2022

# Multi-variate regression analysis of lake level variability: A case of semi-closed, shallow rift valley lake in Northern Tanzania

Okwi, Gustavio

Elsevier

https://doi.org/10.1016/j.envc.2022.100533

Environmental Challenges 7 (2022) 100533

Contents lists available at ScienceDirect

Environmental Challenges

journal homepage: www.elsevier.com/locate/envc

# Multi-variate regression analysis of lake level variability: A case of semi-closed, shallow rift valley lake in Northern Tanzania

Gustavio Okwir[a,d,*], Sharma Pramod Kumar[b], Hongkai Gao[c], Juma Rajabu Selemani[a], Karoli N. Njau[a]

[a] School of Materials, Energy, Water and Environmental Sciences, Nelson Mandela African Institution of Science and Technology, P.O. Box 447, Arusha, Tanzania
[b] Indian Institute of Technology Roorkee, Haridwar, Uttarakhand 247667, India
[c] School of Geographical Sciences, East China Normal University, Shanghai, China
[d] Department of Biosystems Engineering, Faculty of Agriculture and Environment, Gulu University, P.O. Box 166, Gulu, Uganda

ARTICLE INFO

*Keywords:*
Lake level variability
Mann-Kendall
Sensitivity
Stepwise regression

ABSTRACT

Lakes are very important for domestic use, commercial purposes, and ecosystem sustenance; nevertheless, studies on how different stressors influence water resources are limited, posing challenges in the planning and managing of water resources. Lake Babati is one of the important lakes in the East African rift valley as a source of fish, drinking water, and hippopotamus habitat. The lake level has consistently declined, threatening the life of organisms depending on the lake. Nevertheless, no study postulated the reasons for its decline. Therefore, this study used statistical methods, regression analysis and HEC-HMS hydrological model to investigate the association and sensitivity of the hydrological components driving the lake level variability. Results showed that the Lake Babati level was significantly declining ($p < 0.01$) at a rate of 0.025 m yearly. The lake level variability was most sensitive to inflow, while outflow and evaporation had almost equivalent magnitude in driving the lake level variability and direct rainfall had the least influence. Although the lake level variability corresponded to changes in the basin supply components, the declining lake level trend was neither directly related to lake evaporation nor inflow as both parameters showed no significant trends. Furthermore, groundwater abstraction within the lakeshore is smaller than lake evaporation and is unlikely the main driver of the decline in lake level. Therefore, the declining lake level seemed related to the spillway outflow, which had been improved to avoid lake flooding resulting in large outflow during the peak seasons and probably the reason for the significant decline in the lake levels even in the rainy seasons. This study benchmarked the most sensitive hydrological drivers influencing lake level whose accurate monitoring and management could rescue the situation.

## 1. Introduction

Globally, numerous factors influence water resources, including climate change and anthropogenic factors (Herrnegger et al., 2021). These factors threaten to distort the vital roles water resources play in the sustenance of eco-hydrological systems and their co-evolution (Gilfedder et al., 2012; Gao et al., 2014). Therefore, studies on the responses of the water resources systems to the climatic and anthropogenic stresses are very significant for the planning and management of the water resources. The variability of water resources among the East African rift valley lakes under different stressors is widely studied, e.g. (Kebede et al., 2006, Olaka et al., 2010, Hassan and Jin, 2014, Deus et al., 2013, Swenson and Wahr, 2009, Darling et al., 1996). However, the studies have shown different trends for lakes in the same region. For example, Herrnegger et al. (2021) showed that Lake Baringo, Nakuru, Naivasha, Solai, Elementia (rift valley lakes in Kenya) have recently experienced level rises.

In contrast, Lake Manyara, a rift valley lake in northern Tanzania, showed variations on temporal scales. Lake Manyara declined in level after 2002 but experienced a level increase after 2006 – 07 due to the influence of Indian Ocean Dipole (Deus et al., 2013). While Lake Manyara trends are linked to rainfall variations, its neighbouring Lake Babati, an upstream lake within Lake Manyara catchment, consistently recorded a decline in water level. This unique behaviour of lakes within the East African rift valley has motivated several studies (Deus et al., 2013, Darling et al., 1996, Kumambala and Ervine, 2010, Weitz and Demlie, 2013, Mbanguka et al., 2016). However, the paucity of observed data and associated data quality and accessibility issues have limited the

* Corresponding author at: School of Materials, Energy, Water and Environmental Sciences, Nelson Mandela African Institution of Science and Technology, P.O. Box 447, Arusha, Tanzania.

*E-mail address:* okwirg@nm-aist.ac.tz (G. Okwir).

https://doi.org/10.1016/j.envc.2022.100533
Received 28 August 2021; Received in revised form 13 April 2022; Accepted 20 April 2022

understanding of some water systems (Duan et al., 2018). This problem is more common, especially among the East African rift valley lakes (Guma et al., 2019), despite their pivotal roles in providing water for domestic use, commercial purposes, and ecosystem sustenance.

Most studies focusing on lake water systems (Hassan and Jin, 2014, Darling et al., 1996, Weitz and Demlie, 2013, Missi and Atekwana, 2020, Duan et al., 2018) have used hydrochemistry, stable isotopes, and satellite-based observation methods to assess sources and variations of groundwater and surface water. However, pinpointing the actual drivers of lake level variability has been challenging since these methods cannot explicitly quantify the climatic and anthropogenic influences in catchments where the conjunctive use of groundwater and surface water is adopted. Yet, unsustainable water abstractions around such lakes could alter groundwater interactions, leading to further declines in lake levels and their associated secondary impacts (Zohary and Ostrovsky, 2011).

Lake Babati is small and shallow, with a reported maximum depth of 7.0 m. The semi-closed rift valley lake in Northern Tanzania is a source of fish and a habitat for hippopotamus. It has deep wells drilled along its shores to supply water to the burgeoning population of Babati town. Due to its small size, the lake responds rapidly to climatic variations and has flooded several times following heavy rainfall episodes, especially in 1964, 1979, and 1990 (Stromquist, 1992). The construction of an artificial outlet (spillway) at the northeastern part of the lake minimised the extreme flooding incidences observed before 1990. Although heavy rainfalls caused the flooding episodes, they may have been exacerbated by deforestation and land degradation (Sandstrom, 1995). Mbanguka et al. (2016) propound that cloudiness is a sensitive parameter to Lake Babati level variability due to its significant influence on evaporation. While the study by Mbanguka et al. (2016) applied integrated water balance modelling, it was based on a rough estimate of the lake size (drawn from only five transects of the lake bathymetry). However, lake size and morphology greatly influence lake sensitivity to climatic forcing (Olaka et al., 2010). The water resources managers are concerned about the consistent decline in the Lake Babati level, but none of the previous studies has addressed their concern. They suspected that the increasing number of boreholes around the lake shores could be driving the observed declines.

The poor understanding of the actual drivers of the lake water level variability and management concerns highlighted above, the importance of the lake to the communities and ecosystem, and availability of relatively quality secondary data motivated the choice of the study area. Hence, the present study applied statistical methods and regression analysis to determine the association and sensitivity of basin water supply and outflow components on the lake Babati levels to inform management of drivers to prioritise for intervention. Furthermore, it applied hydrologic modelling to assess the temporal variability of the other lake water balance components. The results of this study could offer new insights into the drivers of lake level variability and aid the management of water resources. Furthermore, this study could benchmark the most sensitive hydrological drivers for accurate monitoring and prioritising management interventions.

## 2. Materials and methods

### 2.1. Study area

The present study was conducted within Lake Babati catchment lying between Latitude 4°10′00′ S and 4°30′00″ S and Longitude 35°30′00″ E and 36°00′00″ E. The 390 km$^2$ catchment is located within the Manyara region of Tanzania, with the spread and extents as shown in Fig. 1. Gentle slopes characterise the terrain in the lake valley, while steep slopes define highlands at the southern catchment boundaries. The catchment has a mixed geological formation with superficial deposits of mainly dark brown soils derived from weathered rocks and red and brown soils near the lake (Driessen et al., 2001). The catchment is mainly com-

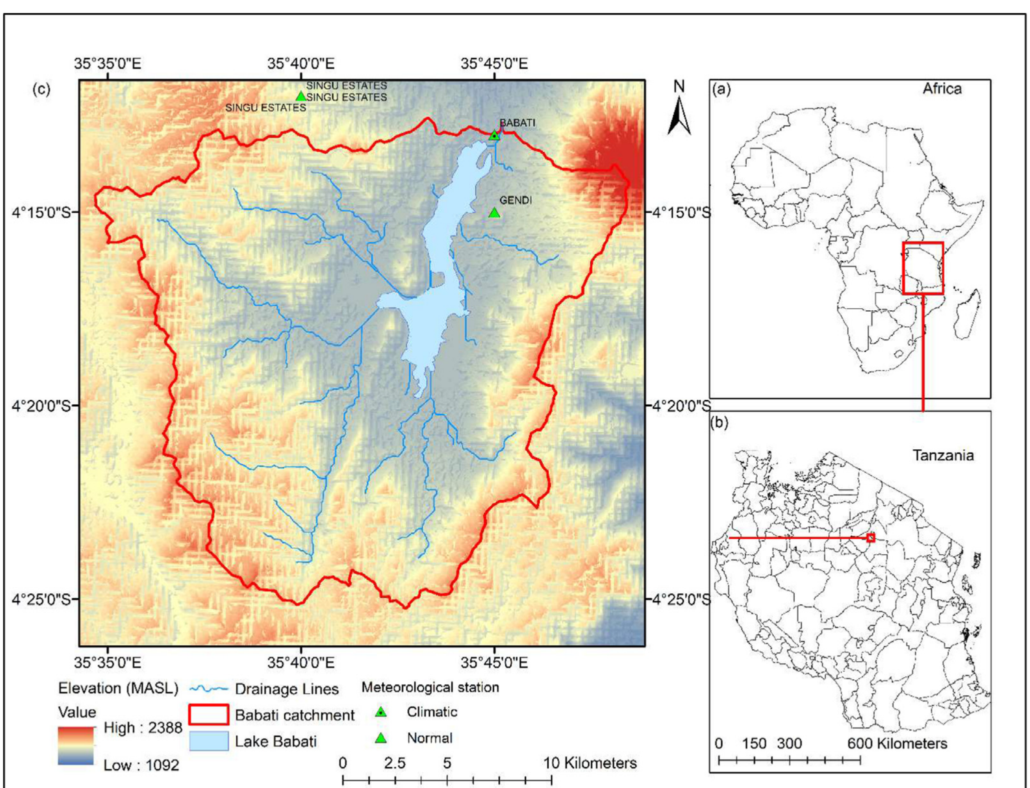


**Fig. 1.** Study area (a) Location of Tanzania on the African map, (b) location of Lake Babati catchment in the Tanzania map (c) Location of Lake Babati in the catchment.

