ORIGINAL PAPER

Developing monthly hydrometeorological timeseries forecasts to reservoir operation in a transboundary river catchment

Hammad-ur-Rehman[1] · Ijaz Ahmad[1] · Faraz-ul-Haq[1] · Muhammad Waseem[1] · Jinxin Zhang[2]

Abstract

Forecasting of hydrometeorological timeseries data plays a vital role in the flood forecasting and predicting the future water availability for various purposes such as irrigation, hydropower generation, industrial, domestic, etc. Therefore, the present study aims to forecast the hydrometeorological timeseries data, i.e., river inflows, precipitation, and evaporation for the improved reservoir operation of a transboundary Mangla catchment by using the auto-regressive integrated moving average (ARIMA) model. Prior to applying the ARIMA model, stationarity of hydrometeorological timeseries data was tested. Moreover, autocorrelation function (ACF) and partial autocorrelation function (PACF) of timeseries were determined to calculate the "p" and "q" terms of the ARIMA model. The best fitted structure of ARIMA was selected by evaluating the coefficient of determination (R^2), mean square error (MAE), and root mean square error (RMSE) to forecast the hydrometeorological timeseries. The seasonal ARIMA structure of $(1,0,0)(2,1,2)_{12}$ was found to be best fitted for the inflow timeseries whereas ARIMA structures of $(14,1,15)$ and $(9,1,19)$ were considered for forecasting the precipitation and evaporation timeseries, respectively. An average water shortage of 14% was detected by using these forecasted hydrometeorological timeseries in the reservoir operation for the period of 2016–2030. It was also observed that the seasonal effect for the reservoir inflows was more pronounced compared to the evaporation and precipitation timeseries. However, variations in the precipitation timeseries were found more abrupt than the inflows timeseries. It is believed that the results of this study may support reservoir operators and managers for developing the efficient real-rime reservoir operation policies and strategies based on the variations in the future water availability.

1 Introduction

In ancient times, human civilization built the towns and cities at places where water was easily accessible for their survival. However, due to rapid increase in population growth, water requirements have been kept increasing and to resolve this issue, artificial water conservation structures, i.e., dams and reservoirs were constructed. These reservoirs can be used for different purposes such as irrigation, recreation, groundwater recharging, hydropower generation, etc. (Sarwar 2013). The economy of many countries in the world has been based on these reservoirs especially in the arid and semi-arid regions where the water availability is decreasing continuously to meet the ever-increasing agricultural demands.

As the water availability is decreasing at an alarming rate to fulfill the water demands of various sectors, i.e., domestic, industry, agriculture, etc. and necessitates the efficient use of the limited water resources (Pulido-Calvo et al. 2003), better understanding of variability in the hydrometeorological timeseries data can play an important role in performing the efficient reservoir operation. The phenomena of global warming and climate change are affecting the trends of hydrometeorological timeseries data and can further alter the future of water resources and its water availability (Musa 2013). The impact of climate change on water resources and the hydropower generation was studied by Zaman et al. (2018) for Xin'anjiang watershed in China.

The variations in the streamflow are very complex in nature. Different models were used to determine these variations in previous studies (Qin et al. 2015; Chang and Yeh

✉ Ijaz Ahmad
dr.ijaz@uet.edu.pk

[1] Centre of Excellence in Water Resources Engineering, University of Engineering and Technology, Lahore 54890, Pakistan

[2] College of Water Resources and Environmental Engineering, Zhejiang University of Water Resources and Electric Power, Hangzhou 310018, China

2016). Integrated Flood Analysis System (IFAS) hydrological model has been applied to observe variations in the transboundary Jhelum river basin (Umer et al. 2021). Similarly, a flood forecasting system was developed using the IFAS model for the river basins of river Jhelum and Chenab (Shahzad et al. 2018). The behavior of precipitation and extreme events was studied using precipitation indices for the northern highlands of Pakistan (Haider et al. 2020; Zaman et al. 2020).

The forecasting models play a pivotal role in the water management to fulfill the crop water requirements and develop the flood mitigation strategies. These models are also used to predict the future water availability and to extend the data trends and detecting the missing values (Zamani et al. 2017). In case of reservoir operation, the future projected inflows have key importance because accurate inflow prediction is not only an important non-engineering measure to ensure flood-control safety and increase water resource use efficiency, but also can provide guidance in reservoir planning and management. The design of water resources projects is mainly dependent on the streamflow and its duration. Therefore, researchers have more interest in streamflow forecasting. Rainfall-runoff models, general circulation models and stochastic models can be used for the prediction of stream flows (Wang et al. 2008, 2019; Bahremand and de Smedt 2010; Xu et al. 2014).

Several studies used the stochastic models for the prediction of streamflow and analyzing the variations in river runoff (Lawrance and Kottegoda 1977; Koutsoyiannis 2000). Among the different stochastic models, autoregressive integrated moving average (ARIMA) stochastic models and autoregressive moving average (ARMA) models have been widely used especially for the river flow forecasting (Adnan et al. 2017). These stochastic models are also known as Box-Jenkins linear stochastic models (Box et al. 2015). These models are widely used because of their capability to model and forecast a high variety of stationary and non-stationary timeseries data (Brown and Mac 2002). These models are also capable to estimate serial correlations and trends in the timeseries data (Gocheva-Ilieva et al. 2014). Moreover, these models can be incorporated into many software packages including SPSS, MINITAB, STATA, R, Python, MATLAB and Mathematica (Ozgur et al. 2015).

Adhikary et al. (2012) used the seasonal ARIMA model to predict the groundwater variations in a shallow unconfined aquifer of Bangladesh. The results indicated that ARIMA model can generate reasonably forecasts in terms of numerical accuracy. In another study, Valipour et al. (2013) used the combination of SARIMA and ARMA models to forecast the inflows into the Dez reservoir in Khuzistan, Iran and concluded that SARIMA model yields better results compared to the ARMA model. Modarres and Ouarda (2013) demonstrated the heteroscedasticity in the streamflow

timeseries by using the ARIMA and Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) models and stated that the ARIMA model performed better than the GARCH model. Kurunç et al. (2005) used the ARIMA model for forecasting the streamflow data of Yesilurmah River at Durucasu monitoring station. The time series data of 13 years was used for the forecasting of water quality and streamflow data. Statistical modeling was performed of the monthly streamflow using timeseries models for the Hindiya Barrage in Iraq (Al-Saati et al. 2021). ARIMA model was used for the generation of streamflow synthetically for the Gomti river basin in Uttar Pradesh, India (Singh and Ray 2021). Various other studies also proved the effectiveness of ARIMA model compared to the other models (Ahmad et al. 2001; Kurunç et al. 2005; Tayyab et al. 2016; Sentas et al. 2016; Yeh and Hsu 2019; Hendikawati et al. 2020).

