Research Article   **Open Access**

Water Management in Tarbela Dam By using Bayesian Stochastic Dynamic Programming in Extreme Inflow Season

Ayesha Nayab* and Muhammad Faisal

Department of Statistics, Quaid I Azam University, Islamabad, Pakistan

Abstract

Existing method of forecasting inflows at Tarbela have some limitations, also system needs an adequate operating policy model to deal with highly volatile inflow of summer months of June, July, August and September. In this paper, historical data of inflows from 1986 to 2014 have been used to forecast upcoming inflows at dam. Bayesian predictive distribution is used to predict future inflows. These forecasted inflows were further incorporated into operating policy model to determine the optimal release during the prescribed months. Weather volatility is a major factor causing unstable inflows. High temperature during summer period cause high inflows at dam. Considering weather volatility, this policy model is proposed for the flood season (15th June to 30th September), in which inflows and outflows are higher than rest of the year. This model maximizes the expected profit from hydro power production, minimizes the expected loss from flood damage and updates the proper estimate of current stage of reservoir storage.

Keywords: Temperature volatility and inflow forecast; Release policy; Bayesian stochastic dynamic programming

Introduction

Water inflow is the uncertain element which directly affects the release policy of the dam. Inflow into the reservoir is usually high in summer due to snowmelt from high temperature. Also rain in the upstream catchment area is the most critical flood-producing factor of the basin. Monsoon caused heavy rain during summer and water elevation at dam surface often rose to the peak. To manage available water to generate maximum hydropower and prevent from flood and damage in downstream area is a big challenge in summer because of high volatility of temperature.

Tarbela is a multipurpose reservoir, firstly it provides water for irrigation and almost 50% of the country irrigation system is run by Tarbela dam. Secondly, it stores water during flood season to escape from high damage at downstream areas. Thirdly, it offers generation of hydropower to satisfy 30% of the energy needs. To achieve these objectives at their best, system needs a reliable forecast for the inflows and operating policy model. Table 1 presents the water discharge limits at different levels of floods.

At minimum flood limit release estimate is 7000 m³/second and at maximum this release estimate goes to 23000 m³/second. In flood season, when the water rises beyond the capacity of the dam, we need to appropriately utilize the substantial amount of available water in such a way that we can generate maximum hydropower with minimum damage at downstream area. Therefore, we need to account uncertainty of the inflow in optimal operating policy.

The concept of addressing forecast uncertainty along with inflow uncertainty was presented by many authors in the past. Bayesian Stochastic Dynamic Programming model incorporates Bayesian approach within classical Stochastic Dynamic Programming (SDP) formulation [1]. Therefore we used Inflow data as prior knowledge and forecasted data as sample information. Posterior distribution was a conditional distribution of forecast over actual inflow to the reservoir. Use of Bayesian Decision Theory (BDT) and Bayesian inference for predicting upcoming flows of water along with classical forecasting methods take into account relatively complex problems of natural resource management [2]. Bayesian Stochastic Dynamic Programming (BSDP) is also used as stochastic optimization tool for the development of operating policy model, which incorporates a Bayesian approach within the classical Stochastic Dynamic Programming (SDP) formulation (Table 2) [3].

Flood Stage	Discharge (m³/second)
Low	7000
Medium	10500
Average	14000
High	18500
Very High	23000

(Source: WAPDA Pakistan, Flood Management Manual 2011)

Table 1: Flood limits at Tarbela.

Notation	Description
I_t	Inflow at time t
nh_t	Net head available at time t
η	Plant efficiency
C_t	Cost at time t due to excess outflow in downstream area
S_t	Storage at time t
R_t	Rain at time t
E_t	Evaporation at time t
O_t	Outflow at time t
λ	Estimated threshold level for damage

Table 2: Notations used in BSDP model with description.

The use of linear autoregressive (AR) model and linear autoregressive-moving-average (ARMA) model in reservoir optimization is available in literature. The multi lag autocorrelation model by a single hydrologic variable, the value of which changes from day to day and is equal to the conditional mean of the daily inflow [4]. Algorithm which can take into account the inflow stochasticity for calculating optimal operating strategies in a multi-reservoir hydroelectric system does not require discretization of the state space [5].

