Modern Education
and Computer Science
PRESS

Trend Analysis and Forecasting of Water Level in Mtera Dam Using Exponential Smoothing

Filimon Abel Mgandu
College of Business Education, P. O. Box 2077, Dodoma, Tanzania.

Mashaka Mkandawile , Mohamed Rashid
University of Dar es salaam, P. O. Box 35091, Dar es salaam, Tanzania.

Abstract: This study presents trend analysis and forecasting of water level in Mtera dam. Data for water level were obtained from Rufiji Basin Development Authority (RUBADA). The study analyzed trend of water level using time series regression while forecasting of water level in Mtera dam was done using Exponential smoothing. Results revealed that both maximum and minimum water level trends were decreasing. Forecasted values show that daily water level will be below 690 (m.a.s.l) which is the minimum level required for electricity generation on 2023. It was recommended that proper strategies should be taken by responsible authorities to reduce effects that may arise. Strategies my include constructing small dams on upper side of Mtera dam to harvest rain water during rainy season as reserves to be used on dry season. In long run Tanzania Electric Supply Company (TANESCO) should invest into alternative sources of energy.

Index Terms: Trend Analysis, Time series, Forecasting, Exponential smoothing, Mtera dam.

1. Introduction

1.1 Background Information

Water resource refers to the surface water or groundwater available in a given area [18]. Some sources of water include rivers, reservoirs, lakes and groundwater. Although the role of water in human development is important, observed trends and climate forecasting provides evidence that water resources are vulnerable and will be strongly affected by climate change and industrialization which place key challenges to their existence [5]. Intergovernmental Panel on Climate Change (IPCC) refer to climate change as any change in average weather conditions over a long period of time as a result of natural factors or human activities [8]. On contrary, the United Nations Framework Convention on Climate Change (UNFCCC) refers to climate change as change in average weather conditions over a long period of time that is caused directly or indirectly by human activities [6].

Global climate change is the main cause of frequent droughts and floods events and their profound impacts are huge particularly on water resources. According to [8], the frequencies of floods and droughts are certain to increase in many parts of the world a fact that call for a longterm strategic planning of water resources and water utilization. In the same context, [1] suggested that quantification of the impacts of climate change on water resources is important as it provide useful input data for planning and management of water resources.

According to [17], around 1,800 million people will be living in places with total water scarcity and two-thirds of the world population could be under water stress situations by year 2025 developing countries being most affected. For the case of African continent, about 25% of her population is suffering water scarcity situation [5]. It is predicted that about 65% of the African population will experience water stress condition by the year 2025 which demand more strategies to overcome the situation [4]. Tanzania in particular, has observed a number of climate associated disasters including floods, droughts, failure of agricultural production in some years and increase of climate related diseases like malaria and others [10]. Study by [19] reported that in East African, Tanzania has been severely impacted by climate change where decreasing water quantity in water resources has been observed and more effort is required to quantify the impacts of climate change on water resources and appropriate strategies be taken.

Global climate changes and increasing demands of water resource as a result of population increase have forced water level regimes of lakes and reservoirs to change [8]. Study by [20] indicated that man-made reservoirs tend to experience greater water level fluctuations than natural lakes causing annual water level fluctuations of tens of meters in electricity

generating dams. It is important to understand the impacts of climate change in Tanzanian's hydropower dams to prepare appropriate measures.

Water resources in Tanzania are earmarked to face non-uniform climate change impacts that will lead to fluctuations water levels in most basins [12]. Increasing river flow in some basins, drying - up of some water bodies and intrusion of sea water into fresh water bodies including groundwater are expected [16]. In the recent past, two - thirds of rivers in Tanzania experienced reduced water volume due to decreased rainfall and increased temperature, which are key indicative parameters in climate change.

Time series forecasting refers to the process of predicting the future values using mathematical models by understanding the behavior of historical values in a time series [13]. A time series is a set of data points measured over successive time intervals and is mathematically defined as a set $x(t), t = 1,2,3, ...$ where t denotes time.

This study investigated trends of water level in Mtera dam. It also forecasted future values of water level to understand future scenarios. Data for water level in Mtera dam were obtained from the Rufiji Basin Development Authority (RUBADA). Water levels were measured in meter above sea level (m.a.s.l).

