Assessment and Forecasting of Water Level of Hatia Dam of Ranchi, Jharkhand

Chandan Kumar Pandit

Research Scholar, University Department of Mathematics, Ranchi University, Ranchi, Jharkhand, India

Abstract: ***This paper presents an analysis of the water level of Hatia Dam, Ranchi, based on historical data from 2009 to 2020. The study identifies a positive trend in water levels during several months (i.e. for February, March, April, May, July, November, December) and seasons (i.e. for Summer, Monsoon, Post-Monsoon), with significant correlations observed between the water level and monsoon rainfall. Temperature, however, shows no significant correlation with water levels. The water level of Hatia Dam reaches its peak during the post-monsoon season meters and the lowest water level is observed during summer season. The ARIMA (0,1,1) model was employed for forecasting, predicting that the water level will reach 590.131 meters by 2030.***

Keywords: Water level, Reservoir management, Forecasting, ARIMA model, Hatia Dam

1. Introduction

Water shortage is one of the main problems in food security management and sustainable economic growth. For life to exist on Earth, sustainable water supply is essential [1,2]. Natural river and stream flows fluctuate significantly with the seasons, resulting in high flows (floods) or low flows (droughts) in different regions [3,4]. By holding water during periods of high flow and releasing it steadily during low flow periods, water reservoirs significantly contribute to mitigating these natural disasters.

Reservoirs have helped manage water resources and improved economic gains, improved ecological habitats. The water level of reservoir is influenced by the regional climatic conditions [5]. Climate change is affecting the rainfall pattern and influencing the water level of reservoir. The scarcity of rainfall also causing irrigation problem and farmers have to rely more on reservoir water for irrigation purpose [6]. Reservoirs are also very important for the generation of electricity. The safely operation of reservoir is very crucial for the security of livelihood of the region. The water level of a reservoir is very important parameter to analyze the safety of reservoir. The fluctuations in water level causes problem in implementation of services including flood control, water supply, and power production, as well as overall reservoir safety [7]. As of the end of 2020; 2068 dam breaches had been reported worldwide in fifty-seven countries (China excluded), according to data on dam failures provided by the International Commission on Large Dams (ICOLD) [8]. The rate of accuracy in forecasting of temporal rainfall is very important for the sustainable reservoir management strategy [9]. Therefore, maximizing the numerous water management challenges requires a high degree of accuracy in reservoir water level forecast. The water level of reservoir can be forecasted using data-driven approach in which historical water level of reservoir is used to train the model.

The purpose of this study is to analyze and forecast the water level of Hatia Dam using historical data and to assess the impact of climatic factors such as rainfall and temperature on water level variations. This study is significant as it provides insights into the water level trends of Hatia Dam, which is crucial for water resource management and disaster preparedness in the region. The forecasting model offers valuable predictions for future water levels, aiding in strategic planning.

2. Study Area

Ranchi is the capital of state Jharkhand and is situated on the Chota Nagpur plateau of eastern India. It falls under the subtropical climatic region but due to surrounded by dense forests, the climate of Ranchi is pleasing. The average annual rainfall of Ranchi is 1300 mm and average annual temperature is 24ºC [10]. Ranchi is situated on a height of mean sea level of 629 meter. The population of Ranchi, according to Census 2011, is 2914253 [11]. The huge population of city depends upon three major dams i.e. Hatia dam, Kanke dam and Getalsud dam, for its water consumption needs. For this study, only water level data of Hatia dam is considered. Hatia dam is situated on the Subarnarekha River and is built in 1963. The water of Hatia dam is used for Irrigation and water Supply. The reservoir has a full reservoir level of 596 meter.

3. Methodology

The daily rainfall and temperature data for the period 2009-20 were collected from website of National Centre For Environmental Information. The daily water level of Hatia dam for the period 2009-20 was obtained from the concerned authority. The daily data are then transformed into monthly, seasonal and annual data. The seasons are classified as Summer (March-May), Monsoon (June-September), Post-Monsoon (October-December) and Winter (January-February). The water level data were analyzed for trends using the Mann Kendall test and Sens slope method. Pearson’s correlation analysis of water level of dam was then analyzed against rainfall and temperature.