**Table 1**
Meteorological stations that provided rainfall records with station elevation in metres above sea level (masl).

| No | Longitude | Latitude | Station ID | From | Name | Elevation (masl) | Timestep |
|---|---|---|---|---|---|---|---|
| 1 | 35.55°E | 3.86 °S | 9335001 | Jun 1990 to Jan 2017 | Mbulu District Office | 1737 | Monthly rainfall |
| 2 | 35.38 °E | 4.05 °S | 9435003 | Jan 1960 to Dec 2017 | Dongobesh Sec. School | 2042 | Daily record |
| 3 | 35.75 °E | 4.22 °S | 9435030 | Jan 1980 to May 2020 | Babati | 999 | Daily record |

posed of luvisols which are porous, well-aerated and well-drained soils (Driessen et al., 2001).

Lake Babati is a freshwater lake within the East African rift valley, and the Internal Drainage Basins Authority of Tanzania manages it. The lake has flooded on several occasions in the past, and its size varied many times. Mbanguka et al. (2016) estimated the current lake size to be about 7.00 $km^2$ which is an underestimate compared to 15.90 $km^2$ by Ministry of Water, 2014. The lake is semi-closed because of a spillway built in its northeastern part. When the floods reach the spillway crest at the stage of 4.74 m, the lake drains excess water into River Kiongozi, a physical link between Lake Babati and Lake Manyara, whose catchment includes the Lake Babati catchment.

The catchment experiences a semi-arid climate with a bimodal rainfall distribution. The major rainfall season occurs from February to May, while a minor rainfall season runs from October to January of the following year. It receives an annual average rainfall of 878 mm with a standard deviation of 273 mm based on the 1980 – 2020 rainfall records (from Babati Meteorological station with ID 9435030) maintained by the Tanzania Meteorological Authority. A dry but colder period occurs from June to September, with a short dry spell in February separating the main and minor rainy seasons.

### 2.2. Climatic and water level data sources and analysis

Rainfall records from meteorological stations (Table 1) within and in the neighbourhood of the Babati catchment were gathered from Tanzania Meteorological stations and the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) data from NASA's Global Modeling and Assimilation Office, 2021. The MERRA – 2 data was chosen because it provided a continuous record of data since 1980 and had a closer association with the instrumental data from the study area. The daily rainfall from MERRA – 2 and Babati had a Spearman rank correlation coefficient of 0.45, which increased to 0.85 when the rainfall data were aggregated to a monthly timestep. Therefore, in the absence of *in-situ* observations of temperature and other climatic parameters, trend analysis was based on MERRA- 2 data (Global Modeling and Assimilation Office, 2021). The Internal Drainage Basin Authority of Tanzania provided the water level records of Lake Babati observed from November 1976 to 2020.

We eliminated apparent errors such as outliers and erroneous entries and tested consistency, homogeneity and associations among different station records using Spearman's Rank test for data quality analysis. We preferred the Spearman's Rank method (a non-parametric method) over Pearson's Rank Correlation (a parametric method) to analyse the association between climatic data from different stations. The wide application of Spearman's rank method for handling data that are not normally distributed, such as climatic data (Byakatonda et al., 2018; Onyutha, 2017; Anand et al., 2018) motivated the choice. We applied the Mann-Kendall method to test monotonic trends among the climatic data and the lake water level (Mann, 1945), (Kendall, 1975). The non-parametric methods of Mann – Kendall and Spearman Rank were preferred for their ability to handle outliers because they consider ranks of measurements instead of the actual values (Mann, 1945, Kendall, 1975, Byakatonda et al., 2018). The methods are well elaborated in Mann (1945), Kendall (1975), Byakatonda et al. (2018) and thus have been applied in this study without repeating for brevity. The Pettit test as applied by Gebremicael et al. (2013), Samy et al. (2014), and Abungba et al. (2020) was performed on the water level records using the XLSTAT software. The XLSTAT takes a random and probabilistic approach (Monte Carlo simulations) to quickly detect the point of change in the data series.

### 2.3. Analysis of anthropogenic factors

The numerous anthropogenic activities identified to influence the hydrological regimes in the catchment included groundwater and surface water exploitation that may increase groundwater discharge, runoff harvesting, and alterations of the natural land cover. No water transfers occur from or into this catchment by any known means except when the lake itself outpours into the Kiongozi River during high floods. An accurate water abstraction rate was unavailable, but the Babati Water Supply Authority (BAWASA) provided 50,000 $m^3$ per week as the best estimate of water production/abstraction. The projections were based on the study area's population growth data to gain insights into the scale of water abstraction.

### 2.4. Hydrologic modelling

Changes in lake levels may be due to evaporation, precipitation, leakage in outlets, or changes in upstream river basin conditions. Similarly, anthropogenic factors can modify climatic parameters and responses of the natural hydrological systems (Onyutha, 2017). Water abstraction, however, directly reduces the amount of water in the reservoir. Different approaches are applied to understand the relationships among different hydrological parameters. Water balance modelling of the Lake Babati catchment was performed using the Hydrologic Engineering Centre – Hydrologic Modelling System (HEC-HMS) computer program version 4.8 by the U.S. Army Corps of Engineers (HEC, 2021 HEC, 2000). The HEC-HMS has several packages (models) which conceptualise rainfall-runoff response in both events and continuous storms in a fully lumped or distributed way. The possibility to model and calibrate an HEC – HMS model using the reservoir levels and its wide application in flood studies (HEC, 2021) motivated its choice in this study.

#### 2.4.1. Modelling equations

The deficit and constant rate method (HEC, 2021 HEC, 2000) was chosen to model the runoff volume, i.e. to model the loss of water to interception, evaporation, and infiltration before the runoff, while the Snyder's Unit Hydrograph (UH) method (Chow et al., 1988) was preferred to simulate the direct runoff on the ground surface. Although the Soil Moisture Accounting method (HEC, 2000) is also suitable for continuous modelling, the study chose the deficit and constant rate loss method to maintain parsimony, i.e. perform the continuous rainfall-runoff modelling with fewer inputs and be calibrated with the available data. HEC (HEC, 2000) provides the detailed modelling equations and computation of the initial and constant loss methods.

The baseflow (groundwater) was routed using the Linear Reservoir model, and no channel flow was routed because all the rivers are seasonal with unknown channel characteristics. Instead, we assumed the sub-basins drain directly into the lake since each sub-basin ended in the lake, as depicted in Fig. 2.

The Snyder UH has only two parameters, i.e. the $t_p$, which is the basin lag between the rainfall peak and the hydrograph peak and $C_p$, a peaking coefficient that varies from 0.4 to 0.8 (HEC, 2000). The $t_p$ was

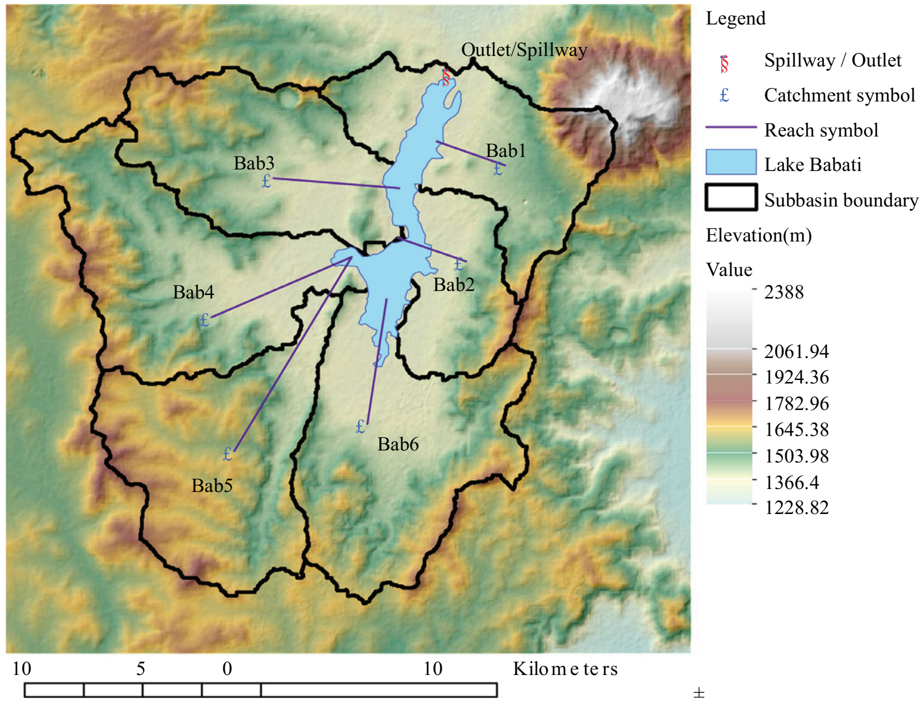


**Fig. 2.** Delineation of Lake Babati catchment into sub-basins and its setup within the HEC – HMS model.

determined from the basin parameters using Eq. (1) (Chow et al., 1988).

$$t_p = CC_t(LL_c)^{0.3} \tag{1}$$

Where $C_t$ is the basin coefficient, $L$ is the length of the mainstream from the outlet to the catchment divide; $L_c$ is the length along the mainstream from the outlet to a point nearest to the centroid, and $C$ is the conversion factor which is 0.75 for the SI units. The $C_t$ is not a physically based parameter. It was determined through calibration but the study ensured it ranged between 1.8 and 2.2, as Bedient et al. (2013) reported

The USACE (USACE 1994) provides an alternative method for estimating the basin lag $t_p$ as shown in Eq. (2).

$$t_p = CC_t\left(\frac{LL_c}{\sqrt{S}}\right)^N \tag{2}$$

Where $S$ is the overall slope of the longest watercourse from the point of concentration to the boundary of the drainage basin and $N$ is an exponent commonly considered as 0.33. Both methods (*Eqs. (1) and (2)*) were applied to determine the probable ranges for the basin lag. Other studies also estimated the basin lag as 50 – 75% of the time of concentration.