1.2 Statement of the problem

Water resources and water volumes are critical inputs to Tanzania major socio-economic activities including food security through irrigation agricultural sector performance, water for industries and water for people (clean and safe water for good health workforce). Furthermore, water in rivers and reservoirs is pivotal to electric powering industrial and household activities [7]. Therefore, changes in water flow affects Tanzania Electric Supply Company (TANESCO) capacity to produce and supply electricity hence hinders Tanzania government machineries capacity to support socio-economic activities of her fast-growing populations [116].

Mtera dam has been subject to fluctuations of water level consequently causing problems to Tanzania Electric Supply Company (TANESCO) and to people's properties. Therefore, forecasting water levels of the dam has attracted attention of the researchers in the country. This study investigated the trend of water level and forecasted water level in Mtera dam to understand the situation for five years to come.

1.3 Related studies

Different studies have been done in different areas of the World on forecasting water level on lakes and reservoirs. Foristance, the study by [11] employed Neural Networks and six feature groups comprising of water levels, rainfall, evaporation rate, discharges for rivers Malewa and Gilgil and one pair of time harmonics were used to develop neural network models to forecast water levels for Lake Naivasha in Kenya. The neural network models developed were able to forecast effectively the reservoir levels for the lake for four consecutive months after a given month and given data for six consecutive months prior to the month. It was found that the more the number of feature groups used, the higher the ability of neural networks to forecast accurately the reservoir levels. Data compression generally reduced the size and computation time of the models.

The study by [15] forecasted groundwater level at Maheshwaram watershed, Hyderabad, India. The study employed artificial neural network (ANN) model to forecast water levels. The model efficiency and accuracy were measured based on the root mean square error (RMSE) and regression coefficient (R2). The model provided the best fit and the predicted trend followed the observed data closely (RMSE = 4.50 and R2 = 0.93). It was observed that ANN was precise and accurate groundwater level forecasting. Also, [3] forecasted surface water level fluctuations of lake Van by artificial neural networks. It is concluded that ANN and ARMAX models are very useful for the short-term predictions of the time series data. It was seen that ANN models outperform ARMAX models.

This study used Exponential smoothing approach in forecasting of water level. Exponential smoothing assigns exponentially decreasing weights where greater weight is given to most recent observations and less weight is given to oldest observations in water levels time series data.

2. Research methods

2.1 Trend analysis

A trend is a tendency or movement of the data at a specific period of time. In this study line charts were used to show how water levels changed over time. To test whether a linear trend occurs, a time series regression equation 1 was used with time t as independent variable and water level y as dependent variable.

$$y = \beta_0 + \beta_1 t + \varepsilon \qquad (1)$$

Where β_1 is the slope coefficient and β_0 is the least square estimate of the intercept. The following hypothesis was used to test the linear trend:

H0: $\beta_1 = 0$ and

H1: $\beta_1 \neq 0$.

If the null hypothesis is rejected, the slope is significantly different from zero and the linear trend is significant at α=0.05

I.J. Mathematical Sciences and Computing, 2020, 4, 26-34

level of significance meaning that there are changes over time in water level.

2.2 Forecasting

On other hand, exponential smoothing approach was used in forecasting of water level. Exponential smoothing assigns exponentially decreasing weights where greater weight is given to most recent observations and less weight is given to oldest observations [7]. Most recent data are seen as more relevant to the model and are given more weight compared to the oldest observations. Triple exponential smoothing is an addition to double exponential smoothing to capture seasonality in time series data. Three smoothing parameters are used: alpha denoted by α for level equation, beta denoted by β to capture trend and gamma denoted by γ to capture seasonality in time series. m denote the frequency of the seasonality, k is the integer part of $\frac{h-1}{m}$ and h= 1,2,…. The additive model equation 2 is desired when the seasonal trend is of the similar magnitude throughout the data set, while the multiplicative model equation 3 is desired when the magnitude of seasonality changes with time [7].