Mann-Kendall Test:
Mann-Kendall test is a non-parametric trend test which is used to determine whether or not there is a linear monotonic trend in a given time series data. It does not require the data to be normally distributed. The null hypothesis states that there is no monotonic trend, and this is tested against one of

three possible alternative hypotheses: (i) there is an upward monotonic trend, (ii) there is a downward monotonic trend, or (iii) there is either an upward monotonic trend or a downward monotonic trend.

Let $x_1, x_2, x_3, \ldots, x_n$ be data set of length n in a time series data. The indicator function Sgn$(x_i - x_j)$ is calculated as follows:

$$\text{Sgn}(x_i - x_j) = \begin{cases} 1, & x_i - x_j > 0 \\ 0, & x_i - x_j = 0 \\ -1, & x_i - x_j < 0 \end{cases}$$

Mean S and variance Var(S) of Sgn$(x_i - x_j)$ are calculated as follows:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(x_i - x_j)$$

And $\text{Var}(S) = \frac{[n(n-1)(2n+5) - \sum_t t(t-1)(2t+5)]}{18}$

Where t is the extent of any given tie.

Mann-Kendall test statistic Z is then computed as follows:

$$\mathbf{Z} = \begin{cases} \frac{S-1}{\sqrt{Var(S)}} & if\ S > 0 \\ 0 & if\ S = 0 \\ \frac{S+1}{\sqrt{Var(S)}} & if\ S < 0 \end{cases}$$

The value of Z>0 represents monotonic upward trend in data series whereas Z<0 represents monotonic downward trend.

Sen's Slope Method:

It is a non-parametric test which is highly efficient to find out linear trend in univariate data series. This can be applied to a data set having missing value and outliers in the data series.

The Sen's estimator β of slope is calculated as follows:

$$\beta = \text{Median}\left(\frac{x_j - x_i}{j - i}\right)$$

Where x_j and x_i are the data values at time j and i (j>i) respectively.

β>0 indicates an upward trend whereas β<0 indicates downward trend in the data series.

Pearson's Correlation Test

It is a bivariate correlation which measures the linear correlation between two sets of data. It is the ratio between the covariance of two variables and the product of their standard deviation. Hence it is essentially a normalized measurement of the covariance, such that the result always has a value between −1 and 1. Since correlation is a measure of linear relationship, a zero value does not mean there is no relationship. It just means that there is no linear relationship, but there may be a quadratic or any other higher degree relationship between the data points. Also, the correlation between one data point and another will now be explored. This is quite different from correlation between variables. The Pearson correlation between two data points, X and Y, is calculated as:

$$\text{Correlation}(X, Y) = \frac{Cov(X,Y)}{\sigma_X \times \sigma_Y}$$

Where Cov(X ,Y) is covariance of X and Y and is given by

$$\text{Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

Where $\bar{x}$ and $\bar{y}$ are mean of data set X and Y respectively and n is the sample size.

And σ_X and σ_Y are standard deviation of X and Y respectively.

The Pearson Correlation is the actual correlation value that denotes magnitude and direction, the Sig.(2-tailed) is the *p*-value that is interpreted to check the significance of correlation. If the *p*-value is less than 0.05, then correlation is statistically significant between the two data set and the *p*-value is more than 0.05, then the correlation is not statistically significant association between the two data set.

ARIMA model

ARIMA means Autoregressive Integrated Moving Average. It is used in the time series analysis for the forecasting upcoming series points. ARIMA consists of three components: Autoregressive (AR), Integrated (I) and Moving Average (MA). The autoregressive (AR) component indicates that the evolving variable is regressed on its own lagged values. The Moving Average (MA) component shows that the regression error is essentially a linear combination of error components, the values of which happened simultaneously and at distinct points in the past. The Integrated (I) component indicates that the values of the data have been replaced by the difference between those values and the previous values.

