Prediction
SARIMA

ABSTRACT

Rainfall is an important parameter for water resources application particularly in rainfed agricultural system. Rainfall in Himalaya varies from palace to place due to complex topography and it is hard to predict based on empirical formula. The objective of the present study is to develop seasonal auto regressive integrative moving average (SARIMA) model for Almora, Hawlbagh and Mukteshwar station located in Kumaon region of Uttarakhand and to determine the accuracy of the developed model in the same region. The present research utilized univariate time series rainfall model to forecast rainfall in Kumon region. We have selected the best SARIMA model fitted to our data which exhibited the least akaike information criterion (AIC) and bayesian information criterion (BIC) values. Finally, we identified the best model separately for each location after following three Box-Jenkings methodologies (model identification, elimination of parameters and diagnostic checking). The selected model is evaluated for residual normality test and observed data. The performance statistics of the developed SARIMA models for monthly scale were found for Almora (RMSE = 26.55 mm, MAE = 22.77 mm, R^2 = 0.86), Hawalbagh (RMSE = 33.66 mm, MAE = 25.63 mm, R^2 = 0.82) and Mukteshwar (RMSE = 35.62 mm, MAE = 28.45 mm, R^2 = 0.81). The performance parameter of the developed models showed that the forecast result mimics well with the observed rainfall data and also captures the extreme events well.

1. INTRODUCTION

For efficient water resource management, accurate forecasting of rainfall over an area or station is an important step (Kumar *et al.,* 2021; Kumar *et al.,* 2021a; Kumar *et al.,* 2022). For Indian agriculture, southwest monsoon (June-September) plays a crucial role for an agrarian economy. A good amount of rainfall is very important for good crop production (Kumar *et al.,* 2021c). Among various hydrological parameters, rainfall plays a pivotal role in irrigation planning, runoff modelling, drought and flood management. Time series models have been routinely implemented in various scientific domains, including meteorology and hydrology. Various stochastic models have been evolved with time (Box and Jenkins, 1994). These covers auto-regressive (AR) models of discrete order, moving average (MA) models for various orders (Gupta and Kumar, 1994 and Verma, 2004), auto regressive moving average (ARMA) models of discrete orders (Katz and Skaggs, 1981; Chhajed, 2004; Katimon and Demon, 2004) for annual stream flow. However, for monthly or intra seasonal stream flows, autoregressive integrated moving average (ARIMA) model (Bender and Simonovic, 1994; Montanari *et al.,* 2000; Trawinski and Mackay, 2008), Thomas- Fiering models (Srinivasan, 1995) were used. Some of the crucial benefits of time series model, their standardized search process, capability for distinguishing, evaluation and diagnostic checking, which makes the process transparent unlike data driven black box technique. Climate changes are mainly derived by the changes in temperature and rainfall. Researches focusing on the estimation of climatic changes are frequently formulated on the analysis of time series of rainfall and temperature (Babazadeh and Shansnia, 2014; Balibey and Serpil, 2015). Both these factors play a pivotal role in blue print for the management of natural calamity such as flash flood and dry spells. Examination of time series universally presumes the computation of trend and seasonality in the data. Long-term changes in the time series is described as trend, whereas seasonality refers to the dissimilarity in the data at defined shorts periods such as weekly, monthly, biyearly quarterly, etc. (Pazvakawambwa and Ogunmokun, 2013; Wang *et al.,* 2013). Prognoses of climate variable are frequently needed for selection of efficient water management practices at command scale.

For prediction of weather at local and state scale, several rainfall forecasting models are present. Regression analysis, auto regression integrated moving average (ARIMA), genetic algorithm, adaptive splines threshold autoregressive (ASTAR), support vector machines (SVMs), K-nearest neighbour (K-NN) are among the popular models available for weather prediction. Regression analysis establishes the strength of association between a dependent variable and a series of independent predictor variables by fitting a regression model. The regression analysis is known as multiple regression analysis, if it involves more than two independent variables. However, this regression analysis is not suggested for most of the practical problems as it tends to simplify the complexities of the real-world problem. ARIMA models predict meteorological variable which are kind of time series data by linearly integrating their historical values. ARIMA model as an instrument deals with different dimensions related to univariate time series model selection and its parameter optimization and prediction. ARIMA model has few limitations such as, over-fitting and misidentification if not used wisely. Genetic algorithm methodology utilizes the concept of natural evolution in problem solving. It involves a random large selection of solution within a specified population. After that it tries to get best-fit solution by repetitively simulating the process of evolution. This whole process finally gives the appropriate number of generations. The succeeding generation is an improvement upon the preceding one and in the end, best solution to the problem is obtained. Genetic algorithm is proficient in finding the best quality solution in a short calculation period, but it cannot assure for an optimal solution to a particular problem. Another data investigation methodology is ASTAR modelling. This method is focused on past and present data to forecast short term prediction for particular time interval. It selects the optimal-fit model from a given sets of data. Support vector machine (SVM) which has established by Vapnik and his colleague uses computer algorithms from supervised learning. It finds the best-fit function by learning through trainings. Limited researchers have got satisfactory results for rainfall forecasting using this method. However, its performance varies from problem to problem. A representative KNN involves sampling with replacement at the same time, the meteorological parameter mostly rainfall and temperature (minimum and maximum) from the historical observed data. Various researchers have applied K nearest neighbour model for predicting climatic variable conveniently. The major advantage of using KNN model is that it neither requires sample nor the training data but simply uses the sample data itself for analysis. Another method is shared nearest neighbour (SNN) cluster algorithm which has been applied to monsoon (June-September) rainfall departures in India to predict land temperature and precipitation. (Rajagopalan and Lall, 1999; Sumi *et al.,* 2008; Salahi *et al.,* 2016; Dhawal and Mishra, 2016; Kakade and Kulkarni, 2016 & 2017; Mahmud *et al.,* 2017; Sharif and Azhar, 2017; Akinbobola *et al.,* 2018).