The component form for additive method is:

Forecast equation: $\qquad \hat{y}_{t+h|t} = l_t + hb_t + s_{t+h-m(k+1)}$

Level equation: $\qquad l_t = \alpha\,(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1})$

$$\tag{2}$$

Trend equation: $\qquad b_t = \beta\,(l_t + l_{t-1}) + (1 - \beta)b_{t-1}$

Seasonality equation: $\qquad s_t = \gamma\,(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)\,s_{t-m}$

The component form for multiplicative method is:

Forecast equation: $\qquad \hat{y}_{t+h|t} = (l_t + hb_t)\,s_{t+h-m(k+1)}$

Level equation: $\qquad l_t = \alpha\,\frac{y_t}{s_{t-m}} + (1 - \alpha)(l_{t-1} + b_{t-1})$

$$\tag{3}$$

Trend equation: $\qquad b_t = \beta\,(l_t + l_{t-1}) + (1 - \beta)b_{t-1}$

Seasonality equation: $\qquad s_t = \gamma\,\frac{y_t}{l_{t-1}+b_{t-1}} + (1 - \gamma)\,s_{t-m}$

2.3 Model selection criteria

Akaike's information criterion (AIC) and Bayesian Information Criterion (BIC) are the statistics that helps in model selection. They compare the superiority of a set of statistical models to each other and lower AIC or BIC value shows a better fit [14]. AIC and BIC are defined in equation 3 to equation 5 and in each case y is the observed data, n is the number of observations, k is the number of parameters to be approximated and L is the maximized value of the likelihood function for the estimated model.

$$AIC = -2 \cdot log\,L + 2k \tag{3}$$

For small sample $\left(\frac{n}{k} < 40\right)$ AIC_C is used

$$AIC_C = -2 \cdot log\ L + 2k + \frac{2k+1}{n-k-1} \tag{4}$$

$$BIC = -2 \cdot ln\ L + k \cdot ln(n) \tag{5}$$

2.4 Forecast Performance Measures

[2] suggested various performance measures including Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Percentage Error (MPE), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) to approximate forecast accuracy and to compare different models. According to [2] performance measures are defined in equation 6 to equation 9 where in each definition y_t is the actual value, f_t is the forecasted value, $e_t = y_t - f_t$ is the forecast error and n is the size of the test set, $\hat{y} = \frac{1}{n}\sum_{t=1}^{n} y_t$ is the test mean and $\sigma^2 = \frac{1}{n-1}\sum_{t=1}^{n}(y_t - \hat{y})^2$ is the test variance.

Mean Absolute Error (MAE)

$$MAE = \frac{1}{n} \sum_{t=1}^{n} |e_t| \tag{6}$$

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left|\frac{e_t}{y_t}\right| . 100 \tag{7}$$

Mean Percentage Error (MPE)

$$MPE = \frac{1}{n} \sum_{t=1}^{n} \left(\frac{e_t}{y_t}\right) . 100 \tag{8}$$

Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} e_t^2} \tag{9}$$

3.Results and discussion

3.1 Trend analysis of water level.

Maximum and minimum water levels were considered and the results are summarized in Table 1. Trend for minimum water level was significant while trend for maximum water level was not significant at $\alpha = 0.05$.

Table 1: Trends of water level

Variable	Trend	p value	Status
Maximum water level	Decreasing	0.117	Not significant
Minimum water level	Decreasing	0.011	Significant

Figure 1 shows a decreasing trend of maximum water level in the dam. However, the trend is not significant at $\alpha=0.05$ with p > 0.05 as shown in Table 1. Figure 2 shows a decreasing trend of minimum water level in the dam. The trend is significant at $\alpha =0.05$ with $p \le 0.05$ as shown in Table 1.
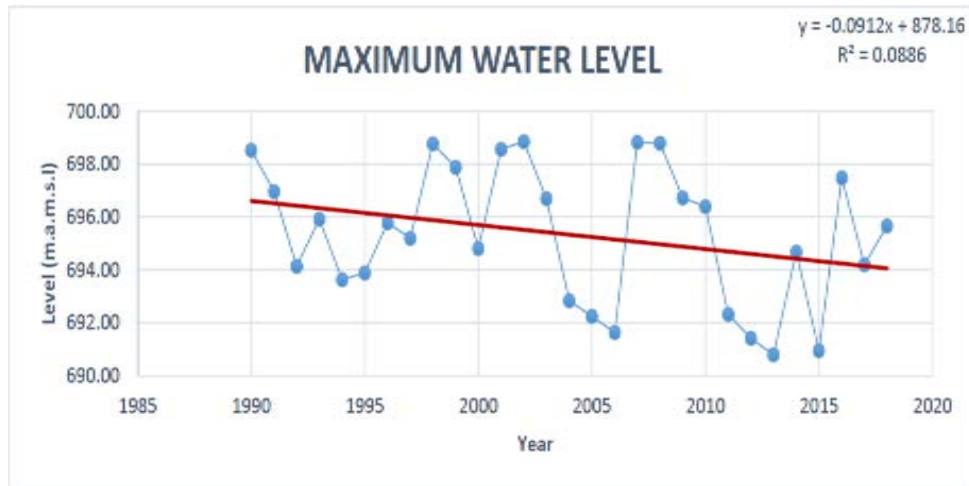
Figure 1: Maximum water level trend.

