



Multivariate framework for the assessment of key forcing to Lake Malawi level variations in non-stationary frequency analysis

Cosmo Ngongondo  · Yanlai Zhou · Chong-Yu Xu

Received: 4 December 2019 / Accepted: 23 July 2020 / Published online: 21 August 2020
© The Author(s) 2020

Abstract Lake Malawi in south eastern Africa is a very important freshwater system for the socio-economic development of the riparian countries and communities. The lake has however experienced considerable recession in the levels in recent years. Consequently, frequency analyses of the lake levels premised on time-invariance (or stationarity) in the parameters of the underlying probability distribution functions (pdfs) can no longer be assumed. In this study, the role of hydroclimate forcing factors (rainfall, lake evaporation, and inflowing discharge) and low frequency climate variability indicators (e.g., El Nino Southern Oscillation-ENSO and the Indian Ocean Dipole Mode-IODM) on lake level variations is investigated using a monthly mean lake level dataset from 1899 to 2017. Non-stationarity in the lake levels was tested and confirmed using the Mann-Kendall trend test ($\alpha = 0.05$ level) for the first moment and the F test for the second moment ($\alpha = 0.05$ level). Change points in the series were identified using the Mann-Whitney-Pettit test. The study also compared stationary and non-stationary lake level frequency during 1961 to 2004, the common peri-

od where data were available for all the forcing factors considered. Annual maximum series (AMS) and peak over threshold (POT) analysis were conducted by fitting various candidate extreme value distributions (EVD) and parameter fitting methods. The Akaike information criteria (AIC), Bayesian information criteria (BIC), deviance information criteria (DIC), and likelihood ratios (RL) served as model evaluation criteria. Under stationary conditions, the AMS when fitted to the generalized extreme value (GEV) distribution with maximum likelihood estimation (MLE) was found to be superior to POT analysis. For the non-stationary models, open water evaporation as a covariate of the lake levels with the GEV and MLE was found to have the most influence on the lake level variations as compared with rainfall, discharge, and the low frequency climatic forcing. The results are very critical in flood zoning especially with various planned infrastructural developments around the lakeshore.

Keywords Non-stationarity · Extreme lake level · Frequency analysis · Flood zoning · Lake Malawi

C. Ngongondo
Department of Geography and Earth Sciences, University of Malawi, Chancellor College, P.O. Box 280, Zomba, Malawi

C. Ngongondo (✉) · Y. Zhou · C.-Y. Xu
Department of Geosciences, University of Oslo, PO Box 1047, Blindern, 0316 Oslo, Norway
e-mail: cngongondo@cc.ac.mw

e-mail: cosmon@geo.uio.no

Introduction

Lakes are critical to the environment, biosphere, and human populations especially in sustaining the socio-economic livelihoods of many rural and poor communities' worldwide (Dubois et al. 2018, Sayer et al. 2018). Benefits from lakes include food, water for domestic use

and agriculture, reservoirs for hydropower generation, and transportation as well as providing a high recreation value for humans (Vainu and Terasmaa 2014, Kafumbata et al. 2014). However, these lakes have not been spared from the impacts of climate change and variability (CCV) and human-induced effects such as land-cover and land-use changes, urbanization, changes in impervious surfaces and drainage network, deforestation, and mining (Bayazit 2015). Such impacts are well exhibited through lake level variations, which have profound primary influences on the productivity and structure of lake ecosystems (Gownaris et al. 2018) thereby bringing a myriad of environmental and socio-economic problems (Ye et al. 2017). In addition, such variations have considerable influence in fields such as engineering design, ecological conservation, and environmental management around a lake region (Cui and Li 2016).

Globally, many studies have reported on decreasing lake levels, both natural and man-made. For instance, Fathian et al. (2016) reported on a persisting trend in the levels of Lake Urmia in Iran between 1966 and 2012, propagated by both natural forces (i.e., climate change) and human forces (i.e., diversion of surface water for upstream use, construction of dams, drought, and mismanagement). The human-made factors alone in Lake Urmia basin have been found to account for the loss of over 70% of the lake area (Garousi et al. 2013). Similarly, Okonkwo et al. (2014) found decreasing levels of Lake Chad, which were negatively associated with key climate forcing such as rainfall and inflowing river discharge and were closely associated with the El Niño Southern Oscillation (ENSO), low-frequency large-scale climate variability forcing. In addition, Yang and Lu (2014) noted on the complete disappearance of more than 350 lakes of greater than 1 km² in size across China, largely due to man-made influence such as water over-exploitation, land reclamation, and urbanization. The result by Yang and Lu (2014) is in contrast with that of Han et al. (2016) in their study in Lake Dongting who found increasing patterns in the maximum lake water level, annual mean lake water level, and annual minimum lake water level from 1961 to 2014 that were largely driven by the change of precipitation and the operation of reservoirs. Other studies that have reported on changes in lake levels (i.e., either decreases or increases) in various regions include the following: Úbeda et al. (2013) in the Ibera wetland of Argentina; Stefanidis et al. (2016) in their study of two

Mediterranean karstic lakes, namely Vegoritis and Petron in northern Greece; and Song et al. (2014) in the Tibetan Plateau. Awange et al. (2008) reported of recessions of Lake Victoria in east Africa between 1977 and 2001. These studies have associated such changes to either climate, man-made, or both, with non-uniform and non-region-specific associations on the lake levels across globally (Šraj et al. 2016).