Cyclic, variability and serial correlation are prominent properties of various time series, which can only be adequately captured using a stochastic model like SARIMA. Time series model like ARIMA, which is established on assumption of serial linear correlation among observations, whereas SARIMA models can effectively explain time series that express non-stationary pattern both within and across season. The major benefits of SARIMA models are that it needs few parameters for explaining time series. In this study, an effort was made to develop seasonal rainfall forecasting model and predict monthly rainfall time series of Hawalbagh, Almora and Mukteshwar Uttarakhand, India using monthly rainfall data of year 1980-2019 (40 yrs).

2. MATERIAL AND METHODS

Study Area

Uttarakhand is the twenty seventh state in India lies between 28°44' and 31°28'N latitudes and 77°35' and 81°01'E longitudes covering an area of 53483 km^2. The terrain and topography of the state is mostly hilly with large area under snow cover and steep slopes. Two districts of Uttarakhand state, India under Kumaon region, namely, Almora and Nanintal were selected (Fig. 1) for the present study. The total geographical area of Almora district is 3,083 sq km and that of Nainital district is 3860 km^2. As of 2011 census the population of Almora district is 621,972 with population density 202 persons per $km^2$and that of Nainital district is 955,128 with population density 247 persons km^{-2}. The climate of Almora is humid sub-tropical and that of Nainital is subtropical highland according to Köppen climate classification. The normal annual rainfall of the state is 1500 mm. The important crops harvested in both the district are finger millet, rice, black soyabean, barnyard millet, horse gram (*kharif*) while, wheat, barley and lentil (*rabi*). Finger millet, rice, wheat and lentil are the major crops harvested in the study area (Kumar *et al.,* 2021b).

Model Description

Auto regressive (p) integrated (d) moving average (q) ARIMA model is mainly used for better understanding time series data. The trend of time series is determined by differencing so that stationary is produced is the basic principle behind ARIMA model (Bahadir 2012, Wang *et al.,*

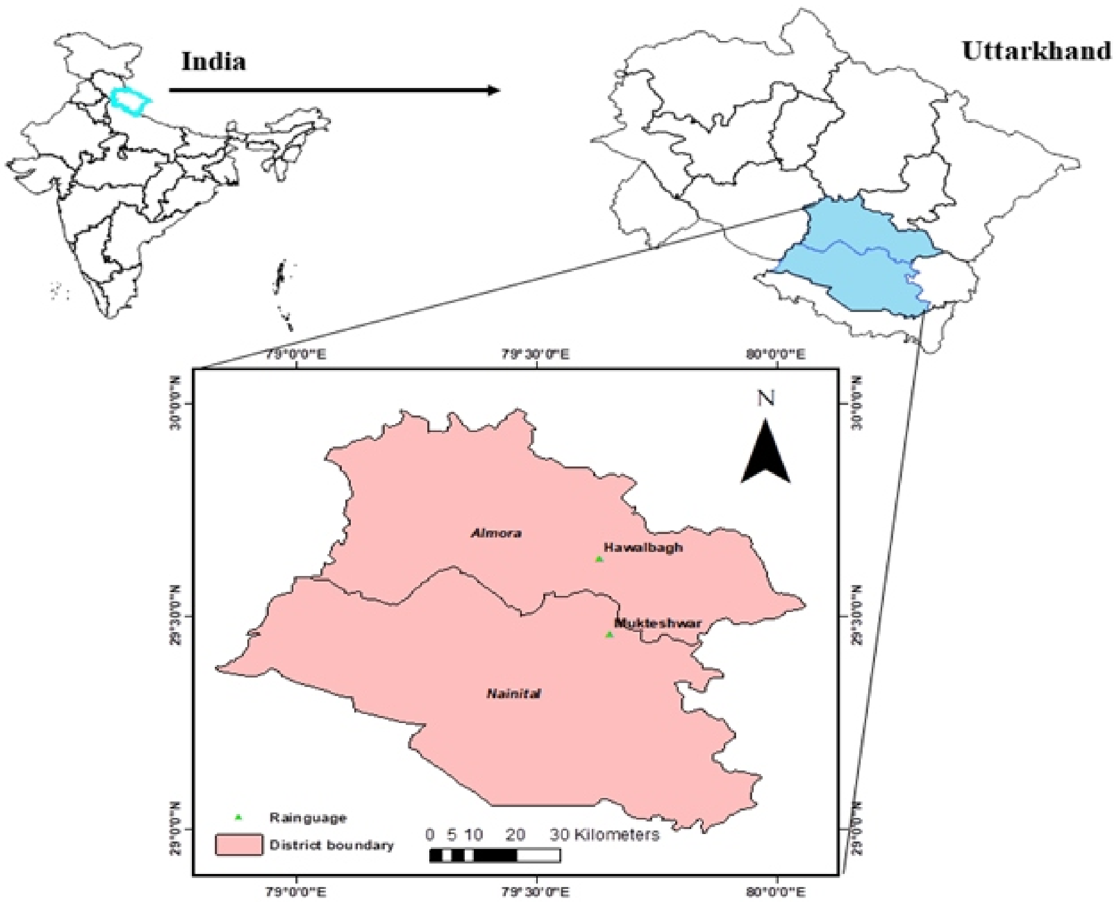

Fig. 1. Location of the study area and rain guage

2014; Afrifa-Yamoah, 2015). This ARIMA model is mainly based on Box-Jenkins approach. The AR part of the ARIMA model defines the variability under consideration which is regressed on its own preceding values. The other part MA defines that the regression error is linear combination which occur at several time period in the past. The number of times subtraction has been done is denoted as I. The whole purpose of selecting an appropriate AR, I and MR terms is to deseasonalize the model before it fits to the data. The general symbol of a non-seasonal ARIMA model is ARIMA (p, d, q), where p is the lag order, d is the order of differencing, and q is the order of moving average. A seasonal ARIMA model is donated as ARIMA (p,d,q), $(P,D,Q,)_M$, where m is the number of periods in each season, and the uppercase P,D,Q refer to the autoregressive (AR), differencing (I), and moving average (MR) terms, respectively, M for seasonal part of the ARIMA model. The limitation of ARIMA model is that it depends upon on past data; however, it performs better for historical stationary time series. It simply explains the past pattern and does not describe the structure of underlying data (Balibey and Serpil, 2015).