Pakistan is an agricultural country and most of its lands are irrigated with the canal irrigation systems. Two major reservoirs (i.e., Mangla and Tarbela) are considered as the backbone of Pakistan's agriculture-based economy. The prediction of water availability can play an important role to regulate the river flows efficiently. For instance, a reservoir should be empty or partially empty before the occurrence of any extreme inflow event. Otherwise, water storage level may exceed the safe limits and causes floods. To avoid these situations, a reliable forecasting of hydrometeorological data is essential. The forecasted hydrometeorological timeseries are considered helpful in managing the reservoir operation and satisfy the water demands of various sectors.

Therefore, this study aims to forecast the hydrometeorological data being used in reservoir operation model to develop future water allocation scenarios. The statistical forecasted models are more famous in research community due to its accuracy and efficiency. Especially ARIMA model is considered most suitable for linear and seasonal timeseries forecasting. Therefore, in this study hydrometeorological timeseries is forecasted by using ARIMA model. The forecasted timeseries were used in reservoir operation to determine the fluctuations in the reservoir storage. It is believed that the results of the present study may guide the reservoir operators and managers to predict the future uncertainties in hydrometeorological timeseries data.

1.1 Study area

Mangla reservoir is the second largest reservoir of Pakistan. It is located in Mirpur district of Azad and Jammu Kashmir as shown in Fig. 1. Its 45% catchment lies in Pakistan administered Kashmir and the remaining is the territory of Indian side of Kashmir. The Mangla reservoir covers an area of 329.7 km^2 with gross storage capacity of 9.22 km^3. The agricultural command area of Mangla reservoir is 60,000 km^2. The primary purpose of the Mangla reservoir is to irrigate the agricultural lands. The Mangla reservoir supplies water to the upper and
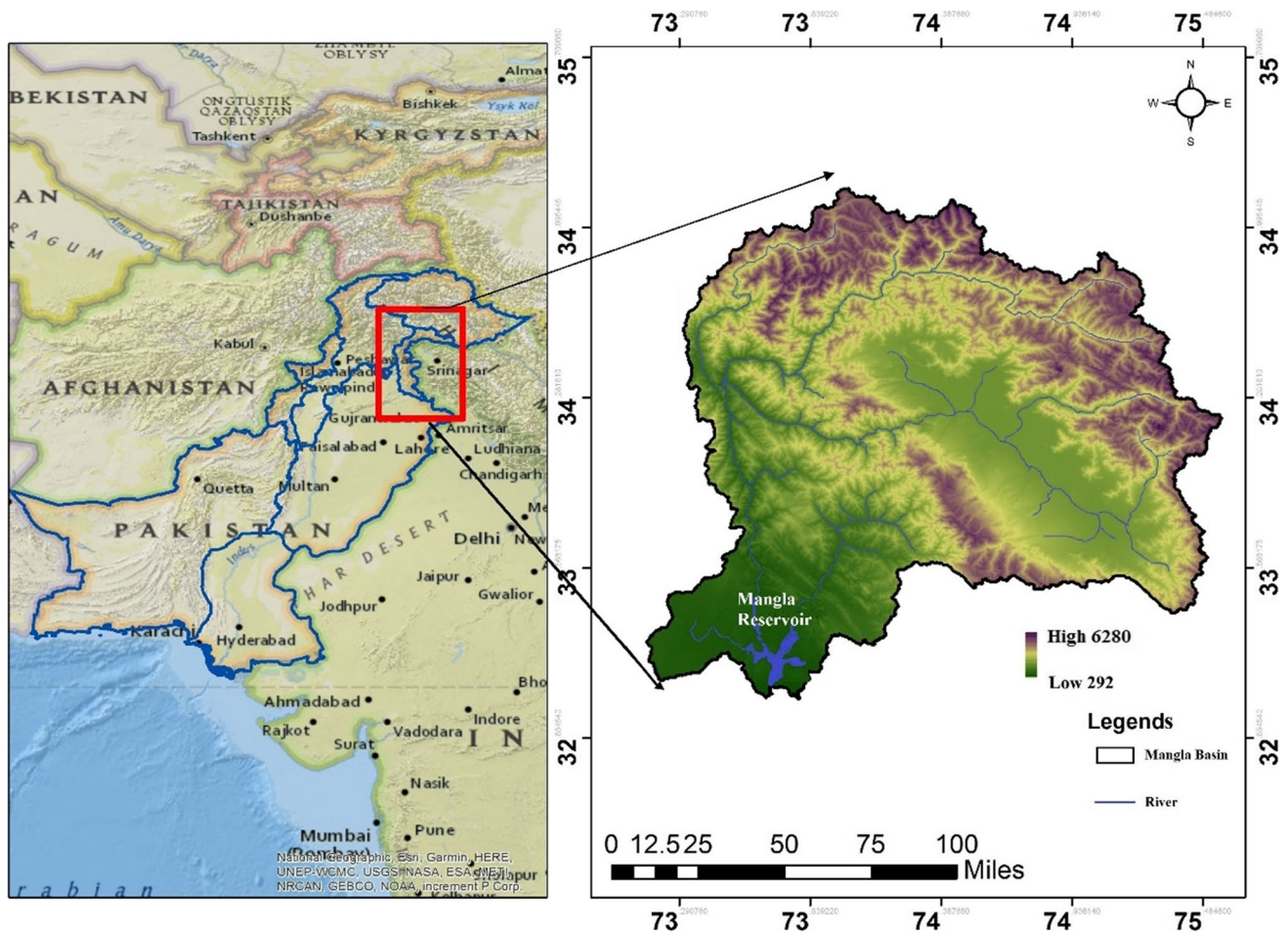
Fig. 1 Mangla watershed and reservoir

lower Chasma-Jhelum link canal regions. During the period of December to March (Rabi seasons) the irrigation demands are at their peak from for upper Jhelum region for which the water supplied from Mangla reservoir. Mean monthly variations in the inflows and precipitation are shown in Fig. 2. The excess water from the upper Jhelum River basin irrigates the agriculture lands of the lower Indus plains through a network of canals. Moreover, hydropower of 1000 MW (megawatt) is produced as by-product when the water released from the Mangla reservoir for any purpose passes through the turbines.

2 Materials and methods

The ARIMA model was used to forecast the hydrometeorological timeseries data. Prior to the application of ARIMA model, the hydrometeorological timeseries data was tested for the stationarity. For this purpose, model identification was performed, and parameters p, d, and q were estimated. These estimated parameters were used to forecast the timeseries. The

forecasted timeseries incorporated into reservoir operation model to estimate the future water availability.