ARIMA models are denoted by ARIMA (p, d, q) where p is the number of auto-regressive order, d is the order of differencing applied to the time series and q denotes the number of moving average order of the data series. The parameters p, d and q are non-negative integers.

ARIMA models require the data set to be stationary. It can also be applied on non-stationary data set by eliminating the non-stationarity of the mean function by introducing an initial differencing step one or more times. ARIMA models for different p, d and q values are developed and then the best model is selected by using the Akaike information criterion (AIC) and Bayesian information criterion (BIC).

ARIMA (p, d, q) is given as follows:

$$\bar{z}_t = \emptyset_1 \bar{z}_{t-1} + \emptyset_2 \bar{z}_{t-2} + \cdots + \emptyset_p \bar{z}_{t-p} + a_t - \theta_1 \bar{z}_{t-1} - \theta_2 \bar{z}_{t-2} - \cdots \theta_q \bar{z}_{t-q} \quad (1)$$

Where $\bar{z}_t = z_t - \mu$ and a_t is the shock.

Equation (1) can be applied after finding the backward shift operator (B) as follows:

$$\emptyset(B)(1 - B)^d z_t = \theta(B) a_t \quad (2)$$

Normalized Bayesian information criterion (BIC)

It is used for the selection of model. It is developed by Gideon E. Schwarz in 1978. It is closely related to Akaike information criterion (AIC). It is possible to increase the maximum likelihood by adding parameters when fitting a model, but this can result in overfitting. BIC tries to resolve this problem by introducing a penalty term for the number of parameters in the model same as in AIC. However, the penalty term in BIC is larger than AIC.

The BIC is given as follows:

$$BIC = k\ln(n) - 2\ln(\hat{L})$$

Where $\hat{L} = p(x\,|\hat{\theta}, M)$ is the maximize value of the model M and $\hat{\theta}$ is the parameter value that maximize the likelihood function; x is the observed value and n is the sample size of the data set; and k= the number of parameters estimated by the model

4. Result

The annual mean water level of Hatia dam is 585.835 meter [Table 1]. The least water level is observed in June (584.003 meter). The south-west monsoon then kicks in mid of June and the water level starts to rise and reaches its maximum level of 587.818 meter in September. The south-west monsoon starts to move out of the region in mid-September. The water level then starts to decline due to uses of water and it reaches its lowest level in June. Seasonally, the highest water level is observed in post-monsoon (587.059 meter) and the lowest level is observed in Summer (584.667 meter).

Trend analysis of water level data of Hatia dam emphasizes significantly positive trend for February, March, April, May, July, November and December month. Seasonally, the positive trend is observed for Summer, Monsoon and Post-Monsoon. No significant trend is observed in the water level in winter season and annually [Table 2].

Correlation analysis between water level and monsoon rainfall shows significant positive correlation [Table 3]. However there is no correlation observed between temperature and water level.

ARIMA model is used to forecast the water level of Hatia dam for next 10 years i.e. upto year 2030. The water level data has seasonal pattern and hence data is non-stationary. The differencing of 1 makes the said data stationary. Hence differencing of 1 is fixed in the ARIMA model and the value of autoregressive and moving average parameter are changed to obtain the most accurate model for forecasting. The best suitable model is selected on the basis of normalized BIC value of model. The ARIMA (0,1,1) has the least normalized BIC (1.059) among the tested various models [Table 4]. The normalized BIC for each season has also been observed the least for ARIMA (0,1,1). The forecasted value of water level of Hatia dam is provide in Table 5.