The general notation of an AR (p) model is given as:

$$Y_t = C + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + ... + \varphi_p Y_{t-p} + Z_t \quad ...(1)$$

Where, $Z_t \sim (0, \sigma^2)$, c is an arbitrary fixed term, which is equal to 1, are the parameters of the AR model.

Similarly, MA (q) model can be denoted as:

$$Y_t = C + Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q} \quad ...(2)$$

Where, $Z_t \sim (0, \sigma^2)$, c is an arbitrary fixed term, and θ_1, θ_q are the parameters of MA model. A stationary time series Y_t is called autoregressive moving average of order (p), ARMA (P,), if for every *t*,

$$Y_t - \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + ... + \varphi_p Y_{t-p} = c + Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q}, Z_t \sim (0, \sigma^2) \quad ...(3)$$

The produced polynomials $(z) = 1 - \varphi_1 z - ... - \varphi_p z^p$ *and* $(z) = 1 + \theta_1 z + ... + \theta_q Z^q$ have no common roots. If we combine the differencing with ARMA model we will get the autoregressive intergrated moving average model, *i.e.* the ARIMA (p, d, q), where d is the order off differencing. So, an ARIMA model corresponds to an ARMA after differencing Y_{td} times. This means that Y_t satisfies the difference equation:

$$(1 - \varphi_1 B - \ldots \varphi_p B^p)(1 - \mathrm{B})^d Y_t = c + (1 + \theta_1 B + \ldots \theta_q B^q) Z_t \qquad \ldots(4)$$

$$\varphi(B)(1\text{-}B)^d Y_t = c + \theta(B) Z_t, Z_t \sim (0, \sigma^2) \qquad \ldots(5)$$

For, d and D are non-negative integers, the time series is a seasonal ARIMA, SARIMA (p,d,q), $(P,D,Q)_m$ process with period m if the differenced time series X_t: $= (1\text{-}B)^d(1\text{-}B^M)DY_t$ is a casual ARMA process $\{X_t\}$ is casual, if there exits constant $\{\psi_j\}$ such that:

$$\sum_{j=0}^{\infty} |\psi_j| < \infty, \text{ and } X_t = \sum_{j=0}^{\infty} |\psi_j| Z_{t-j}, \qquad \ldots(6)$$

$$\varphi(B)\phi(B^m)X_t = 0\ (B)\Theta(B^m)Z_t,\ Z_t \sim WN(0, \sigma^2),\ Z_t \sim (0, \sigma^2) \qquad \ldots(7)$$

Where, $\varphi(Z) = 1 - \varphi_1 z - \ldots - \varphi_p z^p)$, $\varphi(z) = 1 - \varphi_1 z - \ldots - \varphi_p z^p$, $\theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q$, $\Theta(z) = 1 + \theta_1 z - \ldots - \theta_q z^q$. $\{X_t\}$ is casual which is equivalent to the condition that $(z) \neq 0$, $\phi(z) \neq 0\ for\ |z| \leq 1$ for complex (z). Methodology Adopted in the present study is shown in flow chart in Fig. 2.

SARIMA Model and Selection

The property of seasonal variability in time series can be bifurcated into simple model and multiple models by utilizing SARIMA model. Normally the observed time series (Y_t) integrates a lag operator B to process SARIMA (p,d,q) (P,D,Q)s. A typical seasonal ARIMA model may be described as:

$$\varphi_p(B)\phi_p(B^s)^d(1 - B^s)^D Y_t = \theta_q(B)\Theta_q(B^s)\varepsilon\iota \qquad \ldots(8)$$

In eq. 8 B, is lag operator (define as $B^k Y_t = Y_{t-k}$), $\varphi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$, $\varphi_p(B^s) = 1 - \phi_s B^s - \phi_{2s} B^{2s} - \ldots - \phi_{ps} B^{ps}$, $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$, $\Theta_Q(B^s) = 1 - \Theta_{2s} B^{2s} - \ldots - \Theta_{Qs} B^{Qs}$.

Where, ϕ(B) and θ(B) are polynomials of order p and q. respectively; $\phi(B^S)$ and $\Theta(B^S)$ are polynomials in B of degree p and Q, respectively; p is not order of non-seasonal auto regression; d is the number of fixed differences; q is the order of non-seasonal moving average; p is the order of seasonal auto regression; D is the number of seasonal differences; Q is the order of seasonal moving average; and S is length of season. Box and Jenkins proposed building a SARIMA model which needed at least 50 observation samples (the best would be 100).

Selecting Best SARIMA Model

When, all the model parameter are selected (p, d and q P, D and Q) most appropriate model is found. Well-performed residual model is selected after checking of residual of evaluated SARIMA model. Prediction is done for 2016-2019 (4 yrs). The predicted results from the selected model are evaluated with observed data. The accuracy assessment of best performing SARIMA model for rainfall is done for