Mathematically, the hydrological model in a continuous simulation computes the lake water levels at an instantaneous time (t) by solving the general water balance equation derived from the continuity equation, which for a lake system is expressed as in Eq. (3) (Duan et al., 2018; Chow et al., 1988).

$$\frac{dh}{dt} = P(t) - E(t) - W_{ab} + \left(\frac{R_{in}(t) - R_{out}(t) + GW_{in}(t) - GW_{out}(t)}{A(h)}\right) + \varepsilon_t \tag{3}$$

Where $h$ represents the lake level and $dh$, the change in the lake level, A(h) is the lake surface area corresponding to a lake level h, and P(t) is the precipitation received over the lake area. E is the lake evaporation rate, $R_{in}$ and $R_{out}$ are the lake's surface water inflow and outflow, respectively. $W_{ab}$ stands for the water abstractions from the catchment and $GW_{in}$ and $GW_{out}$ groundwater inflow and outflow of the lake, respectively. $\varepsilon_t$ is an error term representing errors in data and unaccounted for water losses. All the parameters above are instantaneous and depend on the time step considered.

### 2.4.2. The model set up

Based on a 30 m resolution Digital Elevation Model (DEM) from Shuttle Radar Topographic Mission (USGS, 2018), six small sub-basins (Fig. 2) were delineated using the Geographic Information System package of HEC – HMS to represent the different hydrological responses. The catchment characteristics derived from the delineated sub-basins (see Table 2) were used to estimate realistic ranges of hydrological basin parameters.

A bathymetric survey of the lake was undertaken using an echosounder at a transect spacing of 100m and depths measured at intervals of 5 s within the transect. Overall, lake depths were measured at 59,060 different points within the lake. The sounding depth and lake boundaries set at 0 m depths were interpolated to generate the bathymetric chart. The lake volume – depth (Eq. (4)) and Lake surface area – depth relationship (Eq. (5)) derived from the chart were used to model lake outflow and lake evaporation.

$$V = 100000h^2 + 5000000h + 4000000 \tag{4}$$

$$A = 4000000 + 3000000h - 139383h^2 \tag{5}$$

Where h is the lake depth (stage) in metres, V is lake volume in $m^3$, and A is the lake surface area in $m^2$. The spillway span of 16m and discharge coefficient of 0.4 was specified to model lake outflow in peak levels.

### 2.4.3. Model calibration and validation

In this study, automatic calibration did not improve the pool level shape since it was restricted to measuring only the goodness-of-fit with the maximum peak pool elevation. Therefore, we applied automatic calibration for initial calibration. After that, the manual calibration was adopted to fine tune and preserve the pool level hydrograph shape, optimise the root means squared errors (RMSE) and the Nash Sutcliffe Efficiency (NSE) objective function calculated as in *Eq. (6)*. Therefore,

**Table 2**
Sub-basin properties which applied for estimation of the sub-basin hydrological parameters.

| | Area (km$^2$) | Longest flow path Length (km) | Longest Flow path slope | Centroidal Flow path Length (km) | Centroidal Flow path Slope | 10 - 85 Flow path Length (km) | 10 - 85 Flow path Slope | Basin Slope | Basin Relief (m) | Relief Ratio | Elongation Ratio | Drainage Density (km/km$^2$) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bab1 | 58.48 | 18.26 | 0.0572 | 10.60 | 0.0055 | 13.70 | 0.0251 | 0.1326 | 1047.0 | 0.0573 | 0.4725 | 0.1563 |
| Bab2 | 33.2 | 12.03 | 0.0211 | 5.76 | 0.0010 | 9.02 | 0.0023 | 0.1164 | 467.0 | 0.0388 | 0.5404 | 0.0750 |
| Bab3 | 62.86 | 16.70 | 0.0238 | 7.34 | 0.0077 | 12.52 | 0.0199 | 0.1302 | 409.0 | 0.0245 | 0.5357 | 0.1614 |
| Bab4 | 80.55 | 23.25 | 0.0139 | 12.35 | 0.0034 | 17.43 | 0.0092 | 0.1422 | 448.0 | 0.0193 | 0.4356 | 0.1835 |
| Bab5 | 72.97 | 20.42 | 0.0235 | 15.26 | 0.0165 | 15.31 | 0.0188 | 0.1430 | 483.0 | 0.0237 | 0.4721 | 0.1695 |
| Bab6 | 81.89 | 19.68 | 0.0174 | 10.51 | 0.0035 | 14.76 | 0.0133 | 0.1367 | 458.0 | 0.0233 | 0.5188 | 0.1565 |

**Table 3**
Goodness of fit measurements of the model at the calibration and validation phases.

| The goodness of fit statistics | Calibration phase | Validation phase |
|---|---|---|
| Sum of Absolute errors (m) | -29.22 | -256.05 |
| Sum of Squared Residuals (m$^2$) | 212.04 | 50.40 |
| Simulated Peak Level | 5.45 | 4.936 |
| Observed Peak Level | 5.56 | 4.93 |
| The Percent error in peak level (%) | -2.01 | 0.12 |
| Mean Observed Level (m) | 4.757 | 4.701 |
| Nash Sutcliffe Efficiency | 0.95 | 0.71 |

model calibrations involved a trial and error approach to match the prediction of model outputs with the observations or the field-measured data (HEC, 2000) by varying the parameter sets.

$$NSE = 1 - \frac{\sum_{i=1}^{n} \left(q_{s(i)} - q_{o(i)}\right)^2}{\sum_{i=1}^{n} \left(q_{o(i)} - \overline{q_o}\right)^2} \tag{6}$$

Where $q_{o(i)}$ is the observed output at time step $i$, $q_{s(i)}$ is the simulated output (simulated lake level) at time step $i$, $\overline{q_o}$ is the average observation (lake levels) for the time series considered and $n$ is the number of computed hydrograph ordinates.

The validation dataset was the rainfall and evapotranspiration data from July 19, 2020, to March 31, 2021. The lake level was measured at the 30 min interval using an automatic pressure transducer (diver) installed within the lake.

In addition to NSE and the RMSE, we used several goodness-of-fit measures, including a scatterplot of predicted lake pool level against the observations, the sum of absolute errors, the sum of squared residuals (HEC, 2000), and the percentage error in peak lake level (HEC, 2000). The goodness-of-fit measures (presented in Table 3) were computed from outside the model using Microsoft Excel. The calibrated parameters of the model are presented in Appendix A, B and C

### 2.5. Sensitivity analysis of drivers of lake levels

The Grey Relational Analysis (GRA) step proposed by Wong et al. (2006) to measure the degree of influence of one sequence over a reference sequence was applied to analyse the sensitivity of natural and anthropogenic factors that drive the water level variability of Lake Babati. GRA determines the geometrical proximity between different discrete sequences and at least one comparison sequence in a system (Li et al., 2014). The proximity, expressed in grey relational grade, measures the similarities between discrete data arranged in sequential order. Thus, a higher grey relational grade implies a higher similarity between the sequential parameters (Wong et al., 2006; Li et al., 2014). GRA has been widely used to analyse uncertainties in systems with imprecise information, including financial and hydrology analyses (Wong et al., 2006; Li et al., 2014; Kung and Wen, 2007), and optimisation of the manufacturing process and quality (Tzeng et al., 2009).

### 2.6. Stepwise regression analysis

Multi-variate regression analysis using quantifiable independent variables such as precipitation, lake evaporation, and runoff was employed to determine the factors which influence the lake levels the most. The study followed stepwise regression analysis of multiple variables described in Draper and Smith (1998) and Rawlings et al. (1998) to determine the partial regression coefficients of independent variables responsible for lake water level variations. Stepwise regression models relating lake levels to the different variables were developed by sequential replacement option, combining forward selection and backward elimination. F- Tests were done on each partial regression coefficient to determine the most significant variable to retain or the least significant

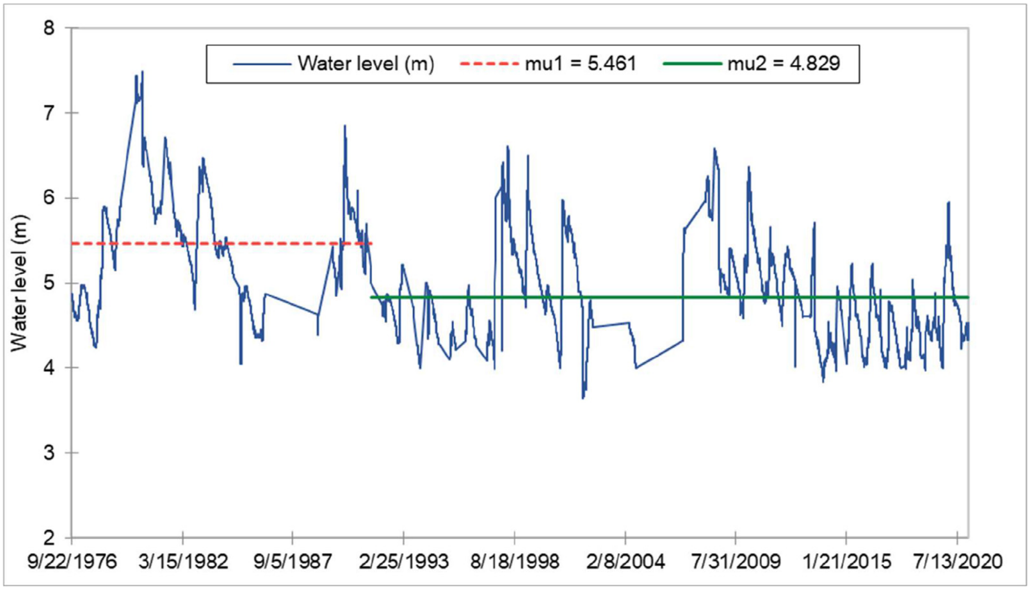


**Fig. 3.** Homogeneity test of the water level records of Lake Babati. The mu1 is the mean daily lake level before July 8, 1991, while mu2 is the mean daily lake level from April 1992 to 2020.

variable for elimination to simplify the model. In addition, the Akaike Information Criterion (AIC) (Akaike, 1973) was used to measure the quality of the resulting model relative to other models after adding or eliminating parameters. Since the optimal parameter sets give the lowest AIC, a smaller AIC was a stopping criterion in stepwise model regression and several "goodness-of-fit" tests, including the sum of squares, RMSE, and mean average error were done. Records of the daily precipitation, lake inflow, lake evaporation, and lake level from August 2019 to March 2021 were used to ensure all the data had the same length.