Case studies with the Brazilian systems give the comparison of

*Corresponding author: Ayesha Nayab, Department of Statistics, Quaid I Azam University, Islamabad, 45320, Pakistan, Tel: +92519064000; E-mail: ayesha.nayab@mail.au.edu.pk

two dynamic programming models, one is deterministic and the other stochastic dynamic programming model. Deterministic model gives good results for reservoirs ranging from medium to large, while SDP gives better results for small reservoirs. The deterministic model (DPR) consists of an algorithm that cycles through three components: a dynamic program, a regression analysis, and a simulation. While the stochastic dynamic program (SDP) describes stream flows with a discrete lag-one markov process [6]. A stochastic dynamic programming model which defines a reservoir release policy for the dam at Aswan in the Nile River Basin where the best forecast of the current period's inflow is employed to define a reservoir release policy calculates the expected benefits from future operations [7].

Weather and climatic extremes can have serious and damaging effects on human society and infrastructure as well as on ecosystems and wildlife. Thus, they are usually the main focus of attention of the news media in reports on climate. Climate models show how they could change in the future either due to natural climate fluctuations or under conditions of greenhouse gas—induced warming. These observed and modeled changes relate directly to the understanding of socioeconomic and ecological impacts related to extremes [8]. A hydro-economic model was developed for the Lake Naivasha Basin in Kenya by simulating the local hydro-economic system under different water institutions [9].

Methodology

Inflow prediction

In this paper we used two classical forecasting methods to generate inflow forecast. First one is Autoregressive process of order one (AR 1) and second one is artificial neural networks (ANN). The forecasting results from AR (1) are closest to actual values of inflows than ANN. On comparing both forecasting methods we choose AR (1) forecast as sample information for predictive distribution. We take inflow forecast as sample information and past inflow data as prior information. The posterior predictive mean is used as forecast.

Bayesian estimation

Bayesian approach is relatively more efficient than classical methods of estimation because Bayesian Decision Theory (BDT) utilizes the current information and updates the prior probabilities turning them into posterior probabilities hence reducing the forecast uncertainty. Furthermore, the classical approach does not take into consideration the uncertainty that results from the forecast error [3]. We take the classical forecast of inflows as our sample observations and actual data of inflows is taken as prior information. The distribution of forecasted data set comes out to be normally distributed as depicted by the distribution fitting software Easyfit. For prior information we find the distribution of actual inflows, that is again, normally distributed as shown by the Easyfit. Normal distribution is a conjugate prior, so the predictive distribution also comes out to be a normal distribution.

Proposed Bayesian Stochastic Dynamic Programming (BSDP) Model

The proposed Bayesian Stochastic Dynamic Programming (BSDP) model considers the forecast for future inflow, storage of dam and aggregate inflow. The objective is to maximize power units produced by water discharge with minimum damage, caused by outflow which exceeds the threshold level. Also to maintain risk free storage level, we put maximum as well as minimum storage levels in constraint to the objective function.

Model formulation

Continuity equation: The continuity equation estimates current level of storage by adding new available estimates of inflow, rain forecast, evaporation rate, preceding day storage and current release. It updates storage level for each current state of the system. For the time period t continuity equation is as under

$$S_t = I_t + R_t - E_t + S_{t-1} - O_t \qquad (1.1)$$

Release constraint: To avoid excess release than a threshold level, we have put release limits insuring safety constraint to the objective function. The release constraint is as under

$$0.89 \leq O_t \leq 2.8 \qquad (1.2)$$

Whereas Ot specifies the release or outflow at time t, 0.89 billion cubic meter describes minimum release level and 2.8 billion cubic meter specifies maximum release. This constraint prevents from heavy release which can cause damage in the form of floods and also from insufficient release to fulfil the water requirements in downstream areas.