For areas around lakes, there is an apparent need for estimating design values of lake level quantiles (Q_T) with their associated return periods (T) in order to inform various aspects such as engineering infrastructural-related safety measures, flood hazard assessments, flood zoning, and subsequent mitigation measures (Botero and Frances 2010). The common approaches are through flood frequency analysis (FFA) using historical records (Jacobellis et al. 2011; Machado et al. 2015). As opposed to the use of physically based models which have various levels of data needs and the concepts of probable maximum precipitation or flood (PMP or PMF) (François et al. 2019), FFA is quiet convenient as it applies statistical methods to derive these quantiles, as long as the records are long enough. Conventional FFA has over the years been dependent on the assumption of a stationary time series, which can be defined as those time series whose probability distribution function (pdf) is independent of time (Bayazit 2015) or a series whose marginal distribution remains invariant with time (Salas et al. 2018). Under stationarity, all realizations of a random variable X are assumed to be independent and identically distributed (*iid*) with a probability distribution function $F_X(x/\varnothing_-)$ and a time invariant parameter vector \varnothing_- (De Luca and Galasso 2018). With the documented evidence of natural and man-made and induced changes in lake levels globally, the concept of stationarity has been considerably challenged (Milly et al. 2008; Debele et al. 2017). Consequently, FFA approaches that incorporate time variance of the pdfs (or non-stationarity) is now being widely explored. Non-stationarity assumes that realizations of a random variable X are independent but not identically distributed (*inid*) with a probability distribution function $F_X(x/\varnothing_-(t))$ having a time variant parameter vector $\varnothing_-(t)$ that is a function of some covariates (De Luca and Galasso 2018). Both approaches have their own sources of uncertainty, with stationary FFA (SFA hereafter) deriving most of uncertainty from the estimation of flood quantiles with return periods beyond the observed values. On the other hand, uncertainty in non-stationary FFA (NFA hereafter) arises from the assumed function(s) linking parameters and covariates. In addition to a plethora of available

studies on both SFA and NFA, the merits and demerits of both approaches have also been reviewed by many authors (e.g., Bayazit 2015; Salas et al. 2018; François et al. 2019). What is clear from the literature is that the non-incorporation of non-stationarity in FFA may lead to over (or under) designing, with related negative societal and economic consequences (Salas et al. 2018). In addition, most of the studies have focused on SFA (and NFA) of either river discharge or precipitation series (e.g., Hounkè et al. 2014, Šraj et al. 2016, Ajami et al. 2017, De Paola et al. 2018, De Luca and Galasso 2018). This is despite there being a lot of studies providing evidence of lake level changes and an obvious need for design lake levels that take into consideration NFA. There is also an apparent need for an understanding of the role of key drivers of the lake level changes in NFA (Lopez and Frances 2013; Su and Chen 2019).

Lake Malawi is a critical transboundary water resource between Malawi, Mozambique, and Tanzania at the southern end of the Great East Africa Rift Valley system (Drayton 1984). According to Jury (2014), Lake Malawi's levels are quite sensitive to variations in the balance between mean annual rainfall over the lake (about 1.3 m), mean annual inflows from 4 main rivers (about 0.9 m), mean annual lake or open water evaporation (1.8 m), and mean annual outflow to the only outlet, the Shire River (0.4 m). The lake is the second largest and deepest among the Great East Africa Rift Valley lakes, after Lake Tanganyika (Neuland 1984). A long-term history of the lake levels from various data sources (proxy, in situ, and modeling) by Owen et al. (1990) and Nicholson (1998) suggests a complex variation pattern mainly attributed to climatic controls. The lakeshore areas are also prone to the extreme lake level variations at both low and high frequency, which affects both infrastructure and socio-economic livelihoods of the riparian communities. Since the peak of 1979, the lake levels have been undergoing a considerable recession (Fig. 1). Makwinja et al. (2017) developed a stochastic model for forecasting lake levels based on historical levels. Calder et al. (1994) attributed lake level changes from 1896 to 1967 to rainfall alone, but assumed constant evaporative demand and forest cover. To our knowledge based on available literature, a frequency analysis of the lake levels both under stationary and non-stationary conditions has not been undertaken. With various proposed infrastructural and major water abstraction projects (e.g., Salima-Lilongwe Water Project), there is a critical need for understanding the nature

of lake levels in terms of their magnitudes, return periods, and the role various forcing (e.g., low-frequency climate variability like El Niño Southern Oscillation (ENSO) on the design extreme lake levels). This aspect has motivated our present study. Therefore, this study is aimed at examining the nature of extreme lake levels under stationary and non-stationary conditions, using the case of Lake Malawi in East-Southern Africa. Specifically, this was achieved through the following: (1) temporal analysis of the lake levels for evidence of non-stationarity; and (2) developing models for frequency analysis of the lake levels under stationary and non-stationary conditions. The results are very critical in flood zoning especially with various planned infrastructural and related developments around the lakeshore. To our knowledge, there are no many studies (e.g., Drayton 1984, Neuland 1984, Calder et al. 1994, Kumambala and Ervine 2010a, b) that have been attempted to account for changes in the levels of Lake Malawi. In addition, there is no documented study on SFA or NFA of levels of Lake Malawi.

Study area, data, and methods

Study area

Lake Malawi is a sub-basin of the Zambezi River basin located in Southern Africa (Fig. 2).

The lake covers a total mean surface area of 28,760 km², total basin area of 150,000 km², volume of 7725 km³, 560 km long, maximum width of 75 km, maximum depth of 695 m, and mean depth of about 474 m and is Africa's third largest lake (Drayton 1984; Neuland 1984; Jury 2014; Sene et al. 2017). About a quarter of the total global freshwater is held by Lake Malawi together with Lakes Tanganyika and Victoria (Bootsma and Hecky 1993). The Lake Malawi basin experiences a tropical climate with a strongly seasonal rainfall pattern. Over 80% of the total annual rainfall occurs during the Austral summer months from November to April when the Intertropical Convergence Zone (ITCZ) and the Congo Air Boundary (CAB) are active in the region. The mean annual rainfall ranges from less than 800 mm/year on the southern lake shore to more than 2400 mm/year in the northern highlands. On the eastern shore of the lake, the Ruhuhu River basin in Tanzania with a catchment area of 14,070 km² is the sole major inflow and contributes about 20% of the total