## 3. Results and discussions

### 3.1. Lake level variability

The available records showed that Lake Babati water levels have fluctuated over the period, varying between 3-9 m, as shown in Fig. 3. Some years such as 1997 and 2006, experienced exceptionally high water levels because of the unusually high rainfall occasioned by the Indian Ocean Dipole (Deus et al., 2013). The general pattern, however, showed a consistent decline in the lake level. A Pettit test for the water level homogeneity also confirmed that a significant change ($p < 0.01$) occurred in the records with July 14, 1991, as the date of shift (Fig. 3). Indeed the mean daily lake level shown in Fig. 3 was 5.461 m between 1976 and 1991, but it dropped to 4.829 m between 1991 and 2020. The shift was preceded by a heavy 1990 flooding of lake Babati that damaged many properties. After the flooding, an artificial outlet of Lake Babati was improved (Sandstrom, 1995). The improvement of the outlet by expansion and lowering the crest level is thought to have increased the outflow of Lake Babati, which resulted in the lake level decline observed in the subsequent years. Indeed, the current lake crest stage of 4.740 m is much lower than 6.15m reported by Sandstrom (1995) as the lake level above which it would flood before 1990. Therefore, the subsequent analysis in this study used the lake level records from 1991 to 2020 as they represented the current lake level scenario.

The Mann-Kendall analysis of the lake level records revealed that the decline of Lake Babati level was significant ($p < 0.05$) (see Fig. 4). In addition, Sen's slope (shown in Appendix D), a non-parametric method that accounts for the effects of the outliers and gross errors on the trend, also indicated an overall lake level decline of 0.025 m yearly.

The Mann-Kendall trend results in Appendix D showed significant lake level decline in all months and seasons. Contrary to expectations, the lake level declined significantly both in the dry and wet seasons. The minor wet season (from October to January) was characterised by higher temperatures. However, as given by Sen's slope, the lake level decline of 0.021 m in the warmer minor wet season was comparable to the major wet season lake level decline of 0.022 m. This implied that the lake level decline was not strongly dependent on temperatures. This contradicts the findings of Mbanguka et al. (2016), who reported a higher sensitivity of the lake level to cloudiness (a factor that results in elevated temperatures and lake evaporation when low and vice versa). Lake Babati exhibited a different behaviour from its downstream neighbour; Lake Manyara was more tied to the rainfall. Lake Manyara level declined in 2002 but maintained a higher level after a sharp increase in 2006-07 due to the high rainfall occasioned by the Indian Ocean Dipole (Deus et al., 2013). The rise in Lake Babati level after 2006 – 07 high rainfall was momentary, as shown in Fig. 3.

### 3.2. Climatic variability and lake levels

#### 3.2.1. Rainfall data quality assessment

We found very weak correlations/associations between rainfall records from nearby climatic stations and Babati at daily time steps. Rainfall records at Mbulu station were the most strongly correlated to Babati with the spearman rank correlation coefficient ($r_s$) = 0.4716 at a daily time step. The results concurred with Mbanguka et al. (2016), who similarly observed a weak correlation of rainfall measured at the Babati meteorological station with its neighbouring stations. This suggests a complex spatial rainfall distribution occasioned by the local influence of orographically induced precipitation (Mbanguka et al., 2016) and the Indian Ocean Dipole (Deus et al., 2013; Awange et al., 2016).

The associations and correlations improved significantly at monthly time steps in all stations. The Spearman rank correlation coefficient ($r_s$) of Mbulu and Babati improved to 0.8807, implying that Mbulu rainfall records explained 77.57% ($r_s^2$) of the observed rainfall variations at Babati. However, aggregating data at a time step longer than a month resulted in a significant decrease in the association; thus, the gaps in the monthly rainfall record of Babati were filled with monthly rainfall from Mbulu station. The improvement in associations with longer time

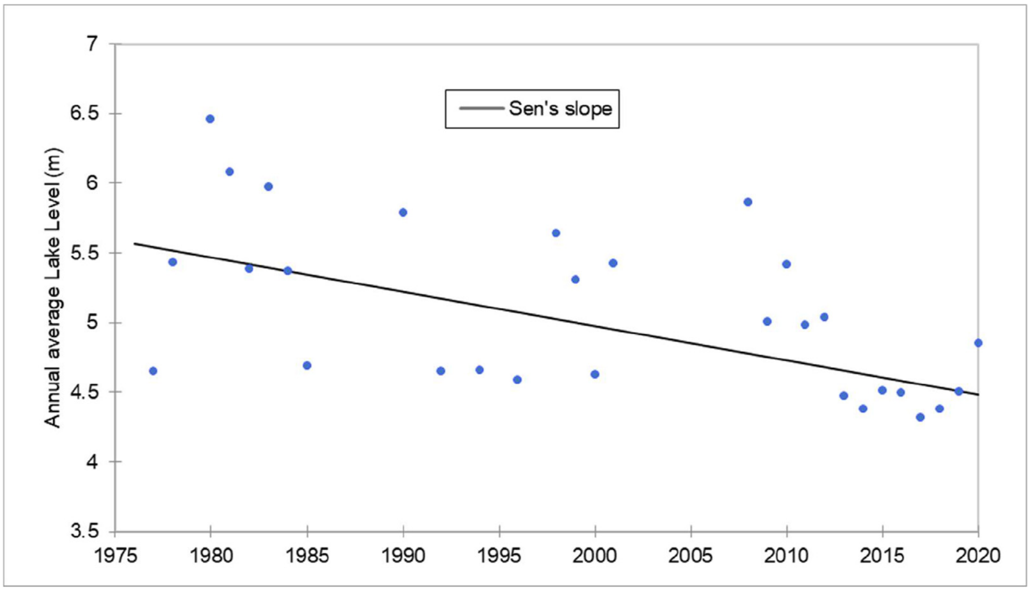


**Fig. 4.** The trend of average annual lake levels suggests a decrease in lake levels from 1976 to 2020. The dots are average annual lake levels.

steps suggests that regional factors drive monthly or seasonal rainfall while location-specific factors such as topography may modify the daily rainfall distribution. Mbanguka et al. (2016) estimated the topographic effect on the precipitation to be 3.6mm/year/100 m of topographic rise. The data analysed supported the interpretation that regional factors drive seasonal rainfall. All stations analysed experienced and recorded the extreme rainfall of 1997 and 2006 attributed to the positive Indian Ocean Dipole (Deus et al., 2013), (Awange et al., 2016). This was also corroborated by the characteristically very high lake levels (shown in Fig. 3) in 1997 and 2006.

### 3.2.2. *Influence of rainfall on lake level variability*

The lake level increased in the rainy season and receded in the dry season, as shown in Fig. 5, suggesting that rainfall is responsible for lake level variations. However, a more in-depth assessment of the rainfall variability using Mann-Kendall trend analysis revealed no significant trends in the rainfall. The Mann-Kendall statistics and Sen's slope presented in Appendix D indicate an insignificant decline ($p$>0.05) in the rainfall received in April, May, and November. Thus, the major wet season rainfall declined, although insignificantly.

Although Spearman's Rank correlation showed a positive association of rainfall with lake level, it was insignificant at 95% confidence intervals. However, averaging the precipitation over varying periods resulted in its positive and significant correlation with the lake level at nine and 18 months. This, therefore, implied that hydrological drought might take 9 to 18 months after the onset of meteorological drought. Similarly, Byakatonda et al. (Byakatonda et al., 2018) reported a 6 month lag between the onset of meteorological drought and hydrological drought in Okavango River systems in Botswana.

### 3.2.3. *Temperature and other climatic data*

The temperatures of Babati obtained from MERRA 2 showed a similar pattern with that of Mbulu. However, the minimum and maximum temperatures at Babati were about 2°C higher than their counterparts in Mbulu observed between 1923 and 1947. Furthermore, both stations' maximum and minimum temperatures were highest from August to March of the following year, while the April to July period had the least temperatures coinciding with the dry season. Therefore, the close association between temperatures observed earlier at Mbulu and the current Babati temperature justifies using MERRA 2 temperatures to represent the current situation at Babati.

Although no significant increase was observed in the maximum temperature annually, it significantly increased ($p$ < 0.05) in July, August, September, and October, which fall in the dry season. In contrast, the minimum temperature increased significantly ($p$ < 0.01) in all months except May. Consequently, all the seasons (both wet and dry) showed an increase in the minimum temperature, which may potentially increase the evaporative power of the atmosphere and the drought severity of the area leading to the decline in lake level.

For most of the year, the wind speed significantly reduced (P<0.05) in January, March, April, May, July, October, November, and December, implying that it reduced annually. The relative humidity remained constant over the years except in March when a significant increment was observed. The potential evapotranspiration showed no significant change in trend, except in March when a significant decline in potential evapotranspiration was registered. The detailed Mann-Kendall parameters for the climatic and hydrological parameters of the catchment area are captured in Appendix D.

### 3.2.4. *Lake evaporation*

The lake evaporation was computed using the Penman-Monteith formula based on the MERRA 2 climatic data (Global Modeling and Assimilation Office, 2021) from 1982 - to 2021. The calculated minimum monthly lake evaporation was 158.4 mm observed in June 1990 (a dry winter season), while the maximum monthly lake evaporation was 286.3 mm, observed in October 1987 (a hot and dry season). The mean and median evaporations stood at 216.6 and 218.0 mm, respectively, implying that lake evaporation had a central tendency with minimum outliers.

The Mann-Kendall tests statistics (Appendix D) showed that lake evaporation declined significantly (p < 0.05) during the wet seasons. Specifically, the significant decrease occurred in March, November, and December but insignificantly decreased in January and February. No single month expressed a significant increase in lake evaporation. Generally, no significant changes occurred in the lake evaporation over the years since 1982.