Storage constraint: Storage constraint fixes the maximum and minimum storage levels at 1550 and 1378 respectively. Storage less than 1378 stops the functioning of dam, because less than minimum storage level causes heavy loss to the agriculture system which entirely depends on water release from this upstream area. In addition, public needs of water are also not fulfilled in this situation while, the situation where the water level increases the storage capacity results in damage to dam infrastructure. Furthermore, excess storage, more than a threshold level can result in destruction of spillways, which surely can cause a worst flood with unbearable losses to livestock and entire system of agriculture in downstream areas. Storage constraint is given below in section 1.3.

$$1378 \leq S_t \leq 1550 \qquad (1.3)$$

Objective function: The objective of the proposed model is to maximize the power generation with minimum damage to the livestock. The function of energy generation in kilowatt hour during time t is [3]:

$$KWH_t = 2725.I_t.nh_t.\eta \qquad (1.4)$$

The relationship described in 1.4 used in formulating the first part of objective function, which maximizes the energy generation. Complete objective function is as below:

$$\text{Maximize } [\{2725((I_t.nh_t.\eta) + \Sigma.nh_t.\eta \, E(I_t))\} - C_t \, e^{-\lambda}] \qquad (1.5)$$

First part in 1.5 is further subdivided into two categories, one being deterministic and other stochastic. Deterministic part uses the current period inflow to estimate the power generation in kilowatt for time t and stochastic part uses expected future inflow which is estimated by Bayesian prediction. Second part contains a cost function which estimates the cost of damage in result of excess release, in downstream areas. Ct is a fixed cost assigned to each unit of excess release. The cost function is exponential in nature because increasing each unit of excess release more than a certain level cause much more damage to livestock in downstream areas. Whereas, in the event of outflow exceeding 500,000 cusecs, it is apprehended that the villages situated on right and left banks of the Indus River for about 40 Km downstream Tarbela would be inundated [10]. So we fixed λ as described in 1.6.

$$\lambda = \max (0, Ot-2.88) \qquad (1.6)$$

Complete Operating Policy Model

$$\text{Maximize } \quad 2725[(I_t.nh_t.\eta) + \Sigma.nh_t.\eta \, E(I_t)] - C_t \, e^{-\lambda}$$

Subject to:

