M.Sc. Thesis

FORECASTING OF HYDRO-METEROLOGICAL TIME-SERIES USING ARIMA MODEL FOR RESERVOIR OPERATION

Advisor

Dr. Ijaz Ahmad

Submitted By:

HAMMAD UR REHMAN
2016-MS-WRE-13

CENTRE OF EXCELLENCE IN WATER RESOURCES ENGINEERING
University of Engineering and Technology, Lahore, Pakistan

2019

ABSTRACT

Pakistan is an agricultural country and about 70% of its populations is directly linked with agriculture sector. This sector is also considered one of the biggest contributors in the country's Gross Domestic Product (GDP). The water storage reservoirs are considered essential for improving the production of agricultural sector to meet the ever-increasing food and fibre requirements. Traditionally, historical hydro-meteorological timeseries data generally used in conducting the reservoir operation studies in Pakistan to estimate the future water availability for various purposes such as irrigation, industry, domestic and hydropower etc. which leads to many issues such as inaccurate estimation of water availability, water shortage and excess periods. Therefore, this study aims to use the forecasted hydro-meteorological timeseries data for conducting the reservoir operation study at Mangla reservoir.

Many statistical models such as auto regressive (AR), auto regressive integrated moving average (ARIMA), artificial neural networks (ANN), etc. are being used in many studies round the world to forecast the hydrometeorological timeseries data. ARIMA model is considered one of the most suitable models for linear and seasonal forecasting of timeseries because it uses the simple linear regression model for forecasting. Hence, ARIMA model was used in this study to forecast the hydrometeorological timeseries data, i.e., inflows, precipitation and evaporation to estimate the future water shortage and excess periods. Before applying the ARIMA model, stationarity of hydrometeorological timeseries data was checked. After this, ACF and PACF of timeseries were determined to determine the "*p*" and "*q*" parameters of the ARIMA model. The best fitted structure of ARIMA model was used to forecast the hydrometeorological timeseries. The calibration and validation of

ARIMA model were performed by evaluating the R^2, MAE and RMSE. Finally, the future predicted hydrometeorological timeseries data were used in the reservoir operation to determine the water shortage and excess periods.

The seasonal ARIMA structure of $(1,0,0)(2,1,2)_{12}$ was found best fitted for the inflow timeseries during model calibration and validation. Whereas, ARIMA structures of (14,1,15) and (9,1,19) were considered for forecasting the precipitation and evaporation timeseries. These forecasted hydrometeorological timeseries were used in the reservoir operation for the period of 2016-2030. The R^2 values of inflows, precipitation and evaporation timeseries were found 0.85, 0.88 and 0.83 respectively. The inflows of Mangla reservoir have seasonal effect more prominent compared to climatic time-series of evaporation and precipitation, whereas precipitation timeseries of Mangla reservoir has many steep peaks. The variations in the precipitation timeseries was found less smooth than the inflows timeseries. These forecasted hydrometeorological timeseries data were used in conducting the reservoir operation and found an average water shortage of 14% during 2016-2030. It is believed that the results of present study may guide the reservoir operators and managers to predict the future uncertainties in hydrometeorological timeseries data.

ACKNOWLEDGMENT

All the admires are for **ALMIGHTY ALLAH (The Most Beneficent and Merciful)**, Who is creator of the world. I pay my humble gratitude to the **HOLY PROPHET HAZRAT MUHAMMAD (Peace Be Upon Him)**, Who is forever a model of gratitude and minaret of knowledge for humanity.

I am thankful to Prof. Dr. Noor Muhammad Khan (Director), Centre of Excellence in Water Resources Engineering, UET Lahore for providing an opportunity to conduct my research work in a supportive environment and I learnt a lot from it.

I am cordially thankful to my research supervisor Dr. Ijaz Ahmad. His valuable suggestions guided me to handle the difficult situations during my research work. His guided path leads me to complete the research work of this study. I would also like to acknowledge the other faculty members of CEWRE, Dr. Ghulam Nabi, Dr. M. Kaleem Sarwar, Dr. Muhammad Waseem and Dr. Muhammad Masood for their support and cooperation throughout my research period.

I am also thankful to all my family members who always supported me with love and pray for my success. It is all his encouragement and support which made me able to complete this research.

Engr. Hammad Ur Rehman

TABLE OF CONTENTS

Table of Contents (Continued)

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBRIVATIONS

ACF	Auto Correlation Function
ANFIS	Adaptive Nuer Fuzzy Interference
ANN	Artificial Neural Network
AR	Auto Regressive
ARIMA	Auto Regressive Integrated Moving Average
BPN	Back propagation network
EMD	Empirical Mode Decomposition
GCM	General Circulation Models
GP	Genetic Programming
IRSA	Indus River System Authority
MAE	Mean Absolute Error
MAPE	Mean Absolute Percent Error
MATLAB	Matrix Laboratory
MLR	Multiple Linear Regression
NSEC	Nash-Sutcliffe Efficiency Coefficient
NWP	Numerical Weather Prediction
PACF	Partial Auto Correlation Function
RMSE	Root Mean Squared Error
SARIMA	Seasonal Auto Regressive Integrated Moving Average
WANN	Wavelet Analogue Model
WMAM	Weighted Mean Analogue Method

DEDICATION

I dedicate this thesis to my beloved parents ,respected teachers and my best friend

Basharat Ali Khan.

Chapter I
INTRODUCTION

1.1 GENERAL

In ancient times, human civilization builds the towns and cities at places where water is easily accessible for survival. However, water requirements increase day by day due to rapid increase in population. To resolve this issue, artificial water conservation structures, i.e., dams and reservoirs are often built to fulfil the needs of human activities. These reservoirs are used for different purposes such as irrigation, recreation, ground water recharging and hydropower, etc. (Sarwar, 2013). The economy of many countries in the world has based on these reservoirs especially in arid and semi-arid regions. In arid and semi-arid regions, water availability is decreasing continuously to meet the ever-increasing agricultural demands. Pakistan lies in arid to semi-arid region with average annual rainfall of less than 240 mm (Ahmed et al., 2007). Water requirements related with the all stakeholders; especially, domestic, agricultural and industrial, etc (Majeed and Piracha, 2011).

According to water conservation policy of Pakistan, agricultural demands are considered most important as this sector contributes 22% in the country’s gross domestic product (GDP). The water requirements to meet the demands of agricultural sector depend on Indus basin irrigation system (IBIS). The Indus River System Authority (IRSA) formulated the water policy of Pakistan. In this system the major part of water demands is fulfilled by two reservoirs, i.e., Mangla and Tarbela. The primary purpose of these reservoirs is to release water for irrigation purpose; however, hydropower is also generated from these reservoirs as by-product as all the water released for irrigations passes through the turbines for the production of hydropower (Bogacki and Ismail, 2016).

In most of the cases, available water resources are limited and decreasing at alarming rate to fulfil the water demands of various sectors, i.e., domestic, industry, agriculture, etc. Therefore, efficient use of water is necessary to meet these ever-increasing demands (Lo, 2003). Better understanding of variations in the hydrometeorological time series data can play an important role in performing the efficient reservoir operation. The phenomena of global warming and climate change is effecting the trends of hydrometeorological time series data, i.e., streamflow, precipitation and evaporation (Musa, 2013). The variations in streamflow are very complex in nature. Different models were used to determine these variations in previous studies. The forecasted models are most important for water management to fulfil the needs of crops and mitigation of floods. The forecasted models are used to predict the future water availabilities and to extend the data trends and complete the missing values (Zamani Sabzi et al., 2017). In case of reservoir operation, the forecasted 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, because streamflow had the major input into reservoirs. The design of water resources projects is mainly dependent on the stream flows and its duration. Therefore, researchers have more interest in streamflow’s forecasting. Mostly rainfall-runoff models, GCM models and statistical model were used for the prediction of stream flows in several previous studies (Bahremand and de Smedt, 2010; Xu et al., 2014; Zhang et al., 2011).

The efficient management of available water resources required better reservoir operation policies. Therefore, the uncertainties and fluctuations in the hydrometeorological data should be incorporated in the reservoir operation studies. In the last few decades, ARIMA model was extensively used to forecast the water quality parameters, rainfall, evaporation, wind speeds, runoff and other hydrological timeseries. Mostly probability distributions are used for the assessment of extreme events of future (Hamidi and Telvari, 2018).

1.2 PROBLEM STATEMENT

Pakistan is an agricultural country. Most of its land is irrigated with the canal irrigation systems. Two major reservoirs (Mangla and Tarbela) are considered as the backbone of this irrigation system. A reservoir is a man-made structure built to store and release the water on timely basis for various purposes, i.e., irrigation, industry, hydropower, domestic use, etc. Traditionally, guided curves are generally used to operate these reservoirs. These guided curves describe the relationship between the stage versus releases based on historical data which has many issues such as inaccurate water estimation, water shortage and excess periods. Therefore, this study aims to forecast the hydrometeorological data of the reservoir and used in reservoir operation model to develop future water allocation scenarios. The prediction of water availability is helpful to regulate the flows efficiently. The reservoir should be empty or partially empty before the extreme inflow events. Otherwise water storage level exceeds the safe limits and causes floods. To avoid these situations a reliable forecasting if hydrometeorological data is essential. Now a days, many forecasting techniques are being used such as ARIMA, ANN, frequency analysis, etc. The forecasted hydrometeorological timeseries are considered very helpful in managing the reservoir operation and satisfy the water demands of various sectors. 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. Whereas, ANN is considered suitable for nonlinear timeseries. Therefore, in this study hydrometeorological timeseries is forecasted by using ARIMA model. The forecasted timeseries are used in reservoir operation to determine the fluctuations in the reservoir storage.

1.3 STUDY AREA

Mangla reservoir is located in Mirpur district of Azad Kashmir. It is a large water storage reservoir on the Jhelum river. The Jhelum river originates from the snow-covered Himalayan and the Karakoram ranges. The Mangla reservoir covers a catchment area of 329.7 km^2. The gross storage capacity of reservoir is 9.220 Mm^3. The command area of Mangla reservoir is 60,000 km^2. The temperature variation of air is 18 °C to 43 °C. The maximum elevation of Mangla is 630 m M.S.L. The primary purpose of Mangla reservoir is to supply water for irrigating the agricultural lands. The Mangla reservoir supplies water to the upper and lower Chasma-Jhelum link canal regions. During the period of December to March (Rabi season) the irrigation demands are at peak from for upper Jhelum region. The excess water from lower and upper Jhelum regions irrigates to the lower Indus basin. Hydropower is produced as by product as the water released from the reservoir for any purpose passes through the turbines and its hydropower capacity is 1000 MW (Mega Watt).

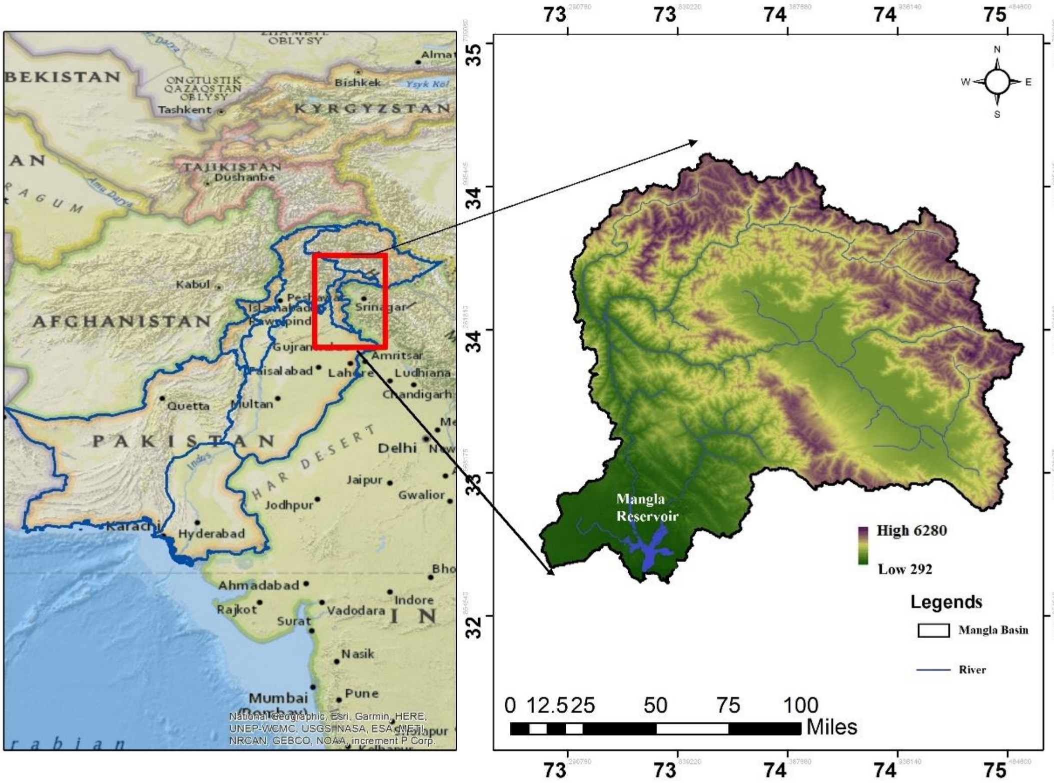


Fig. 0.1 Location of study area

1.4 OBJECTIVE

- Analysis of hydro meteorological timeseries to select the best fit structure of ARIMA model.
- Forecasting of hydro meteorological timeseries to estimate the water availability.
- To perform the reservoir operation to identify the water shortage periods.

1.5 UTILIZATION OF RESEARCH

The forecasting of hydrometeorological timeseries is keen interest of many researchers. Rapid increase in population required better understanding of future scenarios of water availability to fulfil their water requirements. One way to forecast the hydrometeorological data used statistical models. After the forecasting of hydrometeorological, an efficient reservoir operation can be performed.

The proposed research uses the hydrometeorological timeseries data of Mangla reservoir and forecast this to determine the future water availability. The future water availability will decide the water shortage or excess periods. Both conditions are managed by some suitable decisions and action. In many of the advanced regions, reservoir operation is performed on the basis of future water availability. However, there is hardly any research is available which considered the forecasted hydrometeorological data in conducting the reservoir operation studies. Therefore, in this study a statistical model ARIMA is used to forecast the future water availability in Mangla reservoir. The forecasted hydrometeorological timeseries data is used to perform the reservoir operation studies.

Chapter II
LITERATURE REVIEW

2.1 GENERAL

Irrigation water requirements for the agricultural production is largely dependent on the water supplies from western rivers of Indus basin. The flows in these rivers are changing drastically in which the floods and droughts have been often observed. To cope with these variations a reliable information source of hydrometeorological timeseries is compulsory. Therefore, several forecasting models were used in previous studies to predict the future inflows, rainfall and evaporation trends. Forecasting of inflows timeseries has more importance compared to precipitation and evaporation. Inflows forecasting is considered as a key step in the design of water resources projects (dams and reservoirs), maintenance and operation. In previous studies, various statistical models were used to forecast the data such as AR, MLR, ANN, ARIMA, hybrid models, etc. Some uncertainties and errors are always lie in these models. Therefore, after elimination of these uncertainties up to acceptable limit, streamflow timeseries were forecasted (Hamidi and Telvari, 2018).

2.2 ARIMA

Different statistical models ARIMA, AR and ANN were used for forecasting hydrometeorological timeseries data. ARIMA model can make a better prediction since it considers the correlation between pervious timeseries data. Therefore, the irregular component of timeseries can take non-zero autocorrelation. The ARIMA model is usually defined for stationary timeseries. A stationary timeseries is the one whose statistical properties such as mean, variance, autocorrelation etc. are all constant over time. If timeseries is not stationary, then it has to be transformed into a stationary timeseries using any appropriate transformation technique, for instance, differencing technique which determines the differencing order "*d*" to makes it stationary. Differencing removes the trend component of a timeseries leaving the irregular component (Kasyoki, 2015).

Álvarez-díaz (2015) used the ARIMA model in forecasting the hydrometeorological timeseries data for developing the future water management practices. The results of this study showed that ARIMA model can be a helpful tool for forecasting the linear timeseries data as it is consisted of auto-regressive and moving average terms. It was also concluded that the ARIMA model is suitable for linear and seasonal timeseries.

Musa (2013) had used the three statistical models AR, ARMA and ARIMA for the streamflow forecasting of Shiroro river located in the Nigeria by using the MATLAB. The results of this study showed that the ARIMA model performed better as compared to the AR and ARMA models. The forecasted streamflow of ARIMA model were closer to the observed flows. His study concluded that the streamflow of Shiroro river was sufficient enough for optimum usage of irrigation if properly regulated.

Singh and Singh (2011) studied the stream flows of Kangsabati Dam project in India. Seventeen years (1983-2003) stream flows data were used for forecasting. Two forecasted softwares ARIMA and X-12 ARIMA were used for modelling. The timeseries were converted into stationary after first differencing. The results showed that seasonal ARIMA $(2,1,1)(2,1,2)_{12}$ was found best fitted model structure. The observed streamflow was compared with simulated stream flows. The model results are presented in the Table 2.1.

Table 2.1 Comparisons of ARIMA and ARIMA-X-12

MODEL STATISTICS	ARIMA	ARIMA-X-12
R SQUARE	0.753	0.825
RMSE	42.837	30.625
MAE	23.376	17.248
AIC	1822.76	1804.8
BIC	1846.42	1830.30

Hamidi and Telvari (2018) used the 48 years of mean annual discharge of Karkheh river to develop three ARIMA model structures, i.e., (0,1,1), (1,1,1) and (4,1,1). The (4,1,1) structure was found best fitted and used for forecasting the inflows for the period of 2006-2015 at Karkheh dam in the west of Iran. This result showed that the model forecasted values and the observed values were consistent each other and ARIMA (4,1,1) structure was best fitted and had minimum value of Akaike Information Criterion (AIC). Moreover, Yalcin and Tigrek (2017) studied the water requirements from the Garzan reservoir located in Turkey. ARIMA model used to evaluate the water requirements. This study showed that ARIMA model forecasted values were very helpful tool to evaluate the future water requirement.