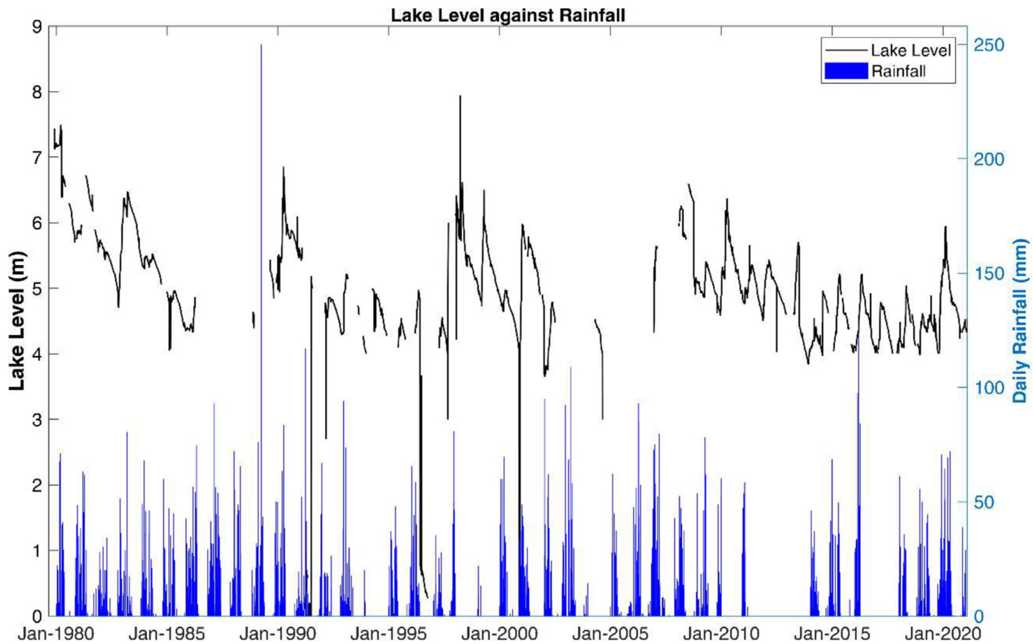


**Fig. 5.** A zoomed-in plot showing a comparison of water level against rainfall variability.

*3.2.5. Population growth and water abstractions*

Lake Babati catchment covers eleven wards (administrative units) of urban and rural settings in the Babati District, where groundwater is a water source for domestic and non-domestic uses. Unfortunately, water consumption data was unavailable. However, the rapid population explosion within the catchment, from 54,864 people in 2002 to 137,357 people in 2012 (150% increase), suggested a corresponding increase in water demand and abstraction. Such a population explosion is likely to outstrip water supply (abstraction) in the catchment whose expansion of social services, if any, might be tagged to the moderate population growth rate of only 3.2% recorded within the Manyara region in the same period (National Bureau of Statistics, 2013)

Using a per capita water consumption of 50 litres per day per person recommended for a low income group with no inhouse sanitary installation, but a metered water installation (Ministry of Water, 2020) as average water consumption, the estimated water consumption within the Babati catchment by 2012 was an upward of 6,688 $m^3$ per day. However, by 2019, BAWASA had already exceeded that and was supplying 50,000 $m^3$ per week (about 7,142 $m^3$ per day) to households within their service area, which covers only the urban and peri urban areas. About 78% of the water supplied is from groundwater sources, while spring water on Girala Mountain (a surface water source) contributes 22% of the supply. Already, BAWASA is engaged in numerous projects to expand its water supply coverage to meet the increased demand. However, rural communities are still using shallow wells for domestic purposes. Some vegetables are grown under irrigation with agriculture and livestock in the district. However, minimal information is available on the extent of irrigated land or the amount of water used explicitly for irrigation. Based on the described scenario, it is evident that the water abstractions have tremendously increased (more than doubled) to accommodate the spiralling urban population and economic activities.

### 3.3. Lake water balance simulation and regression

*3.3.1. Lake water balance simulation*

Using the available lake levels, the lake bathymetric data, the rainfall and the lake evaporation, an HEC-HMS model was built, calibrated, and its performance validated. The graphical evaluation of the calibrated model and the statistical measure of its performance are as shown in Fig. 6 and Table 3, respectively.

The graphical plot and the statistical goodness-of-fit in Table 3 indicated that the model is very good for prediction purposes with a Nash Sutcliffe Efficiency (NSE) of 0.95 (Moriasi et al., 2015). As shown in Fig. 6, the model is more accurate in predicting the low lake levels. Its performance during the peak level seasons, however, harbours some errors. Generally, the model underestimates the lake level during the rainy season and offsets the peak level in time compared to the observed lake level. Although several error sources, including model inefficiency, could offset the peak lake level, the failure of the cumulative daily rainfall depths to capture the exact time when the peak rainfall occurs is believed to have caused this one. Since each sub-basin has a time of concentration of less than 12 h, the exact peak time of the rainfall is essential for predicting the lake levels. The arbitrary frequency duration curve of the synthetic unit hydrograph could also result in the mismatch between the modelled and observed lake level peaks. The improvement of the model calibration for flood studies requires rainfall captured at timesteps shorter than the shortest time of concentration of the sub-basin.

The model matches the low-flow conditions well, making it most suitable for low-flow studies and relevant for water supply and drought studies. However, since the model does not capture the peak flows accurately, its application for flood studies may be limited. The calibrated parameters of the model are summarised in Appendixs A, B and C.

*3.3.2. Grey relational order*

The factors driving the lake water level or storage variability were assessed and ranked using GRA (see Table 4). The most important driver of the lake level variability was found to be the lake inflow (runoff and baseflow), followed by direct rainfall. Lake Evaporation came in third while lake outflow was the least important parameter as captured by their least grey relational grades shown in Table 4. When GRA was performed with the lake storage as the reference series, the lake inflow and rainfall maintained their first and second order of importance. However, the lake evaporation and lake outflow tied in the third position. The grey relational grade (shown in Table 4) of lake evaporation and lake outflow are insignificantly different ($p$ = 0.97), thus implying both have equal

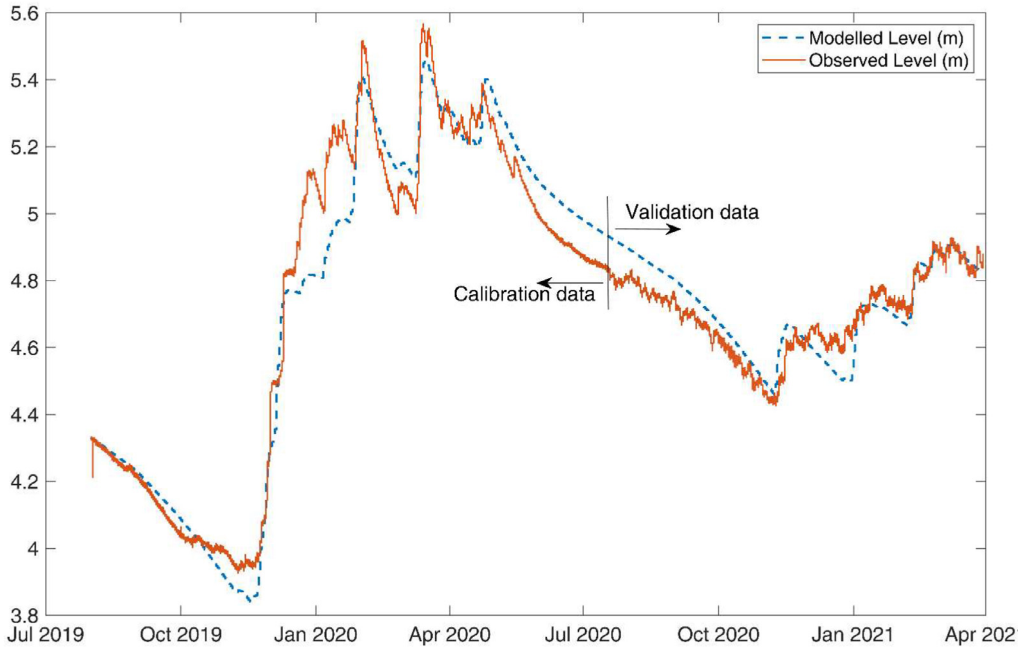


**Fig. 6.** The general agreement between the observed and simulated lake levels during the calibration and validation phases.

**Table 4**
Comparison of the grey relational grades of different parameters based on the lake level and lake storage as reference series.

| Reference series | Inflow ($m^3/s$) | Direct Rainfall (mm) | Lake Evaporation (mm) | Computed Outflow ($m^3/s$) |
|---|---|---|---|---|
| Lake Level (m) | 0.8591 | 0.8652 | 0.6320 | 0.6120 |
| Lake Storage ($m^3$) | 0.8590 | 0.8658 | 0.6152 | 0.6155 |

magnitude/influence in the control of Lake Babati level variability when based on the lake storage.

The evaporation was expected to have a higher weighting (grey relation grade) than the direct rainfall because the rainfall over the lake is smaller than evaporation. However, the high correlation of direct rainfall to the lake inflow may have biased the analysis. The lake outflow had more control over the lake level variability during the rainy year (2019 -2020 hydrological year) when more water was removed from the lake during peak seasons than evaporation. Increased lake outflow, either due to improvement or expansion in lake outlets, have been observed elsewhere to result in lake level drop. For example, Lake Victoria, with a surface area of about 67,000 km$^2$ (Tate et al., 2004), once experienced a 2 m drop in lake level between 1999 – 2006 when a second hydropower facility increased the lake outflow (Swenson and Wahr, 2009). The increased outflow and climatic factors such as drought, accelerates the lake level decline. As observed, Lake Babati experienced more evaporation than lake outflow in the hydrological year 2020 – 2021 when rainfall was generally low. This agreed with Kumambala and Ervine (2013), who suggested that lake evaporation is often the most significant outflow component since the amount of water available does not limit the evaporation rates. Further, it partly agreed with Mbanguka et al. (2016), who reported that lake evaporation and runoff controls the hydrological balance of Lake Babati. Mbanguka et al. (2016), however, missed recognising the influence of the large outflows during high rainfall seasons. Currently, the abstraction of groundwater for urban water supply is far less than the lake's evaporation rates. Therefore, even if groundwater abstraction were directly drawing lake water, its control of the lake level would be less than that of evaporation which was ranked third in this study. Therefore, water abstraction still plays but a very marginal role in controlling the lake level variability.