2.1 Formulation of ARIMA model

2.1.1 Data analysis

Timeseries data analysis was performed to check its trend, extremes and mean monthly values. These factors are also included into the simulated timeseries. The timeseries data analysis performed to evaluate the reliability of the data. In the ARIMA model, the stationary condition of timeseries is necessary before forecasting the hydrometeorological data.

2.1.2 Stationary test

The timeseries data should be stationary prior to the application of ARIMA model. The timeseries is considered stationary if its mean E(at), variance Var(at) and covariance Cov (at, at-1) are constant.

Fig. 2 a Mean monthly inflows at Mangla reservoir. **b** mean monthly precipitation at Mangla
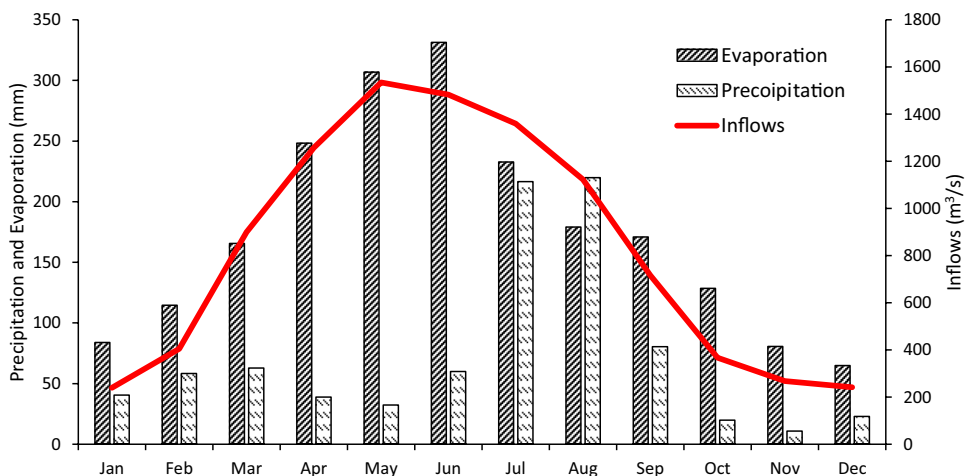
2.1.3 Differencing

If the timeseries was not found stationary, then by taking the differencing it can be converted into stationary timeseries. The trend component of a timeseries will be removed with the help of differencing, and irregular component will be left. The irregular component (sometimes also known as the residual) is what remains after the seasonal and trend components of a time series have been estimated and removed. After first-order differencing, the timeseries will be converted into stationary and if this does not happen, differencing will continue until it converts into stationary, e.g., second-order differencing, third-order differencing and so on.

2.2 Model identification

The model identification depended on the three terms "*AR*" (*p*), order of differencing (*d*) and "*MA*" (*q*) terms. The value "*p*" is identified by the autocorrelation function (ACF). If the value of ACF of timeseries is equal to zero or near to zero, "*p*" term is considered satisfied. Similarly, the PACF is used to determine the "*q*" term. "*d*" represents the degree of differencing to make the timeseries stationary.

2.2.1 Auto Correlation Function (ACF)

A timeseries "a_n" lagged by "*k*" times. Its autocorrelation function is defined as.

$$ACFr_k = \frac{\sum_{t=1}^{n-k} (a_t - \bar{a})(a_{t-k} - \bar{a})}{\sum_{t=1}^{n} (a_t - \bar{a})^2} \tag{1}$$

2.2.2 Partial autocorrelation function (PACF)

Partial autocorrelation function (PACF) is used to determine the degree of association between the "a_t" and "a_{t-k}" and it can be evaluated by using the following formula.

$$r_{kk} = \begin{cases} r_1 \ k = 1, \\ \dfrac{r_k - \sum\limits_{j=1}^{k-1} r_{k-1,j} \cdot r_{k-j}}{1 - \sum\limits_{j=1}^{k-1} r_{k-1,j} \cdot r_k} & \text{if } k = 2, 3, \dots \end{cases} \tag{2}$$

Whereas

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j} \ for \ j = 1, 2, \dots., k-1$$

It means that "r_{kj}" represented ACF of the timeseries by lag "*j*" time units. Therefore, the intervening observation between "*k*" and "*j*" is eliminated, whereas the standard error of PACF is given in the following equation.

$$S_{r_{kk}} = \sqrt{\frac{1}{n}} \tag{3}$$

2.3 Parameter estimation

The timeseries parameter (p, d, q) can be estimated by the sample moment estimation, linear least square method, max likelihood estimation (generalized), method of moments, Bayesian estimation or Kalman filtering. However, the least square method is frequently used to estimate the moment average (*MA*) component. Sometimes *max likelihood function* is used to define the probability of observed data.

2.4 Diagnostic checking

After the model identification mostly Akaike's information criterion (AIC) and Bayesian information criterion (BIC) are

performed and the least value of AIC and BIC is considered as the best fitted model (Reza Ghanbarpour et al. 2010).

$$AIC = -2\ln(L) + 2k \qquad (4)$$

$$BIC = -2\ln(Max.Likelihood) + Tp\ \ln(n) \qquad (5)$$

where L is the likelihood function, k represents the number of parameters to be estimated and denotes the number of terms.

2.5 Forecasting

The h-period ahead forecast based on an ARIMA (*p, d, q*) model where the d = 0 is given in Eq. (5).

$$\hat{y}_{t+h} = \hat{\delta} + \hat{\varphi}_1 y_{t+h-1} + \cdots \hat{\varphi}_p y_{t+h-p} + e_{t+h} - \hat{\theta}_1 e_{t+h-1} - \cdots \hat{\theta}_q e_{t+h-q} \qquad (6)$$

Y_{t+h} terms are replaced with the estimated values until the last observed values.

2.6 Reservoir operation management

The future predicted hydrometeorological timeseries data was used to perform the reservoir operation on monthly time step. The forecasted reservoir operation represented the future expected reservoir water levels, reservoir storages and reservoir shortage/excess periods. This information can be helpful to develop and improve the reservoir operation policies. The inflows and precipitation play important role to raise the water levels where as outflows, evaporations and water demands decreased the water levels in the reservoir. Reservoir release curve represents the reservoir outflows and are dependent on the downstream water demands. Sometimes, inflows were high and water levels were increased abruptly. To avoid such situations, excess water can be released compared to the downstream water demands. Therefore, Güntner et al. (2009) have proposed the following method to conduct the reservoir operation studies.

$$S_t = S_{t-1} + I - O - U_{RL} + (P - E)A_{RL} \qquad (7)$$

where S_t represents the reservoir storage at timestep t, S_{t-1} is the reservoir storage at time step t-1, P denotes the precipitation at water surface area RL, E is the evaporation at water surface area RL and U_{RL} represents the water withdrawal from the reservoir.