Reza Ghanbarpour et al. (2010) studied the Sangsoorakh basin (Iran) located in the southwest sub basin of karkheh. Its inflows were forecasted with ARIMA and de-seasonalised ARIMA. The results were evaluated for the period of 1999-2004 and compared with the observed inflows. The values of MAE, RMSE, NSE and R^2 were found as 1.41, 1.82, 0.49 and 0.52, respectively. This study showed that ARIMA model performed better than the de-seasonalised ARIMA. In another study, Kasyoki (2015) discussed the methodology and performance of ARIMA model. His study concluded that the ARIMA model performed better on linear timeseries. For this purpose, different software could be used especially R package.

Salami and Salami, (2018) studied the Kainji reservoir located in the Niger river of west Africa. The Kainji reservoir used for irrigation, flood control and hydropower. The inflows Kainji reservoir (West Africa) of 27 years were analysed and developed the seasonal ARIMA model for inflow forecasting. Three seasonal ARIMA models were selected for forecasting. But the best fitted model was ARIMA $(2,1,1)(2,1,2)_{12}$. The simulated inflows were used to develop the elevation rule curves and compared with the levels of reservoir by using the satellite images. The R^2 value of actual and predicted data is 0.98. This study showed that the ARIMA model could be used for forecasting of reservoir inflows.

Valipour et al. (2013) developed three different statistical models AR, ARMA and ARIMA for forecasting the inflows of Dez reservoir of Iran. Monthly inflows of 42 years (1960-2002) were used for the model calibration. After this the next five-year (2002-2007) inflows were forecasted by using the AR, ARMA and ARIMA models. The forecasted values compared with the observed values and concluded that the inflows were predicted by ARIMA model more accurate than AR and ARMA.

A reservoir operation model needs extensive hydrometeorological information, i.e., past inflows, current inflow, future predicted inflows into the reservoir in addition to other climatic timeseries data. All these datasets are used for developing the reservoir operating policies and reservoir optimization. The downstream water requirements have to be fulfilled by applying these policies (Martins et al., 2011).

Wang et al. (2019) used the ARIMA model to predict the water quality of Shihe reservoir of China. Water pollution is the major source of environment pollution. ARIMA and Holt winters seasonal model was used for the future prediction of water quality parameters. Total nitrogen and Phosphorus amount present in the water was used as the water quality prediction parameters. This study concluded that ARIMA model with the optimization of Holt winter seasonal model performed better and its accuracy reached up to 97.5%.

Adenan and Noorani (2013) used ARIMA and SVM models for forecasting the streamflow of Tanjung and Tualang river basins of china. The simulated model results were compared with the observed flows. The RMSE values of ARIMA and SVM models were found as 21.78 and 16.71, respectively. Therefore, it was evident that ARIMA model performed better than the SVM model.

Lee and Ko (2011) had studied the timeseries of short-term load of Taipower company by using the ARIMA model. These timeseries were divided into subseries. After this, the subseries was used for forecasting with suitable model structure of ARIMA. The lifting scheme in ARIMA model for forecasting the short-term load of Taipower company was used. One-day ahead load forecasting attained by using the embedded model of ARIMA and lifting scheme. The results showed that the ARIMA model approach is better than Back Propagation network (BPN) algorithm.

Graham et al. (2019) used ARIMA model to forecast the global temperature and rainfall trends. The model simulated results were compared with the observed data and concluded that the ARIMA model performed better than other models. Because its accuracy could be achieved up to 92%.

Balasmeh (2019) forecasted the precipitation timeseries data of five-gauge stations in the eastern part of Jordan by using ARIMA model. The ARIMA model was used to forecast the precipitation data up to year of 2026. Three (03) different type of ARIMA model structure was found best fitted. For monthly data ARIMA (3,1,3) was found best fitted for forecasting. For average annual data ARIMA (4,1,3) was found best fitted and for seasonal timeseries ARIMA (4,2,4) were best fitted. The best-fitted ARIMA models, validated with 10 years of data (2007–2016) were used for predicting the precipitation up to year 2026. The future trend shows that the high level (heavy rain) is decreasing at all stations and low level (normal rain) is increasing, except in the month of December, which showed an increasing trend. This observed pattern warrants effective water management strategies for already water-stressed area of Wadi Shueib catchment of Jordan.

Zamani et al. (2017) studied the streamflow forecasting by using the four statistical models, i.e., ARIMA, ANN, Hybrid ARIMA-ANN and ANFIS at Butte and Caballo reservoirs. ARIMA model was found most suitable for the linear timeseries whereas ANN was found good for nonlinear timeseries forecasting. This study showed that ARIMA model performed better than other models and its R^2 value was found 0.97. This study provides the conceptual procedures of non-seasonal ARIMA model, and since the model is univariate, it demonstrates a strongly reliable inflow prediction when existing information is limited to streamflow data as a predictor.

Kavasseri and Seetharaman (2009) used the f-ARIMA model for forecasting of four wind generation in North Dakota. A day ahead wind speed was forecasted and correlated with the power generation curve of an operational turbine. A day ahead of wind power production was forecasted by function ARIMA model. The results of the function ARIMA showed that the significance improvement of forecasted values. Utilizing these forecasted values, they could take better decision for the distribution of energy and load.

2.3 OTHER STATISTICAL MODELS

General Circulation Models (GCMs) have also extensively used for forecasting the hydrometeorological timeseries. However, these have many problems such as mismatch scale, vegetation cover and time span issues.

Tian et al. (2018) used the mesoscale Numerical Weather Prediction (NWP) model for forecasting the rainfall. The catchment Zijingguan of Northern China was selected as a case study. Different heights of radars were used for this purpose. Three-dimensional (3D) variational data assimilation (3-DVar) technique adopted to assimilate the Doppler radar data. Radar reflectivity and radial velocity were assimilated separately and jointly. Each type of radar data was divided into seven data sets according to the observation heights: (1) 2000 m; (7) all heights. Results showed that the assimilation of radar reflectivity gave better results. The accuracy of the predicted rainfall deteriorated as the rise of the observation height of the assimilated radar data. In this study provide a reference for efficient utilisation of the Doppler radar data in numerical rainfall prediction for hydrological use.

Lee and Tong (2011) used the ANN, ARIMA and combined ARIMA-GP (Genetic programming) models for forecasting of China's energy consumption and Canadian lynx data. Mostly nonlinear data forecasting used ANN algorithms. However, nonlinear timeseries data forecasting could be achieved by combining the ARIMA and genetic programming (GP) models. Model forecasted values compared with the original values and determined the statistical parameters RMSE, MAPE and MAE and have concluded that the hybrid ARIMA-GP model had efficiently forecasted the non-linear timeseries data.

Accuracy of the forecasted timeseries was the main task of these models. But it is not that simple especially for different types of timeseries models. The ANN model considered better for nonlinear timeseries; whereas, ARIMA model is suitable for linear timeseries. Sometime composite or hybrid model is used for complex timeseries. The timeseries were decomposed into linear and nonlinear. Then hybrid model of ARIMA and ANN which easily forecasted the timeseries values of linear and nonlinear data. The hybrid model was used for forecasting the three sets of real known data of sunspots and improved the accuracy of forecasting values (Khashei and Bijari, 2011).

Forecasted streamflow affected the decision-making of reservoir operation. Because forecasted streamflow gave the useful information for water availability in subsequent periods. The water release decision depended on the forecasted streamflow and reservoir storage. If storage capacity of reservoir had sufficient enough then water released from the reservoir otherwise stream-flows regulated in the short time framework. Release pattern was affected by the forecasted inflows uncertainty. To improve the forecasting different composite models such single spectrum analysis (SSA) and ARIMA is used. For this purpose, the timeseries were divided into subseries. And this subseries was analysed by suitable model. After the forecasting, these subseries were combined with each other (Zhang et al., 2011).

In another study, pair model of ARIMA and EEMD were used for forecasting the hydrological timeseries data of Biliuhe reservoir, Dahuofang reservoir and Mopanshan reservoir in China. The Annual inflows timeseries data decomposed into the finite and intrinsic mode functions (IMF). After this the annual inflows were forecasted by using the EEMD-ARIMA model. The forecasted values were compared by different statistical measures, e.g., RMSE, MAPE, R and NSEC. The results showed that the hybrid ARIMA model should be used for forecasting the inflows of reservoir (Lohani et al., 2014).

Khanal et al. (2014) investigated the Kulekhani catchment of NEPAL for forecasting the hydro-meteorological time-series data by using HBV (hybrid varied) model. The seven consecutive days forecasted inflows of reservoirs and meteorological data were used for reservoir operation. The advanced seven days of energy production were estimated. This future production is very helpful to fulfil the demands of energy sector. Another benefit was that the advance seven days levels were also determined.

Wang and Qiu (2018) studied to decompose the timeseries by using Empirical Mode Decomposition (EMD) and Ensemble Empirical Mode Decomposition (EEMD methods. The hybrid model of EEMD/EMD ARIMA was used to forecast the long-term timeseries of upper stream of Yellow river basin. The observed data was collected six year of daily inflows at the Tangnaihai station. After this, timeseries were decomposed into subseries for further analysis of model fittings of hybrid model.

Landeras et al. (2009) analysed the weekly evapotranspiration on the Álava in the Basque (northern Spain) by using the ARIMA and ANN models. They compared the one week ahead model results with observed values by finding the MAD (mean absolute difference) and RMSD (root mean square difference). The result values showed that ANNs model performs better during the months (May–August), and ARIMA models forecasted better from September to November. There were no significant differences between the models in January and December.

Jothiprakash and Magar (2012) applied ARIMA model for forecasting the inflows of koyna reservoir, India. The computation time of model was fourteen years and the model results compared with different statistical tests, e.g., Nash–Sutcliffe efficiency (NE), Root mean square error (RMSE), Akaike information criterion (AIC) and Bayesian information criterion (BIC). The results of these tests showed that the ARIMA model performs better in forecasting compared to AR and MLR models.

Zhao et al. (2012) forecasted the inflows of Danjiangkou reservoir China. The stream flow forecasting uncertainty and stream flow forecasting limitations were the major constraints in the reservoir operation. The stream flow forecasting had two major issues, i.e., forecast uncertainty and limited forecasted period. These uncertainties increase with the increase of time period limit.

Somchit et al. (2018) used wavelet analogue model (WANN) and weighted mean analogue method (WMAM) for the inflow forecasting of Sirikit Dam located in Thailand. These forecasted inflows were very helpful tool for reservoir operators. In this study these forecasted inflows were managed on the basis of downstream water requirements and an efficient reservoir operation was performed.

Schwanenberg et al. (2015) used the hydrological modelling and data assembling techniques to forecast the inflow timeseries. The study has been done on the Tres Marias hydropower reservoir. Streamflow of 15 days were forecasted for the analysis. The forecasted values were validated on the basis of Nash-Sutcliffe efficiency. In the first stage, inflows were forecasted, and second stage take the decision on the basis of forecasted flows. Third stage was to manage the constraints of downstream demands.

Nohara and Saito (2018) used the Monte Carlo simulations method. In this method multi-purpose reservoir releases the water before flood events. The flood events were predicted and integrated in the reservoir operation. It was applied on the Nagayasuguchi reservoir in the Naka river basin, located in the southwest part of Japan. The case study revealed that this method provides the quantitative information about taking the decision of water release for reservoir managers. Therefore, ensembled hydrological prediction reservoir operation is very helpful and beneficial.

Boucher and Ramos (2018) described the ensemble prediction of streamflow forecasting in reservoir operation. A reservoir is useful many purposes for flood management, hydropower generation and water conservation. To improve their efficiency requires different span forecasting. A flood management protection requires one-week information forecasted inflows. Whereas, optimization, managing and planning of reservoir operation requires seasonal forecasting.

Hejazi and Cai (2011) studied the 79 reservoir operation systems in California, the downstream water demands of reservoir were also determined. The water releases from the reservoirs according to the water demands. After the development of this system the water shortage could be minimum. This study showed that combined reservoir operation reduced the water shortages more.

2.4 CONCLUDING REMARKS

The following concluding remarks were drawn from the extensive literature review presented in this chapter.

- Timeseries should be stationary for the analysis of ARIMA model. Otherwise convert into stationary by differencing or log transformation.
- ARIMA model was suitable for linear timeseries because it resembled with the simple regression.
- All type of hydrometeorological timeseries data can be forecasted by using the ARIMA model.
- The reservoir operation was performed by using the forecasted hydrometeorological data.
- ARIMA model was performed better than ANN in linear timeseries.

Chapter III
METHODOLOGY

3.1 GENERAL

This chapter presents the basic concept of timeseries and its types. To Forecast hydrometeorological timeseries ARIMA model was used and its methodology is presented in this chapter. Firstly, hydrometeorological timeseries should be stationery otherwise required to be converted into stationary. After this, model identification had been performed and estimated the parameters p, d, and q. These estimated parameters were used to forecast the timeseries. The forecasted timeseries incorporated into reservoir operation model. The future water availability scenarios were determined.

3.2 TIMESERIES

A timeseries is a set of observations with a corresponding time span. Time interval may be daily, weekly, monthly or annual. Timeseries forecasting is now a very important research area owing to the importance of prediction in various applications. For instance, forecasting internet traffic helps the service providers to enhance their services. Forecasting future water availability scenarios helps to improve the agricultural sector. Forecasting disasters aid in taking necessary precautions and helps mankind to be prepared. Forecasting financial data helps investors to invest safely in the market. However, timeseries data do not always have the same characteristics. Some timeseries are seasonal; for example, road traffic is high at some particular times in the day, climate variations repeat according to the seasons, etc. Some other timeseries are nonseasonal, such as financial and stock market data. Some timeseries are highly volatile, such as windspeed data, and some are less volatile, such as global temperature and annual rainfall. Some data are almost linear in nature, such as the growth of an animal, plant, or human being. Various techniques are available in the literature to forecast the timeseries data. In this study, ARIMA model in being used to forecast the hydrometeorological timeseries data for efficient reservoir operation.

3.3 ARIMA

ARIMA model used in the field on hydrology and especially in river flow forecasting. It is suitable for linear stationary timeseries otherwise timeseries is required to be converted into stationary (Babu and Reddy, 2014). ARIMA model consisted of the three terms. The first is the autoregressive terms, i.e., “AR” terms. The second is the moving average terms, i.e., “MA” terms and third is order of differencing or integrated terms “I”.

ARIMA model performs the following steps in forecasting the timeseries data and are discussed in subsequent sections.

1. Data Analysis
2. Model formulation
3. Parameter Estimation
4. Diagnostic checking
5. Forecasting

3.3.1 Data Analysis

Timeseries data analysis was performed to check its trend, extreme and mean monthly values. These factors are also included into the simulated timeseries. Basically, timeseries analysis is the understanding of the data either it is reliable or not. In the ARIMA model, the stationary condition of timeseries is necessary before forecasting the hydrometeorological data.

3.3.1.1 Data Stationarity

Data stationarity is an essential property of all timeseries of ARIMA otherwise it is required to be converted into stationary timeseries. It means that the ARIMA model only works on the stationary timeseries. The timeseries is considered stationary if its mean E(at), variance Var(at) and covariance Cov (at,at-1) are constant.

$$E(a_t) = \mu_a \text{ for all } t \tag{3.1}$$

$$Var(a_t) = E\left[(a_t - \mu_a)^2 \right] = \sigma^2{}_a \tag{3.2}$$

$$Cov\ (a_t, a_{t-k}) = \gamma_k \text{ for all "}t\text{"} \tag{3.3}$$

The white noise series "ε_t" satisfied the stationary condition;

$$E(\varepsilon_t) = 0 \tag{3.4}$$

$$Var(\varepsilon_t) = \sigma^2 \tag{3.5}$$

$$Cov\ (\varepsilon_t.\varepsilon_{t-1}) = 0 \text{ for all } s \neq 0 \tag{3.6}$$

The random disturbance term is typically assumed to be "white noise"; i.e., it is identically and independently distributed with a mean of "0" and a common variance across all observations.

A stationary timeseries is correlated with the backshift timeseries etc.

$$\begin{aligned} a_t &= \varphi_1 a_{t-1} + \varepsilon_t \\ &= \varphi_1(\varphi_1 a_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_1^2 \varepsilon_{t-2} + \varphi_1^3 \varepsilon_{t-3} + ... + \varphi_1^t a_o \end{aligned} \tag{3.7}$$

Without loss of generality, assume that $y_o = 0$. Then $E(y_t)=0$

Assuming that "*t*" is large, i.e., the process started a long time ago, then

$$Var(a_t) = \frac{\sigma^2}{(1-\varphi_1^2)}, \text{ provided that } |\varphi_1| < 1 \tag{3.8}$$

It can also be shown that provided that the same condition is satisfied if;

$$\text{cov}(a_t a_{t-s}) = \frac{\varphi_1^s \sigma^2}{(1-\varphi_1^2)} = \varphi_1^s \text{ var}(a_t) \tag{3.9}$$

Special case: $\emptyset_1 = 1$

$$a_t = a_{t-1} + \varepsilon_t . \tag{3.10}$$

It is a "random walk" process, now,

$$a_t = \sum_{j=0}^{t-1} \varepsilon_{t-j} . \tag{3.11}$$

Then the all- timeseries

$$E(a_t) = 0 \tag{3.12}$$

$$Var(a_t) = t\sigma_z \tag{3.13}$$

$$Cov(a_t, a_{t-s}) = (t-s)\sigma_z{}^2 \tag{3.14}$$

If the auto regressive terms AR (2) then

$$a_t = \varphi_1 a_{t-1} + \varphi_1 a_{t-1} + \varepsilon_t \tag{3.15}$$

If α_t is a stationary term, then

$$\varphi_1 + \varphi_2 < 1, \ \varphi_2 - \varphi_1 < 1 \text{ and } |\varphi_2| \ < 1 \tag{3.16}$$

"*AR*" term is stationary if it satisfied the above conditions.