Fig. 1 The extent of Lake Malawi recession showing the exposed lake bed in February, 2017. The boats' docking area used to be close to the exposed house terrace on the right



annual lake inflow, whereas the other major inflows each account for less the 10% of the inflows (Lyons et al. 2011 and references within). The second largest inflow is South Rukuru in Malawi with a total catchment area of 12,110 km². The Bua and Linthipe rivers, all in Malawi, are also significant tributaries with

respective catchment areas of 10,700 km² and 8560 km². The mean annual temperature during 1992 to 2007 is 22.4 °C over the land part of the basin and 24.8 °C over the water (Lyons et al. 2011). Over many parts of Malawi, Ngongondo et al. (2011, 2015) reported of rainfall declines that were not statistically

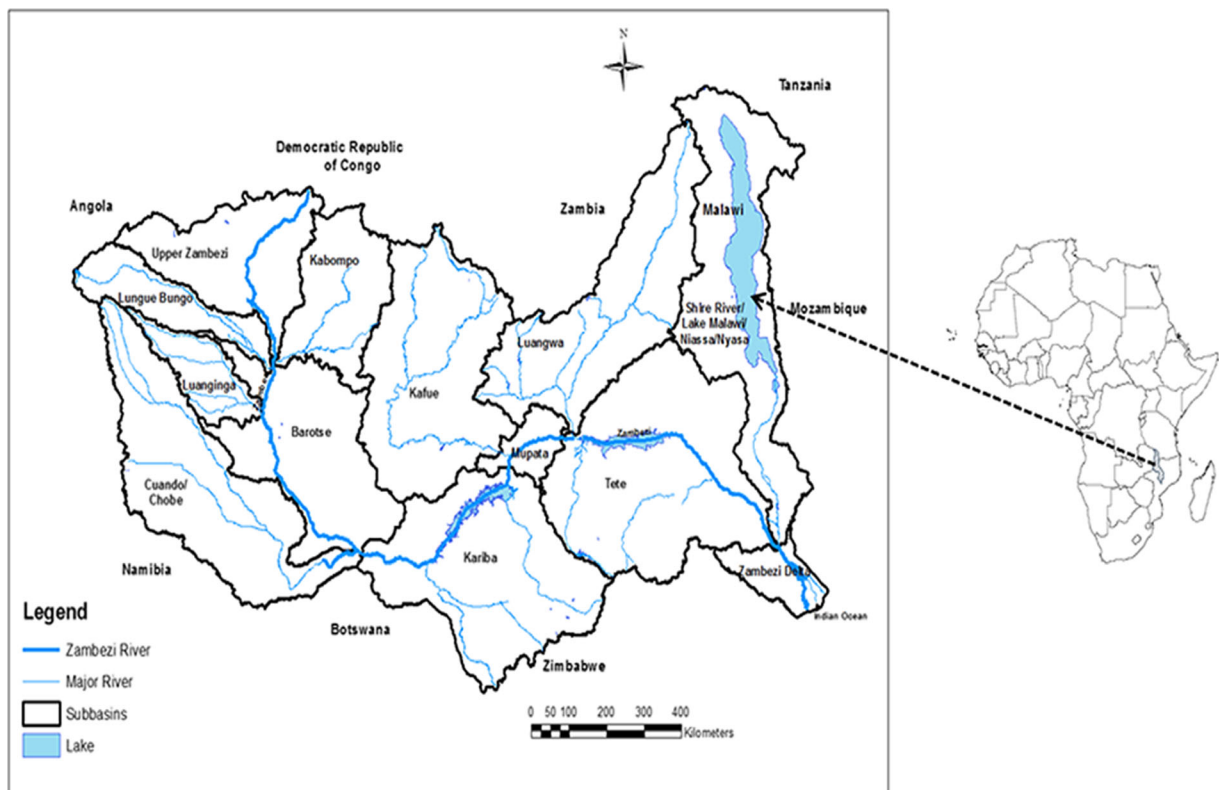


Fig. 2 Map of Africa and Zambezi basin showing location of Lake Malawi

significant, coupled with significant increases in mean temperature during 1960 to 2007 and 1971 to 2001, respectively. The lake provides a mean annual discharge of about $480 \text{ m}^3 \text{ s}^{-1}$ flows to the Shire River at Mangochi, the only outlet from the lake (Jury 2014). Over 90% of Malawi's hydropower installation is located at a series of rapids in the middle Shire River basin, downstream of the outlet. The Kamuzu Barrage at Liwonde, about 72 km from the outlet, regulates flows to the hydropower stations. There are also plans for a major water abstraction project from the lake for supply to Lilongwe City.

Lake level data and tests for stationarity

The main data source for the study was a long-term record of mean monthly lake levels, from 1899 to 2017 from which mean annual lake levels, annual maximum series (AMS), and partial duration series (PDS) for the peak over threshold (POT) analysis were derived. The monthly mean lake level is derived from three water level stations on the western shore of the lake at Chilumba (-10.43° S , 34.25° E), Nkhata Bay (11.72° S , 34.33° E) and Monkey Bay (14.08° S , 34.91° E). These data were obtained from the Malawi Department of Water Resources, Surface Water Resources in the Ministry of Agriculture, Irrigation and Water Development, through the Climate Justice Water Futures Programme. The lake levels are presented in meters Above Shire Valley Datum (m ASVD), which is approximately equal to sea level (Drayton 1984). To account for the Lake Malawi level variation, the following parameters were used as covariates: (1) annual rainfall data for three stations located along the Lake Malawi shore, namely Karonga, Nkhata Bay, Nkhotakota, Salima, and Mangochi. These were obtained from the Watch Forcing data (WFD) (Weedon et al., 2011 & 2014). The WFD is a gridded data set at 0.5° by 0.5° ($55 \times 55 \text{ km}$) grid boxes over the lake centered at these stations. These were aggregated into an aerial average by simple arithmetic average. There was no missing rainfall data from the WFD for the period: (2) depth of the mean annual discharge derived from 17 inflowing rivers into the lake on the Malawi part of the catchment as shown in Table 1. No river flow data were available from the Tanzania and Mozambique parts of the catchment on the eastern lake shore, but it was assumed that the 17 inflowing rivers could ably represent all the inflows.