### 3.3.3. *Stepwise regression analysis*

A stepwise regression analysis of the precipitation over the lake area ($P$ in mm), inflow to the lake ($R_{in}$ in m$^3$/s), lake outflow ($R_{out}$ in m$^3$/s) and lake evaporation ($E_L$ in mm/day) in relation to the lake water levels was assessed at daily timesteps. Therefore, the daily lake level in meters ($L_{daily}$) can be predicted using the daily precipitation, lake evaporation, lake inflow and outflow using the regression formula in Eq. (7). All the added parameters improved the model, and the coefficients of the parameters were not zero ($p < 0.01$). The regression had an F-statistics of 273.8 on 4 and 608 degrees of freedom and the multiple $R^2 = 0.6443$.

$$L_{daily} = 5.6403 \ - 0.0017P + 0.0081R_{in} + 0.1761R_{out} - 0.1537E_L \tag{7}$$

The stepwise elimination improved the model by removing the less sensitive or adding the most sensitive parameters. The AIC (Akaike, 1973) was used to evaluate the improvement. The lake level was not very sensitive to the direct rainfall over the lake, and thus it was eliminated, giving the final regression model in Eq. (8). The final model had fewer input parameters but improved with the multiple $R^2 = 0.6437$, an F-statistics of 363.8 on 3 and 604 degrees of freedom and $p < 0.01$.

$$L_{daily} = 5.6478 \ + 0.0041R_{in} + 0.1786R_{out} - 0.1547E_L \tag{8}$$

The prediction of the lake storage (given in Eq. (9)) using the same parameters was more accurate than the lake level and had the following goodness-of-fit statistics: a multiple $R^2 = 0.7415$, an F-statistics of 432.4 on 4 and 603 degrees of freedom and $p < 0.01$.

$$LS_{daily} = 67127331 - 34253P + 66873R_{in} + 3105860R_{out} - 2583060E_L \tag{9}$$

A stepwise regression analysis resulted in a more improved model with fewer input parameters since direct rainfall was eliminated for being less important. The resultant model (Eq. (10)) had the following goodness-of-fit statistics; a multiple $R^2 = 0.7407$, an F-statistics of 575 on 3 and 604 degrees of freedom, and $p < 0.01$.

$$LS_{daily} = 67276591 - 13723R_{in} + 3155399R_{out} - 2601460E_L \quad (10)$$

The regression indicates that the lake inflow, lake outflow and evaporation are the most significant parameters for predicting the daily lake levels. Furthermore, the stepwise regression removed the direct rainfall in backward elimination and did not substitute the direct rainfall in the forward elimination, implying an optimal model prediction could be achieved without direct rainfall. Overall, the direct rainfall showed the least influence in driving the lake level variability because it is very minimal due to the small lake surface area. With the catchment-lake surface area ratio of 26:1, 1mm of catchment runoff translates into a 26 mm depth inflow over the lake surface, thus masking the effect of direct rainfall. Furthermore, at an average evaporation rate of 6 mm per day, the lake evaporation is substantial and outstrips direct rainfall as it reaches about 2000 mm per year compared to the annual rainfall.

Evaporation is the main depleting factor for a closed lake (Kumambala and Ervine, 2013). However, it had a minimal variation which could not explain the level decline of Lake Babati, especially during rainy seasons. In years of heavy rainfall, the outflow during peak level seasons withdrew more water from the lake than evaporation. The lake level decline in December – February (dry season) seemed to be attributable to evaporation which was highest in this period with elevated temperatures. The dry winter season (June – September) appeared to dampen the overall effect of evapotranspiration. Subsequently, it implied that lake outflow played almost an equal influence as evaporation in controlling the lake storage/level as a large volume of water is spilt during peak levels.

Therefore, the most important variables in predicting the lake level or storage are the inflow, outflow and evaporation. The GRA corroborated these findings as it ranked the lake inflow as the most sensitive parameter, then the direct rainfall followed by the outflow and the lake evaporation at equal magnitudes of influence. Although GRA ranked direct rainfall as the second most sensitive parameter, the stepwise analysis discounted the influence of direct rainfall in driving the lake level because of its small magnitude compared to lake inflow. The strong correlation of direct rainfall with lake inflow might have influenced the GRA to rank it as the second most sensitive parameter.

## 4. Conclusions

The study analysed trends of hydrological components of Lake Babati and modelled its variability with the different water balance components. The results found that the Lake Babati level is significantly declining ($p < 0.01$) at a rate of 0.025 m yearly after accounting for the effects of the outliers and gross errors on the trend using Sen's slope. Furthermore, the seasonal analysis indicated that all months and seasons experience significant ($p < 0.05$) lake level decline.

The lake level has often varied, reflecting the cycles of the wet and dry seasons. Although the lake level peaks have corresponded to the rainfall, Spearman's Rank correlation indicated no significant association between the lake level and rainfall. However, the lake level decline occurred when evaporation remained constant while rainfall did not show significant changes seasonally or annually. Therefore, the direct attribution of the lake level decline to the rainfall variability was not significant.

An HEC-HMS hydrological model of the Lake Babati catchment was built, calibrated and validated. The model showed high accuracy in predicting the low-lake levels, making it an excellent tool for studying water supply conditions. The accuracy of low-flow prediction is instrumental for sustainable design, water abstraction and environmental conservation. This model was applied to generate the lake inflow from the available rainfall data.

The GRA of the lake basin components showed that inflow is the primary parameter controlling the lake level, followed by direct rainfall and outflow. On a grey relational scale, lake evaporation and outflow are tied in magnitude as the least important parameters for lake level variability. However, an optimised stepwise regression model indicated that direct rainfall had negligible influence on the lake level. The huge response of the lake level to the rainfall was attributed to inflow (runoff and baseflow) due to the large catchment relative to the lake surface area. Therefore, in relative terms, the direct rainfall had little effect on the lake level than the lake inflow and evaporation. Although the water abstraction is increasing to support the economic boom and population explosion around Babati town, it is still low and falls behind evaporation. However, if the current water abstraction trend continues, it could soon become a sensitive parameter for lake level variability.

The declining lake level seemed to be related to the outflow controlled by a spillway constructed in the northeast of the lake to avoid flooding the lake. The improvements/expansion in the spillway have resulted in large outflow during peak seasons and could be why the lake levels were significantly declining in the rainy seasons. However, the lake level decline in dry seasons was probably related to evaporation loss and the reduced lake inflow.

This study demonstrated that the lake level is controlled by the inflow and outflow of the lake. Therefore, management should prioritise interventions that guarantee lake inflow (surface runoff and baseflow). These interventions may include catchment protection and minimising infield water harvesting. Additionally, optimal regulation of the lake outflow using variable height control gates could make the lake useful for both flood control and reservoir storage. Finally, continuous monitoring of lake levels and rainfall at shorter timesteps is recommended to enrich future studies within this catchment.

## Funding

The WISE – Futures African Centre of Excellence supported this work and sponsored the first author for PhD research at The Nelson Mandela African Institution of Science and Technology, Arusha Tanzania.

## Declaration of Competing Interest

None.

## Acknowledgements

We extend our sincere gratitude to the Internal Drainage Basin Authority of Tanzania and Babati Water Supply Authority for providing lake levels and borehole data. The first author is most grateful to the Indian Institute of Technology, Roorkee, for the Capacity-building exchange fellowship and Gulu University for the study leave and other financial support. This work is part of a PhD research in Hydrology and Climate Studies.

## Appendix A

The initial and calibrated parameters for initial abstraction, runoff volume and catchment runoff routing.

| | Simple Canopy | | | | | | Simple Surface | | | | | | | Snyder Transform | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Initial Storage (%) | Initial Storage (%) | Initial Storage (%) | Maximum Storage (mm) | Maximum Storage (mm) | Maximum Storage (mm) | Crop Co-efficient (-) | Initial Storage (%) | Initial Storage (%) | Initial Storage (%) | Maximum Storage (mm) | Maximum Storage (mm) | Maximum Storage (mm) | Lag Time (HR) | Lag Time (HR) | Lag Time (HR) | Peaking Coeffi-cient | Peaking Coeffi-cient | Peaking Coeffi-cient |
| | Minimum | Maximum | Calibrated | Minimum | Maximum | Calibrated | | Minimum | Maximum | Calibrated | Minimum | Maximum | Calibrated | Minimum | Maximum | Calibrated | Minimum | Maximum | Calibrated |
| Bab4 | 0 | 100 | 0.0267 | 0 | 10 | 1 | 1 | 0 | 100 | 100 | 0 | 100 | 1 | 0.1 | 500 | 7.3743 | 0.4 | 0.8 | 0.5 |
| Bab5 | 0 | 100 | 0.0294 | 0 | 10 | 1 | 1 | 0 | 100 | 100 | 0 | 100 | 1 | 0.1 | 500 | 7.558 | 0.4 | 0.8 | 0.5 |
| Bab6 | 0 | 100 | 0.0315 | 0 | 10 | 1 | 1 | 0 | 100 | 100 | 0 | 100 | 1 | 0.1 | 500 | 6.6849 | 0.4 | 0.8 | 0.5 |
| Bab3 | 0 | 100 | 0.0329 | 0 | 10 | 1 | 1 | 0 | 100 | 100 | 0 | 100 | 1 | 0.1 | 500 | 5.7119 | 0.4 | 0.8 | 0.5 |
| Bab2 | 0 | 100 | 0.0336 | 0 | 10 | 1 | 1 | 0 | 100 | 100 | 0 | 100 | 1 | 0.1 | 500 | 4.8154 | 0.4 | 0.8 | 0.5 |
| Bab1 | 0 | 100 | 0.0338 | 0 | 10 | 1 | 1 | 0 | 100 | 100 | 0 | 100 | 1 | 0.1 | 500 | 6.5522 | 0.4 | 0.8 | 0.5 |

## Appendix B

Initial and calibrated linear reservoir model parameters for the HEC – HMS model of Lake Babati catchment.