Whereas the "*MA*" terms without drift.

$$a_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} \tag{3.17}$$

For stationarity, its mean, variance and covariance should be

$$E(a_t)=0 \tag{3.18}$$

$$Var(a_t)=\sigma^2(1+\theta_1^2) \tag{3.19}$$

$$Cov(a_t, a_{t-s}) = \begin{Bmatrix} -\theta_1\sigma^2 \ \ if \ \ s=1 \\ otherwise \ \ 0 \end{Bmatrix} \tag{3.20}$$

For an MA (moving average two terms) process

$$a_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} \tag{3.21}$$

Whereas its mean, variance and covariance are

$$E(a_t)=0 \tag{3.22}$$

$$Var(a_t)=\sigma^2(1+\theta_1^2+\theta_2^2) \tag{3.23}$$

$$Cov(a_t, a_{t-s}) = \begin{Bmatrix} -\theta_1\sigma^2 \ if \ s=1 \\ -\theta_2\sigma^2 \ \ if \ s=2 \\ otherwise \ 0 \end{Bmatrix} \tag{3.24}$$

3.3.1.2 Differencing

If the timeseries is not stationary, then take the differencing and convert into the stationary. For example, a non-stationary timeseries.

$$a_t = a_{t-1} + \varepsilon_t \tag{3.25}$$

Then taking differencing

$$w_t = a_t - a_{t-1} = \varepsilon_t \ \ stationary \tag{3.26}$$

Differencing continue until it converts into stationary, it means 2nd order differencing, 3rd order differencing

First order differencing $\Delta a_t = a_t - a_{t-1}$ (3.27)

Second order differencing

$$\Delta^2 a_t = \Delta(\Delta a_t) = \Delta(a_t - a_{t-1}) = a_t - 2a_{t-1} + a_{t-2} \tag{3.28}$$

3.3.2 Model Formulation

The model identification depended on the three terms "*AR*" (*p*), order of differencing (*d*) and "*MA*" (*q*) terms.

The value "*p*" is identified by autocorrelation function (ACF). If the value of ACF of timeseries is equal to zero or near to zero, "*p*" value is considered satisfied. Similarly, the PACF is used to determine the "*q*" value. "*d*" represented the degree of differencing to make stationary of timeseries.

3.3.2.1 Auto Correlation Function (ACF)

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

$$ACF\ \ r_k = \frac{\sum_{t=1}^{n-k}(a_t - \bar{a})(a_{t-k} - \bar{a})}{\sum_{t=1}^{n}(a_t - \bar{a})^2} \tag{3.29}$$

3.3.2.2 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 from the following formula;

$$r_{kk} = \begin{cases} r_1 & k = 1, \\ \dfrac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} \cdot r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} \cdot r_k} & \text{if } k = 2,3,\ldots \end{cases} \tag{3.30}$$

Whereas

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j} \qquad \text{for j=1, 2, ..., k-1} \tag{3.31}$$

It means that "r_{kj}" represented ACF of timeseries by lag "j" time units. Therefore, the intervening observation between "k" and "j" eliminated. Whereas standard error of PACF.

$$s_{r_{kk}} = \sqrt{\frac{1}{n}} \tag{3.32}$$

3.3.2.3 Seasonality

In some cases, the seasonal factor is more prominent; therefore, seasonal model structure is considered best fitted. In other words, the seasonal means that regular pattern of changes in timeseries that repeats after certain time periods. If seasonal pattern of *max* and *min* value occurs in the same month. Then its value is 12 and structure is represented as

$$ARIMA \quad \underbrace{(p,d,q)}_{Nonseanonal}\underbrace{(P,D,Q)}_{Seasonal} \tag{3.33}$$

3.3.3 Parameter Estimation

The timeseries parameter is estimated by the following method.

- 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 *MA* component. Sometime *max* likelihood function is used to define the probability of observed data.

3.3.4 Forecasting

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

$$\hat{y}_{t+h} = \hat{\delta} + \hat{\varphi}_1 y_{t+h-1} + ...\hat{\varphi}_p y_{t+h-p} + e_{t+h} - \hat{\theta}_1 e_{t+h-1} - .....\hat{\theta}_q e_{t+h-q} \qquad (3.34)$$

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

3.4 RESERVOIR OPERATION

Reservoir is as artificial structure used to regulate the flows of rivers for various purposes, e.g., irrigation use, industrial use, domestic use, flood control etc. To operate the reservoir is a challenging task. Because there are many restrictions on the basis of downstream conditions. Sometime expected inflows are not well determined. To improve this condition prediction of inflows and evaporation are required. Güntner et al., (2004) have proposed following method to conduct the reservoir operation studies.

$$S_t = S_{t-1} + I - O - U_{RL} + (P - E)A_{RL} \qquad (3.35)$$

Whereas

S_t = reservoir storage at timestep t

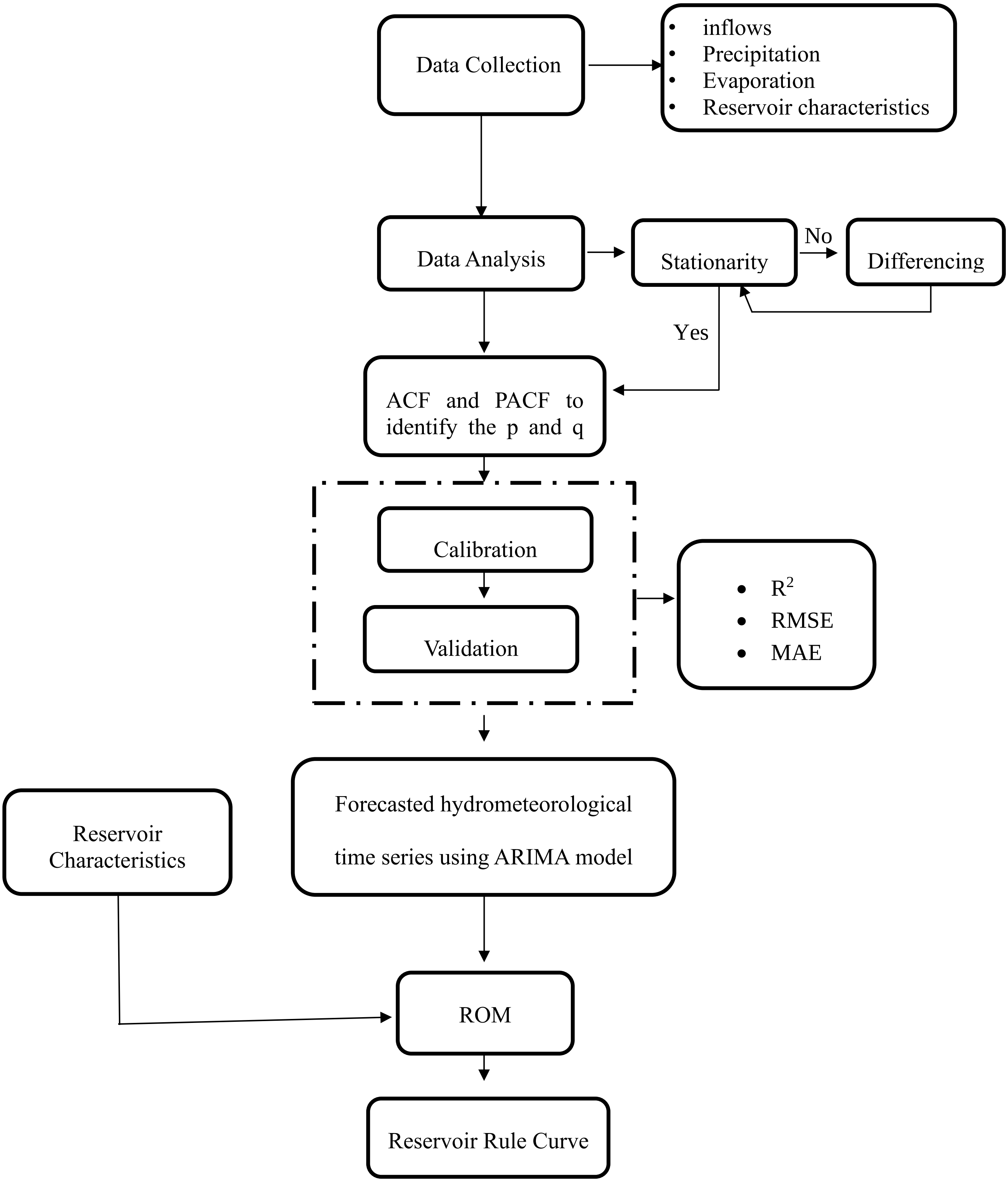
S_{t-1}= reservoir storage at time step t-1

P = precipitation at water surface area RL

E= evaporation at water surface area RL

U_{RL} = water withdrawal

3.5 FLOW CHART



Chapter IV
RESULTS AND DISCUSSIONS

4.1 GENERAL

Pakistan is an agricultural country and its irrigation system relies heavily on availability of fresh surface water resources from the Tarbela and Mangla reservoirs. Pakistan has the world largest continuous irrigation system with two major storages 19 barrages 12 link canals 46 main canals and thousands of hydraulic structures. One of these major storages Mangla reservoir is very important for the irrigation system of Punjab. In this chapter, detailed results of the variations in the hydrometeorological timeseries, calibration and validation, forecasting and reservoir operation are presented in the subsequent sections.

4.2 INFLOWS VARIATION

At Mangle reservoir, inflow peaks lie in the months of May to July, i.e., monsoon season. These flows are stored during the flood season and used in the low flow periods. In the Mangla catchment a major portion of area is covered with the snow. Therefore, the summer seasons snowmelts at higher rate and it contributes larger amount of discharge in Mangla reservoir.

Figure 4.1 showed the mean monthly variations in stream flows. It is clear from Figure 4.1 that the flows were maximum during the monsoon period (Jun-Aug). However, during the dry season, least flows were observed. During these low flow periods the water requirements of downstream were fulfilled by the stored water in the reservoir.

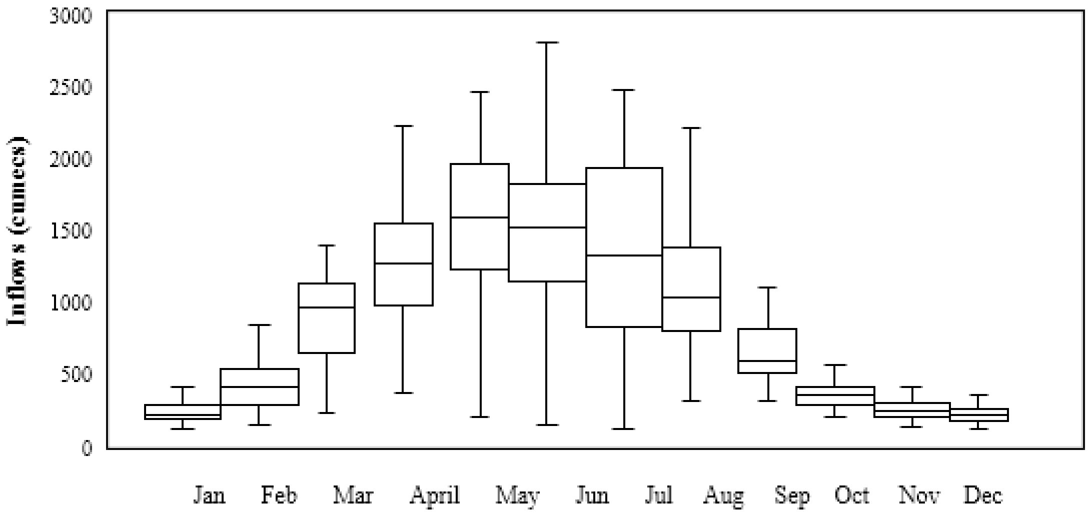


Fig. 4.1 Mean monthly inflows into Mangla reservoir

About 75% of flows were observed in the six months of March to August. Whereas, the 25% flows were observed in the dry season. Figure 4.2 showed the mean annual inflows of Mangla reservoir from 1991-2015. It is cleared from the Figure 4.2 that the mean annul flows were continuously decreasing from 1991-2001. After this the mean annual flows were increased. Highest mean annual stream flows were observed in the years of 1991 and 1996.

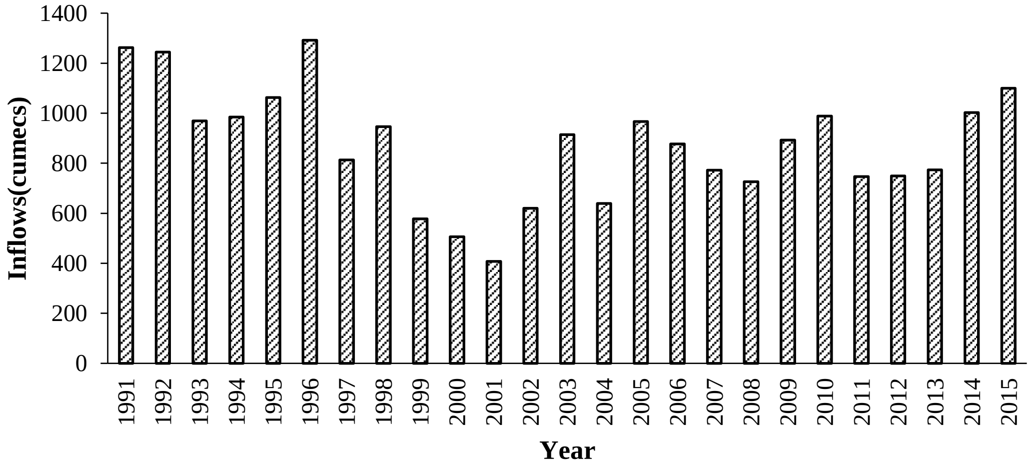


Fig. 4.2 Observed mean annual inflows into Mangla reservoir during the period of 1991-2015

4.3 PRECIPITATION VARIATIONS

In the Figure 4.3 the mean monthly precipitation variation had been shown. It was cleared from the Figure 4.3 that the monsoon period the rainfalls were maximum. Mostly the peak rainfall events were occurred in the month of June or July. Whereas Nov and Dec were the dry months. During this period there was almost no rainfall.

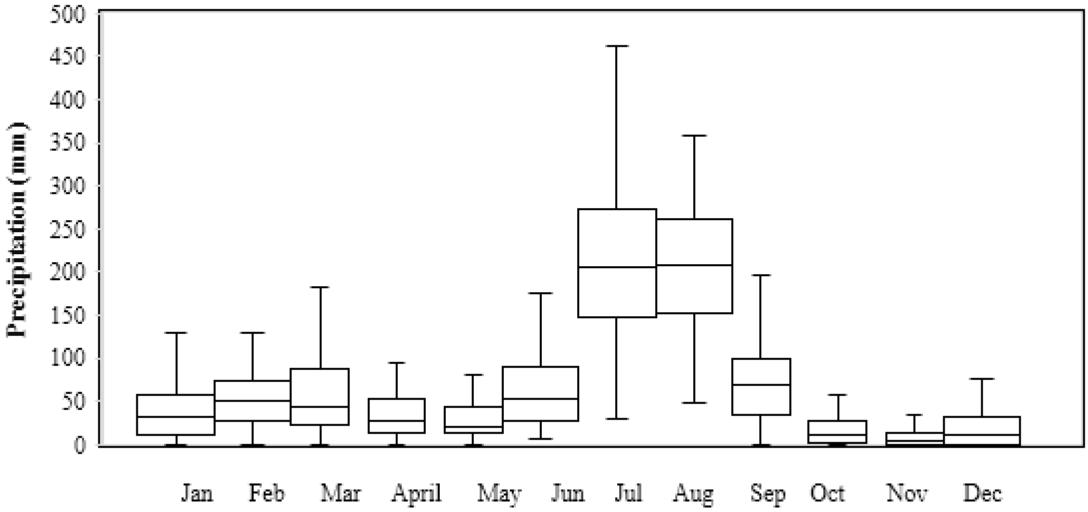


Fig. 4.3 Mean monthly precipitation at Mangla reservoir

The Figure 4.4 were represented the annual precipitaion variation at the Mangla reservoir from 1991- 2015. It is clear from Figure 4.4 that the annul precipitation from 1991-2000 remained almost same And afterwards decreased upto 2012. Maximum precipitation was recorded in the year of 2014.