$$S_t = I_t + R_t - E_t + S_{t-1} - O_t$$

$$0.89 \leq O_t \leq 2.88$$

$$1378 \leq S_t \leq 1550$$

Applications

Study area

Tarbela reservoir is earth and rock filled dam which started functioning in 1974-1975. It is located on the Indus River. It is built over the River Indus near the small town of Tarbela in the Haripur District of the country. It is also the second largest dam in the world in terms of reservoir capacity, which is 11.62 million acre-feet (14.3 billion cubic meters). The dam is 485 feet (148 m) high above the riverbed. The dam forms the Tarbela Reservoir, with a surface area of approximately 250 square kilometers (97 sq. m). The dam was completed in 1976 and was designed to store water from the Indus River for irrigation, flood control, and the generation of hydroelectric power.
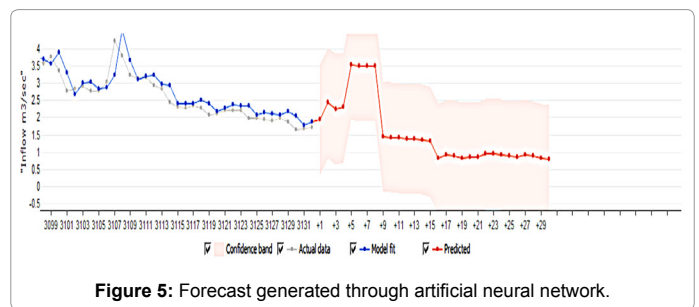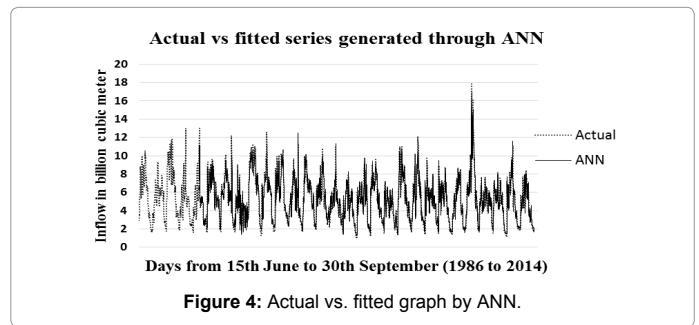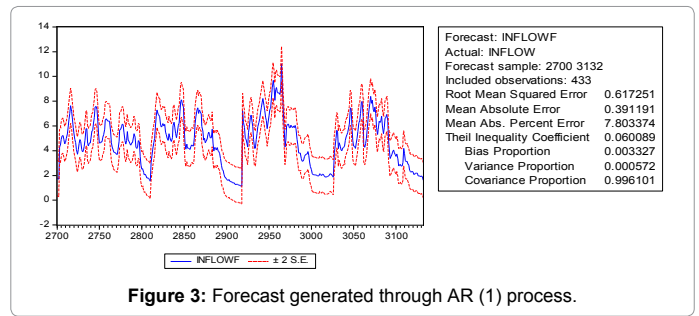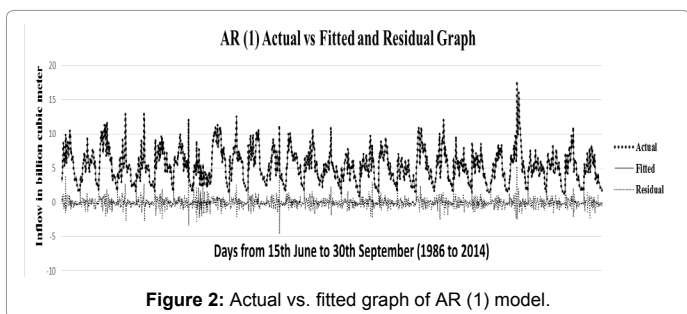
Results and Discussion

The series of historical inflow at dam during past 28 years from 1986 to 2014 is shown in Figure 1. Inflows at dam are measured every six hours. Average value of 24 hours inflow is taken as one day observation.

Water inflow in billion cubic meter is given on y axis and days (from 1986 to 2014) are given on x axis. The series in Figure 1 seems stationary as it moves around the mean line with no trend pattern. We also check the stationary of the dataset by using software package Eviews, before going for any forecasting technique for incoming flow of water in dam. Table 3 is showing the result of Unit Root Test for stationary. Test is

Figure 1: Historical inflows pattern.
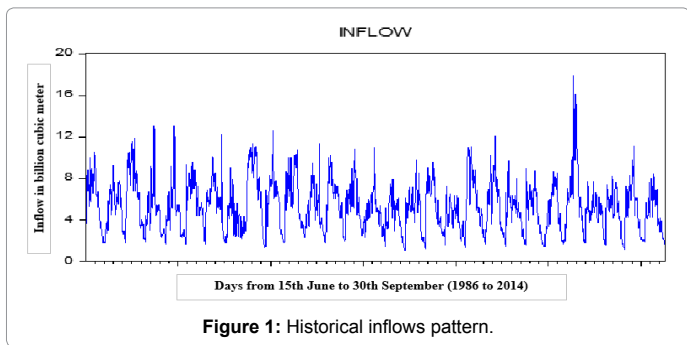
Variables		t-Statistics
Augmented Dickey-Fuller test statistics		-8.785166
Test critical values	1% level	-3.432253
	5% level	-2.862266
	10% level	-2.567201

Table 3: Unit root test critical values.

Figure 2: Actual vs. fitted graph of AR (1) model.

Figure 3: Forecast generated through AR (1) process.

Figure 4: Actual vs. fitted graph by ANN.

Figure 5: Forecast generated through artificial neural network.

applied to original series of data without any differencing to original series. The results shows original series is stationary without any taking differences.