3 Results and discussions

The historical hydrometeorological timeseries data of Mangla reservoir from 1991 to 2015 was used to formulate the ARIMA model. Mangla watershed has widespread snow melting effect in the water availability. The inflow timeseries was found to be nonstationary which implied that its mean, variances, and covariance are not constant. Therefore, it is required to convert the timeseries into stationary as ARIMA model cannot perform analysis on to the nonstationary timeseries.

Several trials were performed to convert the nonstationary timeseries into stationary. However, after the first-, second-, and third-order differencing, timeseries was not found stationary. Therefore, the log transformation was used as the second option. After log transformation, the inflow timeseries was converted into stationary. Further analysis was performed on stationary timeseries.

3.1 ACF and PACF

Different scenarios were created due to abrupt variations in precipitation timeseries. Therefore, autocorrelation function (ACF) and partial autocorrelation function (PACF) were evaluated as shown in Fig. 3. Widespread variations in the precipitation timeseries were also evident from the high values of ACF and PCF. It should also be noted that the nine values were significant out of the first 12 tested values. Afterwards, all the values remained in the limit. The maximum value of ACF was 0.5 which occurred after lag 11. For PACF, most of the values remained in the range of confidence limits in the precipitation timeseries. The maximum value was 0.3 which occurred after the first lag. The ACF and PACF values for the evaporation timeseries are also shown in Fig. 3. The ACF and PACF spikes were changed after high lag values and lie in the limit of 5% confidence interval. Therefore, 9 AR terms and 19 MA terms were selected, respectively. Similarly, ACF and PACF values for the inflow timeseries were evaluated.

3.2 Calibration and validation

The hydrometeorological timeseries of the Mangla reservoir was divided into two periods for the calibration and validation. The calibration period was selected from 1991 to 2005, whereas validation period from 2006 to 2015. In the calibration and validation of the ARIMA model, the simulated hydrometeorological timeseries of inflows, evaporation, and precipitation timeseries were compared with the observed values. The R^2, MAE and RMSE values of inflow timeseries was evaluated and found as 0.89, 113, and 209, respectively. Whereas the observed precipitation timeseries were also comparable with the model simulated values, and their R^2, MAE, and RMSE values were 0.81, 28, and 43, respectively. The values of R^2, MAE, and RMSE of evaporation timeseries were 0.78, 27, and 39, respectively.
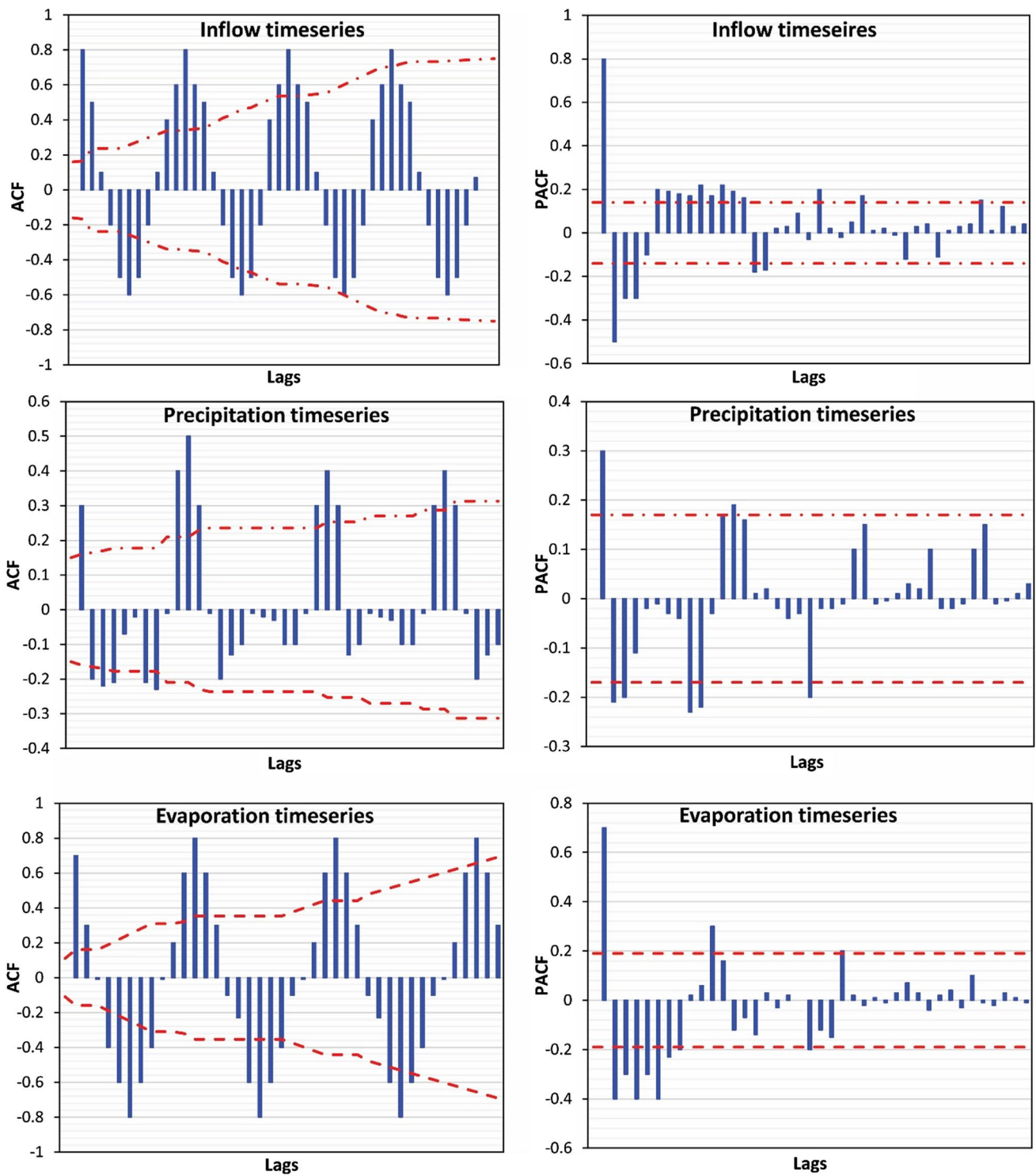
Fig. 3 ACF and PACF of inflow, precipitation and evaporation timeseries

In validation period (2006–2010), the simulated time-series were compared with observed data, and the values of R^2, MAE, and RMSE were evaluated. For the inflows timeseries, the values of R^2, MAE, and RMSE were 0.85, 145 and 195, respectively. Whereas the R^2, MAE and RMSE of precipitation timeseries were 0.83, 14.5, and 31.7,

respectively. The values of R², MAE, and RMSE of evaporation timeseries are 0.88, 21 and 31, respectively.