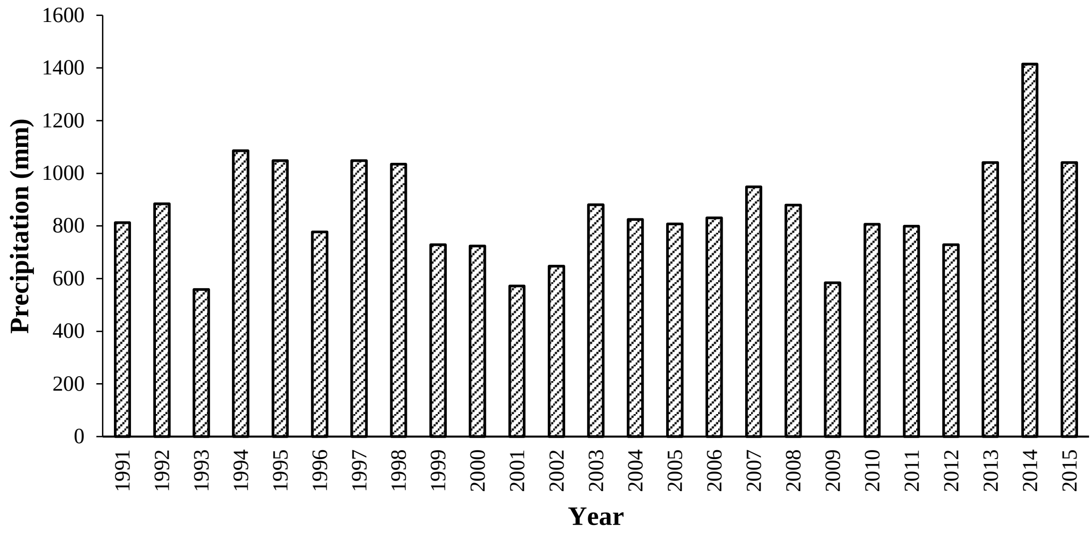


Fig. 4.4 Observed mean annual precipitation at Mangla reservoir during the period of 1991-2015

4.4 EVAPORATION VARAITIONS

The rate of evaporation increases with the increase in temperature. During summer season the rate of evaporation is maximum over Mangla reservoir. Whereas, during winter season, evaporation is minimum because of low temperatures. Figure 4.5 presents the mean monthly evaporation at Mangla reservoir. It is clear from Figure 4.5 that during hot months maximum evaporation was observed. Whereas, during cold months, i.e., Dec and Jan, least evaporation was recorded.

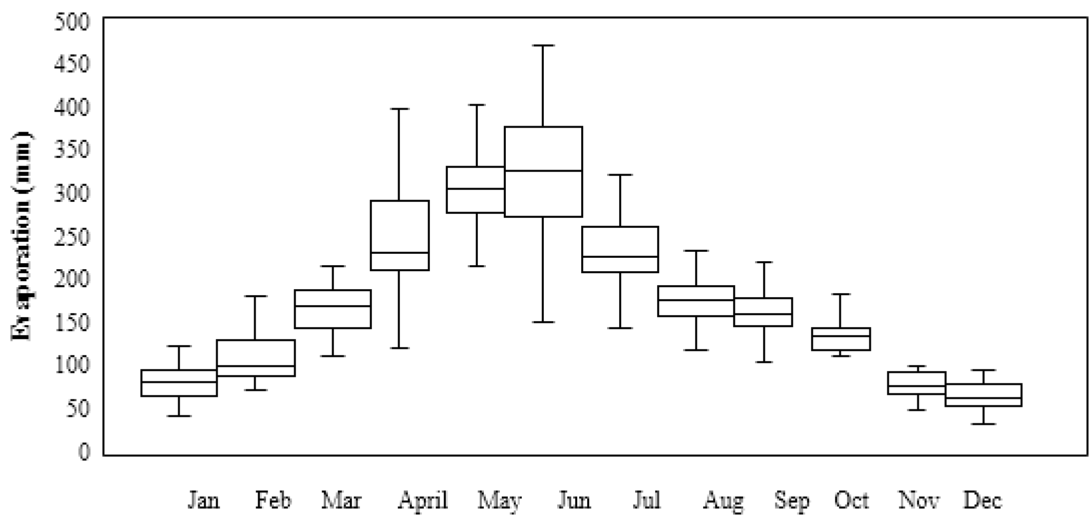


Fig. 4.5 Mean monthly evaporation (mm) at Mangla reservoir

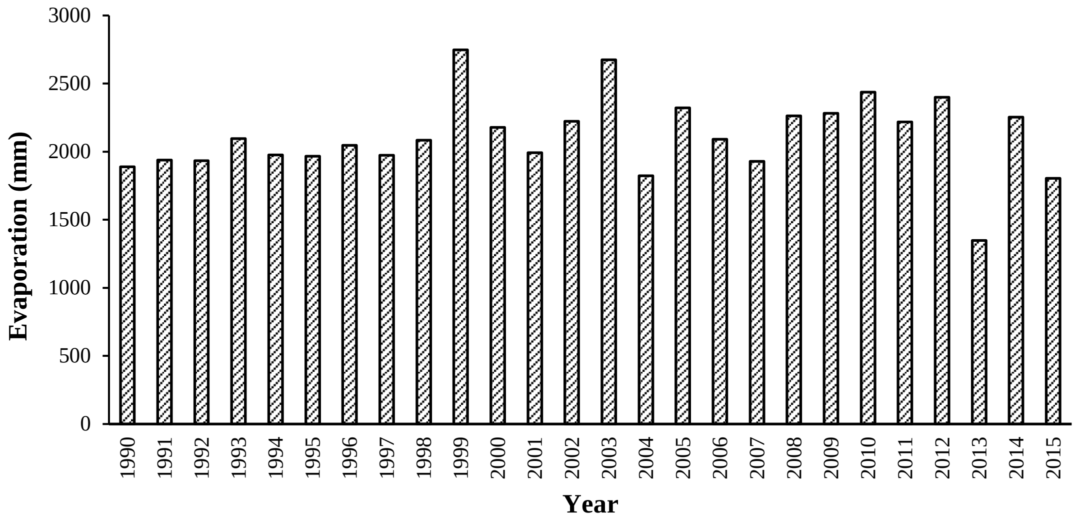


Fig. 4.6 Observed mean annual evaporation (mm) at Mangla reservoir during the period of 1990-2015

In the Figure 4.6 showed that the mean annual evaporation variations during 1990-2015. Figure 4.6 showed that in the hot years the mean annul evaporation was high especially 1991 and average annual evaporation remained the same.

4.5 DATA ANALYSIS

4.5.1 Inflows Timeseries

Mangla inflow timeseries is nonstationary as the mean value for the year 2000 shows large difference from the year 1991. Non stationary timeseries means that mean, variances and covariance of timeseries are not constant. It is compulsory to convert into stationary because ARIMA model could not perform the analysis by using the nonstationary timeseries. Figure 4.7 presented the monthly nonstationary timeseries of inflows at the Mangla reservoir. Figure 4.7 showed that the mean values of nonstationary timeseries has large difference. Such as mean value of year 2001 was 476 cumecs whereas mean value of 1991 was 1440 cumecs.

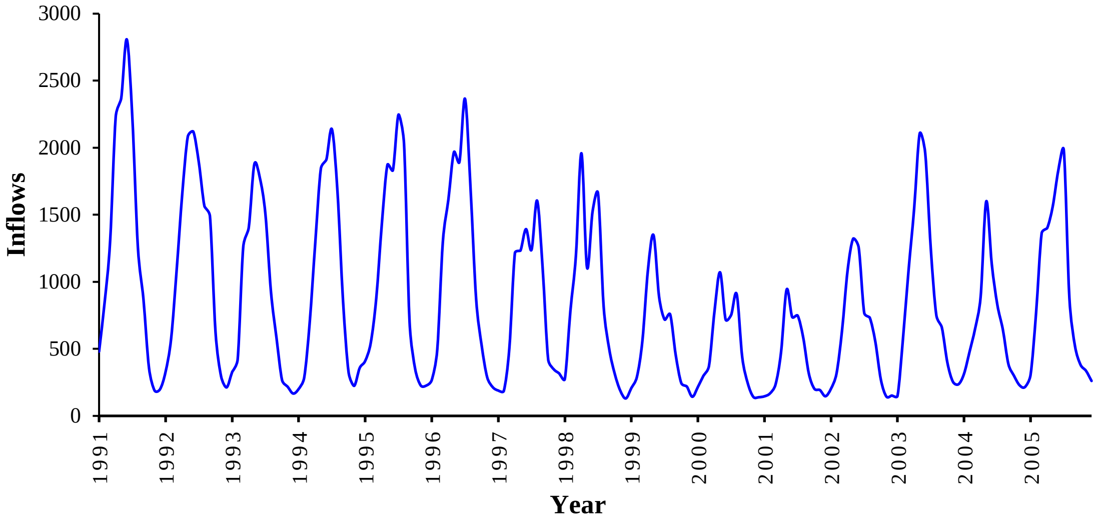


Fig. 4.7 Nonstationary observed inflows timeseries of Mangla reservoir for the period of 1991-2005

A number of trials were performed to convert the non-stationary hydrometeorological timeseries data into stationary. First of all, first difference was adopted, second order difference and so on. However, after differencing it was not converted into stationary. Therefore, the second option of log transformation was used. After log transformation the inflow timeseries was converted into stationary. The model identification and model formulation were performed on stationary timeseries. After forecasting the timeseries, it was reconverted into original form by taking anti-log of forecasted timeseries. The forecasted result was represented in the actual observed timeseries format in the Figure 4.25.

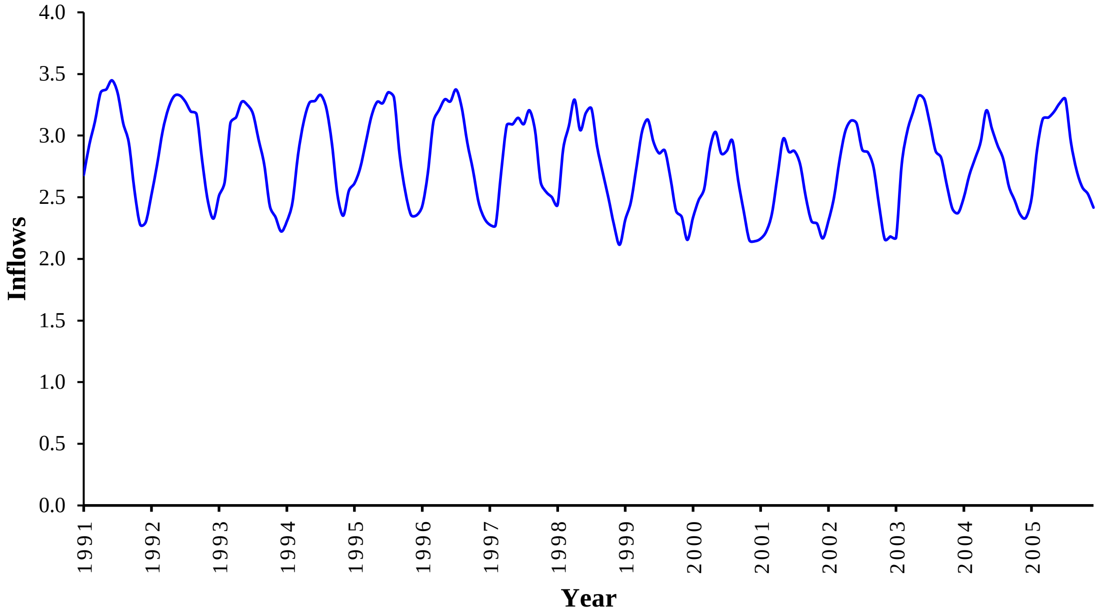


Fig. 4.8 Log transferred stationary observed inflows timeseries of Mangla reservoir for the period of 1991-2005

In the Figure 4.8, the stationary timeseries of inflows is presented. Figure 4.8 clearly described the inflows variation during the period of 1991-2005.

4.5.2 Precipitation Timeseries

Precipitation timeseries of Mangla reservoir is presented in the Figure 4.9. According to the Figure 4.9, the precipitation timeseries were fluctuated between 0 to 500. Peak value of curve shows large variations from year to year. Therefore, the mean values were not constant throughout the timeseries. In the Figure 4.9 the non-stationary precipitation timeseries is presented.

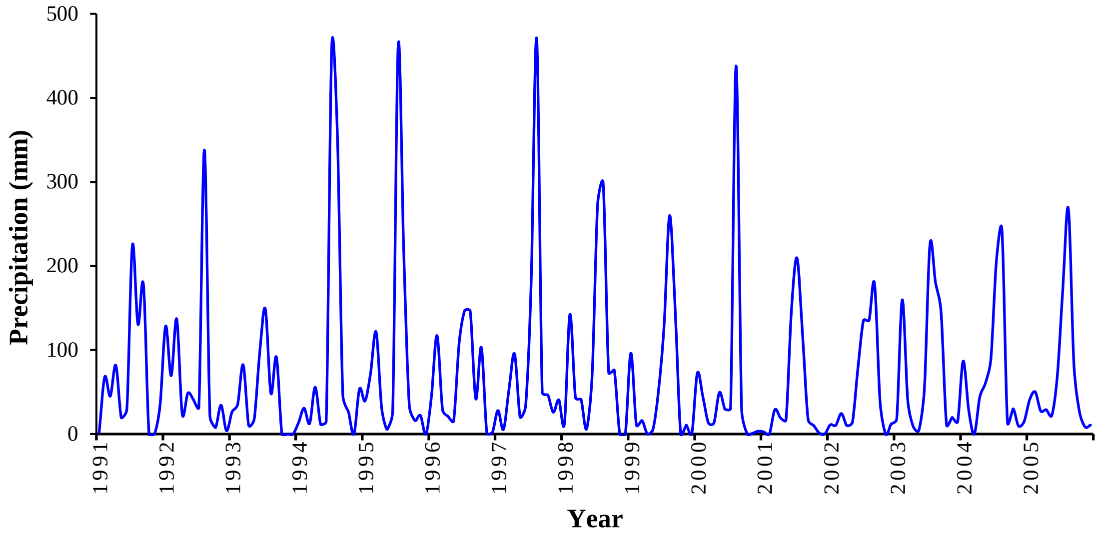


Fig. 4.9 Nonstationary observed precipitation timeseries of Mangla reservoir for the period of 1991-2005

The same steps were performed to convert the timeseries into stationary. Precipitation timeseries were converted into stationary by taking log defencing operation. First of all, timeseries was converted into differencing and then into log form. After this, precipitation timeseries was converted into stationary. Figure 4.10 presents the stationary precipitation timeseries showing that the timeseries mean were constant and it was an example of stationary timeseries.

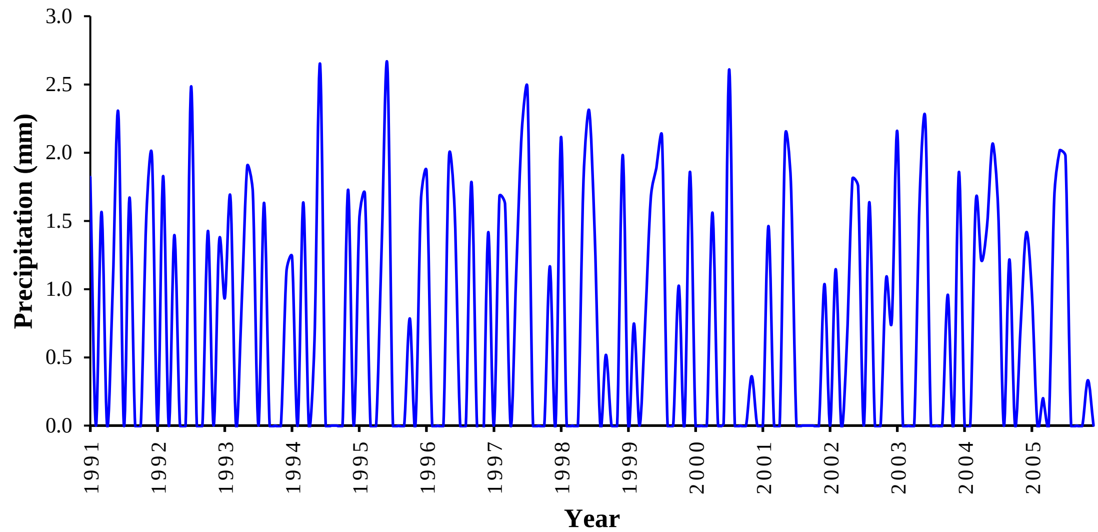


Fig. 4.10 Stationary observed precipitation timeseries of Mangla reservoir for the period of 1991-2005

4.5.3 Evaporation Timeseries

Evaporation is highly depended on the temperature. As soon as the temperature rise the evaporation rate are also increased. In the summer season the evaporation is maximum whereas during the winter seasons it is low. Therefore, evaporation timeseries resembled with the bell-shaped timeseries. It is comparatively easier to convert the evaporation timeseries into stationary because of less randomness.

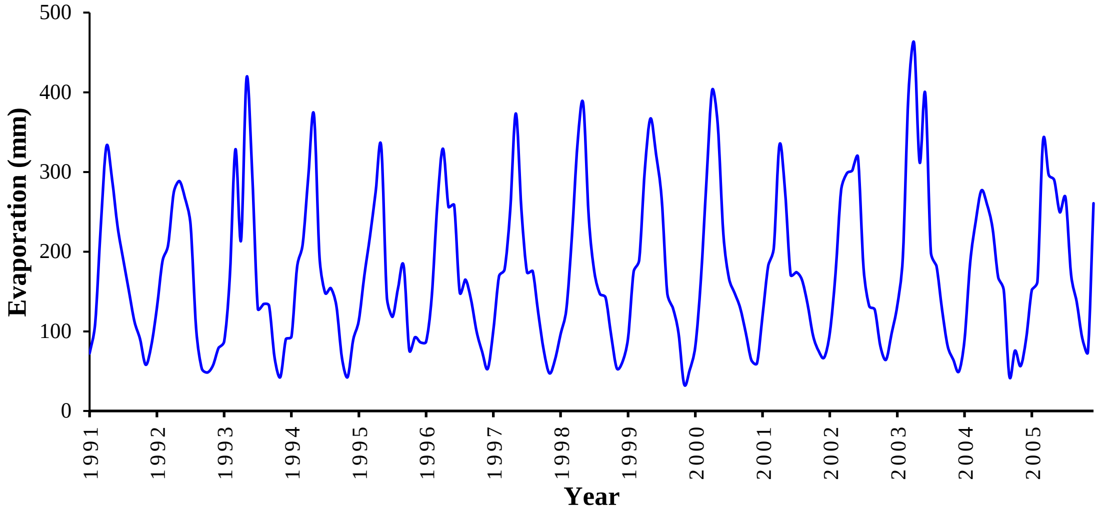


Fig. 4.11 Nonstationary observed evaporation timeseries of Mangla reservoir for the period of 1991-2005

Evaporation timeseries is converted into stationary by taking log transformation. The log transformed evaporation timeseries is a good example of stationary timeseries as the ARIMA model deals with the stationary timeseries because it has constant mean, variance and covariance. The forecasted timeseries are developed on the basis of these parameters. ARIMA model incorporates these parameters into the forecasted results.