From the 17 inflowing rivers, the mean annual depth of inflow into Lake Malawi was calculated as a weighted mean (Q_T) from:

$$Q_T = \frac{1}{A_T} \sum_{i=1}^N Q_i A_i \quad (1)$$

where Q_i and A_i are the annual mean discharges and catchment areas, respectively, of each inflow, and $A_T = 51,692 \text{ km}^2$ is the total catchment area of the $N = 17$ inflowing rivers. The data was already quality controlled by the Malawi Department of Water Resources and each station had less than 5% of missing record: (3) open water evaporation at five weather stations along the Lake Malawi shore, namely Karonga, Nkhata Bay, Nkhotakota, Salima, and Mangochi. This was estimated from the mean temperature and station latitude using Thornthwaite's model (Beguería and Vicente-Serano, 2017). Results of the Thornthwaite's model were also compared with those from pan measurements at the stations, although the later had some gaps. A simple arithmetic average was used to calculate the mean open water evaporation. The monthly temperature data used was already quality controlled by the Malawi Department of Climate Change and Meteorological Services. Missing records accounted for less than 4% at each station and were filled by averaging readings from nearest stations: (4) the Southern Oscillation Index (SOI) sourced from <https://www.ncdc.noaa.gov/teleconnections/enso/indicators/soi/>. The SOI is an index of pressure changes between west and eastern tropical pacific (at Tahiti and Darwin, Australia) and is related to the El Niño Southern Oscillation (ENSO). The ENSO is a low-frequency planetary scale climatic variation pattern (Vasiliades et al. 2015) which has significant influences on the rainfall pattern of southern Africa. Negative values of the SOI indicate El Niño conditions, whereas positive values are associated with La Niña conditions and values around zero suggest normal conditions. Both strong and moderate to weak El Niño events have been found to result in drier than normal conditions over Southern Africa (Pomposi et al. 2018 and references within) whereas La Niña conditions often result in above normal rainfall; (5) and the Indian Ocean Dipole (IOD) Mode Index (DMI), a coupled oceanic-atmospheric inter-annual system characterized by sea surface temperature (SSTs) anomalies in the tropical western and eastern Indian Ocean. Negative (positive) IOD phases are marked by lower (higher) SSTs in the

Table 1 Some major flows (1 to 17) into Lake Malawi on the western shoreline and the only outflow (1)

Serial	River	Station	ID#	Lat (°S)	Long (°E)	Area (km ²)
1	Songwe	Mwandenga	9B7	- 9.59	33.77	2870
3	Lufira	Ngerenge	9A2	- 9.81	33.84	1890
4	Nrukuru	Mwakimeme	8A5	- 9.93	33.79	2091
5	North Rumphu	Chiweta	7H3	- 10.69	34.18	6,78
6	South Rukuru	Mlowe	7G18	- 10.75	34.21	12,027
7	Lweya	Zayuka	16F2	- 11.78	34.20	2320
9	Dwambadzi	Nthanda	16E6	- 12.24	33.98	1770
10	Dwangwa	Rupashe	6D10	- 12.51	34.12	7768
11	Bua	S3RoadBridge	5C1	- 12.79	34.20	10,659
12	Chirua	Mtambe	15A4	- 13.46	34.24	150
13	Lingadzi	Kaniche	15A8	- 13.54	34.24	450
14	Linthipe	M5 Bridge	4B1	- 13.79	34.45	8180
15	Nadzipulu	Mtakataka	3F3	- 14.21	34.51	224
16	Namikokwe	Mua Mission	3E1	- 14.28	34.51	138
17	Livulezi	Khwekwerere	3E3	- 14.44	34.54	477
1	Shire	Mangochi	1T1	- 14.48	35.27	126,500

south western equatorial Indian Ocean, thereby affecting the rain bearing systems of Eastern and Southern Africa (Marchant et al., 2006). Unlike ENSO, whose effect on climate systems is global, the IOD has been found to more localized to the oceanic and land masses around the Indian Ocean. The study further assumed that all lake level variations can be explained by these covariates (rainfall, discharge, evaporation, ENSO, and IOD), independent of time. To ensure that the estimated parameters were comparable in magnitude, the values of each covariate were rescaled (Prosdociami et al. 2015) to the range [0, 1] using:

$$z = \frac{x_i - \min\{x_1, \dots, x_n\}}{\max\{x_1, \dots, x_n\} - \min\{x_1, \dots, x_n\}} \quad (2)$$

where z is the rescaled value of each covariate; x_i is the value of each covariate in each year from 1958 to 2001; $\min\{x_1, \dots, x_n\}$ is the minimum value of each covariate during the period from 1958 to 2001; $\max\{x_1, \dots, x_n\}$ is the maximum value of each covariate during the period from 1958 to 2001. The non-parametric Pettit (Pettitt 1979) was applied to detect abrupt change points in the raw data series. The test assumes that in series of observed data x_1, x_2, \dots, x_n , a change point can be found at an unknown time t , where the series x_1, x_2, \dots, x_t has a distribution $F_1(x)$ that is different from the distribution $F_2(x)$ of the series x_{t+1}, x_2, \dots, x_n (Jaiswal et al., 2015 and

references within). The change point K_T in the Pettit Test is given as:

$$K_T = \max|U_{t,T}| \quad (3)$$

where $U_{t,T} = \sum_{i=1}^t \sum_{j=t+1}^T \text{Sign}(x_i - x_j)$.