3.2.1 Calibration of inflow, precipitation and evaporation timeseries

On the basis of RMSE, R^2 and MAE values, the best fitted seasonal ARIMA model structure of $(1,0,0)(2,1,2)_{12}$ was selected for the inflow timeseries. The model simulated and observed flows are presented in Fig. 4a. The simulated values were compared with the observed flows and most of the peak values were found consistent with the observed ones. Precipitation timeseries has large fluctuations and frequent peaks. Precipitation timeseries was less smooth compared to the inflows and evaporation timeseries. The best fitted model structure for the precipitation timeseries was found as (14,1,15) and during the calibration process, the AR value of 14 was selected. The model simulated and observed flows are presented in Fig. 4b.

ARIMA model structure of (9,1,19) was found to be best fitted for the evaporation timeseries. During the calibration period, i.e., 1990–2005, the simulated and observed data were compared and the results are presented in Fig. 4c. It is clear from Fig. 4 that the first two peaks of observed evaporation were lower compared to simulated timeseries. Afterwards, observed data peaks were higher than the simulated data. The maximum difference for simulated and observed values was observed in the year 1998.

3.2.2 Validation of inflow, precipitation and evaporation timeseries

The period of 1991–2015 was selected as the validation period for the inflow timeseries. In the validation period, most of the peaks of observed inflow timeseries were found slightly above the model simulated values. In the inflow timeseries, the seasonal trend was more prominent. Therefore, seasonal ARIMA model structure was found to be best fitted for the inflow timeseries. It was compared with the observed flows and found the values of R^2, RMSE, MAE as 0.85, 195 and 145, respectively. After the validation, the same ARIMA model structure was used to forecast the inflow timeseries for the period of 2016–2030. In the validation period of 2005–2015, the first peak of simulated flows is slightly higher than that of the observed flows as shown in Fig. 4a. In the year of 2007, maximum value of inflow is 1961 cumecs, whereas in the same year, the simulated flows exhibited a peak value of 1280 cumecs. The next year 2008, the observed inflows are lower than the simulated flows.

In the validation period, precipitation peaks were fluctuated between values of 150 and 330 mm. The precipitation peak was decreased from years of 2006 to 2010 and afterwards again raised. However, in the last 2 years of the

validation period, the simulated peaks were higher than the observed timeseries. The values of 31.7 and 21 were found for RMSE and MAE, respectively during the comparison of observed and model-simulated precipitation timeseries as shown in Fig. 4b.

During the validation period (2006–2015) of evaporation timeseries, the same ARIMA model structure was used as that of during the calibration process and found that the simulated evaporation timeseries was comparable with the observed evaporation data. The first three peaks of timeseries were overlapped with the simulated values. The values of statistical parameters of MAE and RMSE values were found as 14.5 and 31, respectively. Few sudden peaks in the observed values were found slightly deviating with simulated values as shown in Fig. 4c.

3.3 Forecasting of hydrometeorological timeseries data

After the calibration and validation process, the same ARIMA model structure was used to forecast the inflow timeseries for the period of 2015–2030. It is clear from Fig. 5 that the maximum flows are predicted in the year 2018–2019, whereas the lowest flows are expected in the year of 2023. This forecasted inflows timeseries was used as input in the reservoir operation model. Afterwards, the forecasted reservoir rule curves were determined along with the water shortage periods during the study period. These predictions may help the reservoir operators and managers to fulfill the water demands of the various sectors. Similarly, to extrapolate the precipitation timeseries up to the year of 2030, the same ARIMA model structure of (14,1,15) was used as adopted during the model calibration and validation. The simulated results are showing many similarities to the historical precipitation timeseries data. Most of peak values were observed less than 300 mm. The fluctuations in the precipitation timeseries were observed steeper compared to the inflow timeseries and the highest peak was predicted in the year of 2027.