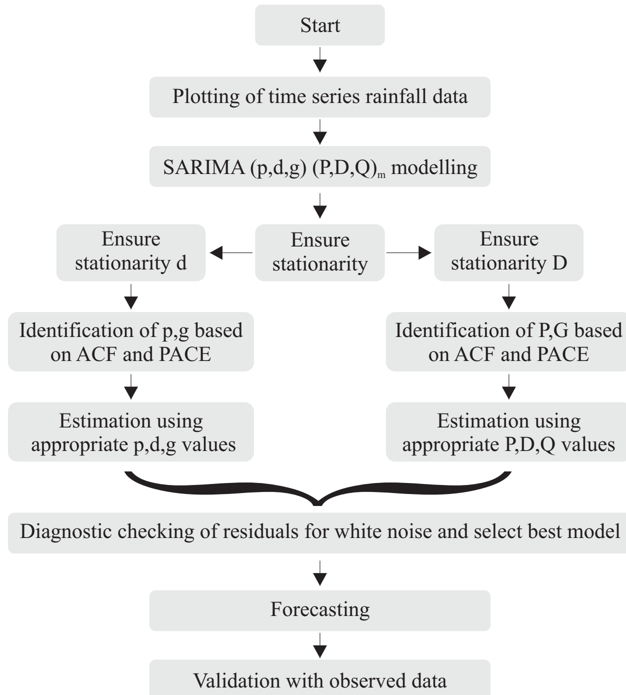

Fig. 2. Flowchart of SARIMA model

each station. The AIC and BIC are followed to check the accuracy for a given dataset. BIC is another important criterion for model identification. The model having the lowest BIC is selected. It is a based on the likelihood function and is closely related to the AIC.

Performance Parameters

The predictive capability of selected SARIMA model is evaluated by comparing the observed and forecasted data. The root mean square error (RMSE), mean absolute error (MAE) and coefficient of determination are employed to determine the efficacy of selected forecasting model, as described in eqs 9-11.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i^o - y_i^f\right)^2} \qquad \ldots(9)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i^o - y_i^f\right| \qquad \ldots(10)$$

$$R^2 = \left[\frac{\sum_{i=1}^{N}\left(y_i^f - \overline{y^f}\right)\left(y_i^o - \overline{y^o}\right)}{\sum_{i=1}^{N}\sqrt{\left(y_i^f - \overline{y^f}\right)^2\left(y_i^o - y^o\right)^2}}\right]^2 \qquad \ldots(11)$$

Where, N denotes the number of datasets; y_i^o represents the observed rainfall; y_i^f represents forecasted rainfall y_{max}^o represents the maximum of observed data; y_{min}^o represents the minimum of observed data; $\overline{y^o}$ represents the average of observed data and $\overline{y^f}$ represents the average of forecasting rainfall. Apart from the above statistical measure two more statistical measures two more statistical measures are BIC and Lungs Box Q statistics can be calculated as follows:

$$BIC = \log\left(\frac{rss}{n}\right) + \frac{k}{n}\log n$$

Where, rss = residual sum of squares, k = no. of coefficient estimated and n = no. of observations.

The expression for Lungs Box Q statics is as follows:

$$Q = n(n+2)\sum_{k=1}^{h}\frac{\rho^2}{n-k}$$

Where, n = number of residuals, h = number of time lags includes in the test, p^2, k = the residual auto correlation at lag k.

3. RESULTS AND DISCUSSION

Rainfall variation in the Study Area

The variation of historical precipitation for Almora, Hawalbagh, Mukteshwar stations is presented in Fig's 3 and 4, respectively. The rainfall data shows that the major portion of the rainfall occurs during monsoon season for all the three stations and nearly no or scanty rainfall in months of November and December. The historical monthly rainfall data for all the three station 40 yrs, starting on January, 1980 to December, 2019 has been employed for Almora, Hawalbagh and Mukteshwar in the present study. We have presented year in X-axis and observed monthly rainfall in Y-axis as shown in Fig. 3.

Decomposition of Time Series

From time series plot it is found that the precipitation data shows a seasonal variation. Rainfall reaches its maximum in monsoons and minimum in winter at all the stations. Therefore, they required to be stationarized and seasonality needs to be removed. The rainfall data were separately analysed using python package in Jupyter notebook. On statistically analysing the rainfall data, it is found that their statistical properties are not constant with time. To make the data stationary, differencing was done separately for each station. Once the time series has been stationarized by differencing, the next step is to find AR or

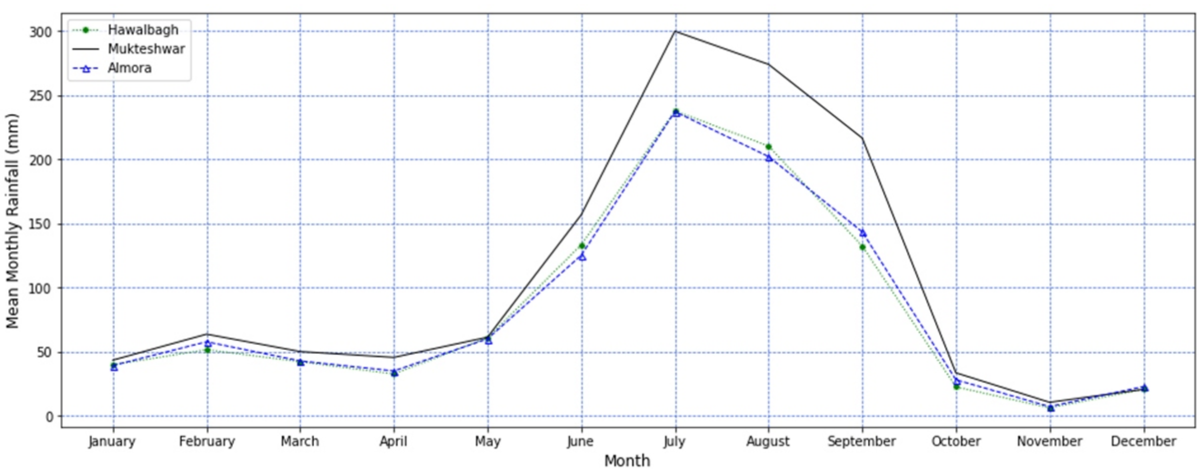

Fig. 3. Monthly average rainfall variation at Almora, Hawalbagh and Mukteshwar station