The significance of K_T at 0.05 level is estimated from $p \approx 2 \exp\left(\frac{-6K_T^2}{T^3 + T^2}\right)$.

The change point model (CPM) framework (Ross 2015) in R software was applied to identify all possible change points in the raw lake levels as well as the first difference lake level series. Furthermore, the Student's t test (Li et al. 2013) was used to test if means of the lake levels before and after a change point are statistically different.

To detect the significance of trends in the first moment of lake levels (mean), the Mann-Kendall (MK) (Mann 1945; Kendall 1975) was used at $\alpha = 0.05$ significance level. Under the null hypothesis (H_0) of no trend in the data, and the alternative hypothesis (H_1) of a trend, the H_0 is rejected when the MK statistic $|uc| > u_{1-\alpha/2}$, corresponding to a $1-\alpha/2$ quantile of the standard normal distribution (Hisdal et al. 2001; Karlsson et al. 2014). The MK is the most widely used rank-based non-parametric test (Kundzewicz and Robson, 2004; Bayazit 2015). The MK test is considered robust as compared with other tests and is recommended by the World

Meteorological Organisation (WMO) for application in the detection of monotonic trends in hydrometeorological variables (WMO 1988).

In addition, the F test was used to determine any differences in the second moment (variance) of the lake level time series. The test was undertaken under the null hypothesis (H_0) of equal variance ($H_0 : \sigma_1^2 = \sigma_2^2$) and an alternate hypothesis (H_1) of unequal variances ($H_1 : \sigma_1^2 \neq \sigma_2^2$) as follows: $F_c = s_1^2/s_2^2, s_1^2 > s_2^2$, and H_0 are rejected if $F_c \geq F_{1-\alpha, n_1-1, n_2-2}$ where $F_{1-\alpha, n_1-1, n_2-2}$ is the critical value from the F-distribution table (Li et al. 2013).

Stationery and non-stationary frequency analysis

Extraction of extreme lake levels

Extreme lake levels were determined using two approaches from the lake levels (Rosbjerg and Madsen 2004, Ngongondo et al. 2015 and references within): through analysis of the annual maximum series (AMS) and partial duration series (PDS), also called peak over threshold (POT). AMS involves a series composed of one maximum mean lake level in each hydrological year (November to October), resulting in a total of 119 values for the period 1899 to 2017 in this study. Apart from rearranging the data into a hydrological year, the extraction of the AMS series from the lake levels data is a relatively straight forward procedure. If (X_1, X_1, \dots, X_n) is a sequence of independent random variables (in this case, lake levels), then $M_n = \max \{X_1, X_1, \dots, X_n\}$ is the maximum lake level from the series in a particular year (De Paola et al. 2018).

For the POT analysis, a mean excess (ME) plot (Ghosh and Resnick 2010) of the entire series was firstly used to pre-determine the threshold (u) lake level (Madsen et al., 1997), above which levels recorded should be considered to be extremes. According to Ghosh and Resnick (2010), the distribution of u is assumed to have a function F where:

$$F_u(x) = P [X-u \leq x | X > u] \tag{4}$$

and the ME function of the form:

$$M(u) = E [X-u | X > u] \tag{5}$$

The assumption in $M(u)$ is that the lowest value of the selected threshold u has an approximately linear relationship with the mean excess $E(X-u)$ (Coles 2001).

Extreme value distributions and parameter fitting

In this study, a two-tier approach was adopted for lake level frequency analysis (FFA) under stationary and non-stationary conditions (SFA & NFA). The first part was the identification of the best probability distribution functions (pdf). The AMS and PDS series fitted candidate pdfs of members from the extreme value distributions (EVD) family that are among those commonly applied in hydrological studies (Debele et al. 2017), namely Gumbel, generalized extreme value (GEV), generalized Pareto (GPA), exponential (Exp), and Pearson type 3 (PP3). There are many parameter fitting methods for probability distributions for hydrological assessments, as summarized by Singh (1998). In this study, the performances of maximum likelihood estimation (MLE), generalized maximum likelihood estimation (GMLE), Bayesian (through Markov chain Monte Carlo (MCMC) simulations), and L-moments (Table 2) on the candidate distributions were tested and compared with select the best. Each of these methods has their own merits and demerits and the reliability often depends on the length of the available record (Martins and Stedinger 2000). A total of 15 different models based on the candidate distributions, series (AMS and POT), and parameter fitting methods were assessed under SFA, from which the best approach was selected.

For NFA, the study followed the generalized additive models for location, scale and shape (GAMLSS) framework approach (Rigby and Stasinopoulos 2005) as in among others, Machado et al. (2015), Prosdocimi et al. (2015), Debele et al. (2017), and Su and Chen (2019). A total of 19 models with different combinations of the covariates in the location ($\hat{\mu}$) and scale parameter ($\hat{\sigma}$) were developed (Table 3). In addition to the stationary model (AMS₀), three further groups of non-stationary models can be identified in Table 3:

1. Those with varying $\hat{\mu}$ but stationary $\hat{\sigma}$. This group also has covariate combined models in $\hat{\mu}$;
2. Those with varying $\hat{\sigma}$ but with stationary $\hat{\mu}$; and
3. Those with both $\hat{\mu}$ and $\hat{\sigma}$ varying.