After the validation of evaporation timeseries, ARIMA model structure of (9,1,19) was adopted. These forecasted timeseries were used for the reservoir operation and are presented in Fig. 5. Highest peak in the evaporation timeseries was observed in the month of June because of the highest temperature being observed in this month. The simulated timeseries was used for reservoir operation and forecasted water shortage and respective rule curves of reservoir were determined. These predictions of future water availability and demands may be helpful for reservoir operators and managers to regulate the reservoir efficiently.
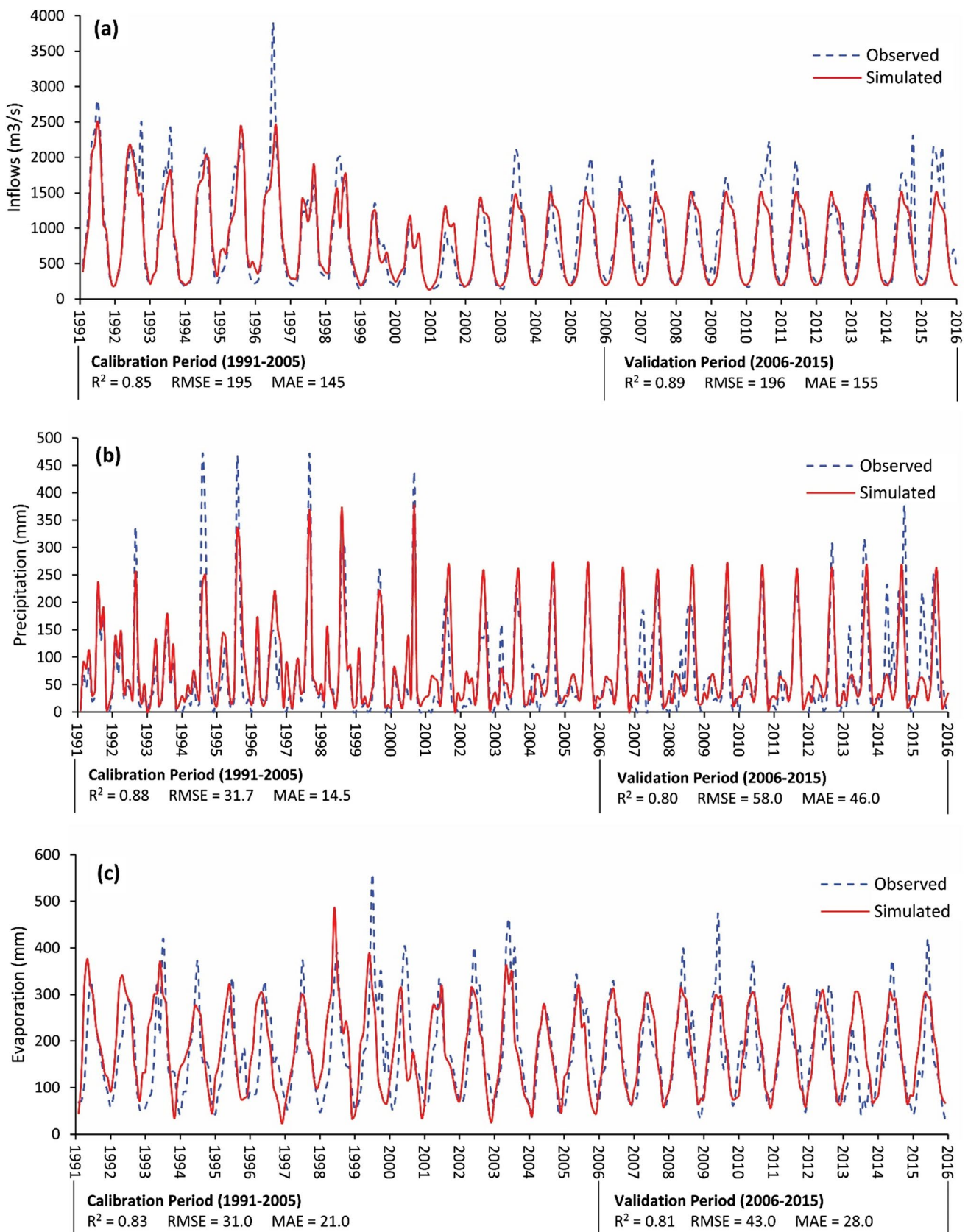
Fig. 4 Calibration and validation of observed and simulated data of ARIMA model during the period 2006–2015
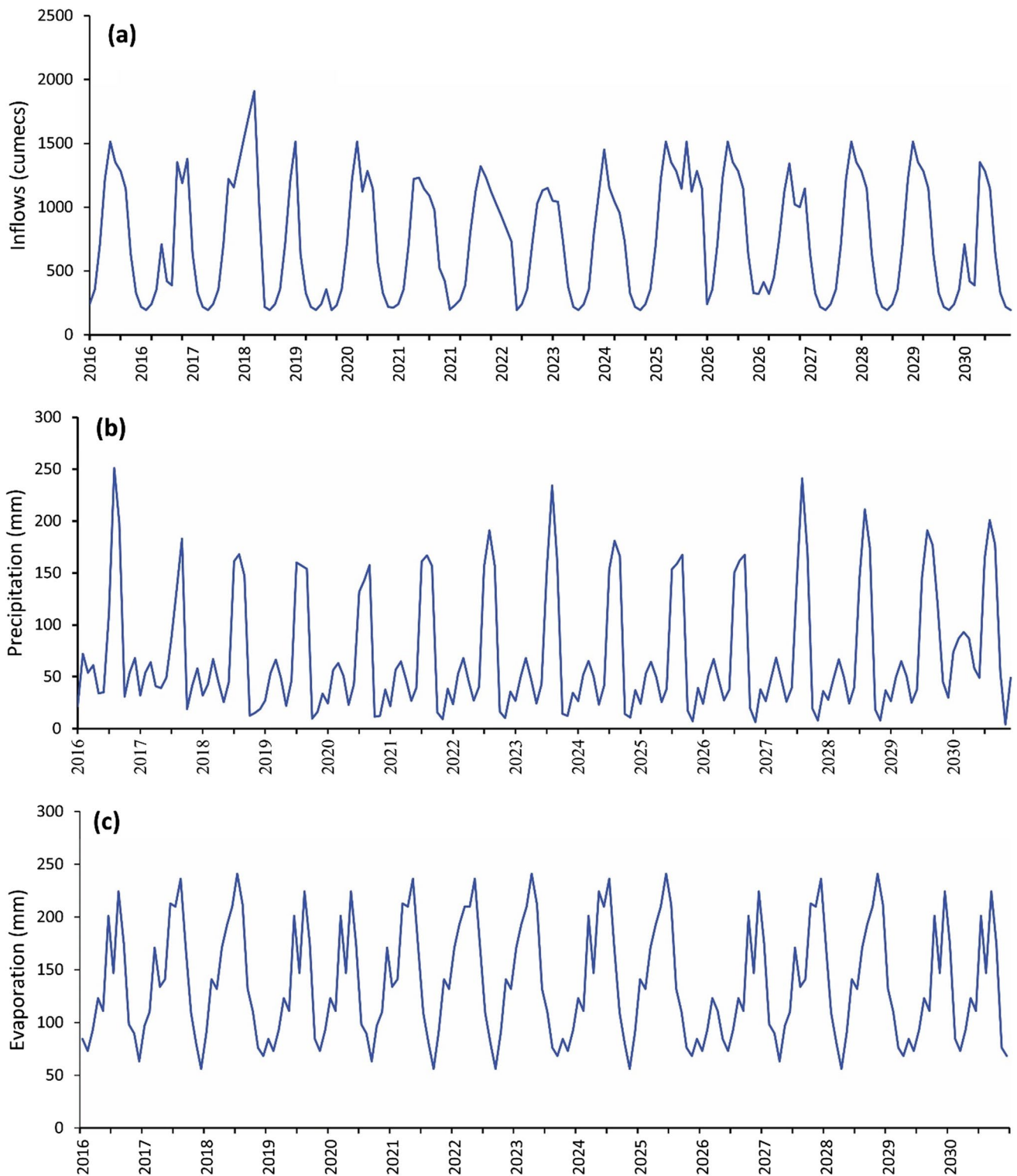
Fig. 5 Simulated data of ARIMA model during the period 2016–2030