Figure 2: Minimum water level trend.

Trend analysis indicates that water level in Mtera dam is decreasing over time. Both maximum and minimum water level trends were negative however for maximum it was not statistically significant at $\alpha = 0.05$ compared to minimum trend which was significant. These results are in line with the study by [9] which indicated that water levels in storage reservoirs have gone down and competition for water between farmers and hydropower generators has increased.

3.2 Forecasting water level in Mtera dam

The models were estimated in R software using the ets() function in the forecast package. ETS stands for Error, Trend and Seasonal respectively. It is possible to form 15 models with additive errors and 15 models with multiplicative errors [7]. To determine the best model out of 30 ETS models, several criteria such as Akaike's Information Criterion (AIC), Akaike's Information Criterion correction (AICc) and Bayesian Information Criterion (BIC) were used and the aim was to find the model with small AIC,AICc and BIC. Figure 3 shows the monthly water level in Mtera dam.
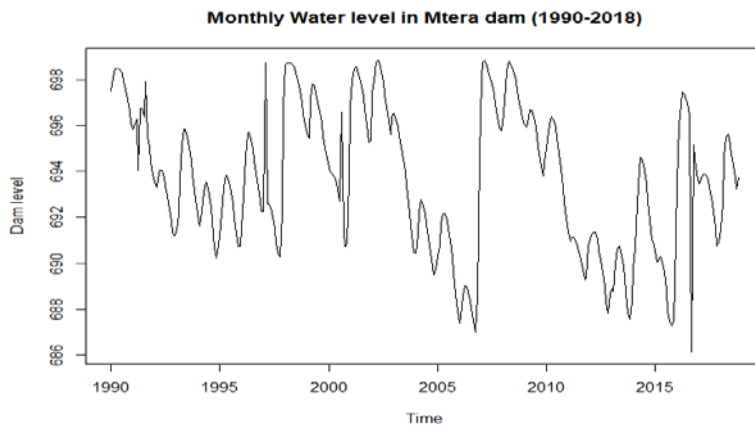
Figure 3: Monthly water level in Mtera dam.

ets() function in forecast package suggest ETS(A,A,A) as appropriate model for monthly water level data in Mtera dam. The model is triple exponential smoothing with trend and additive seasonal component and the smoothing parameters are shown in the Table 2.

Table 2: Smoothing parameters

Alpha (α)	Beta (β)	Gamma (γ)
0.8881	0.0001	0.0001

The AIC, AICc and BIC were used to determine which of the ETS models is most appropriate for a given time series. The AIC, AICc and BIC values for optimal model are shown in Table 3.