Both linear and non-linear link functions between the covariates and the location and scale parameters (μ and σ) were tested to select the best approach. For the three

Table 2 EVD and the parameter fitting methods for SFFA and NFFA

Series	EVD	Parameter fitting methods	Covariates
AMS	Gumbel, GEV	ML, GMLE, Bayesian, L-moments	None
PDS	GP	ML, GMLE, Bayesian, L-moments	None
PDS	PP	ML, GMLE, Bayesian	None
PDS	Exponential	ML, GMLE	None
AMS	Gumbel, GEV	ML, GMLE, Bayesian, L-moments	Rainfall, discharge, evaporation, SOI & IOD

parameter distributions (GEV and PE3), the shape parameter (ξ) was assumed to be constant due to difficulties in the reliability of its estimates (Šraj et al. 2016, Gillend and Katz 2016). The best fitting model was evaluated through a multi-criteria approach comprised of goodness of fit (GOFs) measures, namely the Akaike information criteria (AIC) and Bayesian information criteria (BIC) which should be a minimum and the deviance information criteria (DIC) (for Bayesian parameter estimation only) which should be a maximum.

In addition, likelihood ratios (RL) between the distribution fitting approaches for nested models were assessed at $\alpha = 0.05$ significance level (Coles 2001; Madsen et al. 2014; Prosdocimi et al. 2015; Su and Chen 2019). In the extreme lake level frequency analyses, the package extRemes (Gilleland 2018) was applied in R statistical software (R Core Team, 2013).

P rainfall, *Disch* discharge, *Evap* evaporation, *IOD* Indian Ocean Dipole Index, *SOI* Southern Oscillation Index

Table 3 Models for SFA and NSFA that were tested

Model type	ID	Description	
		μ	σ
Stationary	AMS ₀	μ_0	σ_0
Non-stationary	AMS ₁	$\mu_0 + \mu_1 * P$	σ_0
	AMS ₂	$\mu_0 + \mu_1 * D$	σ_0
	AMS ₃	$\mu_0 + \mu_1 * \text{Evap}$	σ_0
	AMS ₄	$\mu_0 + \mu_1 * \text{IOD}$	σ_0
	AMS ₅	$\mu_0 + \mu_1 * \text{SOI}$	σ_0
	AMS ₆	$\mu_0 + \mu_1 * \text{Evap} + \mu_2 * P$	σ_0
	AMS ₇	$\mu_0 + \mu_1 * \text{Evap} + \mu_2 * \text{Disch}$	σ_0
	AMS ₈	$\mu_0 + \mu_1 * P + \mu_2 * \text{Disch}$	σ_0
	AMS ₉	$\mu_0 + \mu_1 * P + \mu_2 * \text{Disch} + \mu_3 * \text{Evap} +$	σ_0
	AMS ₁₀	μ_0	$\sigma_0 + \sigma * P$
	AMS ₁₁	μ_0	$\sigma_0 + \sigma * D$
	AMS ₁₂	μ_0	$\sigma_0 + \sigma * \text{Evap}$
	AMS ₁₃	μ_0	$\sigma_0 + \sigma * \text{IOD}$
	AMS ₁₄	μ_0	$\sigma_0 + \sigma * \text{SOI}$
	AMS ₁₅	$\mu_0 + \mu_1 * \text{Evap}$	$\sigma_0 + \sigma_1 * \text{Evap}$
	AMS ₁₆	$\mu_0 + \mu_1 * P$	$\sigma_0 + \sigma_1 * P$
	AMS ₁₇	$\mu_0 + \mu_1 * D$	$\sigma_0 + \sigma_1 * D$
	AMS ₁₈	$\mu_0 + \mu_1 * \text{IOD}$	$\sigma_0 + \sigma_1 * \text{IOD}$
	AMS ₁₉	$\mu_0 + \mu_1 * \text{SOI}$	$\sigma_0 + \sigma_1 * \text{SOI}$

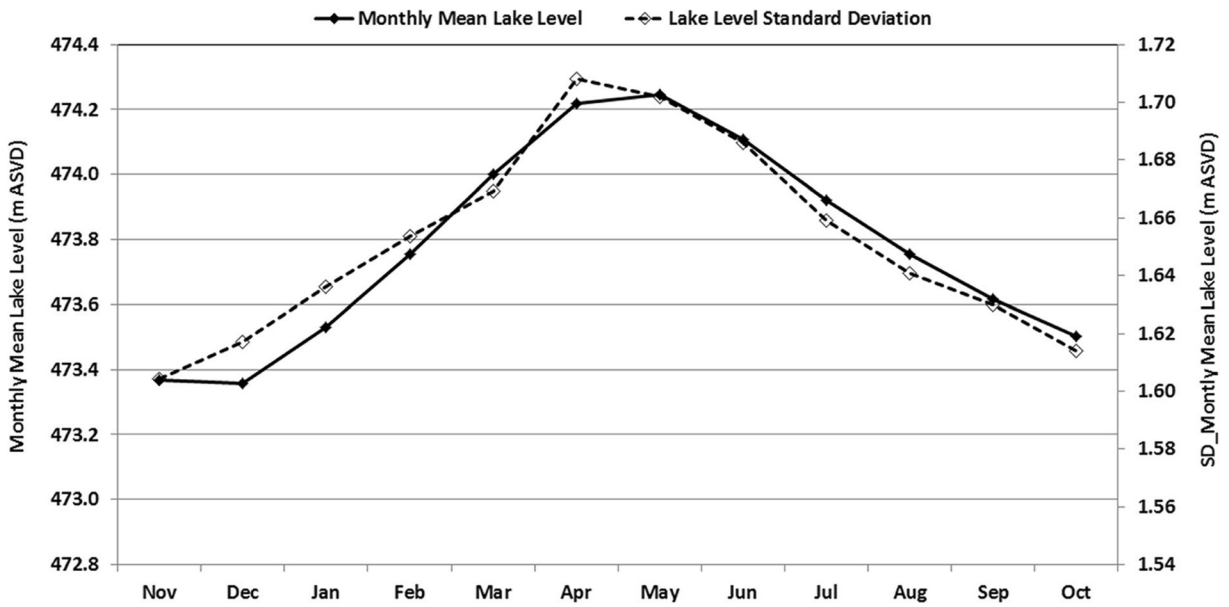


Fig. 3 Monthly mean levels of Lake Malawi and their standard deviations during 1899 to 2017

Results and discussions

Temporal hydro-climatic pattern of Lake Malawi basin (1899 to 2017)

The mean monthly lake levels and their standard deviations from 1899 to 2017 are shown in Fig. 3. The lake levels peak at the end of the rain season between March and May (about 474.25 m ASVD). This peak is followed by a recession with minimum levels occurring

between October and December (about 473.36 m ASVD). These two extremes give lake level range of about 1 m. The standard deviation pattern closely follows the mean lake levels.