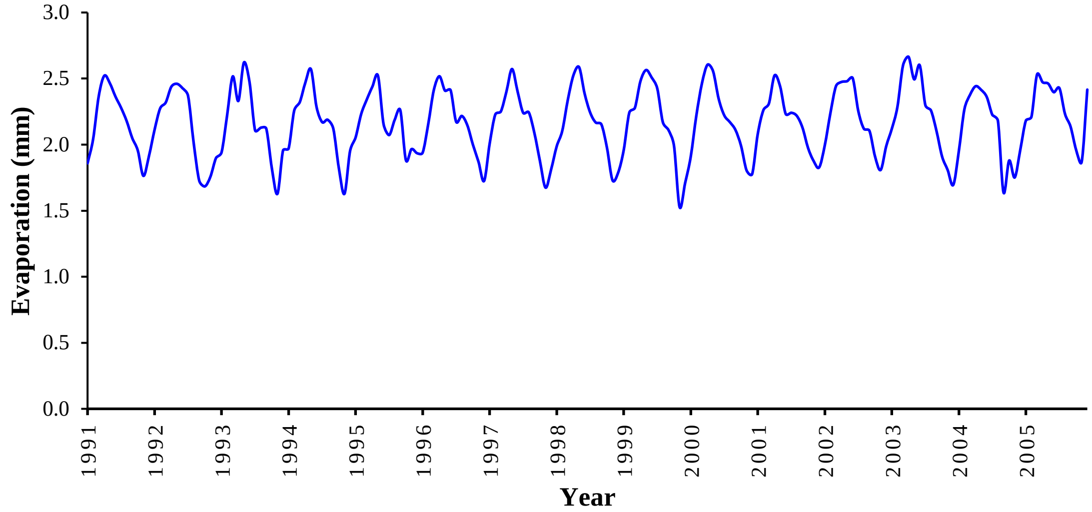


Fig. 4.12 Stationary observed evaporation timeseries of Mangla reservoir for the period of 1991-2005

4.5.4 Preliminary Analysis

ARIMA model has three parameter p, d and q. The parameter "p" was representing the autoregressive (AR) terms of timeseries and parameter "q" represents the moving average (MA) terms of timeseries. Whereas the third parameter "d" depicts the order of differencing.

4.5.1.1 Autocorrelation Function (ACF)

Autocorrelation function (ACF) helps to select the MA terms or p terms of the timeseries. The significant value of ACF represents the value of p terms order. Generally, ACF was used to evaluate the autocorrelation of timeseries with its lagged values. It described how well the present value of the timeseries is related with its proceeding values. A timeseries has components like trend, seasonality, cyclic and residual. ACF considered all these components while finding correlations; hence, it was a 'complete auto-correlation plot'.

4.5.1.2 Partial Autocorrelation Function (PACF)

Partial autocorrelation function (PACF) represents the correlation of the residuals of timeseries. Therefore, any hidden information in the timeseries residuals can be modelled by the next lag until aa good correlation can be attained. In other words, PACF helps to select the AR terms order or “q” terms. The significant value of PACF represents the value of “q” terms order.

Figure 4.13 presents the ACF value of inflow timeseries. This Figure clears that the inflow timeseries had consisted two significant values of ACF at the start. Therefore, AR term order was selected as 2.

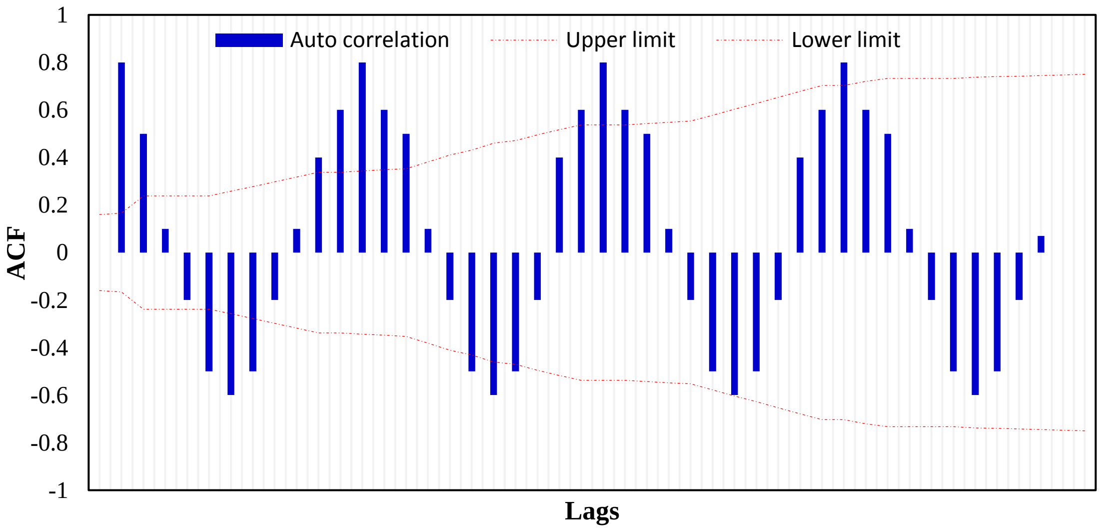


Fig. 4.13 ACF of inflow timeseries.

In the Figure 4.14 the PACF of inflow timeseries were presented. The Figure 4.14 was cleared that the first four values were crossed the confidence limit and all other values were remained in the limits.

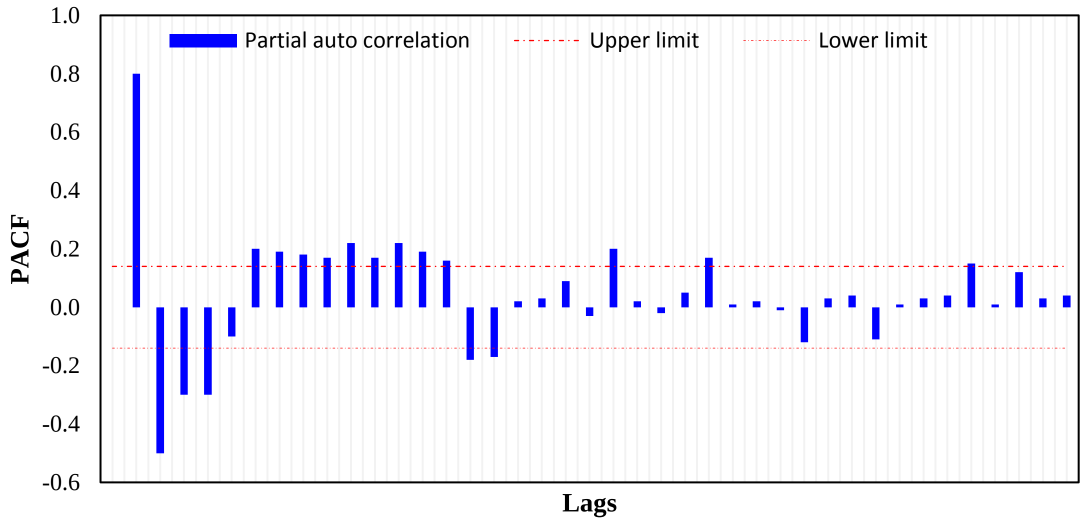


Fig. 4.14 PACF of inflow timeseries

In the Figure 4.15, the ACF value of precipitation timeseries were presented. The Figure 4.15 was cleared that the nine values were significant out of first 12 values. After this all the values were remained in the limit. The maximum value of ACF was 0.5 which occurred after lags 11.

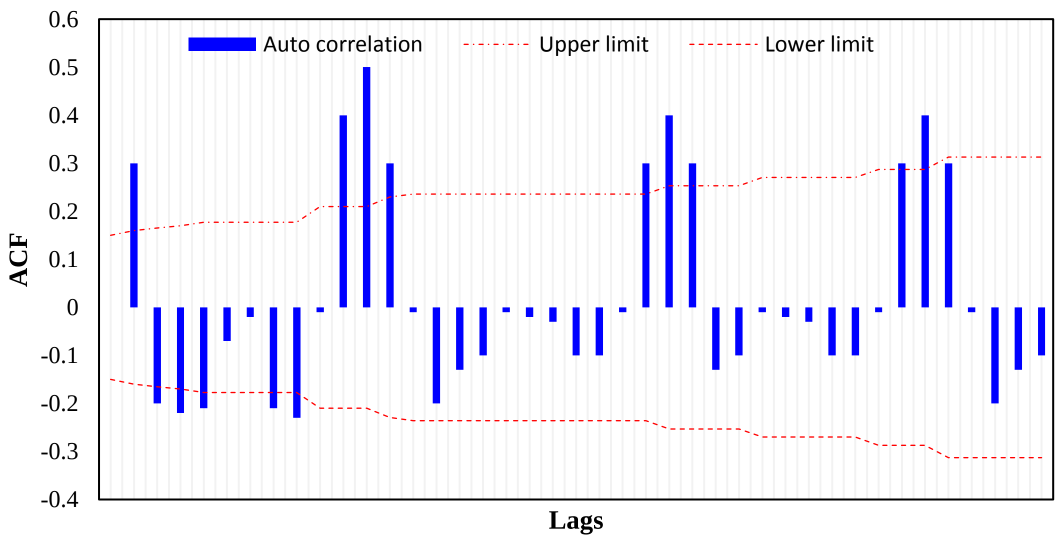


Fig. 4.15 ACF of precipitation timeseries

In the Figure 4.16, PACF of precipitation times series were presented. The Figure 4.16 was cleared that most of values were remained in the range of confidence limit. The maximum value was 0.3 which occurred after lags1.

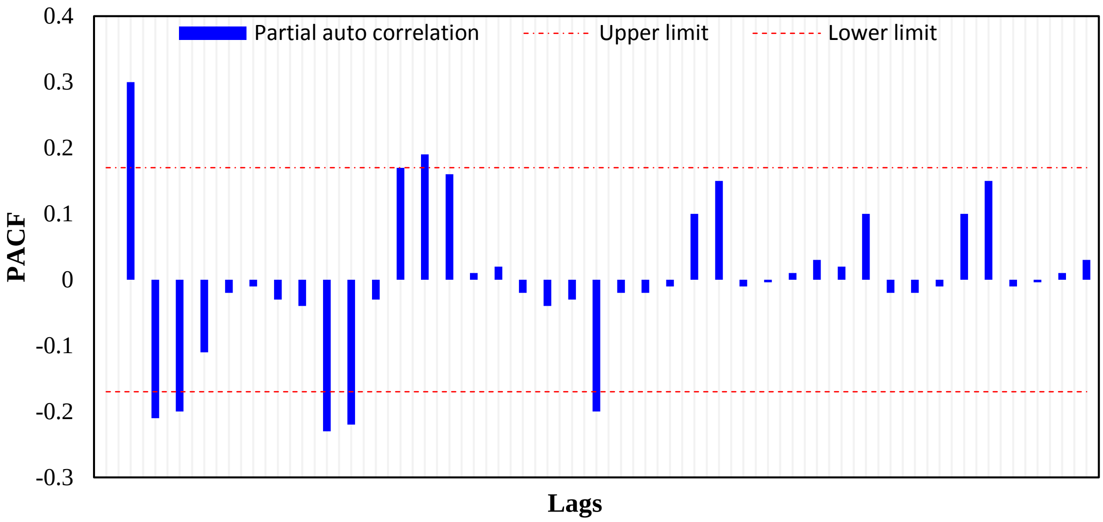


Fig. 4.16 PACF of precipitation timeseries.

Figure 4.17 shows that the maximum value of ACF for evaporation timeseries is 0.8 which is considered a significant value. Therefore, next negative peak values of ACF were changed after 6 lags. However, peak value of 0.8 was again attained which means that evaporation timeseries has a cyclic variation in the ACF values.

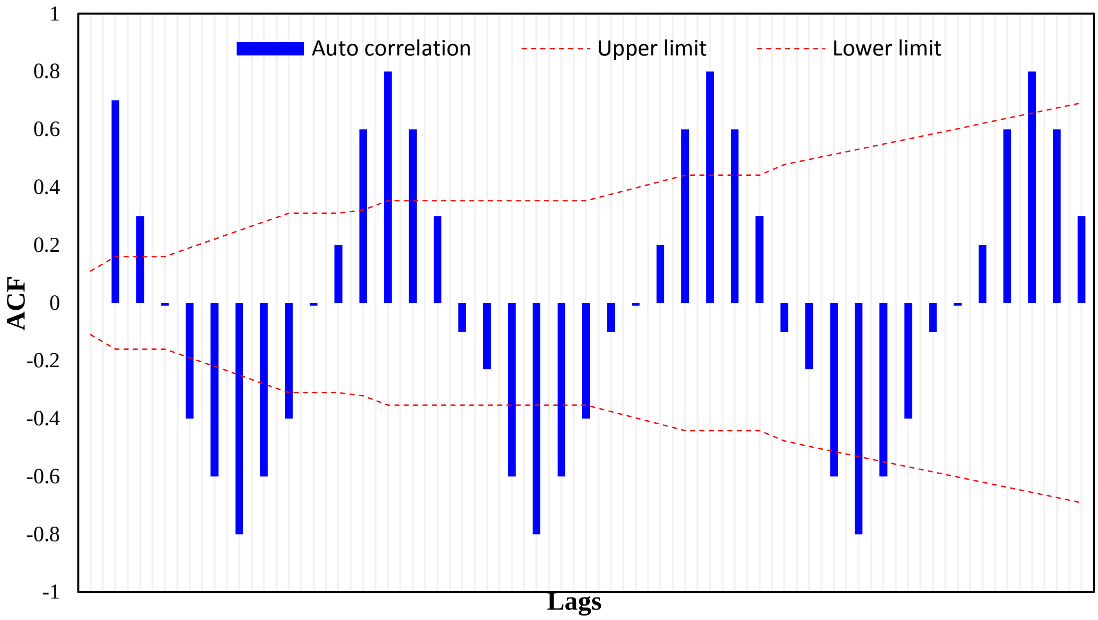


Fig. 4.17 ACF of evaporation timeseries.

Figure 4.18 represents the PACF values of evaporation timeseries. It is clear from Figure 4.18 that the maximum value of PACF of 0.7 was attained which remained within the confidence limit after taking 6 lags. However, the values of PACF are very close to zero; therefore, can be accepted for further analysis

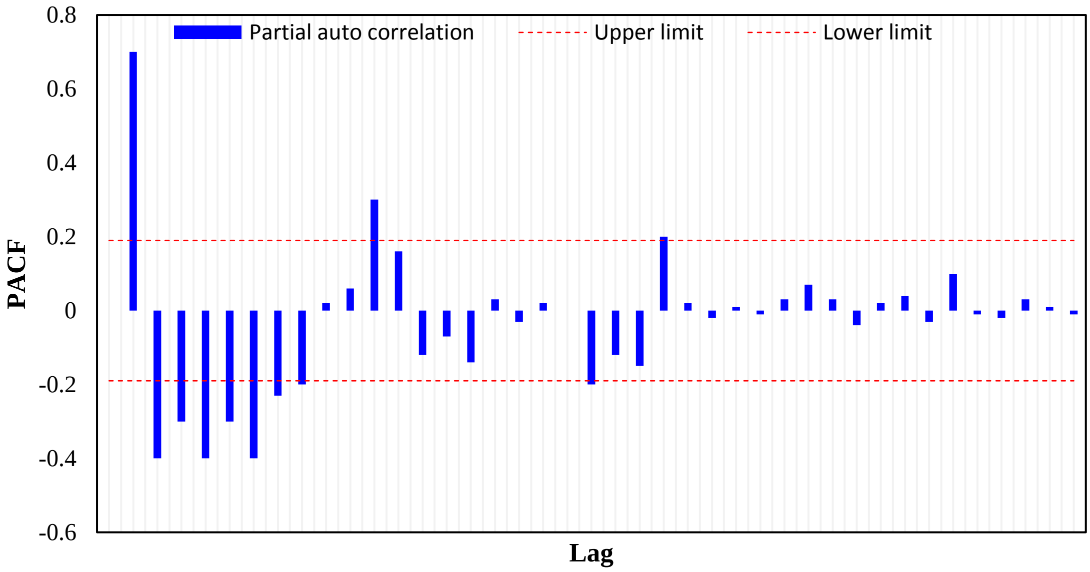


Fig. 4.18 PACF of evaporation timeseries.

4.6 CALIBRATION AND VALIDATION OF ARIMA MODEL

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

In validation period (2006-2010) the simulated timeseries were compared with observed data and the value of R^2, MAE and RMSE were evaluated. For the inflows timeseries, the value 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 value of R^2, MAE and RMSE of evaporation timeseries is 0.88, 21 and 31, respectively.