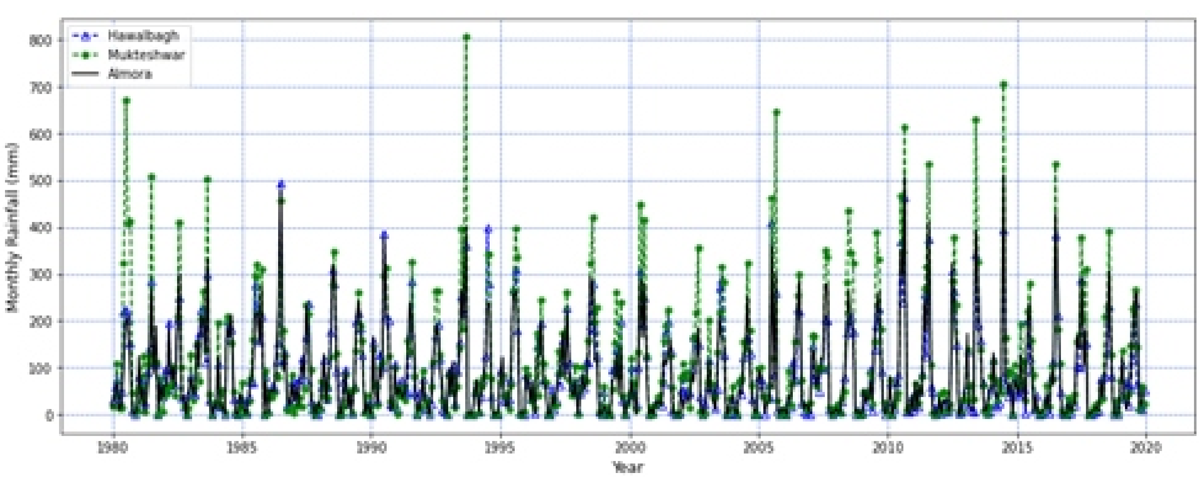

Fig. 4. Monthly rainfall variation at Almora, Hawalbagh and Mukteshwar station

MA terms in order to fit the ARIMA model. Next is the removal of seasonality from the time series data.

The seasonal part of ARIMA model is dealt with in a similar manner as non-seasonal part. The P,D,Q values are found on the same lines as p, d and q. Therefore, the seasonality of the data was removed by finding P, D and Q terms for 12-month time gap and constant coefficients and was found separately for each station. The most appropriate seasonal ARIMA (SARIMA) model for monthly rainfall was developed for each station. Representation of observed, seasonal, trend and residual for each station is shown in Fig. 5.

Diagnostic and Model Selection

To be sure about the identification of the correct model, diagnostic checking with the standardized residuals, ACF of residuals, normal Q-Q plot of the standardized residuals and p-values of the Ljung-Box statistic is required. The diagnostics for each station are displayed in Fig. 6.

Evaluation of Developed Model

The forecast trends for the selected SARIMA model is shown in Fig. 7 while model fit, and model performance statistics are shown in Table 1 for Almora, Hawalbagh and Mukteshwar stations.

The rainfall prediction from the models for all the three stations captures the local and seasonal trend and is slightly smoother in appearance because both the trend and seasonal pattern are effectively arranged over the seasons. From the 20-yrs forecast, last 3 years, *i.e.,* 2016-2019 data were validated by comparing it with the observed data at the respective stations as shown in Fig. 7. The comparison shows that the precipitation forecast fits quite well with the observed with some over-predictions of heavy rainfall events especially during the monsoons. The RMSE value for Almora, Hawalbagh and Mukteshwar are, respectively Almora = 26.55 mm, Hawalbagh = 33.66 mm and Mukteshwar = 28.45 mm). The monthly rainfall forecast fits well with the observed data, following the same pattern as observed data reaching maximum during monsoons and minimum in winters. The disadvantage of the SARIMA model is that it can only extract linear relationships within the time series data. As an extension of the ARIMA method, the SARIMA model not only captures regular difference, autoregressive, and moving average components as the ARIMA model does but also handles seasonal behaviour of the time series.

4. CONCLUSIONS

This study focused on modelling and forecasting a monthly rainfall series using SARIMA model. The results obtained can be applied for various hydrological and water resource management studies. This will certainly assist policy and decision-makers to establish strategies, priorities, and the proper use of water resources in hilly region of Uttarakhand. Rainfall is the main source of irrigation for rainfed agriculture in mountain. In the present study, monthly rainfall time series were studied and best fitted seasonal ARIMA (SARIMA) is found after the removal of seasonality. The forecasting was done using the same model. The forecast results for rainfall are found to be well matches with the observed rainfall data, while it also captured the extreme events. The monthly precipitation series recorded in the stations of the study area show a seasonal behaviour, reflecting the typical unimodal rain regime. Similarly, it was determined that a trend was not maintained throughout the recording period. The SARIMA approach showed an improvement over the integration of