3.4 Reservoir operation

Reservoir operators and managers are very conscious to its safety and efficiency. To enhance its benefits, future aspects are also taken into considerations. Therefore, ARIMA model was used for forecasting of inflows, precipitation, and evaporation timeseries. After forecasting the hydrometeorological time-series, reservoir operation simulation was performed on monthly
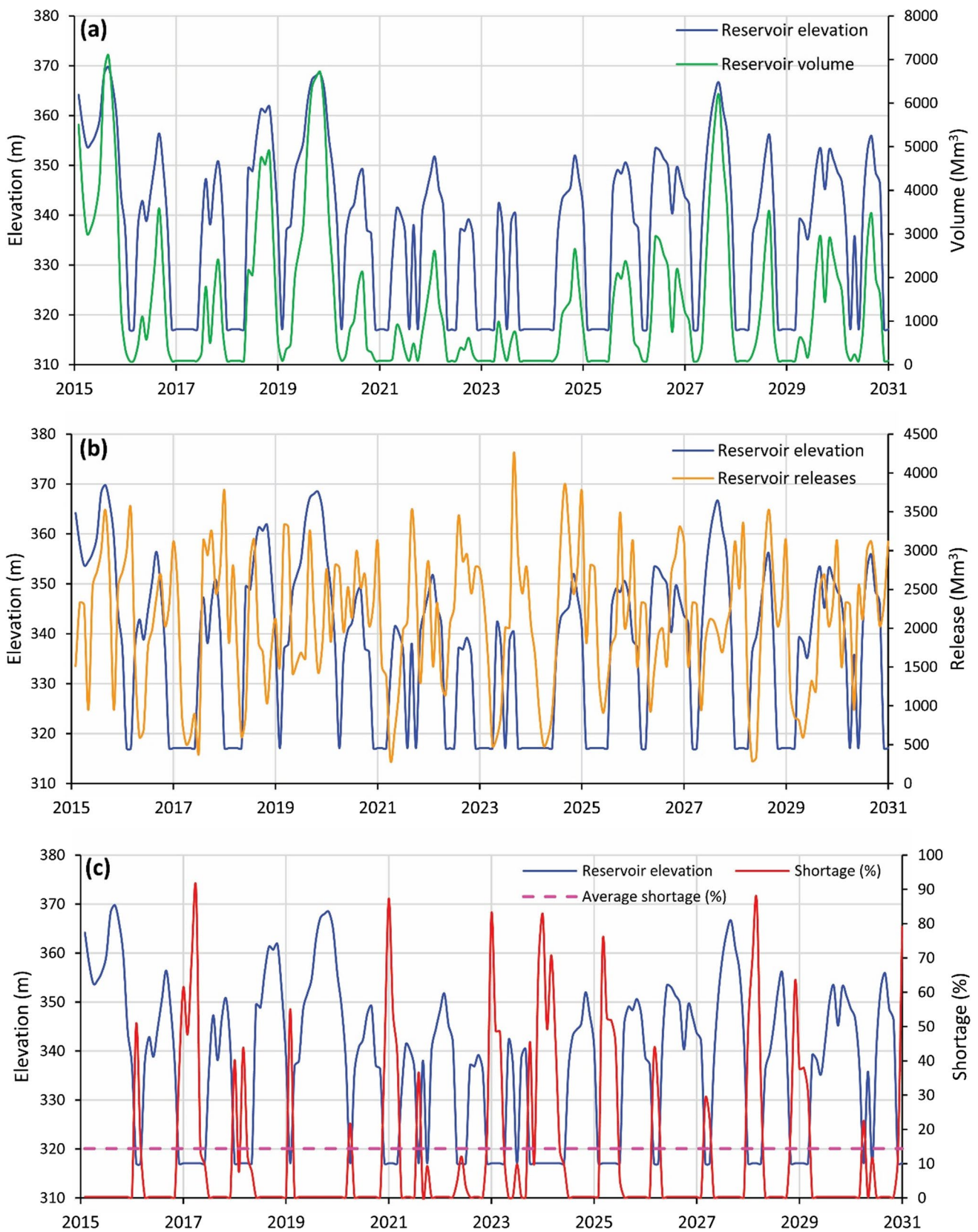
Fig. 6 **a** Mangla reservoir levels. **b** Mangla reservoir volume. **c** Reservoir shortage curve for the period of 2016–2030

time step. The forecasted reservoir operation represented the future expected reservoir water levels, reservoir storages, reservoir shortage and/or excess periods. This information can be helpful to develop and improve the reservoir operation policies.

Reservoir elevation curve represents the water levels in the reservoir at various time periods. The inflows and precipitation play an important role to raise the water levels whereas outflows, evaporation, and water demands result in decreasing the reservoir water levels in the reservoir. Sudden rise and drawdown are considered dangerous for a reservoir. Such conditions may create the cracks in the reservoir body. In Fig. 6a, reservoir elevation curve is presented. Most of the peaks in the curve have 1-year difference which is safe and smooth changes in the water levels.

Reservoir release curve represents the reservoir outflows. In most of the cases, reservoir releases are dependent on the downstream water demands. Sometimes, inflows were high and water levels were increased abruptly. To avoid such situations, excess water can be released than the downstream demands, so that water levels remained in the safe limits. Figure 6b presents the reservoir releases during period of 2016–2030. The y-axis represents the quantity of water in million cubic meters (Mm^3) and x-axis represents the time scale. The reservoir operation was performed on monthly time step; therefore, these values were represented monthly basis.

The quantity of water stored in the reservoir is represented by the reservoir storage curves. When reservoir storages were high, then the water was available to fulfill the downstream water requirements. Figure 6b represents the reservoir storages. During the period of 2016–2020, water is available to fulfill the water demands. However, during the period of 2021–2026, the reservoir remains partially filled. In the year of 2027, excess water is available for storage. When the downstream water demands are not fulfilled, it is represented as the water shortage period in percentage. With the rapid increase in population, water demands for various purposes are increasing at alarming rate in Pakistan, whereas, water storages remained the same, to fulfill these ever-increasing water demands. Therefore, water shortages are increased. Another aspect of sedimentation is also decreasing the water storages. In Fig. 6c, water shortages were presented. During the low flow periods, the water shortages were found higher. Another aspect is the high-water demand periods, when the irrigation demands are higher than the amount of water is required more and causing the water shortages. The average water shortage of forecasted reservoir operation is 14% during the period of 2016–2030.

4 Conclusions

In this study, ARIMA model is used to forecast the hydrometeorological timeseries of the Mangla reservoir. Seasonal ARIMA model $(1,0,0)(2,1,2)_{12}$ is best fitted for the inflow time series. For smoothening the time series (ln) log transformation is used. After comparison of the results with the observed inflows, it is highly accurate based on R^2, RMSE, and MAE, whereas precipitation and evaporation timeseries are not periodic. Therefore, ARIMA model (14,1,15) is best fitted for precipitation and (9,1,19) is selected for evaporation timeseries. After forecasting of these timeseries up to 2030, reservoir operation of Mangla is performed on the monthly time step. The forecasted rule curves are developed and the water shortage up to 2030 is determined. The average water shortage for this period is 14%. The forecasted hydro-meteorological timeseries data may support the reservoir operators and managers for developing the efficient real-time reservoir operation policies and strategies and optimization of reservoir.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

References

Adhikary SK, Rahman MM, Gupta A Das (2012) A stochastic modelling technique for predicting groundwater table fluctuations with time series analysis. Int J Appl Sci Eng Res 1. https://doi.org/10.6088/ijaser.0020101024

Adnan RM, Yuan X, Kisi O, Curtef V (2017) Application of time series models for streamflow forecasting. Int J Civ Environ Eng 9:56–63

Ahmad S, Khan IH, Parida BP (2001) Performance of stochastic approaches for forecasting river water quality. Water Res 35:4261–4266

Al-Saati NH, Omran II, Salman AA et al (2021) Statistical modeling of monthly streamflow using time series and artificial neural network models: Hindiya Barrage as a case study. Water Pract Technol 16:681–691. https://doi.org/10.2166/wpt.2021.012