| | Number of Layers | GW 1 Initial – Discharge per unit area (M3/s/Km$^2$) | | | GW1 Fraction (-) | GW 1 Coefficient (HR) | | | GW 1 Steps | GW 2 Initial Discharge per unit area (M3/s/Km2) | | | GW 2 Fraction (-) | GW 2 Coefficient (HR) | | | GW2 Steps |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Minimum | Maximum | Calibrated | | Minimum | Maximum | Calibrated | | Minimum | Maximum | Calibrated | | Minimum | Maximum | Calibrated | |
| Bab4 | 2 | 0 | 0.4 | 0.0001 | 0.2 | 0.1 | 1000 | 59.407 | 1 | 0 | 0.15 | 0.0015 | 0.15 | 1 | 5000 | 1232 | 2 |
| Bab5 | 2 | 0 | 0.4 | 0.0001 | 0.2 | 0.1 | 1000 | 57.855 | 1 | 0 | 0.15 | 0.0015 | 0.12 | 1 | 5000 | 1448 | 2 |
| Bab6 | 2 | 0 | 0.4 | 0.0001 | 0.2 | 0.1 | 1000 | 52.671 | 1 | 0 | 0.15 | 0.0015 | 0.1 | 1 | 5000 | 1316 | 2 |
| Bab3 | 2 | 0 | 0.4 | 0.0001 | 0.2 | 0.1 | 1000 | 33.658 | 1 | 0 | 0.15 | 0.0015 | 0.1 | 1 | 5000 | 1167 | 2 |
| Bab2 | 2 | 0 | 0.4 | 0.0001 | 0.2 | 0.1 | 1000 | 27.235 | 1 | 0 | 0.15 | 0.0015 | 0.15 | 1 | 5000 | 1120 | 2 |
| Bab1 | 2 | 0 | 0.4 | 0.0001 | 0.2 | 0.1 | 1000 | 36.156 | 1 | 0 | 0.15 | 0.0015 | 0.15 | 1 | 5000 | 1183 | 2 |

## Appendix C

Parameter ranges and calibrated values for the deficit and constant rate methods of accounting infiltration rates.

| | Deficit and Constant | | | | | | | | | Impervious (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Initial Deficit (mm) Minimum | Initial Deficit (mm) Maximum | Initial Deficit (mm) Calibrated | Maximum Storage (mm) Minimum | Maximum Storage (mm) Maximum | Maximum Storage (mm) Calibrated | Constant Rate (mm/HR) Minimum | Constant Rate (mm/HR) Maximum | Constant Rate (mm/HR) Calibrated | |
| Bab4 | 0 | 5 | 0.14196 | 0.001 | 1000 | 5 | 0.001 | 30 | 3 | 0 |
| Bab5 | 0 | 5 | 0.11462 | 0.001 | 1000 | 5 | 0.001 | 30 | 3 | 0 |
| Bab6 | 0 | 5 | 0.087282 | 0.001 | 1000 | 5 | 0.001 | 30 | 3 | 0 |
| Bab3 | 0 | 5 | 0.059943 | 0.001 | 1000 | 5 | 0.001 | 30 | 2.5 | 0 |
| Bab2 | 0 | 5 | 0.032604 | 0.001 | 1000 | 5 | 0.001 | 30 | 2.5 | 0 |
| Bab1 | 0 | 5 | 0.005265 | 0.001 | 1000 | 5 | 0.001 | 30 | 2.5 | 0 |

## Appendix D

Summary of Mann-Kendall trend test of the Lake Babati levels and rainfall received in Babati.[1]

| | Lake Babati Level | | | Rainfall in Babati | | | Minimum temperature | | | Maximum Temperature | | | Lake Babati Evaporation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Series\ Test | Kendall's tau | p-value | Sen's slope | Kendall's tau | p-value | Sen's slope | Kendall's tau | p-value | Sen's slope | Kendall's tau | p-value | Sen's slope | Kendall's tau | p-value | Sen's slope |
| Jan | -0.310 | 0.015 | -0.026 | 0.100 | 0.440 | 1.570 | 0.250 | 0.019 | 0.020 | -0.118 | 0.289 | -0.030 | -0.130 | 0.230 | -0.250 |
| Feb | -0.290 | 0.023 | -0.023 | 0.110 | 0.390 | 0.980 | 0.320 | 0.003 | 0.020 | -0.054 | 0.633 | -0.015 | 0.0025 | 0.990 | 0.010 |
| Mar | -0.300 | 0.026 | -0.020 | 0.070 | 0.610 | 0.580 | 0.260 | 0.018 | 0.017 | -0.210 | 0.057 | -0.060 | -0.283 | 0.010 | -0.470 |
| Apr | -0.310 | 0.006 | -0.021 | -0.120 | 0.400 | -1.460 | 0.270 | 0.013 | 0.015 | 0.018 | 0.870 | 0.003 | -0.030 | 0.730 | -0.070 |
| May | -0.359 | 0.000 | -0.030 | -0.150 | 0.270 | -0.550 | 0.150 | 0.160 | 0.009 | 0.095 | 0.395 | 0.015 | 0.003 | 0.980 | 0.005 |
| Jun | -0.308 | 0.016 | -0.020 | -0.040 | 0.800 | 0.000 | 0.338 | 0.002 | 0.035 | 0.137 | 0.217 | 0.021 | 0.070 | 0.530 | 0.071 |
| Jul | -0.359 | 0.004 | -0.023 | -0.190 | 0.230 | 0.000 | 0.331 | 0.002 | 0.031 | 0.264 | 0.017 | 0.031 | 0.065 | 0.560 | 0.041 |
| Aug | -0.333 | 0.001 | -0.022 | -0.110 | 0.410 | 0.000 | 0.420 | 0.0001 | 0.036 | 0.264 | 0.026 | 0.024 | 0.079 | 0.470 | 0.049 |
| Sep | -0.335 | 0.001 | -0.021 | 0.080 | 0.640 | 0.000 | 0.490 | 0.0001 | 0.032 | 0.279 | 0.011 | 0.021 | 0.074 | 0.500 | 0.038 |
| Oct | -0.382 | 0.007 | -0.021 | 0.000 | 1.000 | 0.000 | 0.443 | 0.0000 | 0.029 | 0.218 | 0.048 | 0.024 | -0.028 | 0.800 | -0.010 |
| Nov | -0.384 | 0.004 | -0.021 | -0.010 | 0.940 | -0.080 | 0.502 | 0.0000 | 0.032 | -0.074 | 0.506 | -0.009 | -0.200 | 0.060 | -0.220 |
| Dec | -0.360 | 0.007 | -0.026 | 0.080 | 0.590 | 1.660 | 0.456 | 0.0000 | 0.021 | -0.087 | 0.435 | -0.013 | -0.209 | 0.059 | -0330 |
| Annual | -0.468 | 0.000 | -0.025 | -0.050 | 0.760 | -2.230 | 0.580 | 0.0000 | 0.025 | 0.064 | 0.574 | 0.005 | -0.148 | 0.180 | -0.970 |
| Minor wet season (ONDJ) | -0.394 | 0.001 | -0.021 | 0.070 | 0.640 | 2.070 | 0.591 | 0.0000 | 0.025 | -0.058 | 0.607 | -0.005 | -0.261 | 0.018 | -0.900 |
| Major wet season (FMAM) | -0.324 | 0.005 | -0.022 | -0.090 | 0.510 | -1.860 | 0.388 | 0.0006 | 0.017 | -0.0544 | 0.632 | -0.007 | -0.120 | 0.278 | -0.430 |
| Wet season (ONDJF-MAM) | -0.387 | 0.000 | -0.020 | -0.09 | 0.53 | -4.340 | 0.595 | 0.0000 | 0.022 | -0.062 | 0.583 | -0.005 | -0.249 | 0.024 | -1.410 |
| Dry season (JJAS) | -0.363 | 0.02 | -0.020 | -0.10 | 0.51 | 0.000 | 0.505 | 0.0000 | 0.036 | 0.253 | 0.023 | 0.025 | 0.084 | 0.450 | 0.180 |

[1] ONDJ is the October, November, December and January period. FMAM is the February, March, April and May period, ONDJFMAM is the wet season from October to May and JJAS is the June, July, August and September.

## References


Abungba, J.A., Khare, D., Pingale, S.M., Adjei, K.A., Gyamfi, C., Odai, S.N., 2020. Assessment of Hydro-climatic Trends and Variability over the Black Volta Basin in Ghana. Earth Syst. Environ. 4 (4), 739–755. doi:10.1007/s41748-020-00171-9.

Akaike, H., 1973. Information theory as an extension of the maximum likelihood principle. In: Proceedings of the Second International Symposium on Information Theory, pp. 267–281.

Anand, J., Gosain, A.K., Khosa, R., Srinivasan, R., 2018. Regional scale hydrologic modelling for prediction of water balance, analysis of trends in streamflow and variations in streamflow: The case study of the Ganga River basin. J. Hydrol. Reg. Stud. 16 (March), 32–53. doi:10.1016/j.ejrh.2018.02.007.

Awange, J.L., Schumacher, M., Forootan, E., Heck, B., 2016. Exploring hydro-meteorological drought patterns over the Greater Horn of Africa (1979—2014) using remote sensing and reanalysis products. Adv. Water Resour. doi:10.1016/j.advwatres.2016.04.005.

Bedient, P.B., Huber, W.C., Vieux, B.E., 2013. Hydrology and Floodplain Analysis, 27, 5th ed. Pearson Education Limited, Edinburgh.

Byakatonda, J., Parida, B.P., Kenabatho, P.K., Moalafhi, D.B., 2018. Infl uence of climate variability and length of rainy season on crop yields in semiarid Botswana. Agric. For. Meteorol. 248 (March 2017), 130–144. doi:10.1016/j.agrformet.2017.09.016.

Byakatonda, J., Parida, B.P., Moalafhi, D.B., Kenabatho, P.K., 2018. Analysis of long term drought severity characteristics and trends across semiarid Botswana using two drought indices. Atmos. Res. 213 (March), 492–508. doi:10.1016/j.atmosres.2018.07.002.

Byakatonda, J., Parida, B.P., Kenabatho, P.K., 2018. Relating the dynamics of climatological and hydrological droughts in semiarid Botswana. Phys. Chem. Earth 105 (March 2017), 12–24. doi:10.1016/j.pce.2018.02.004.