Table 3: AIC, AICc and BIC values

AIC	AICC	BIC
2168.921	12170.781	2234.3604

Figure 4 shows the residual analysis results from ETS(A,A,A) where most ACF values are within the 95%, indicating that there is no correlation amongst the residuals. Also Figure 4 is an indicator of the independence of the residual. It also shows that the residuals follow the normal distribution.
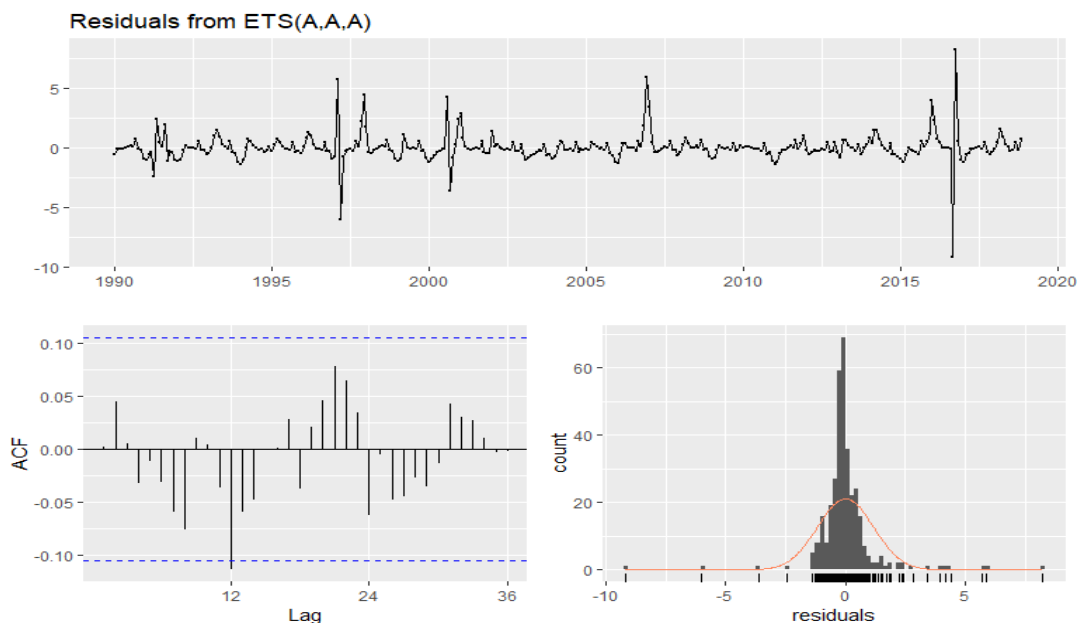

Figure 4: Residual analysis results from ETS(A,A,A)

The predicted values shows a good agreement with the observed values in Figure 5, indicating that ETS(A,A,A) model has satisfactory predictive ability.
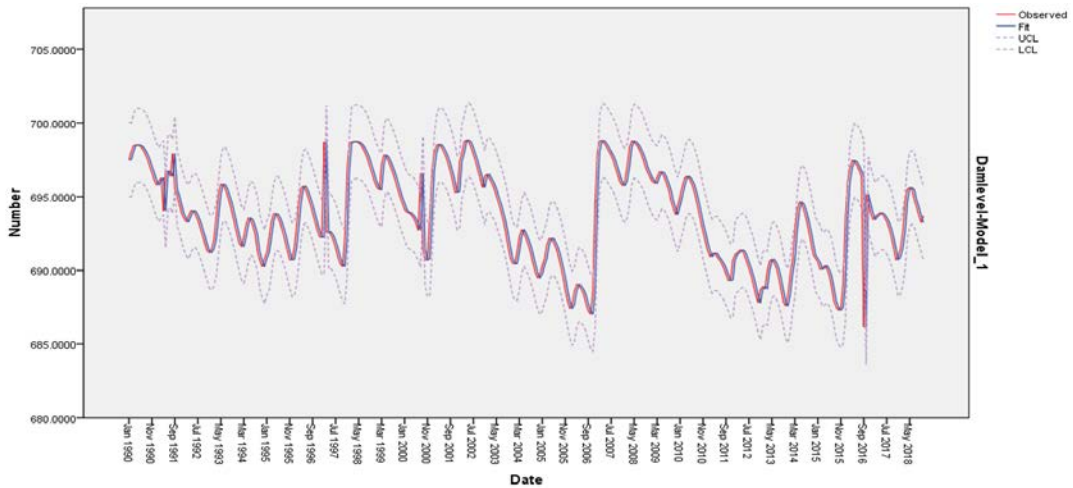

Figure 5: ETS(A,A,A) model validation.

The ETS(A,A,A) model was used to forecast monthly water level in Mtera dam for five years and the forecast are represented by Figure 6.
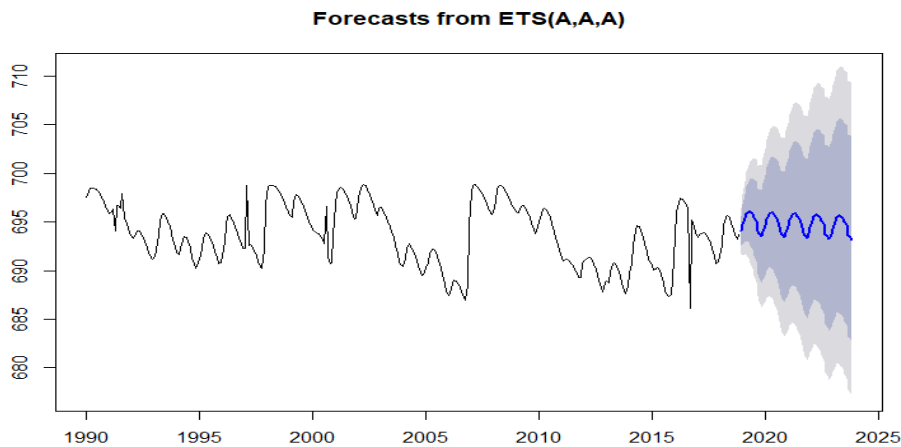