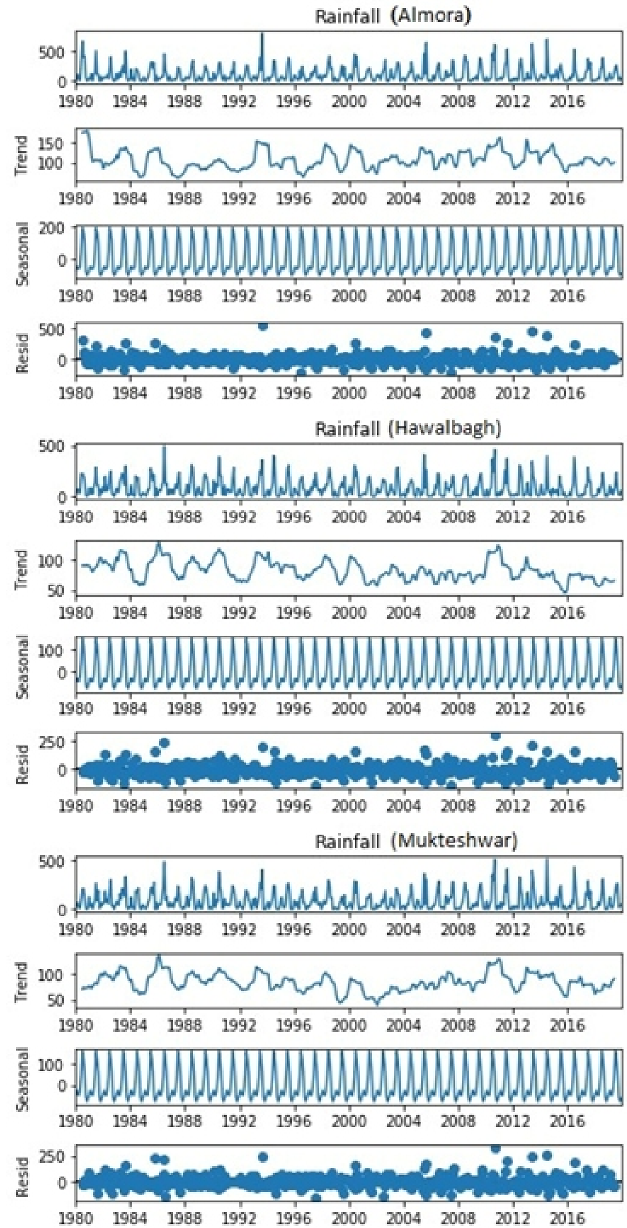

Fig. 5. Decomposition of time series by additive method during 1980-2019 at Almora Hawalbagh and Mukteshwar

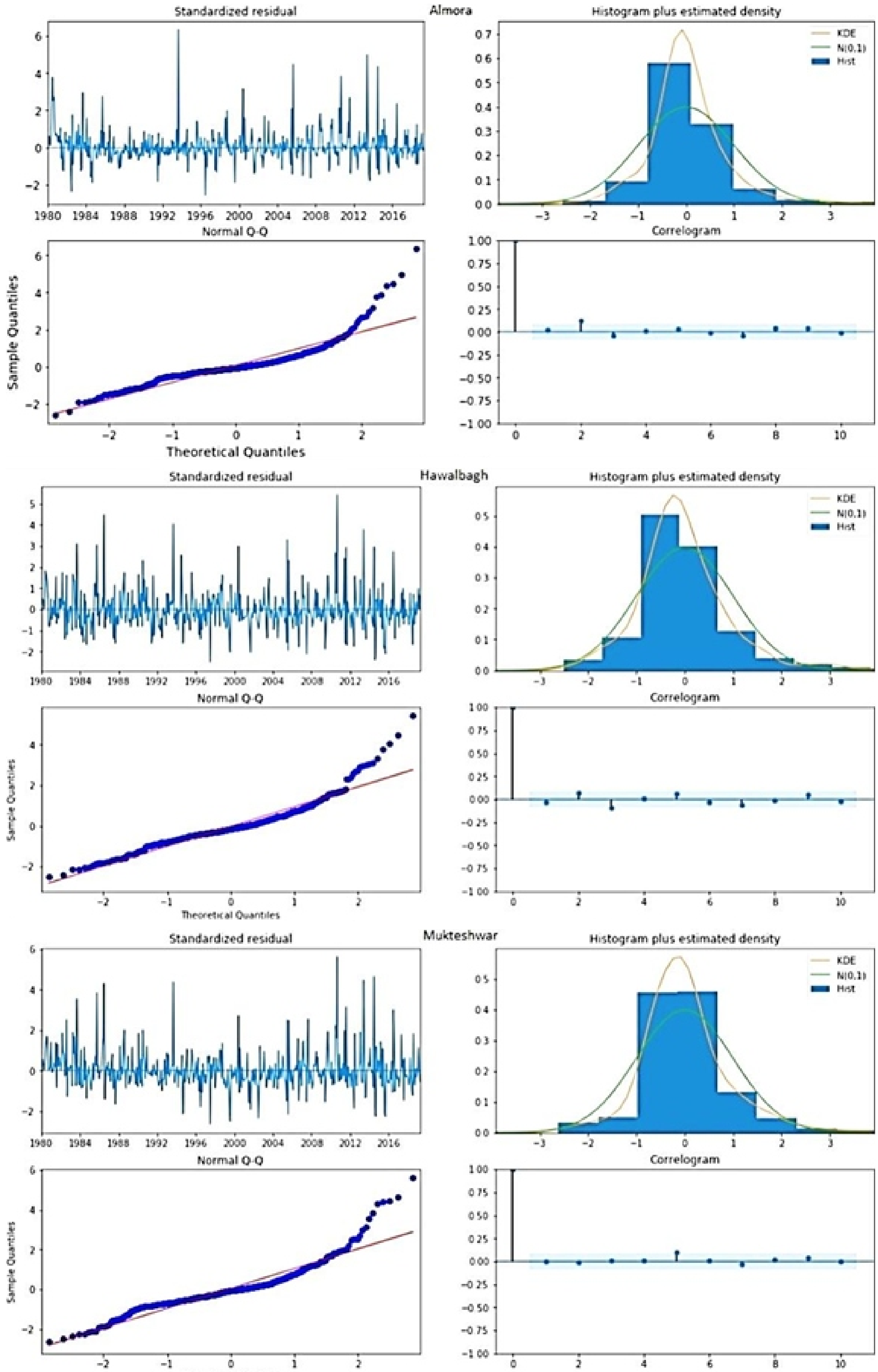

Fig. 6. Diagnostic checking for the selected SARIMA model using ACF of residuals and Normal Q-Q plot of the standardized residuals for Almora, Hawalbagh and Mukteshwar