The t-Statistics value for Augmented Dickey-Fuller test statistic is -8.785166, which is greater than t-Statistic values against all level of significance in absolute form. The coefficient against the variable Inflow (-1) in Table 4 is also negative i.e. -0.048742. These results are showing the stationarity of the original series.

After examining the stationarity of the data, we used AR(1) model to generate forecast for future events. Figure 2 shows the actual vs. fitted graph of AR (1) process along with observed residuals pattern.

In Figure 2 water inflow in billion cubic meter is given on y axis and days (from 1986 to 2014) are given on x axis. AR (1) seems the best fitted to original series. Plane line shows the actual series and series with big dots shows the fitted model to the data. Small doted series shows the residual behavior. Residual series is random in behavior.

In Figure 3 water inflow in billion cubic meter is given on y axis and days are given on x axis. Forecast is generated for last three years data from 2012 to 2014 through Autoregressive process of order one. The forecasted series remains in confidence bands of ± 2 standard error. Errors summary is also shown in the right side of the forecast. Root mean square is 0.62 approximately. Mean absolute error is also very minimum with value 0.391191. Bias proportion is minor and covariance proportion is high. Results show the excellent forecast.

Another classical method used for generating forecast was Artificial Neural Network (ANN). Figures 4 and 5 show the results for ANN mechanism. In Figure 4 water inflow in billion cubic meter is given on y axis and days (from 1986 to 2014) are given on x axis. This figure exhibit the predicted vs. actual inflows generated through ANN mechanism. The plane line shows the model fit and dotted line shows actual inflows in preceding years. Fitted series shown in plane line, started after tracking 500 observations of actual inflows. As in step one the ANN mechanism trains the data and after pattern recognition fit the model. The series seems good fit. To examine the forecast we take a magnified image of generated forecast in Figure 5.

In Figure 5 water inflow in billion cubic meter are given on y axis and days are given on x axis. The figure presents the predicted values more clearly in shaded area. Forecast is generated for the inflows from 15th June to 15th July, in this time inflows are low than other peak times so fitted and forecasted series exhibit a downward trend.

Then third and last method used to generate forecast for prediction of future inflows was Bayesian Predictive distribution. For this purpose first we find the distribution of the actual and forecasted inflows by using software Easy fit. Distribution of actual and predicted inflows was estimated by using statistical software Easy Fit. The goodness of fit tests used by Easy Fit is Kolmogorov Simirnov. The hypothesis made is given below.

H0: The data follow the specific distribution.

HA: The data do not follow the specified distribution.

In Figure 6 f(x) is the frequency of water inflow in billion cubic meters which are shown on y axis and water inflow in billion cubic meter are shown on x axis. The forecasted data distribution by EasyFit is shown in Figure 6 that comes out normal distribution. This distribution
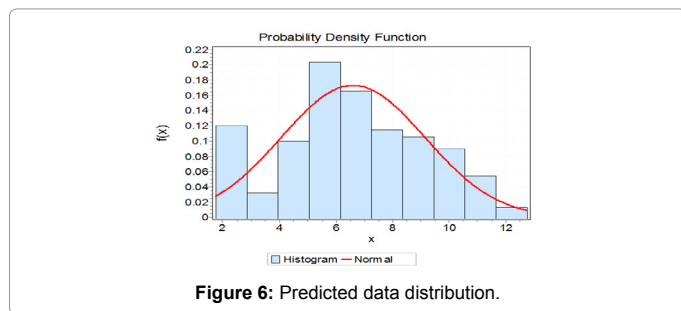

Figure 6: Predicted data distribution.

Variable	Coefficient	Std. Error	t-Statistic
INFLOW(-1)	-0.048742	0.005548	-8.78517

Table 4: Unit root test coefficient values.