The long-term temporal pattern of the levels (Fig. 4) shows a considerably complex variation pattern from the late 19th century to the early 21st century. The lake level had a mean annual lake level of 473.8 m ASVD (Stdev = 1.65 m ASVD) with a coefficient of variability (CV) of 0.349%. The CV suggests relatively low

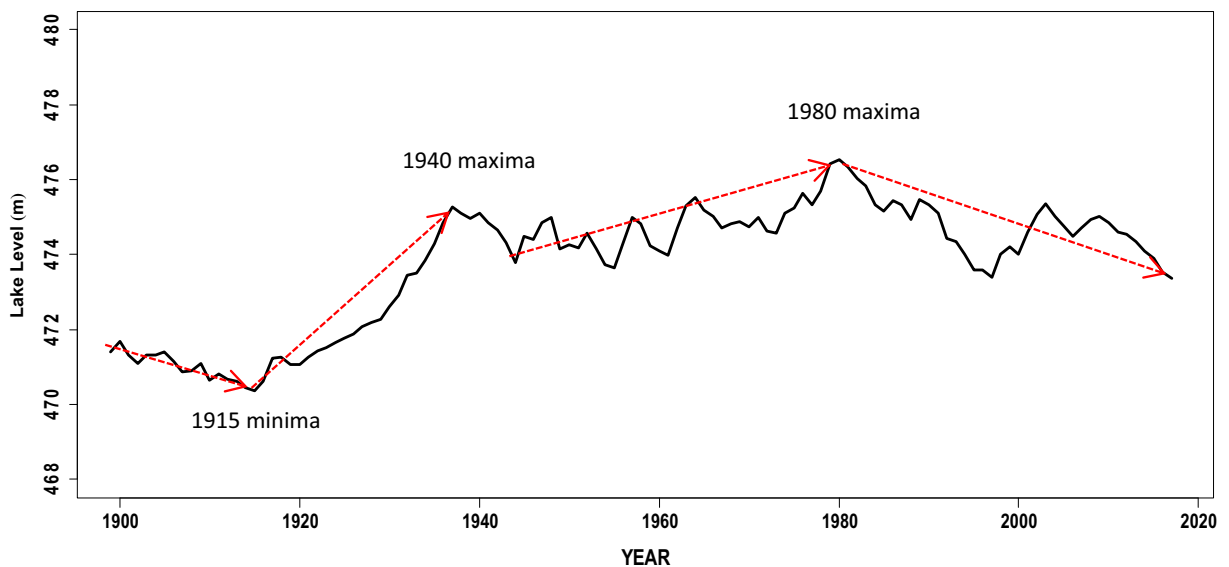


Fig. 4 Hydrograph of mean annual levels of Lake Malawi from 1899 to 2017

Table 4 Mann-Whitney multiple change points and detection years of Lake Malawi levels

*CD_Year	CD_Level	*CP_Year	CP_Level	#Years
1918	471.2592	1906	471.16	12
1926	471.8775	1916	470.62	10
1936	474.8392	1926	471.88	10
1946	474.3908	1934	473.83	12
1976	475.6258	1962	474.72	14
1982	475.3375	1975	475.24	7
1995	473.5967	1983	475.84	12
2005	474.7783	1991	475.09	14
2011	474.5908	2000	474.02	11

*CD year of detecting change in series, CP year the actual change occurred

variation around the mean lake level during the period. Based on the available empirical record from 1899 (Fig. 4), the lake levels had a decreasing trend up to a minima of about 470.4 m ASVD in 1915. Records (e.g., Sene et al. 2017; Nicholson et al. 2018) also show persistent extreme drought conditions over many parts of Malawi due to low rainfall during the late 19th century and early 20th century. The extremely low lake levels exposed a sandbar at the lake outlet which consequently blocked flows into the Shire River between 1915 and 1935. However, Sene et al. (2017) suggest that the Shire River actually stopped flowing in 1908 due to the extremely

low lake levels. The lake levels started to rise almost linearly from 1915, reaching a local peak of 475.1 m ASVD in 1940. Neuland (1984) attributed this steady rise in lake levels to a blockage of the Shire River by the tributary Nkasi River. Flows to the Shire River were restored around 1935 when the lake levels overtopped the sandbar with a lake level of 473 m ASVD. The hydrograph of the nearby Lake Chilwa, located to the south east of Lake Malawi, also shows a similar temporal lake level variation pattern up to around 1980 (Nicholson 2001). After the May 1980 peak, the lake levels have mostly been on a downward trajectory. Local peaks can be seen around 2003 and 2009.

The MK statistic for the annual lake levels during the period 1899 to 2017 had a statistically significant positive trend at $\alpha = 0.05$ level. This trend is obviously influenced by the extremely low level of 1915 and the peak level of 475.8 m ASVD in May 1980. For the period 1935 to 2017, the lake levels did not have a statistically significant MK trend at $\alpha = 0.05$. However, the levels had a negative MK trend during 1980 to 2017 that was statistically significant at $\alpha = 0.05$ level, with a linear regression trend of -0.044 m per year. From the raw lake level series, the Mann-Whitney multiple change point detection test detected and identified nine change points in the annual lake levels during the period 1899 to 2017 (Table 4 and Fig. 5).