Bahremand A, de Smedt F (2010) Predictive analysis and simulation uncertainty of a distributed hydrological model. Water Resour Manag 24:2869–2880. https://doi.org/10.1007/s11269-010-9584-1

Box GEP, Jenkins GM, Reinsel GC, Ljung GM (2015) Time series analysis: forecasting and control, 5th edn. Wiley

Brown LC, Mac BP (2002) Statistics for environmental engineers, 2nd edn. CRC Press

Chang C-M, Yeh H-D (2016) Stochastic modeling of variations in stream flow discharge induced by random spatiotemporal fluctuations in lateral inflow rate. Stoch Environ Res Risk Assess 30:1635–1640. https://doi.org/10.1007/s00477-015-1170-x

Gocheva-Ilieva SG, Ivanov AV, Voynikova DS, Boyadzhiev DT (2014) Time series analysis and forecasting for air pollution in small urban area: an SARIMA and factor analysis approach. Stoch Environ Res Risk Assess 28:1045–1060. https://doi.org/10.1007/s00477-013-0800-4

Güntner A, Krol MS, Araújo JC De, Bronstert A (2009) Simple water balance modelling of surface reservoir systems in a large data-scarce semiarid region. 49. https://doi.org/10.1623/hysj.49.5.901.55139

Haider H, Zaman M, Liu S et al (2020) Appraisal of climate change and its impact on water resources of Pakistan: a case study of Mangla Watershed. Atmosphere (basel) 11:1071. https://doi.org/10.3390/atmos11101071

Hendikawati P, Subanar A, Tarno (2020) A survey of time series forecasting from stochastic method to soft computing. J Phys Conf Ser 1613:012019. https://doi.org/10.1088/1742-6596/1613/1/012019

Koutsoyiannis D (2000) A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. Water Resour Res 36:1519–1533. https://doi.org/10.1029/2000WR900044

Kurunç A, Yürekli K, Çevik O (2005) Performance of two stochastic approaches for forecasting water quality and streamflow data from Yeşilurmak River, Turkey. Environ Model Softw 20:1195–1200. https://doi.org/10.1016/j.envsoft.2004.11.001

Lawrance AJ, Kottegoda NT (1977) Stochastic modelling of riverflow time series. J R Stat Soc Ser A 140:1. https://doi.org/10.2307/2344516

Modarres R, Ouarda TBMJ (2013) Modelling heteroscedasticty of streamflow times series. Hydrol Sci J 58:54–64. https://doi.org/10.1080/02626667.2012.743662

Musa JJ (2013) Stochastic modelling of Shiroro River stream flow process. 49–54

Ozgur C, Kleckner M, Li Y (2015) Selection of statistical software for solving big data problems: a guide for businesses, students, and universities. https://doi.org/10.1177/2158244015584379

Pulido-Calvo I, Roldán J, López-Luque R, Gutiérrez-Estrada JC (2003) Demand forecasting for irrigation water distribution systems. J Irrig Drain Eng 129:422–431. https://doi.org/10.1061/(ASCE)0733-9437(2003)129:6(422)

Qin G, Li H, Zhou Z et al (2015) Hydrologic variations and stochastic modeling of runoff in Zoige wetland in the Eastern Tibetan Plateau. Adv Meteorol 2015:1–6. https://doi.org/10.1155/2015/529354

Reza Ghanbarpour M, Abbaspour KC, Jalalvand G, Moghaddam GA (2010) Stochastic modeling of surface stream flow at different

time scales: Sangsoorakh karst basin, Iran. J Cave Karst Stud 72:1–10. https://doi.org/10.4311/jcks2007ES0017

Sarwar M (2013) Reservoir life expectancy in relation to climate and land-use changes : case study of the Mangla Reservoir in Pakistan The University of Waikato. THE UNIVERSITY OF WAIKATO

Sentas A, Psilovikos A, Psilovikos T, Matzafleri N (2016) Comparison of the performance of stochastic models in forecasting daily dissolved oxygen data in dam-Lake Thesaurus. Desalin Water Treat 57:11660–11674. https://doi.org/10.1080/19443994.2015.1128984

Shahzad A, Gabriel HF, Haider S et al (2018) Development of a flood forecasting system using IFAS: a case study of scarcely gauged Jhelum and Chenab river basins. Arab J Geosci 11:383. https://doi.org/10.1007/s12517-018-3737-6

Singh H, Ray MR (2021) Synthetic stream flow generation of river Gomti using ARIMA Model. pp 255–263

Tayyab M, Zhou J, Zeng X, Adnan R (2016) Discharge forecasting by applying artificial neural networks at the Jinsha River Basin, China. Eur Sci J ESJ 12:108. https://doi.org/10.19044/esj.2016.v12n9p108

Umer M, Gabriel HF, Haider S et al (2021) Application of precipitation products for flood modeling of transboundary river basin: a case study of Jhelum Basin. Theor Appl Climatol 143:989–1004. https://doi.org/10.1007/s00704-020-03471-2

Valipour M, Banihabib ME, Behbahani SMR (2013) Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. J Hydrol 476:433–441. https://doi.org/10.1016/j.jhydrol.2012.11.017

Wang S, Kang S, Zhang L, Li F (2008) Modelling hydrological response to different land-use and climate change scenarios in the Zamu River basin of northwest China. Hydrol Process 22:2502–2510. https://doi.org/10.1002/hyp.6846

Wang J, Zhang L, Zhang W, Wang X (2019) Reliable model of reservoir water quality prediction based on improved ARIMA method. Environ Eng Sci 36:1041–1048. https://doi.org/10.1089/ees.2018.0279

Xu J, Chen Y, Li W et al (2014) Integrating wavelet analysis and BPANN to simulate the annual runoff with regional climate change : a case study of Yarkand River. Northwest China 28:2523–2537. https://doi.org/10.1007/s11269-014-0625-z

Yeh H-F, Hsu H-L (2019) Stochastic model for drought forecasting in the southern Taiwan basin. Water 11:2041. https://doi.org/10.3390/w11102041

Zaman M, Naveed Anjum M, Usman M et al (2018) Enumerating the effects of climate change on water resources using GCM scenarios at the Xin'anjiang watershed. China Water 10:1296. https://doi.org/10.3390/w10101296

Zaman M, Ahmad I, Usman M et al (2020) Event-based time distribution patterns, return levels, and their trends of extreme precipitation across Indus Basin. Water 12:3373. https://doi.org/10.3390/w12123373

Zamani H, Phillip J, Abudu S (2017) Developing an intelligent expert system for streamflow prediction, integrated in a dynamic decision support system for managing multiple reservoirs : a case study. Expert Syst Appl 83:145–163. https://doi.org/10.1016/j.eswa.2017.04.039