Kolmogorov-Smirnov Statistic	P-Value
0.05367	0.0905

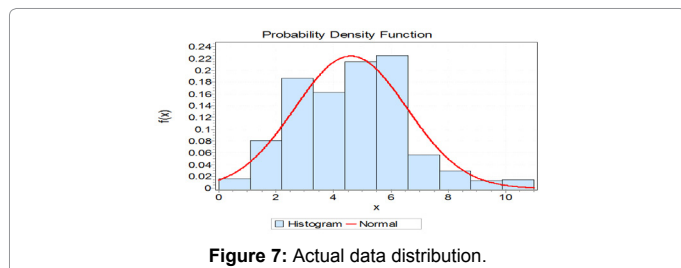**Table 5:** Predicted data goodness of fit test results.

Figure 7: Actual data distribution.

Kolmogorov-Smirnov Statistic	P-value
0.06767	0.01332

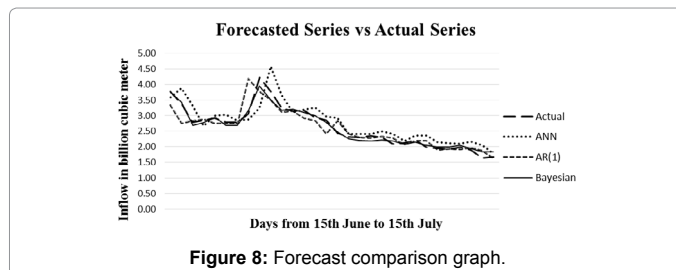**Table 6:** Actual data goodness of fit test results.

Figure 8: Forecast comparison graph.

Day	Actual	ANN	AR(1)	Bayesian	MAE ANN	MAE AR(1)	MAE Bayesian
1	3.77	3.58	3.34	3.74	0.19	0.42	0.03
2	3.37	3.89	2.75	3.44	0.52	0.62	0.07
3	2.78	3.31	2.84	2.69	0.53	0.06	0.09
4	2.86	2.68	2.89	2.78	0.18	0.03	0.08
5	2.92	2.99	2.75	2.95	0.08	0.16	0.03
6	2.78	3.04	2.75	2.69	0.26	0.02	0.09
7	2.78	2.84	2.75	2.69	0.06	0.03	0.09
8	3.06	2.88	4.19	3.17	0.17	1.13	0.11
9	4.22	3.24	3.77	3.95	0.98	0.45	0.27
10	3.79	4.57	3.5	3.5	0.77	0.29	0.29
11	3.23	3.66	3.12	3.18	0.43	0.11	0.05
12	3.14	3.1	3.15	3.2	0.04	0	0.06
13	3.17	3.19	2.92	3.1	0.02	0.25	0.07
14	2.94	3.26	2.84	3	0.31	0.11	0.06
15	2.86	2.97	2.42	2.78	0.11	0.44	0.08
16	2.44	2.94	2.84	2.48	0.5	0.4	0.04
17	2.32	2.41	2.42	2.26	0.09	0.1	0.06
18	2.29	2.42	2.3	2.2	0.12	0.01	0.09
19	2.35	2.42	2.28	2.19	0.07	0.07	0.16
20	2.29	2.5	2.33	2.22	0.2	0.04	0.07
21	2.1	2.41	2.28	2.19	0.32	0.18	0.09
22	2.12	2.18	2.08	2.09	0.06	0.04	0.03
23	2.21	2.36	2.19	2.15	0.15	0.02	0.06
24	1.98	2.36	2.19	2.04	0.38	0.21	0.06
25	1.95	2.16	1.9	2	0.21	0.05	0.04
26	1.93	2.12	1.93	2	0.2	0	0.07
27	1.98	2.1	1.91	2.04	0.11	0.07	0.06
28	1.9	2.17	1.97	1.95	0.27	0.07	0.05
29	1.64	2.05	1.88	1.84	0.41	0.24	0.2
30	1.67	1.79	1.63	1.84	0.12	0.04	0.17
Total					7.88	5.68	2.72

Table 7: Comparison of forecast generated thorough different methods.

is taken as current information in Bayesian estimation. Goodness of fit test results are given below. The goodness of fit test is accepted at the significance level of 5% (Table 5).