Figure 6: Five years monthly water level forecast from ETS(A,A,A)

Exponential smoothing methods were applied on Daily water level data and the results are summarized in Table 4.

Table 4: Smoothing parameters

Alpha (α)	Beta (β)	Gamma (γ)
0.25	0.00	0.25

Substituting the values for alpha(α), beta (β) and gamma (γ) on triple exponential smoothing for additive model (system of equations 2), the forecast model is obtained and was used for forecasting.

Different performance measure such as Mean Absolute Error (MAE), Mean Percentage Error (MPE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) for the forecasting model are shown in Table 5.

Table 5: Performance measures of ETS(A,A,A)

RMSE	MAE	MPE	MAPE
1.16369	0.592	-0.0002	0.0853

Forecasting daily water level data in Mtera dam was done for five years and the results are shown on Figure 7.
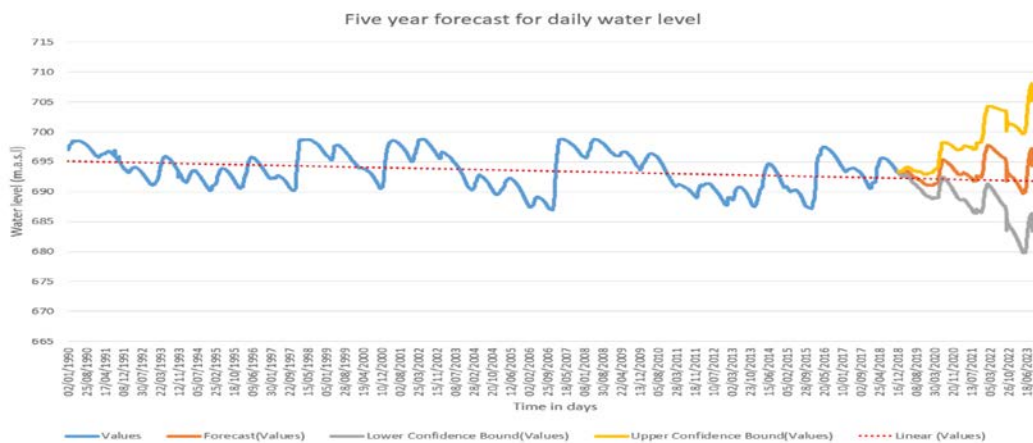
Figure 7: Forecasting with exponential smoothing

Forecasts for five years from exponential smoothing on daily water level data shows that water level will be below the 690 m.a.s.l, which is the minimum water level required for electricity generation on April and May 2023 and therefore TANESCO should prepare strategies to face this challenge.

4.Conclusions and Recommendations

The first aim of this study was to conduct trend analysis on water level. Findings indicate that water level in Mtera dam is decreasing over time. Both maximum and minimum water levels have negative trends. The second aim of this study was to forecast water level in Mtera dam. Forecasts for five years from exponential smoothing on daily water level data shows that water level will be below 690 m.a.s.l, which is the minimum water level required for electricity generation on April and May 2023.

Based on the results from this study that water level in Mtera dam will be below 690 m.a.s.l, which is the minimum water level required for electricity generation on April and May 2023, its recommended that proper strategies should be taken by responsible authorities to reduce the effect that may arise. Strategies may include, constructing small dams on the upper side of Mtera dam on the Great Ruaha river to harvest rain water when the main dam is at its maximum level (698 m.a.s.l). The stored water will be used to fill the main dam when it is approaching its minimum level (690 m.a.s.l). Also, Tanzania Electric Supply Company (TANESCO) should invest on alternative source of energy such as thermal, gas and coal power rather than depending on hydropower which is affected by climate change.