4.6.1 Calibration of Inflow Timeseries

The inflow timeseries consisted the seasonal variations. This effect was also observed in the ACF values. Therefore, the seasonal ARIMA structure was selected for inflow timeseries. The selection of parameter "p" was depended on the ACF value of timeseries. As shown in Figure 4.13, two significant values of ACF were observed. Therefore, the value of "p" was selected as "2". Similarly, the parameter "q" was depended on the PACF value. As shown in Figure 4.14 two significant values of PACF were also observed. Therefore, the value of "q" was selected as "2". The other parameter was selected by hit and trail method. Afterwards, a number of trials were performed to select the $(1,0,0)(2,1,2)_{12}$ model structure. On the basis of RMSE, R^2 and MAE the inflow timeseries has best fitted seasonal ARIMA model structure of $(1,0,0)(2,1,2)_{12}$. The model simulated and observed flows are presented in Figure 4.19. A number of ARIMA model structures were used for trials to find the best fit model structure of ARIMA. After the comparison of observed data, seasonal ARIMA model structure $(1,0,0)(2,1,2)_{12}$ was found best fitted. The simulated values are compared with the observed flows. Most of peaks are found consistent with the observed flows.

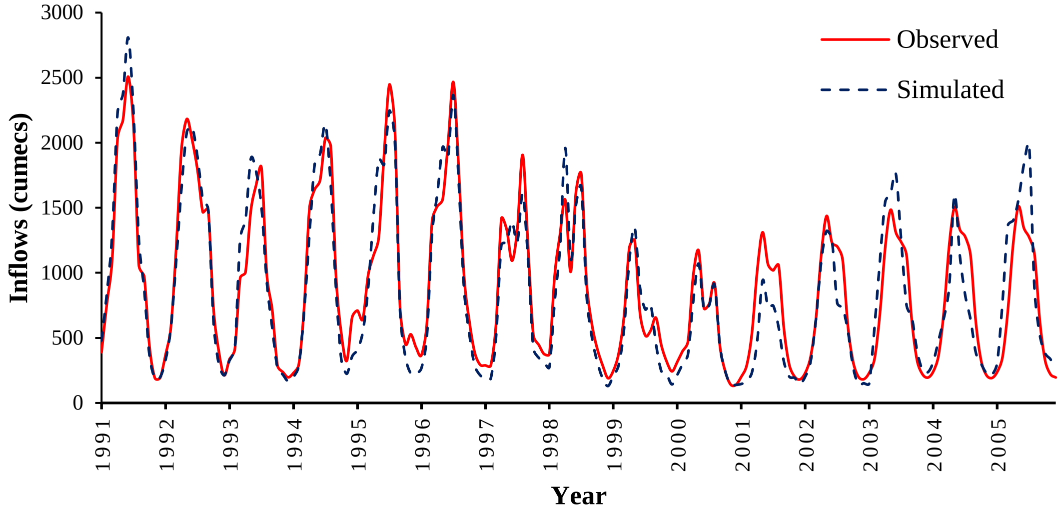


Fig. 4.19 Calibration of observed and simulated inflows into Mangla reservoir using ARIMA model structure (1,0,0) $(2,1,2)_{12}$ during period of 1991 to 2005

4.6.2 Precipitation Timeseries Calibration

Precipitation timeseries has many fluctuations and steep peaks. Precipitation timeseries was less smooth as compared to the inflows and evaporation. Precipitation timeseries consisted of abrupt changes which creates different scenarios such as its parameter "*p*" and "*q*" values were higher. Such the value of "*p*" was 14 and the value of "*q*" was 15. The value of ACF and PACF gives the prejudgment for the selection of "*p*" and "*q*". Therefore, ACF and PACF are drawn and shown in the Figures 4.16 and 4.17. These Figures are consisted of high peak values which causes random changes in the precipitation timeseries.

Consequently, the best fit model structure for the precipitation timeseries was (14,1,15). During the calibration of ARIMA model the AR value of 14 was selected.

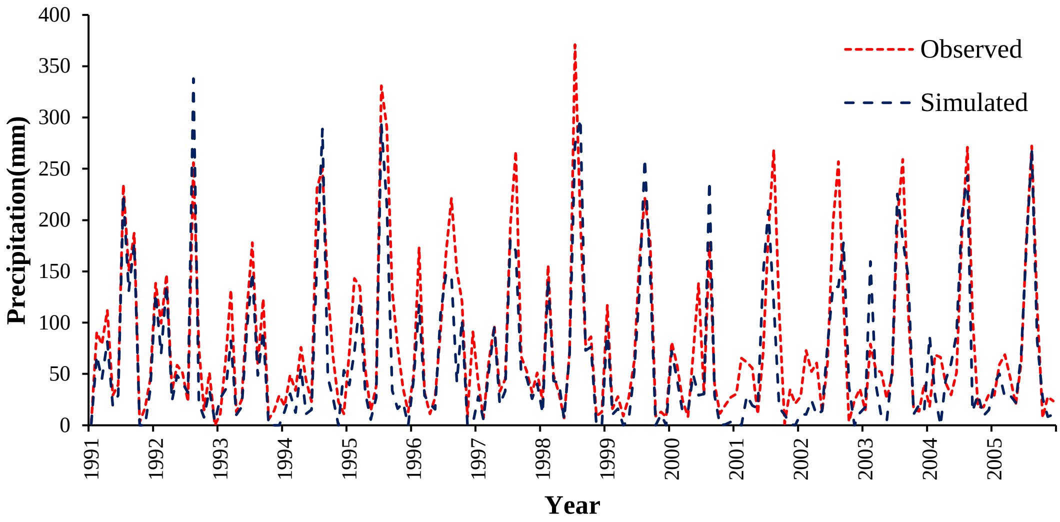


Fig. 4.20 Calibration of observed and simulated precipitation at Mangla reservoir using ARIMA model structure (14,1, 15) during the period (1991-2005)

4.6.3 Evaporation Timeseries Calibration

Evaporation was maximum in the month of June, in this month highest temperature of the year is generally observed. Therefore, the rate of evaporation was maximum in this month. After decreasing the temperature, the rate of evaporation decreased. When the monthly evaporation timeseries were drawn, these variations can be clearly seen in Figure 4.21. The ACF and PACF were used to identify the "p" and "q" terms of ARIMA model. The ACF and PACF values of evaporation timeseries are presented in Figures 4.19 and 4.20. The ACF and PACF values were found comparatively higher. Therefore, a number of trials was performed to select the suitable "p" and "q" values. After the selection of "*p*" and "*q*" parameters, the model results were compared and their RMSE, MAE and R^2 and AR and MA terms were selected as 9 and 19, respectively.

ARIMA model structure of (9,1,19) was best fitted for evaporation timeseries. In calibration period, i.e., 1990-2005, the simulated data and observed data were compared and the results are presented in Figure 4.21. It is clear from Figure 4.21 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 of data was observed in the year 1998. In 1998 the maximum simulated value was 456 mm whereas the observed value was 343 mm.

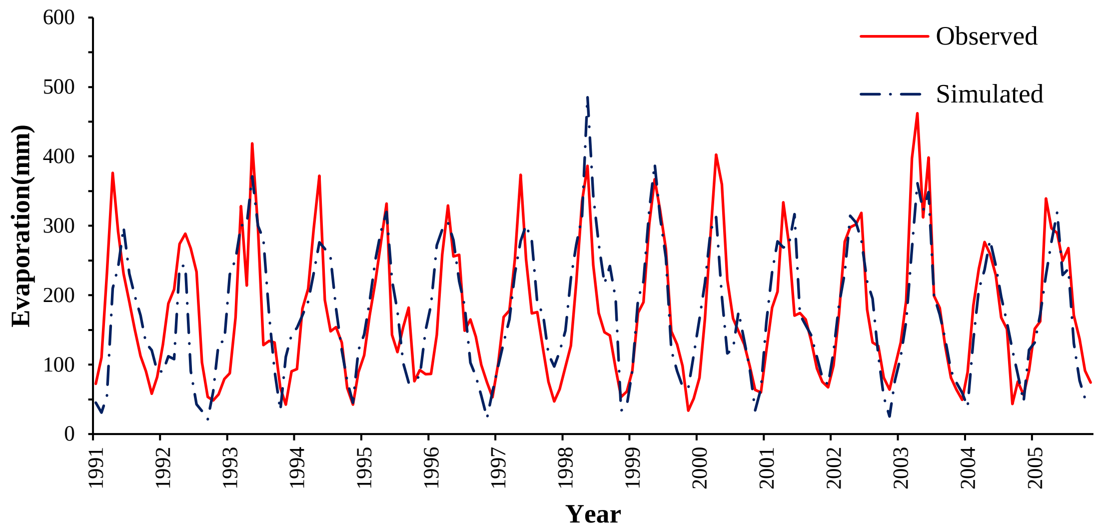


Fig. 4.21 Calibration of Observed and Simulated Evaporation (mm) at Mangla Reservoir using ARIMA Model structure (9, 1, 19) during the period (1991-2005)

4.6.4 Validation of Inflow Timeseries

The period of 1991-2015 was selected as the validation period. In the validation period, most of the peaks of observed data were slightly above the simulated timeseries. In the inflow timeseries, the seasonal trend was more prominent. Therefore, seasonal ARIMA model structure was found best fitted for the inflow timeseries. It was compared with the observed flows and found the values of R^2, RMSE, MAE are 0.85, 195 and 145, respectively. After the validation of these results the same ARIMA model structure is used to forecast the inflow timeseries for the period of 2016–2030.

In the validation period 2005-2015, the first peak of simulated flows is slightly higher than the observed flows which is shown in the following Figure 4.22. In the 2007 the max value of inflow is 1961 cumecs whereas in the same year the simulated flows max value is 1280. The next year 2008 the observed inflows are lower than the simulated flows.

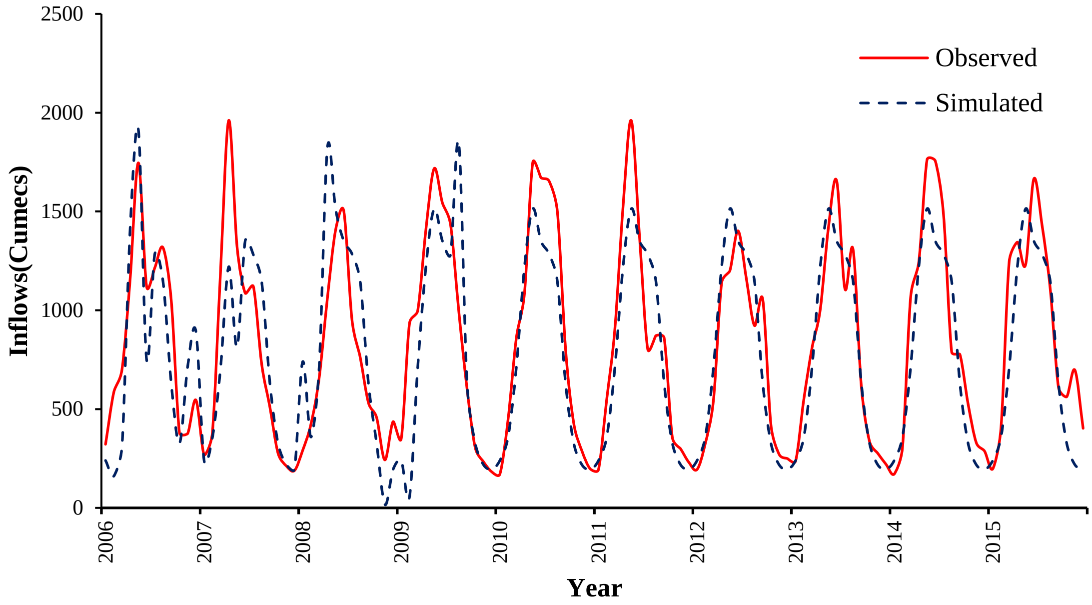


Fig. 4.22 Validation of Observed and Simulated Inflows into Mangla Reservoir using ARIMA Model Structure $(1,0,0)(2,1,2)_{12}$ during the Validation period (2006-15)

4.6.5 Precipitation Timeseries Validation

In the validation period, precipitation peaks were fluctuated between values of 150 to 330. The extreme trend of precipitation was decreased from 2006 to 2010 and after this it was increased. Whereas, the last two years of validation period the simulated peaks were higher than the observed timeseries. After the comparison of the value of RMSE and MAE were 31.7 and 21, respectively.

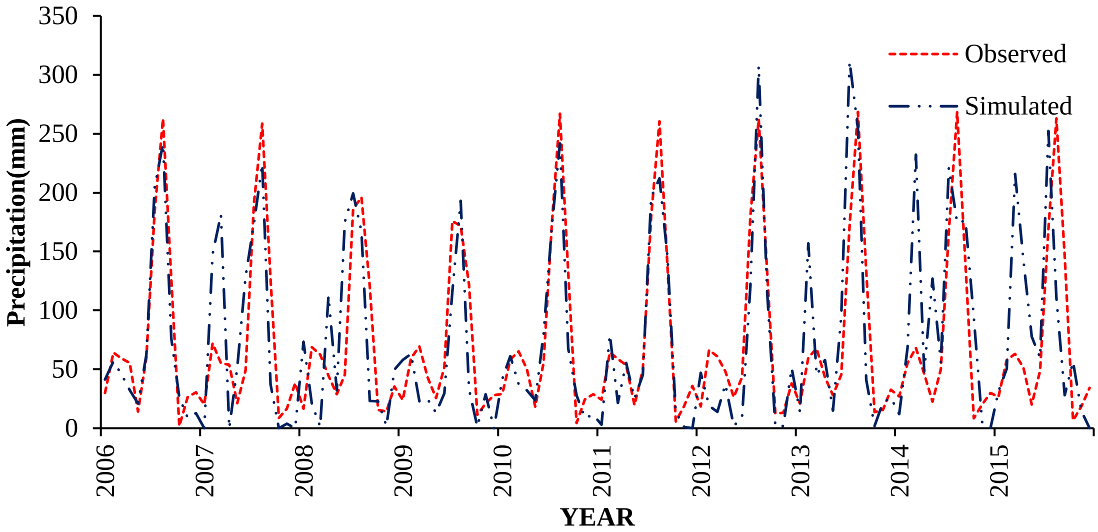


Fig. 4.23 Validation of Observed and Simulated Precipitation at Mangla Reservoir using ARIMA Model structure (14,1,15) during the period (2006-2015)

4.6.6 Evaporation Timeseries Validation

During the validation period (2006-2015), the same ARIMA model structure was used as that of during the calibration process and found that the simulated evaporation timeseries were comparable with the observed evaporation timeseries. 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. Some sudden rise of curve in the observed values were found slightly deviating with simulated values. In the validation period the results are compared and represented in the Tabel 4.1.

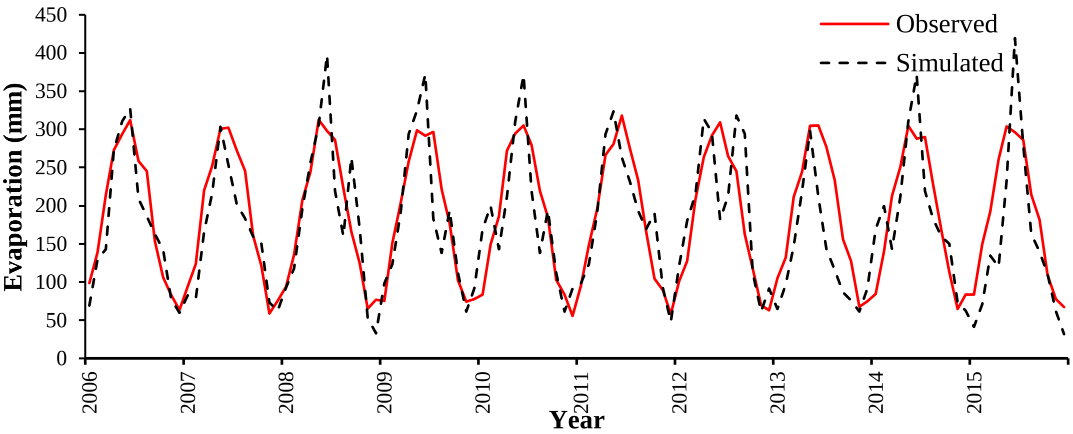


Fig. 4.24 Validation of Evaporation (mm) at Mangla Reservoir using ARIMA Model structure (9,1,19) during the period (2006-2015)

Table 4.1 Statistcal summary of validation period

Variable	R^2	RMSE	MAE
Inflows	0.85	195	145
Precipitation	0.83	31.7	21
Evaporation	0.88	31	14.5

4.7 FORECASTING

4.71 Inflows Timeseries Forecasting

After the calibration and validation process, same ARIMA model structure was used to forecast the inflow timeseries for the period of 2015–2030. It is clear from Figure 4.25 that the maximum flows are predicted in the year 2018-19. Whereas, the lowest flows are expected in the year 2023. This forecasted inflows timeseries was used in the reservoir operation. After this, the forecasted rule curve of reservoir was determined, and water shortage was also determined. These predictions may help the reservoir operators and managers to fulfil the water demands.