Table 1: Mean water level of Hatia dam

Months/ Seasons	Mean Water level (in meter)
January	585.9835
February	585.6087
March	585.1722
April	584.6418
May	584.1597
June	584.0031
July	584.7195
August	586.737
September	587.8186
October	587.6696
November	586.9777
December	586.5322
Summer	584.6579
Monsoon	585.8196
Post-Monsoon	587.0598
Winter	585.7961
Annual	585.8353

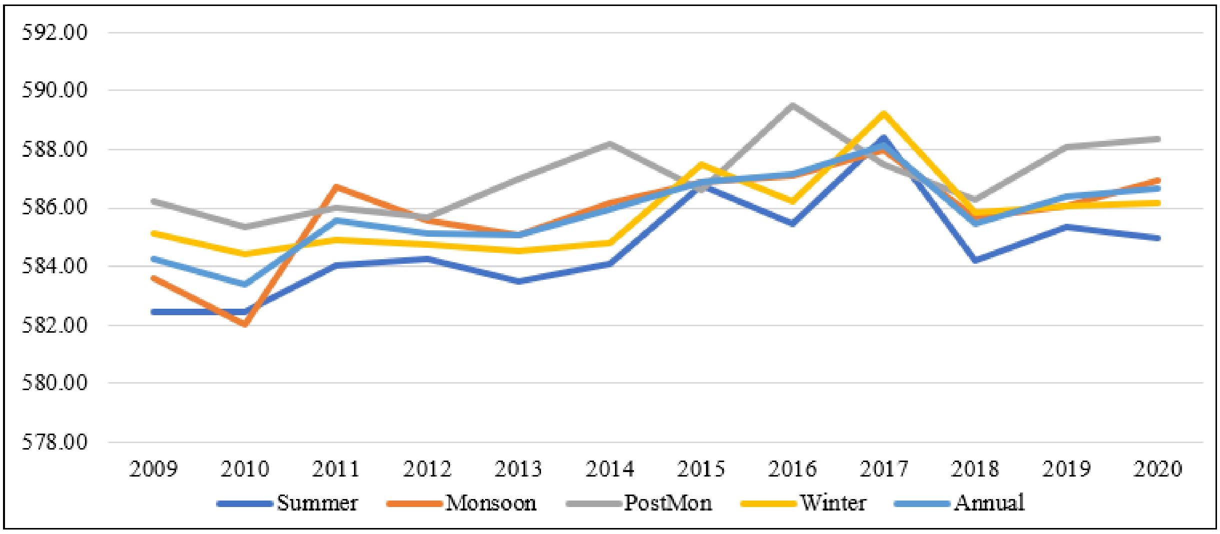

Figure 1: Seasonal Water Level of Hatia Dam

Table 2: MK test Z-statistics value for trend analysis

Months/Seasons	Z Statistics	Sen's slope method	Trend
January	1.85	0.158	
February	1.99	0.185	+
March	2.54	0.218	+
April	2.13	0.287	+
May	2.40	0.308	+
June	1.71	0.289	
July	2.26	0.342	+
August	1.71	0.292	
September	0.89	0.109	
October	1.30	0.164	
November	2.40	0.272	+
December	2.40	0.253	+
Summer	2.26	0.272	+
Monsoon	2.13	0.296	+
Post-Monsoon	2.26	0.247	+
Winter	1.85	0.179	
Annual	0.78	0.285	

Table 3: Seasonal correlation co-efficient

Seasons	Correlation Co-efficient Against Rainfall		Correlation Co-efficient Against Temperature	
	Co-efficient	p-value	Co-efficient	p-value
Summer	-0.198	0.538	0.087	0.787
Monsoon	0.623	0.030	-0.233	0.466
Post-Monsoon	-0.111	0.732	0.414	0.181
Winter	0.272	0.392	0.494	0.102
Annual	0.335	0.287	0.229	0.473

Table 4: Normalized BIC value for various ARIMA models

Models	**Normalized BIC**					
	Overall	Summer	Monsoon	Postmonsoon	Winter	Annual
ARIMA (0,1,1)	1.059	1.357	1.409	0.751	1.108	0.670
ARIMA (1,1,0)	1.211	1.429	1.514	1.123	1.197	0.793
ARIMA (0,1,0)	1.226	1.592	1.427	1.087	1.330	0.695
ARIMA (0,1,2)	1.382	1.682	1.734	1.076	1.429	0.986