References

[1] Abbaspour KC, Faramarzi M, Ghasemi SS and Yang H. Assessing the impact of climate change on water resources in Iran. Water resources research. 2009; 45:10.
[2] Adhikari, R. and Agrawal, RK. An introductory study on time series modeling and forecasting. arXiv preprint arXiv: 2013; 1302.6613.
[3] Altunkaynak, A. Forecasting surface water level fluctuations of Lake Van by artificial neural networks. Water resources management. 2007; 21(2):399-408.
[4] Ashton PJ. Avoiding conflicts over Africa's water resources. AMBIO: A Journal of the Human Environment. 2002;31(3): 236-242.
[5] Bates BC, Kundzewicz ZW, Wu S and Palutikof JP. Climate change and water. Technical paper of the intergovernmental panel on climate change, IPCC secretariat, Geneva. Climate Change Policy with a Renewed Environmental Ethic. 2008; 21:85-101.
[6] Bodansky D. The United Nations framework convention on climate change: a commentary.Yale J. Int'l l. 1993;18:451.
[7] Hyndman RJ and Athanasopoulos G. Forecasting: principles and practice. 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. Accessed on May 2019. 2018
[8] IPCC. Report of the nineteenth session of the intergovernmental panel on climate change (IPCC) Geneva, 17-20 (am only) April 2002. 2007.
[9] Mdemu MV and Magayane MD. Conflict of water use between hydropower and irrigation in Tanzania: The conundrum of sectoral policy approaches to water resources development. 2005.
[10] Omambia CS and Gu. The cost of climate change in Tanzania: impacts and adaptations. Journal of American Science. 2010;6(3).
[11] Ondimu, S. and Murase, H. Reservoir level forecasting using neural networks: Lake Naivasha. Biosystems engineering. 2007: 96(1):135-138.
[12] Orindi VA and Murray LA. Adapting to climate change in East Africa: a strategic approach (No. 117). International Institute for Environment and Development. 2005.

[13] Raicharoen T, Lursinsap C and Sanguanbhokai P. Application of critical support vector machine to time series prediction. In Proceedings of the 2003 International Symposium on Circuits and Systems. 2003;5.

[14] Snipes, M. and Taylor, D.C. Model selection and Akaike Information Criteria: An example from wine ratings and prices. Wine Economics and Policy. 2014; 3(1):3-9.

[15] Sreekanth, P.D., Geethanjali, N., Sreedevi, P.D., Ahmed, S., Kumar, N.R. and Jayanthi, P.K. Forecasting groundwater level using artificial neural networks. Current science. 2009; 933-939.

[16] United Republic of Tanzania (URT). National adaptation programme of action (NAPA). Vice president's office, division of environment. Government printers, Dar es Salaam. 2007.

[17] UN-WATER. Water a shared responsibility. The United Nations, World Water Development Report 2 (601). UN-WATER/WWAP/2006/3. 2006.

[18] Water words dictionary. Division of Water Planning. 1998.

[19] Yanda PZ, Kangalawe RY and Sigalla RJ. Climatic and socio-economic influences on malaria and cholera risks in the Lake Victoria region of Tanzania. 2005.

[20] Zohary T and Ostrovsky I. Ecological impacts of excessive water level fluctuations in stratified freshwater lakes. Inland Waters. 2011;1(1):47-59.

Authors' Profiles

Filimon Abel Mgandu works at College of Business Education (CBE)- Department of ICT and Mathematics. He was awarded Bsc. With Education (Informatics and Mathematics) in 2016 from Sokoine University of Agriculture and MSc in Mathematical Modelling in 2019 from University of Dar es salaam. His research area includes Data science, Statistics and Computer programming.

Dr. Mashaka Mkandawile works at the University of Dar es Salaam as a Lecturer and Coordinator of Actuarial Science Programme, Department of Mathematics, College of Natural and Applied Sciences (CONAS). He obtained his PhD in Operations Research in 2015 at the University of Dar es Salaam with research work and licentiate at Umea University in Sweden. His research area includes Actuarial Science, Statistics and Probability Theories, Numerical Optimization and Computer Programming.

Dr. Mohamed Rashid works at the University of Dar es Salaam as a Lecturer, Department of Mathematics, College of Natural and Applied Sciences (CONAS). He has Masters and PhD both obtained at the University of Dar es Salaam. His research area includes Statistics and Probability Theories and Numerical Optimization.