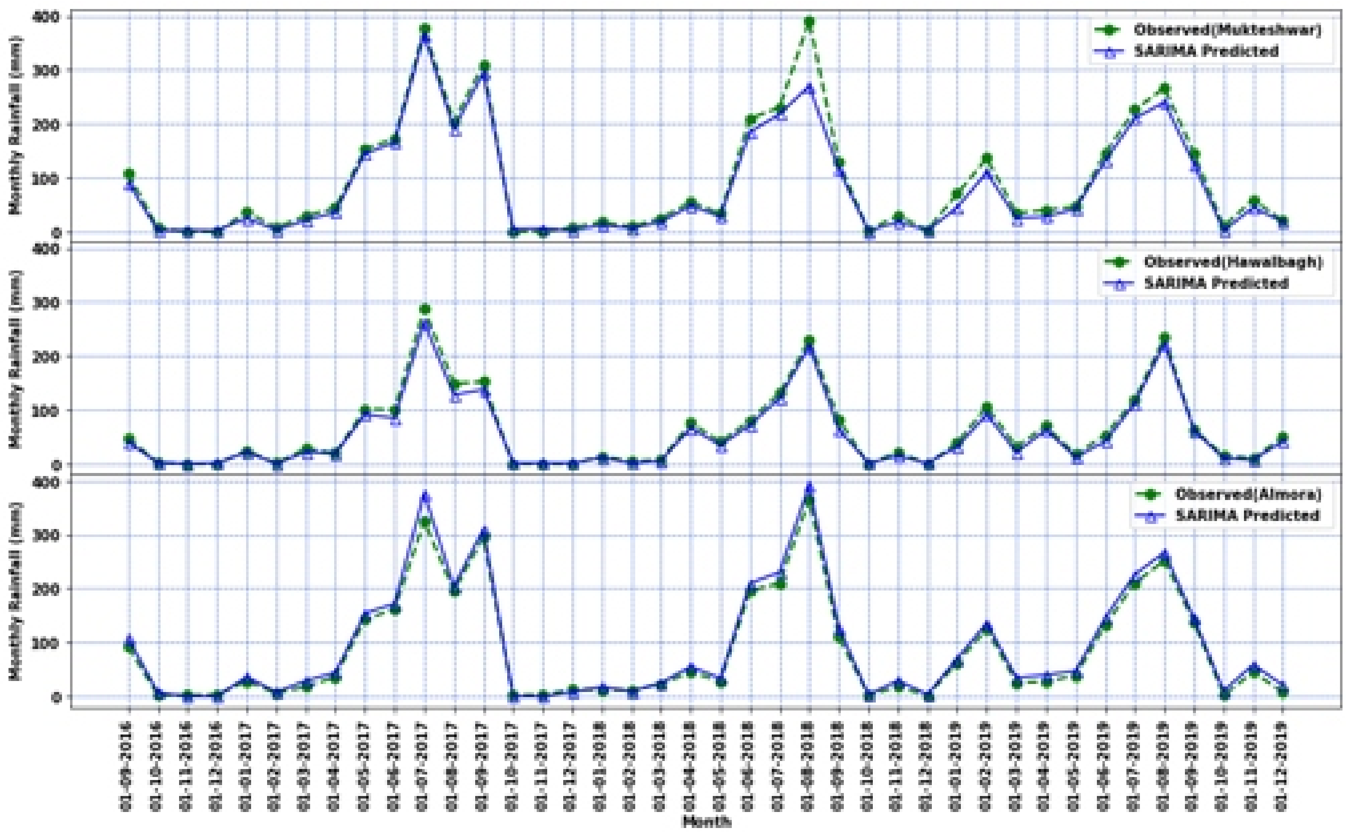

Fig. 7. Observed and SARIMA predicted rainfall values at Mukteshwar station (September 2016 - December 2019) for Almora, Hawalbagh and Mukteshwar

Table: 1
SARIMA model performance for rainfall at Almora, Hawalbagh and Mukteshwar stations

Station	SARIMA	RMSE	MAE	AIC	BIC	R^2
Almora	$(0,0,3) \times (1,0,1,)_{12}$	26.55	22.77	5373	5743	0.86
Hawalbagh	$(1,0,1) \times (1,0,1,)_{12}$	33.66	25.63	5309	5334	0.82
Mukteshwar	$(1,0,1) \times (1,0,1,)_{12}$	35.62	28.45	5718	5402	0.81

the removal of the trend, periodicity, and stochastic components approach in time series modelling. The different performance statistics in the model development and validation phase of this study also confirmed higher prediction accuracy using the SARIMA models. It is concluded from the study that SARIMA models found in the present study will assist scientists and decision makers to develop policies for efficient scheduling of irrigation, urban planning and environmental planning. The forecast can assist as a prediction tool for development of better management practices. It is important to highlight that although the forecasts obtained with the models do not allow predicting the exact precipitation amount, they can reveal the probable trend of future rains and provide information that can help decision-makers to establish strategies in areas such as agriculture, where it is of utmost importance to know the beginning and end of the rainy seasons the planning of civil works; and the preparation of mitigation plans for natural dangers, such as floods and droughts. Finally, it is worth noting that rational planning and comprehensive management of water resources require forecasting events that may occur in the future, bearing in mind at the same time that it is usually based on past events. For this reason, time series analysis is a valuable tool since it allows making inferences about the future.

ACKNOWLEDGEMENTS

Authors are thankful to all the field and laboratory staff for their help in this study. We are very much thankful to the ICAR-VPKAS, Almora as well as Division of Agricultural Engineering, ICAR-IARI, New Delhi for providing financial support during the course of investigation. We acknowledge the efforts of Mr Sanjay Kumar Arya in compiling the weather data of Headquarter Office, Almora.

REFERENCES

Afrifa-Yamoah, E. 2015. Application of ARIMA models in forecasting monthly average surface temperature of Brong Ahafo Region of Ghana, *Int. J. Stat. Appl.,* 5(5): 237-246.

Akinbobola, A. Okogbue, E.C. and Ayansola, A.K. 2018. Statistical modelling of monthly rainfall in selected stations in forest and savannah eco-climatic regions of Nigeria, *J. Climatol. Weather Forecasting*, 6: 226.