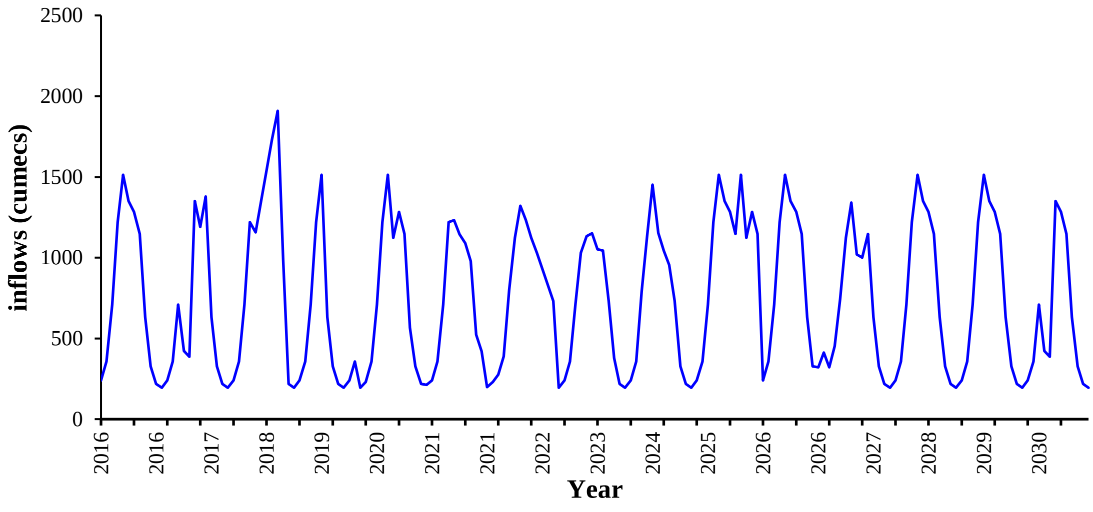


Fig. 4.25 Simulated inflows at Mangla reservoir using ARIMA model structure (1,0,0)(2,1,2)12 during the period (2016-2030)

4.7.2 Precipitation Timeseries Forecasting

To extrapolate the precipitation timeseries up to 2030, 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. The fluctuation of timeseries were steep as the historical data of precipitation timeseries has many peak values. The highest peak was expected in the year 2027.

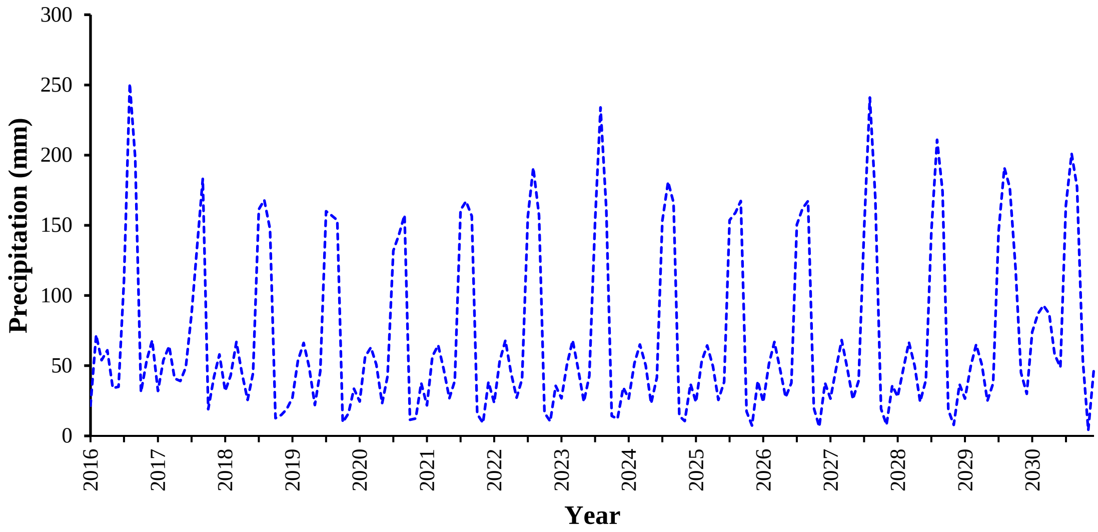


Fig. 4.26 Simulated precipitation at Mangla reservoir using ARIMA model structure (14,1,15) during the period (2016-2030)

4.7.3 Evaporation Timeseries Forecasting

After the validation of evaporation timeseries, ARIMA model structure of (9,1,19) was adopted. These forecasted timeseries were used for the reservoir operation. These forecasted timeseries were presented in the following Figure 4.27. Highest peak of in the evaporation timeseries was observed in the month of June because of the highest temperature being observed in this month.

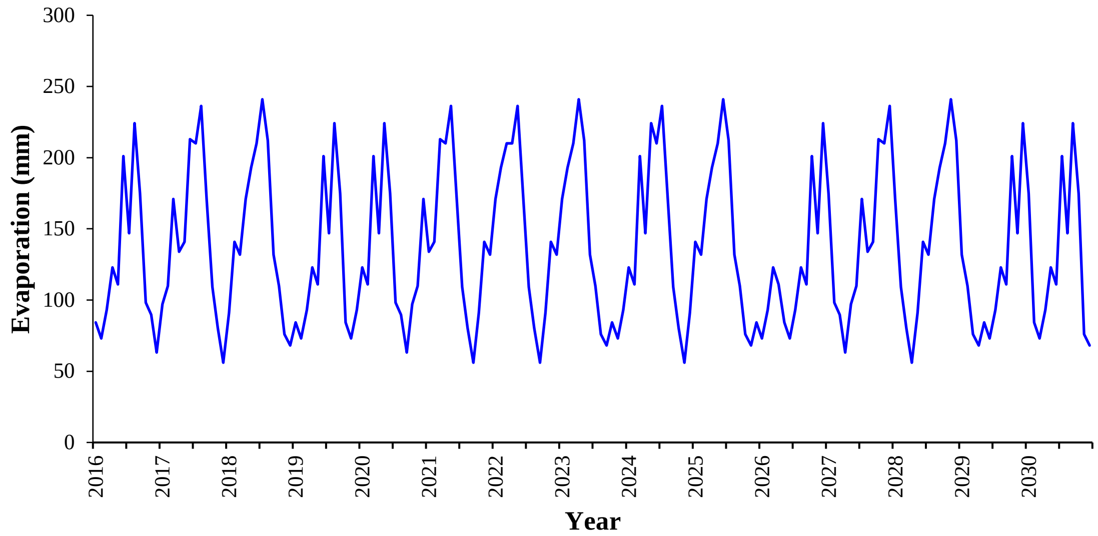


Fig. 4.27 Simulated evaporation at Mangla reservoir using ARIMA model structure (9,1,19) during the period (2016-2030)

The simulated timeseries was used for reservoir operation and forecasted water shortage and 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.

4.8 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 timeseries, reservoir operation was performed on monthly time step. The forecasted reservoir operation represented the future expected reservoir water levels, reservoir storages, reservoir shortage excess periods. This information can be helpful to develop and improve the reservoir policies.

Reservoir elevation curve represents the water levels of reservoir. 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. Sudden rise and drawdown are considered dangerous for reservoir. Such conditions may create the cracks in the reservoir body.

In the following Figure 4.28, reservoir elevation curve is presented. Most of peaks in the curve have one-year difference which is safe and smooth changes in the water levels.

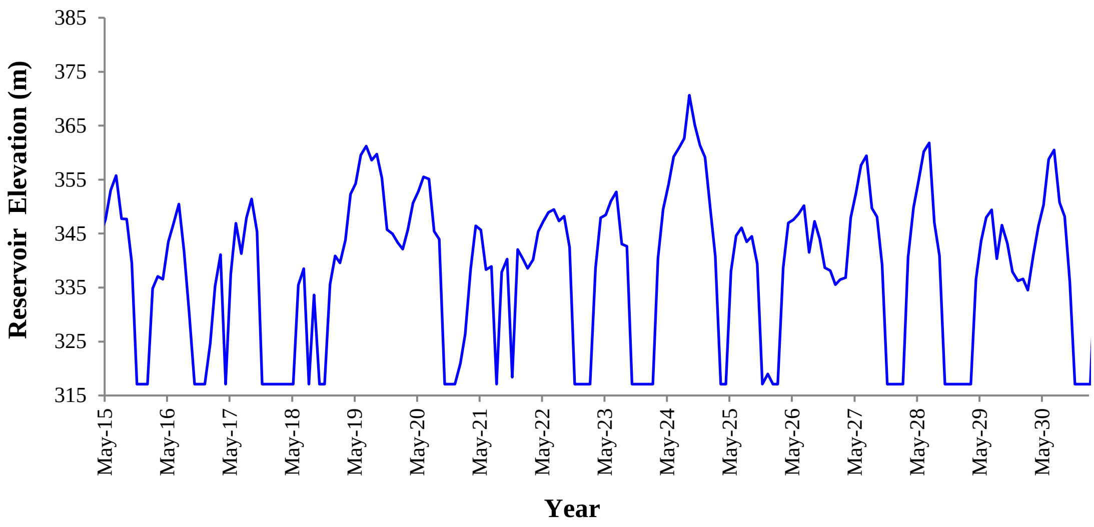


Fig. 4.28 Reservoir elevation curve for the period 2016-2030

Reservoir release curve represents the reservoir outflows. In most of cases, reservoir releases are dependent on the downstream water demands. Sometime 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 conditions.

Figure 4.29 presents the reservoir releases during period of 2016–20130. 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.

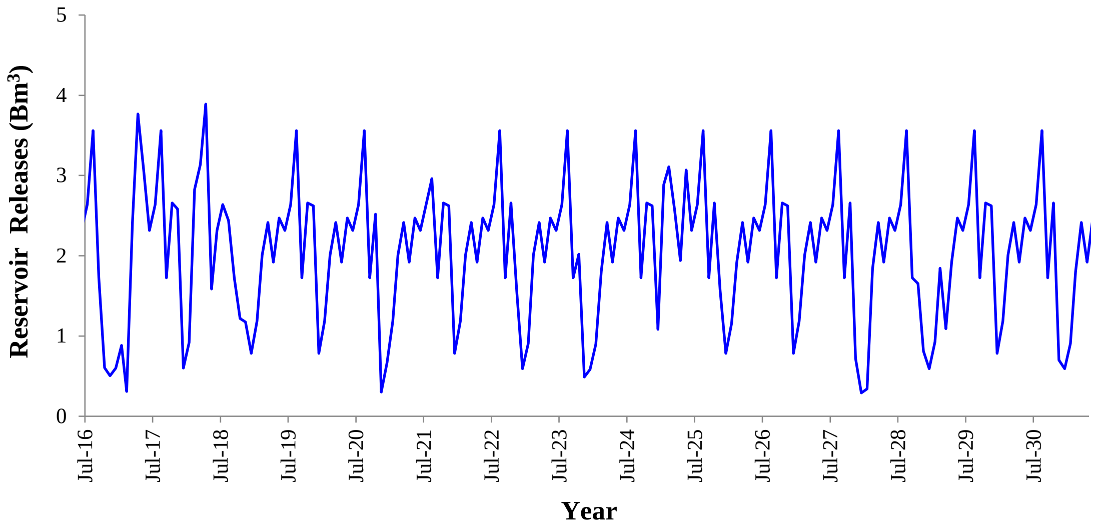


Fig. 4.29 Reservoir Release curve for the period 2016-2030

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 fulfil the downstream water requirements.

Figure 4.30 represents the reservoir storages. During the period of (2016-2020) water is available to fulfil 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 storages.

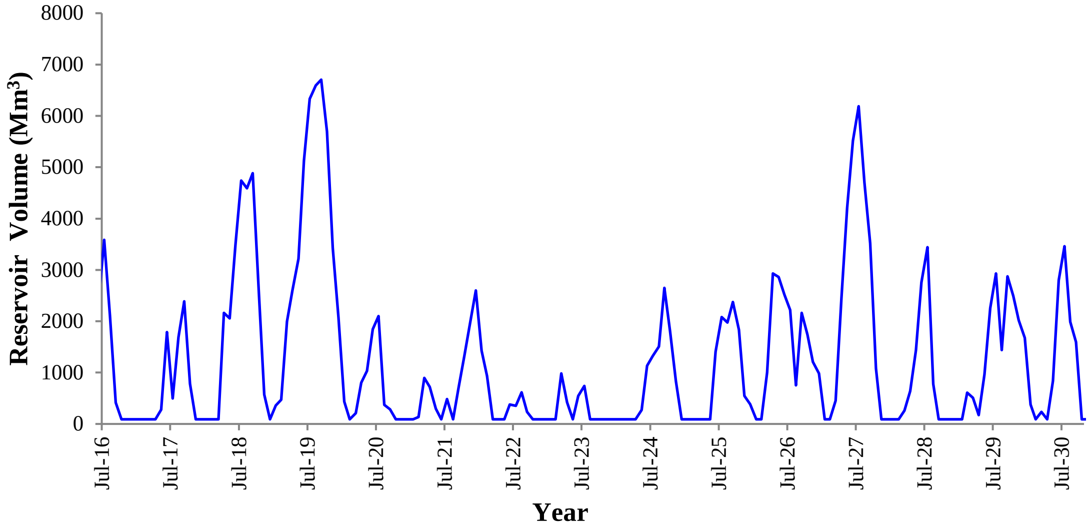


Fig. 4.30 Reservoir volume curve for the period 2016-2030

When the downstream water demands are not fulfilled it is represented the as 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 are remained same, to fulfil these ever-increasing water demands. Therefore, water shortages are increased. Another aspect of sedimentation is also decreasing the water storages.

In the following Figure 4.31, water shortages were presented. During the low flow periods the water shortage 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.

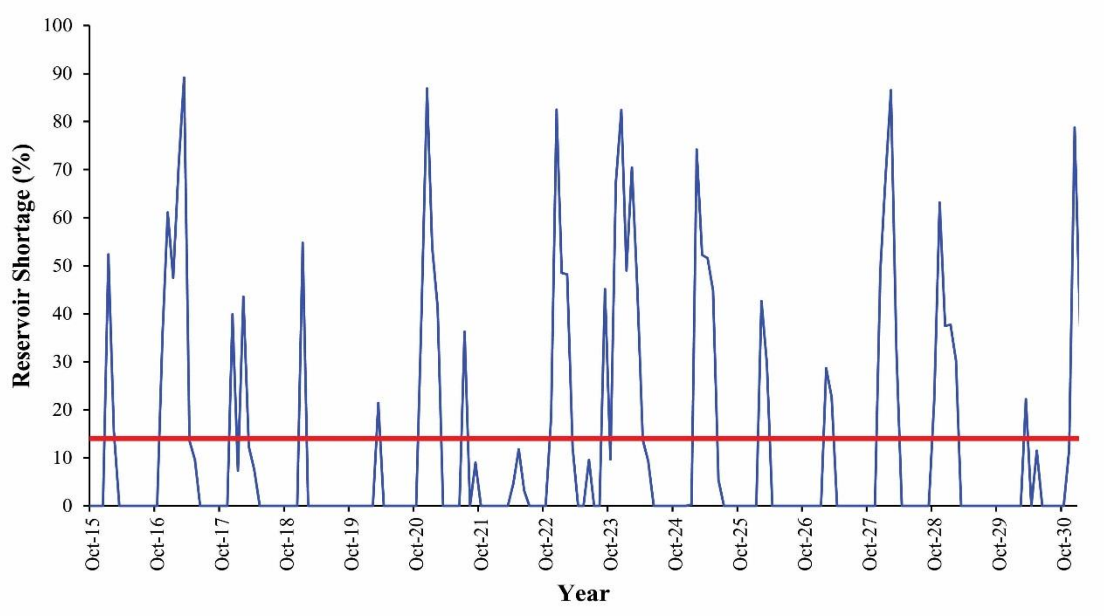


Fig. 4.31 Reservoir shortage curve for the period 2016-2030

Chapter V
CONCLUSIONS AND RECOMMENDATIONS

5.1 CONCLUSIONS

ARIMA model is suitable for the linear timeseries forecasting. The hydro-metrological timeseries of Mangla reservoir are forecasted by using the ARIMA model. The simulated results are compared with the observed values. The R^2 values of inflows, precipitation and evaporation timeseries are found as 0.85, 0.88 and 0.83, respectively. The results of ARIMA model showed that it can be used for developing the policies for reservoir operation.

- The inflows of Mangla reservoir has seasonal effect more prominent compared to climatic time-series of evaporation and precipitation. Therefore, the seasonal ARIMA $(1,0,0)(2,1,2)_{12}$ was found best fitted for inflows timeseries.
- The precipitation timeseries of Mangla reservoir has many steep peaks. The smoothness of precipitation timeseries is less than the inflows timeseries. Therefore, non-seasonal ARIMA structure (14,1,15) was found best fitted for forecasting of precipitation timeseries.
- The evaporation timeseries changes with the change of temperature as in summer months the temperature is high, and rate of evaporation was also found high. The ARIMA (9,1,19) was found best fitted model for forecasting of evaporation timeseries of Mangla reservoir.
- On the basics of forecasted hydrometeorological timeseries of ARIMA model, 14% of water shortages were expected in Mangla reservoir during the period of 2016-2030.

5.2 RECOMMENDATIONS

- The forecasted hydro-meteorological timeseries data may support the reservoir operators and managers for developing the efficient real-time reservoir operation policies and strategies.
- The water demands are also increased day by day which increases the water shortages. Therefore, additional water storages are required to fulfil these demands and another aspect of efficient irrigation system may be helpful to decrease the water shortages.
- The sedimentation deposition in the reservoir decreases the storage capacity of reservoir and future studies may be conducted to save the water storage capacity from sediment depositions.

REFERENCES

Adenan, N.H., Noorani, M.S.M., 2013. River Flow Prediction Using Nonlinear Prediction Method 7, 1589–1592.

Álvarez-díaz, M., 2015. Forecasting the US CPI: Does Nonlinearity Matter ? Rangan Gupta.

Ayaz Ahmed, HENNA IFTIKHAR, and G.M.C., 2007. Water Resources and Conservation Strategy of Pakistan 997–1009.