In Figure 7 Actual data distribution also comes out normal. Table 6 shows the goodness of fit results calculated through Kolmogrov Simirnov test. Then taking normal distribution as sample information with conjugate prior, posterior distribution also comes out normal. We use posterior mean value as perdition for upcoming inflows at dam. Then we compare the forecast generated by all three methods in Table 7.

In Figure 8 and Table 6 forecast comparison is shown for one month time period. It can be seen that Bayesian forecast is very close to actual inflow values as compared to other forecasting methods. ANN

mechanism gives poor forecast and forecast generated through AR (1) is better than but not as accurate as Bayesian forecast. Also mean absolute error of all three techniques is calculated. Bayesian process predicts inflows with minimum MAE i.e., 2.72. So we chose Bayesian forecast to put in BSDP model.

Then we put all estimated and actual data values in BSDP model. And to run this BSDP model we use programming software GAMS. At different values of inflows, GAMS generate output for outflow, damage cost, storage and profit i.e., at inflow value 2.48. Outflow is 1.2284500 billion cubic meter. Storage is 1402 ft. as λ is zero it means there is no excess outflow at June 15 hence no damage cost. The profit generated by hydropower generation through release is 5.537 million USD. Continuing the same way proposed policy model gives the best estimates of storage, outflow, profit and damage cost against all predicted inflows.

Conclusion

Study presents the inflow prediction and development of operating policy model in extreme weather conditions. Application of this model made the reservoir to function in a systematic way. Model results specify volume of release and level of storage at given inflow levels. Also it gives the cost and profit on current policy. The proposed model takes daily inflow forecast values as input and gives estimate of release, risk, storage and profit from release. This estimate leads to optimum decision in release of water from the dam. This study can be extended by adding connected reservoirs network information in proposed BSDP model. Outflow of previous reservoir becomes inflow to the next and looking at lag time for the inflow all connected nodes can be best operated through a single policy model. Reservoir sedimentation is also a major issue of Tarbela reservoir; it can be studied with proposed policy model by adding sedimentation constraint. As reservoir sedimentation decreases storage level of water and dam needs complete run off of it. So water in peak time of inflow can be release in such a way to run off reservoir sedimentation.

References

1. Karamouz M, Haralambos VV (1992) Bayesian stochastic optimization of reservoir operation using uncertain forecasts. Water Resour Res 28: 1221-1232.

2. Karamouz M, Azadeh A, Ali M (2009) Probabilistic reservoir operation using Bayesian stochastic model and support vector machine. Adv Water Resour 32: 1588-1600.

3. Mujumdar PP, Nirmala B (2006) A Bayesian stochastic optimization model for a multi-reservoir hydropower system. Water Resour Manag 21: 1465-1485.

4. Turgeon A (2005) Solving a stochastic reservoir management problem with multilag autocorrelated inflows. Water Resour Res 41: 12-14.

5. Pereira MVF (1989) Optimal stochastic operations scheduling of large hydroelectric systems. International Journal of Electrical Power & Energy Systems 11: 161-169.

6. Karamouz M, Mark HH (1987) Comparison of stochastic and deterministic dynamic programming for reservoir operating rule generation. J Am Water Resour Assoc 23: 1-9.

7. Stedinger JR, Bola FS, Daniel PL (1984) Stochastic dynamic programming models for reservoir operation optimization. Water Resour Res 20: 1499-1505.

8. Meehl GA, Karl T, Easterling DR, Changnon S, Pielke JR, et al. (2000) An introduction to trends in extreme weather and climate events: observations, socioeconomic impacts, terrestrial ecological impacts and model projections. Bull Am Meteorol Soc 81: 413-416.

9. Arnim K, Wolfgang B, Daniel KW, Pieter VO (2016) Simulating the viability of water institutions under volatile rainfall conditions – The case of the Lake Naivasha Basin. Environ Modell Softw 75: 373-387.

10. Water and Power Development Authority (2011) Flood Management Manual: Tarbela Dam Project. WAPDA, Pakistan. pp: 1-22.