Babazadeh, H. and Shamsnia, S.A. 2014. Modeling climate variables using time series analysis in arid and semi-arid regions, *Afr. J. Agr. Res.*, 9(26): 2018-2027.

Bahadir, M. 2012. The analyses of precipitation and temperature in Afyonkarahisar (Turkey) in respect of Box-Jenkins technique, *J. Acad. Soc. Sci. Studies*, 5(8): 195-212.

Balibey, M. and Serpil, T. 2015. A time series approach for precipitation in Turkey, *GU. J. Sci.,* 28(4): 549-559.

Bender, M. and Simonovle, S. 1994. Time-series modeling for longrange stream flow forecasting, *J. Water Resour. Plan. Manage. ASCE,* 120(6): 857-870.

Box, G.E.P. and Jenkins, G.M. 1994. Time series analysis, forecasting and control, Revised Edition, Holden-Day, San Francisco, California, United States.

Chhajed, N. 2004. Stochastic modeling for forecasting Mahi river inflows, M.E. Thesis submitted to MPUAT, Udaipur.

Dhawal, H. and Mishra, N. 2016, A Survey on rainfall prediction techniques, *Int. J. Comput. Appl.,* 6(2): 1797-2250.

Gupta, R.K. and Kumar, R. 1994. Stochastic analysis of weekly evaporation values, *Indian J. Agric. Eng.,* 4(3-4):140-142.

Kakade, S.B. and Kulkarni, A. 2016. Prediction of summer monsoon rainfall over India and its homogeneous regions, *Meteorol. Appl.,* 23: 1-13.

Kakade, S.B. and Kulkarni, A. 2017. Seasonal prediction of summer monsoon rainfall over cluster regions of India, *J. Earth Syst. Sci.*, 126: 34.

Katimon, A. and Demun, A.S. 2004. Water use trend at university technology Malaysia: Application of ARIMA model, *J. Tech.*, 41(B): 47-56.

Katz, R.W. and Skaggs, R.H. 1981. On the use of autoregressive moving average processes to model meteorological time series, *Monthly Weather Rev.*, 109: 479-484.

Kumar, U., Srivastava, A., Kumari, N., Rashmi, Sahoo, B., Chatterjee, C. and Raghuwanshi, N.S. 2021b. Evaluation of spatio-temporal evapotranspiration using satellite-based approach and lysimeter in the agriculture dominated catchment. *J. Indian Soc. Remote Sens.,* 49: 1939-1950.

Kumar, U., Meena, V.S., Singh, S., Bisht, J.K. and Pattanayak, A. 2021a. Evaluation of digital elevation model in hilly region of Uttarakhand: A case study of experimental farm Hawalbagh. *Indian J. Soil Conserv.,* 49: 77-81.

Kumar, U., Panday, S.C., Kumar, J., Parihar, M., Meena, V.S., Bisht, J.K. and Kant, L. 2022. Use of a decision support system to establish the best model for estimating reference evapotranspiration in sub-temperate climate: Almora, Uttarakhand. *Agric. Eng. Int. CIGR J.,* 24(1): 41-50.

Kumar, U., Panday, S.C., Kumar, J., Meena, V.S., Parihar, M., Singh, S., Bisht, J.K. and Kant, L. 2021. Comparison of recent rainfall trend in complex hilly terrain of sub-temperate region of Uttarakhand, *Mausam,* 72: 349-358.

Kumar, U., Rashmi, Chatterjee, C. and Raghuwanshi, N.S. 2021c. Comparative evaluation of simplified surface energy balance index-based actual ET against lysimeter data in a tropical river basin. *Sustainability*, 13: 13786.

Mahmud, I., Bari, S.H. and Rahman, M.T.U. 2017. Monthly rainfall forecast of Bangladesh using autoregressive integrated moving average method*, Environ. Eng. Res.*, 22(2): 162-168.

Montanari, A., Rosso, R. and Taqqu, M.S. 2000. A seasonal fractional ARIMA model applied to the Nile river monthly flows at Aswan, *Water Resour. Res.*, 36(5): 1249-1259.

Pazvakawambwa, G.T. and Ogunmokun, A.A. 2013. A Time-series forecasting model for Windhoek rainfall Namibia, UNAM, pp 1-11.

Rajagopalan, B. and Lall, U. 1999. A k-nearest neighbour simulator for daily precipitation and other variables, *Water Resour. Res.,* 35(10): 3089-3101.

Salahi, B., Nohegar, A. and Behrouzi, M. 2016. The modelling of precipitation and future droughts of Mashhad Plain using stochastic time series and standardized precipitation index (SPI), *Int. J. Environ. Res.,* 10(4): 625-636.

Sharif, M. and Azhar, H. 2017. Simulation of extreme dry and wet spells in Brahmaputra Basin using K-Nearest Neighbour Model, *Int. J. Environ. Sci. Nat. Res.,* 4(5): 170-176.

Srinivasan, K. 1995. Stochastic modeling of Monsoon River flows, *J. Appl. Hydrol.,* 8: 51-57.

Sumi, S.M., Zaman, M.F. and Hirose, H. 2008. A Rainfall forecasting method using machine learning models and its application to Fukuoka City case, *Int. J. Appl. Math. Comput. Sci.*, 22(4): 841-854.

Trawinski, P.R. and Mackay, D.S. 2008. Meteorologically conditioned time series predictions of West Nile virus vector mosquitoes, *Vector-Borne and Zoonotic Diseases*, 8(4): 505-522.

Verma, A. 2004. Stochastic modeling on monthly rainfall of Kota, Rajasthan, M.E. Thesis Submitted to GBPUAT, Pantnagar, India.

Wang, H.R., Wang, C., Lin, X. and Kang, J. 2014. An improved ARIMA model for precipitation simulations, *Nonlin. Process. Geophys.*, 21(6): 1159-1168.

Wang, S., Feng, J. and Liu, G. 2013. Application of Seasonal time series model in the precipitation forecast, Math. *Comput. Model*, 58(3-4): 677-683.