Babu, C.N., Reddy, B.E., 2014. A moving-average filter based hybrid ARIMA – ANN model for forecasting time series data. Appl. Soft Comput. J. 23, 27–38. https://doi.org/10.1016/j.asoc.2014.05.028

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

Balasmeh, O. Al, 2019. Trend analysis and ARIMA modeling for forecasting precipitation pattern in Wadi Shueib catchment area in Jordan.

Bogacki, W., Ismail, M.F., 2016. Seasonal forecast of Kharif flows from Upper Jhelum catchment. Proc. Int. Assoc. Hydrol. Sci. 374, 137–142. https://doi.org/10.5194/piahs-374-137-2016

Boucher, M.-A., Ramos, M.-H., 2018. Ensemble Streamflow Forecasts for Hydropower Systems. Handb. Hydrometeorol. Ensemble Forecast. 1–19. https://doi.org/10.1007/978-3-642-40457-3_54-1

Graham, A., Sahu, J.K., Sahu, Y.K., 2019. Short communication Forecast future rainfall & temperature for the study area using seasonal auto-regressive integrated moving averages (SARIMA) model 7, 894–897.

Güntner, A., Krol, M.S., Araújo, J.C. De, Bronstert, A., 2004. Simple water balance modelling of surface reservoir systems in a large data-scarce semiarid region / Modélisation simple du bilan hydrologique de systèmes de réservoirs de surface dans une grande région semi-aride pauvre en données. Hydrol. Sci. J. 49. https://doi.org/10.1623/hysj.49.5.901.55139

Hamidi, K., Telvari, A., 2018. Forecasting by Stochastic Models to Inflow of Karkheh Dam at Iran Forecasting by Stochastic Models to Inflow of Karkheh Dam at Iran.

Jothiprakash, V., Magar, R.B., 2012. Multi-time-step ahead daily and hourly intermittent reservoir inflow prediction by artificial intelligent techniques using lumped and distributed data. J. Hydrol. 450–451, 293–307. https://doi.org/10.1016/j.jhydrol.2012.04.045

Kasyoki, A., 2015. Simple Steps for Fitting Arima Model to Time Series Data for Forecasting Using R 4, 2013–2016.

Kavasseri, R.G., Seetharaman, K., 2009. Day-ahead wind speed forecasting using f-ARIMA models. Renew. Energy 34, 1388–1393. https://doi.org/10.1016/j.renene.2008.09.006

Khanal, A., Timalsina, N.P., Alfredsen, K., 2014. Runoff Forecasting and Its Application in Reservoir Operation and Flood Warnings in Nepal. Hydro Nepal J. Water, Energy Environ. 15, 23–29. https://doi.org/10.3126/hn.v15i0.11288

Khashei, M., Bijari, M., 2011. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. Appl. Soft Comput. J. 11, 2664–2675. https://doi.org/10.1016/j.asoc.2010.10.015

Landeras, G., Ortiz-Barredo, A., López, J.J., 2009. Forecasting Weekly Evapotranspiration with ARIMA and Artificial Neural Network Models. J. Irrig. Drain. Eng. 135, 323–334. https://doi.org/10.1061/(asce)ir.1943-4774.0000008

Lee, C.M., Ko, C.N., 2011. Short-term load forecasting using lifting scheme and ARIMA models. Expert Syst. Appl. 38, 5902–5911. https://doi.org/10.1016/j.eswa.2010.11.033

Lee, Y.S., Tong, L.I., 2011. Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. Knowledge-Based Syst. 24, 66–72. https://doi.org/10.1016/j.knosys.2010.07.006

Lo, R., 2003. Demand Forecasting for Irrigation Water Distribution Systems 129, 422–431.

Lohani, A.K., Goel, N.K., Bhatia, K.K.S., 2014. Improving real time flood forecasting using fuzzy inference system. J. Hydrol. 509, 25–41. https://doi.org/10.1016/j.jhydrol.2013.11.021

Majeed, Z., Piracha, A., 2011. Water Conservation of Pakistan' s Agricultural , Municipal and Industrial Water. Int. J. Water Resour. Arid Environ. 1, 232–238.

Martins, O.Y., Sadeeq, M.A., Ahaneku, I.E., 2011. ARMA Modelling of Benue River Flow Dynamics : Comparative Study of PAR Model 2011, 1–9. https://doi.org/10.4236/ojmh.2011.11001

Musa, J.J., 2013. Stochastic Modelling of Shiroro River Stream flow Process 49–54.

Nohara, D., Saito, H., 2018. Assessment on Effects of Preliminary Release Operation of a Multi-purpose Reservoir Considering Ensemble Inflow Prediction 3, 1568–1560. https://doi.org/10.29007/9fpp

Reza Ghanbarpour, M., Abbaspour, K.C., Jalalvand, G., Moghaddam, G.A., 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

Salami, Y.D., Salami, A.W., 2018. Volumetric Inflow Estimation in a Hydropower Dam Using Autoregressive Integrated Moving Average (ARIMA) Modelling and Altimetric Lake Levels. Curr. J. Appl. Sci. Technol. 31, 1–6. https://doi.org/10.9734/cjast/2018/45157

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.

Schwanenberg, D., Fan, F.M., Naumann, S., Kuwajima, J.I., Montero, R.A., Assis dos Reis, A., 2015. Short-Term Reservoir Optimization for Flood Mitigation under Meteorological and Hydrological Forecast Uncertainty: Application to the Três Marias Reservoir in Brazil. Water Resour. Manag. 29, 1635–1651. https://doi.org/10.1007/s11269-014-0899-1

Singh, M., Singh, R., 2011. Application of Software Packages for Monthly Stream Flow Forecasting of Kangsabati River in India 20, 7–14.

Tian, J., Liu, J., Li, C., Yu, F., 2018. Doppler Radar Data Assimilation for Mesoscale Numerical Rainfall Prediction 3, 2073–2064. https://doi.org/10.29007/h6dv

Valipour, M., Banihabib, M.E., Behbahani, S.M.R., 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, J., Zhang, L., Zhang, W., Wang, X., 2019. Reliable Model of Reservoir Water Quality Prediction Based on Improved ARIMA Method. Environ. Eng. Sci. 00, ees.2018.0279. https://doi.org/10.1089/ees.2018.0279

Wang, Z., Qiu, J., 2018. Hybrid Models Combining EMD / EEMD and ARIMA for Long-Term Streamflow Forecasting 1–14. https://doi.org/10.3390/w10070853

Xu, J., Chen, Y., Li, W., Nie, Q., Song, C., Wei, C., 2014. Integrating Wavelet Analysis and BPANN to Simulate the Annual Runoff With Regional Climate Change : A Case Study of Yarkand River , Northwest China 2523–2537. https://doi.org/10.1007/s11269-014-0625-z

Yalcin, E., Tigrek, S., 2017. Optimization of the Garzan Hydropower System operations. https://doi.org/10.1007/s12517-017-3166-y

Zamani Sabzi, H., King, J.P., 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

Zhang, Q., Wang, B. De, He, B., Peng, Y., Ren, M.L., 2011. Singular Spectrum Analysis and ARIMA Hybrid Model for Annual Runoff Forecasting. Water Resour. Manag. 25, 2683–2703. https://doi.org/10.1007/s11269-011-9833-y

Zhao, T., Yang, D., Cai, X., Zhao, J., Wang, H., 2012. Identifying effective forecast horizon for real-time reservoir operation under a limited inflow forecast 48, 1–15. https://doi.org/10.1029/2011WR010623

ANNEXURE - A

Table A.1 Mean monthly inflows at Mangla reservoir

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1991	480	845	1307	2230	2371	2808	2227	1259	873	359	190	199
1992	330	578	1091	1654	2082	2120	1889	1565	2488	626	296	212
1993	327	423	1259	1405	1876	1775	2405	928	582	268	218	167
1994	201	286	702	1296	1843	1913	2137	1670	860	326	224	357
1995	410	544	880	1432	1872	1833	2247	2030	700	353	228	226
1996	267	495	1296	1613	1967	3895	2365	1695	875	525	288	214
1997	189	187	506	1215	1233	1393	1239	1606	1102	420	348	316
1998	275	780	1197	1959	2008	1523	1656	819	501	315	189	130
1999	208	290	563	1092	1348	889	718	757	455	245	219	143
2000	216	299	372	782	1071	719	753	907	436	243	142	139
2001	146	167	234	476	941	738	747	583	318	203	193	147
2002	203	316	637	1085	1321	1256	769	732	554	270	144	152
2003	149	582	1089	1545	2106	1960	1250	753	659	397	254	236
2004	314	479	651	890	1599	1134	830	644	387	300	229	216
2005	306	771	1362	1401	1562	1830	1980	881	523	383	336	261
2006	322	584	700	1163	1746	1120	1216	1317	1050	380	376	546
2007	269	378	1200	1961	1326	1087	1120	730	512	276	213	189
2008	291	423	657	1059	1405	1503	952	762	533	457	243	436
2009	352	933	998	1409	1717	1544	1433	996	605	305	234	183
2010	170	447	858	1089	1747	1672	1953	2204	804	427	288	197
2011	187	527	906	1425	1961	1376	804	672	861	356	299	231
2012	194	336	544	1140	1201	1403	1163	923	1056	420	271	250
2013	237	549	810	1004	1420	1656	1110	1308	617	337	277	221
2014	171	301	1065	1246	1763	1758	1489	790	2306	526	330	287
2015	197	404	1247	2141	2033	1666	2133	1108	611	563	698	403

Table A.2 Mean monthly precipitation at Mangla reservoir

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1991	2	68	45	82	19	29	225	131	178	0	0	35
1992	128	69	137	24	49	41	32	338	21	8	34	4
1993	27	35	82	10	18	98	149	49	92	0	0	0
1994	14	31	13	56	11	15	461	359	46	27	0	54
1995	39	72	121	31	5	27	464	216	34	16	22	0
1996	45	117	29	21	16	109	147	146	42	103	1	2
1997	28	6	53	95	20	33	180	470	49	47	26	41
1998	12	142	43	41	6	69	273	299	73	76	0	0
1999	96	11	16	0	6	52	129	260	149	0	11	0
2000	72	44	13	13	50	29	30	438	29	0	1	3
2001	3	0	29	19	16	148	209	118	17	10	1	0
2002	11	10	24	10	14	79	135	135	178	38	0	12
2003	18	160	40	9	4	54	225	180	146	11	20	15
2004	87	28	0	44	60	89	206	243	14	30	9	15
2005	41	50	27	29	22	69	174	268	82	26	9	11
2006	41	57	46	33	21	62	205	241	76	26	9	13
2007	0	148	181	0	51	129	173	224	37	0	4	0
2008	73	21	1	112	32	175	199	167	23	23	2	50
2009	58	63	23	25	13	30	116	193	33	2	29	0
2010	42	61	38	32	24	81	172	242	68	28	9	11
2011	3	79	22	56	25	45	196	212	146	14	1	0
2012	47	19	14	37	2	9	121	307	118	2	2	51
2013	14	157	45	58	15	96	311	255	41	2	22	26
2014	12	68	232	48	127	56	221	174	376	94	5	0
2015	31	51	216	143	78	60	252	111	28	57	15	0

Table A.3 Mean monthly evaporation at Mangla reservoir

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1991	74	74	113	215	306	323	212	170	141	121	86	54
1992	68	72	110	232	332	291	231	190	151	112	90	58
1993	82	129	188	208	274	289	267	234	103	53	48	57
1994	79	87	168	328	214	419	297	128	134	132	66	42
1995	90	93	181	210	295	372	193	148	154	132	66	42
1996	89	113	171	220	275	332	143	118	154	182	76	92
1997	86	87	143	259	329	256	258	149	165	140	100	74
1998	53	102	169	178	252	373	252	174	176	123	75	47
1999	65	97	127	219	335	386	243	174	147	142	95	54
2000	61	91	175	190	303	559	322	267	348	139	198	94
2001	51	81	166	290	403	359	222	167	148	129	98	64
2002	60	120	182	205	334	276	171	174	165	135	94	75
2003	67	99	172	277	402	302	319	179	132	127	81	64
2004	97	132	192	397	462	312	399	199	182	137	91	74
2005	50	89	185	238	277	260	229	168	151	43	76	56
2006	91	152	162	339	297	290	249	268	171	137	91	74
2007	69	130	143	273	311	327	209	186	163	142	77	60
2008	83	78	166	216	304	253	202	183	160	150	73	63
2009	93	118	196	256	306	397	219	160	263	170	53	33
2010	98	123	191	294	325	472	183	138	194	112	61	91
2011	169	200	143	212	311	371	219	178	219	212	81	121
2012	98	123	191	294	324	263	232	193	170	190	93	48
2013	121	180	215	314	296	182	213	318	294	112	61	91
2014	65	97	127	229	168	151	43	76	56	122	71	141
2015	169	200	143	212	311	371	219	183	160	150	73	63

ANNEXURE-B

Table B.1 Model performance evaluation of inflow timeseries during calibration period

Year	Observed	Simulated	RMSE	MAE	NE
1991	1162	1241	157	138	0.668
1992	1198	1131	135	104	0.657
1993	842	883	208	148	0.483
1994	1049	1011	186	161	0.799
1995	1114	1070	250	200	0.984
1996	1165	1124	153	125	0.870
1997	901	834	124	105	0.887
1998	918	869	172	142	0.934
1999	597	574	154	144	0.791
2000	559	510	109	81	0.845
2001	646	403	293	243	0.889
2002	719	620	172	103	0.879
2003	748	857	247	186	0.835
2004	760	640	237	175	0.461
2005	764	966	352	265	0.600
Average	876	849	196	155	0.77

Table B.2 Model performance parameter of inflow timeseries during validation period

Year	Observed	Simulated	RMSE	MAE	NE
2006	877	855	124	87	0.83
2007	772	798	103	71	0.94
2008	727	762	179	135	0.93
2009	892	766	222	162	0.88
2010	904	766	207	172	0.89
2011	827	766	248	195	0.87
2012	740	766	182	138	0.87
2013	795	766	146	113	0.86
2014	875	766	226	189	0.87
2015	908	783	257	190	0.86
Average	832	780	195	145	0.88

Table B.3 Model performance parameter of precipitation timeseries during calibration period

Year	Observed	Simulated	RMSE	MAE	NE
1991	86	71	18	14.91	0.77
1992	79	74	31	21.09	0.91
1993	64	47	24	20.28	0.87
1994	80	60	41	33.78	0.85
1995	111	73	48	38.01	0.82
1996	95	65	50	34.89	0.81
1997	79	62	30	16.82	0.79
1998	91	86	41	23.32	0.79
1999	71	61	20	16.33	0.80
2000	57	44	38	24.41	0.80
2001	76	48	62	47.87	0.78
2002	77	54	52	41.85	0.75
2003	76	73	38	28.34	0.74
2004	76	69	44	34.95	0.73
2005	76	67	19	15.27	0.74
Average	80	64	37	27	0.80

Table B.4 Model performance parameter of precipitation timeseries during validation period

Year	Observed	Simulated	RMSE	MAE	NE
2006	174	173	26	11	0.65
2007	161	176	31	18	0.83
2008	189	195	28	14	0.86
2009	182	187	27	13	0.91
2010	185	192	30	16	0.87
2011	185	194	30	16	0.79
2012	200	199	26	11	0.81
2013	134	146	29	15	0.84
2014	188	189	71	46	0.93
2015	150	180	93	52	0.89
Average	175	183	31.7	21	0.84

Table B.5 Model performance parameter of evaporation timeseries during calibration period

Year	Observed	Simulated	RMSE	MAE	NE
1991	166	152	85	69	0.147
1992	161	109	93	75	0.168
1993	176	216	78	65	0.192
1994	165	174	55	42	0.303
1995	164	174	51	40	0.366
1996	168	172	60	51	0.366
1997	165	181	53	44	0.407
1998	173	233	77	69	0.427
1999	181	187	30	25	0.467
2000	182	161	70	51	0.501
2001	167	189	54	35	0.503
2002	179	177	33	28	0.525
2003	216	192	51	33	0.567
2004	155	156	37	31	0.592
2005	188	159	44	35	0.604
Average	174	175	58	46	0.41

Table B.6 Model performance parameter of evaporation timeseries during validation period

Year	Observed	Simulated	RMSE	MAE	NE
2006	69.20	69.02	7	5	0.72
2007	79.02	83.17	10	7	0.83
2008	73.26	77.08	7	6	0.83
2009	48.68	50.61	11	7	0.92
2010	67.18	68.54	6	5	0.88
2011	66.54	65.48	6	5	0.92
2012	60.74	60.34	4	3	0.90
2013	86.71	88.79	7	6	0.96
2014	101.21	90.19	52	38	0.97
2015	86.77	73.99	83	63	0.88
Average	73.93	72.72	31	14	0.88

ANNEXURE-C

Table C.1 Reservoir operation summary

Year	Net Reservoir Inflows (Mm^3)	Water demands (Mm^3)	Reservoir releases (Bm^3)	Reservoir Shortages percentage	Spillage (Mm^3)
2016	24195	29683	21.46	14	0
2017	21186	30134	21.46	23	0
2018	26531	30765	21.66	9	0
2019	24473	31276	20.33	6	0
2020	28019	31818	21.47	12	0
2021	22135	32359	22.02	12	0
2022	25595	32900	22.58	10	0
2023	25243	33441	23.00	27	0
2024	23060	33982	23.48	16	0
2025	30667	34524	24.12	19	0
2026	24088	35065	23.43	6	0
2027	26680	35606	25.37	8	0
2028	23912	36147	26.53	26	0
2029	21980	36689	26.53	6	0
2030	24772	37230	26.55	10	0