Chow, V. Te, Maidment, D.R., Mays, L.W., 1988. Applied Hydrology. McGraw-Hill Book Co, New Delhi.

Darling, W.G., Gizaw, B., Arusei, M.K., 1996. Lake-groundwater relationships and fluid-rock interaction in the East African Rift Valley: Isotopic evidence. J. Afr. Earth Sci. 22 (4), 423–431. doi:10.1016/0899-5362(96)00026-7.

Deus, D., Gloaguen, R., Krause, P., 2013. Water balance modeling in a semi-arid environment with limited *in situ* data using remote sensing in Lake Manyara, East African rift, Tanzania. Remote Sens. 5 (4), 1651–1680. doi:10.3390/rs5041651.

Draper, N., Smith, H., 1998. Applied-Regression-Analysis-Third-Edition Norman Draper and Hary Smith.pdf, 3rd ed. Wiley and sons, New York.

Driessen, P., Deckers, J., Spaargaren, O., 2001. In: Lecture Notes on The Major Soils of the World," in World Soil Resources Reports : FAO, 94. Food and Agriculture Organisation of the United Nations, Rome, Italy, p. 307.

Duan, Z., Tuo, Y., Liu, J., Gao, H., Song, X., Zhang, Z., 2018. Hydrological evaluation of open-access precipitation and air temperature datasets using SWAT in a poorly gauged basin in Ethiopia. J. Hydrol. 612–626. doi:10.1016/j.jhydrol.2018.12.026.

Duan, Z., Gao, H., Ke, C., 2018. Estimation of lake outflow from the poorly gauged Lake Tana (Ethiopia) using satellite remote sensing data. Remote Sens. 10 (7), 1–21. doi:10.3390/rs10071060.

Gao, H., Hrachowitz, M., Schymanski, S.J., Fenicia, F., Sriwongsitanon, N., Savenije, H.H.G., 2014. Climate controls how ecosystems size the root zone storage capacity at a catchment scale. Geophys. Res. Lett. 1–8. doi:10.1002/2014GL061668.Received.

Gebremicael, T.G., Mohamed, Y.A., Betrie, G.D., van der Zaag, P., Teferi, E., 2013. Trend analysis of runoff and sediment fluxes in the Upper Blue Nile basin: A combined analysis of statistical tests, physically-based models and landuse maps. J. Hydrol. 482, 57–68. doi:10.1016/j.jhydrol.2012.12.023.

Gilfedder, M., Rassam, D.W., Stenson, M.P., Jolly, I.D., Walker, G.R., Littleboy, M., 2012. Incorporating land-use changes and surface-groundwater interactions in a simple catchment water yield model. Environ. Model. Softw. 38, 62–73. doi:10.1016/j.envsoft.2012.05.005.

Global Modeling and Assimilation Office, 2021. Modern-Era Retrospective analysis for Research and Applications, Version 2 (MEERA-2). Global Modeling and Assimilation Office May 1 POWER_SinglePoint_Daily_19820101_20210331_04d24S_035d74E_f606bc19, 2021.

Guma, B.E., Owor, M., Muwanga, A., 2019. Journal of African Earth Sciences Hydrogeological characteristics of the Albertine Graben, Uganda : Evidence from surface geophysics and hydraulic testing. J. Afr. Earth Sci. 150, 224–238. doi:10.1016/j.jafrearsci.2018.11.008, November 2018.

Hassan, A.A., Jin, S., 2014. Lake level change and total water discharge in East Africa Rift Valley from satellite-based observations. Glob. Planet. Change 117, 79–90. doi:10.1016/j.gloplacha.2014.03.005.

HEC, 2000. Hydrologic Modeling System HEC-HMS Technical Reference Manual. HEC, Davis, California [Online]. Available.

HEC, 2021. HEC-HMS User's Manual. HEC, Davis, California [Online]. Available.

Herrnegger, M., Stecher, G., Schwatke, C., Olang, L., 2021. Journal of Hydrology : Regional Studies Hydroclimatic analysis of rising water levels in the Great rift Valley Lakes of Kenya. J. Hydrol. Reg. Stud. 36 (May), 100857. doi:10.1016/j.ejrh.2021.100857.

Kebede, S., Travi, Y., Alemayehu, T., Marc, V., 2006. Water balance of Lake Tana and its sensitivity to fluctuations in rainfall, Blue Nile basin, Ethiopia. J. Hydrol. 316 (1–4), 233–247. doi:10.1016/j.jhydrol.2005.05.011.

Kendall, M.G., 1975. Rank Correlation Methods, 4th ed. Griffin, London.

Kumambala, P. G., Ervine, A., 2010. Water balance model of Lake Malawi and its sensitivity to climate change. The Open Hydrology Journal 4, 152–162. doi:10.2174/1874378101004010152.

Kumambala, P.G., Ervine, A., 2013. Water Balance Model of Lake Malawi and its Sensitivity to Climate Change. Open Hydrol. J. 4 (1), 152–162. doi:10.2174/1874378101004010152.

Kung, C., Wen, K., 2007. Applying grey relational analysis and grey decision-making to evaluate the relationship between company attributes and its financial performance - a case study of venture capital enterprises in Taiwan. Decis. Support Syst. 43, 842–852. doi:10.1016/j.dss.2006.12.012.

Li, X., Li, G., Zhang, Y., 2014. Identifying major factors affecting groundwater change in the North China plain with grey relational analysis. Water 1581–1600. doi:10.3390/w6061581.

Mann, H.B., 1945. Nonparametric test against trend. Econometrica 13 (3), 245–259. doi:10.2307/1907187.

Mbanguka, R.P., Lyon, S.W., Holmgren, K., Lopez, M.G., Jarsjö, J., 2016. Water balance and level change of Lake Babati, Tanzania : sensitivity to hydroclimatic forcings. Water 8. doi:10.3390/w8120572.

Ministry of Water, 2020. Design, construction supervision, operation and maintenance manual, volume I. In: Lema, I.N., Kaseva, M., Sabaya, W. (Eds.), Design of Water Supply Projects, 4th ed.. Mkuki na Nyota Publishers Ltd, DodomaTanzania.

Ministry of Water, 2014. Preparation of integrated water resources management and development plans (IWRMDP) for the Internal Drainage Basin (IDB) final report, Vol. 2 Water Availability, Part A – Surface Water. Dar es salaam.

Missi, C., Atekwana, E.A., 2020. Physical, chemical and isotopic characteristics of groundwater and surface water in the Lake Chilwa Basin, Malawi. J. Afr. Earth Sci. 162 (December 2019), 103737. doi:10.1016/j.jafrearsci.2019.103737.

Moriasi, D.N., Gitau, M.W., Pai, N., Daggupati, P., 2015. Hydrologic and water quality models: Performance measures and evaluation criteria. Trans. ASABE 58 (6), 1763–1785. doi:10.13031/trans.58.10715.

Olaka, L.A., Odada, E.O., Trauth, M.H., Olago, D.O., 2010. The sensitivity of East African rift lakes to climate fluctuations. J. Paleolimnol. 44 (2), 629–644. doi:10.1007/s10933-010-9442-4.

Onyutha, C., 2017. On rigorous drought assessment using daily time scale: Non-stationary frequency analyses, revisited concepts, and a new method to yield non-parametric indices. Hydrology 4 (4). doi:10.3390/hydrology4040048.

Rawlings, J.O., Pantula, S.G., a Dickey, D., 1998. Applied Regression Analysis : A Research Tool, 2nd ed. Springer Texts in Statistics.

Samy, A., Ibrahim, M.G., Mahmod, W.E., Fujii, M., Eltawil, A., Daoud, W., 2014. Statistical assessment of rainfall characteristics in upper Blue Nile basin over the period from 1953 to. Water 11 (3), 2019. doi:10.3390/w11030468, (Switzerland).

Sandstrom, K., 1995. The recent Lake Babati floods in semi-arid Tanzania - a response to changes in land cover? Geogr. Ann. Ser. A 77 A (1–2), 35–44. doi:10.1080/04353676.1995.11880427.

Stromquist, L., 1992. Environmental impact assessment of natural disasters, a case study of the recent Lake Babati Floods in Northern Tanzania. Swed. Soc. Anthropol. Geogr. 74 (2), 81–91. doi:10.2307/521286.

Swenson, S., Wahr, J., 2009. Monitoring the water balance of Lake Victoria, East Africa, from space. J. Hydrol. 370 (1–4), 163–176. doi:10.1016/j.jhydrol.2009.03.008.

Tate, E., Sutclifffe, J., Conway, D., Farquharson, F., 2004. Water balance of Lake Victoria: upadate to 2000 and climate change modelling to 2100. Hydrol. Sci. Sci. Hydrol. 49 (4), 563–574.

Tzeng, C.J., Lin, Y.H., Yang, Y.K., Jeng, M.C., 2009. Optimization of turning operations with multiple performance characteristics using the Taguchi method and Grey relational analysis. J. Mater. Process. Technol. 209 (6), 2753–2759. doi:10.1016/j.jmatprotec.2008.06.046.

USACE, 1994. Engineering and Design: Flood-Runoff Analysis. U.S. Army Corps of Engineers, Washington, D.C.

, 2013. United Republic of Tanzania: 2012 Population and Housing Census. National Bureau of Statistics, Dar es salaam [Online]. Available.

USGS, 2018. Shuttle Radar Topographic Mission. Global Land Cover Facility, University of Maryland, College Park, Maryland [Online]. Available.

Weitz, J., Demlie, M., 2013. Conceptual modelling of groundwater-surface water interactions in the Lake Sibayi Catchment, Eastern South Africa. J. Afr. Earth Sci. doi:10.1016/j.jafrearsci.2013.11.018.

Wong, H., Hu, B.Q., Ip, W.C., Xia, J., 2006. Change-point analysis of hydrological time series using grey relational method. J. Hydrol. 324, 323–338. doi:10.1016/j.jhydrol.2005.10.007.

Zohary, T., Ostrovsky, I., 2011. Ecological impacts of excessive water level fluctuations in stratified freshwater lakes. Int. Waters 1 (1), 47–59. doi:10.5268/IW-1.1.406.