Table 5: forecasted water level of Hatia dam

Water Level (Meter)	Summer	Monsoon	Post Monsoon	Winter	Annual
2021	586.334	587.8532	588.575	586.9273	587.637
2022	586.5919	588.166	588.8082	587.1013	587.9141
2023	586.8497	588.4789	589.0413	587.2753	588.1913
2024	587.1076	588.7917	589.2744	587.4493	588.4685
2025	587.3655	589.1046	589.5075	587.6233	588.7457
2026	587.6233	589.4175	589.7406	587.7974	589.0229
2027	587.8812	589.7303	589.9737	587.9714	589.3
2028	588.139	590.0432	590.2068	588.1454	589.5772
2029	588.3969	590.3561	590.4399	588.3194	589.8544
2030	588.6547	590.6689	590.673	588.4935	590.1316

5. Conclusion

This study confirms that the water level of Hatia Dam shows a positive trend during several months (i.e. for February, March, April, May, July, November, December) and seasons (i.e. for Summer, Monsoon, Post-Monsoon), with significant correlations between the water level and monsoon rainfall. The ARIMA(0,1,1) model proves to be effective in forecasting future water levels, predicting an increase to 590.131 meters by 2030. These findings are crucial for water resource management and planning in the Ranchi region.

References

[1] Farhad Zulfiqar, Mazhar Zubair, Raza Ullah,Chapter 22 - Climate-induced water scarcity and the effectiveness of community-based water resource management,Natural Resource Governance in Asia,Elsevier,2021,Pages 343-351,ISBN 9780323857291. https://doi.org/10.1016/B978-0-323-85729-1.00024-4.

[2] Cosgrove, W. J., and D. P. Loucks (2015), Water management: Current and future challenges and research directions, Water Resour. Res., 51, 4823–4839. doi:10.1002/2014WR016869.

[3] Poff NL, Olden JD, Merritt DM, Pepin DM, "Homogenization of regional river dynamics by dams and global biodiversity implications". Proc Nat Acad Sci, 104(14):5732–5737,2007. https://doi.org/10.1073/pnas.0609812104

[4] Postel, S. and Richter, B. (2012) Rivers for Life: Managing Water for People and Nature. Island Press, Washington.

[5] Salami, A. W., Ibrahim, H., & Sojobi, A. O. (2015). Evaluation of impact of climate variability on water resources and yield capacity of selected reservoirs in the north central Nigeria. Environmental Engineering Research, 20(3), 290-297.

[6] Rymuza, K., Radzka, E., Lenartowicz, T. (2015). THE IMPACT OF PRECIPITATION CONDITIONS ON MEDIUM-EARLY CULTIVARS OF POTATO YIELDING. Journal of Ecological Engineering, 16(3), 206-210. https://doi.org/10.12911/22998993/2958

[7] Coops, H., Beklioglu, M. & Crisman, T.L. The role of water-level fluctuations in shallow lake ecosystems – workshop conclusions. Hydrobiologia 506, 23–27 (2003). https://doi.org/10.1023/B:HYDR.0000008595.14393.77

[8] Wantzen, K.M., Rothhaupt, KO., Mörtl, M. *et al.* Ecological effects of water-level fluctuations in lakes: an urgent issue. *Hydrobiologia* 613, 1–4 (2008). https://doi.org/10.1007/s10750-008-9466-1

[9] Nima Ehsani, Charles J. Vörösmarty, Balázs M. Fekete, Eugene Z. Stakhiv,Reservoir operations under climate change: Storage capacity options to mitigate risk,Journal of Hydrology,Volume 555,2017,Pages 435-446.https://doi.org/10.1016/j.jhydrol.2017.09.008.

[10] Chandan Kumar Pandit, Anamol Kumar Lal, Uma Shanker Singh; "ASSESSMENT OF TEMPORAL RAINFALL AND ITS CORRELATION WITH TEMPERATURE IN RANCHI, JHARKHAND", International Journal of Current Research, Vol. 15, Issue 08, pp. 25519-25522, August, 2023.

[11] District Census Handbook Ranchi, series-21,Part Xii-B; Directorate of census operations, Jharkhand.