However, the annual lake levels suggest high auto-correlations up to lag = 30 years. The large lake

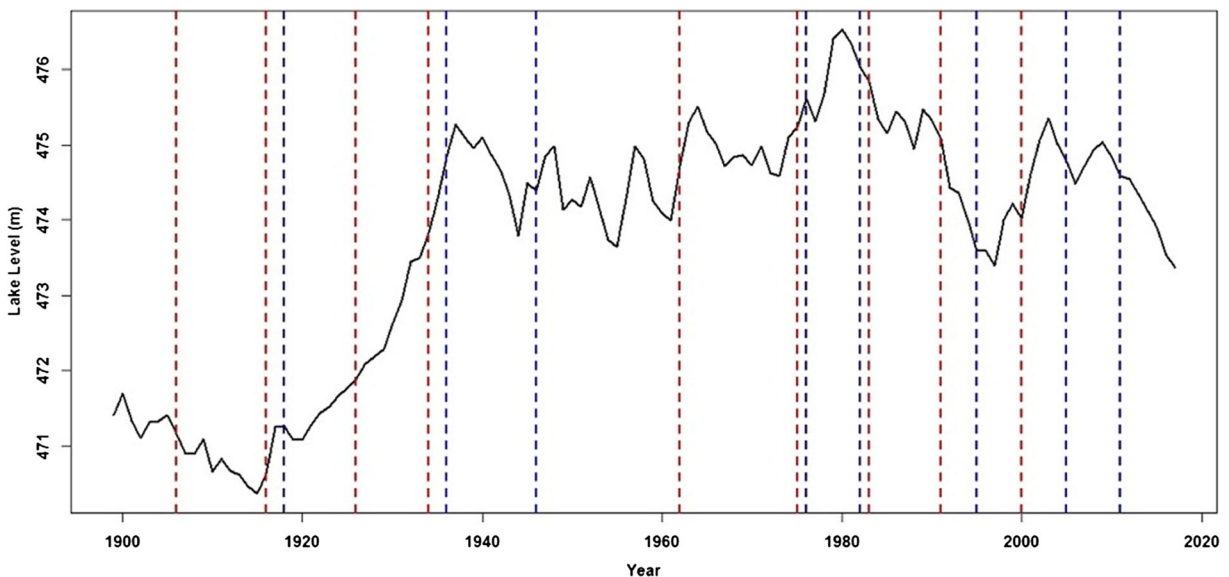


Fig. 5 Multiple change point detection and points from normal lake levels between 1899 and 2017. The dashed vertical blue lines are actual change points, and the dashed redlines line are when the change point was identified

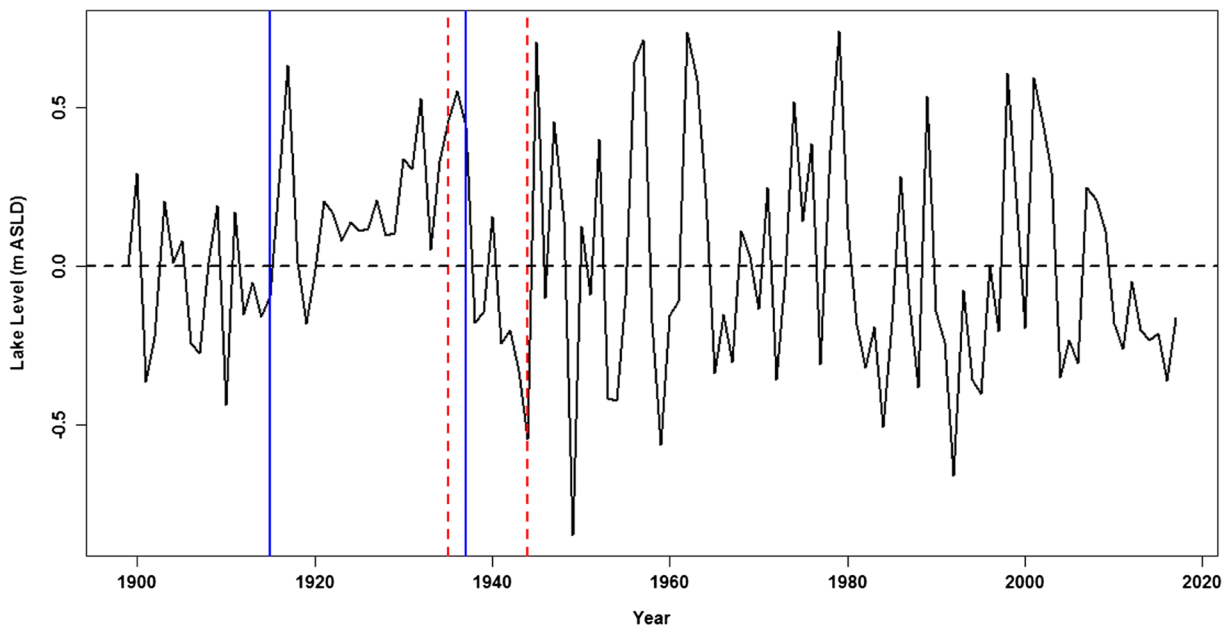


Fig. 6 Multiple change point detection and points from first difference lake levels during 1899 to 2017. The straight vertical line is the actual change point and the dashed line is when the change point was identified

autocorrelations should be expected as the lake has a large residence time (τ) of 114 years (Bootsma and Hecky 1993). This is an indication that the series violated the assumption of *iid* in the tests. To consider the influence of the autocorrelations on the multiple detected and identified change points, a first difference ΔL_t of the lake level series (L_t) was therefore used to identify the change points as follows:

$$\Delta L_t = L_t - L_{t-1} \tag{6}$$

where L_t and L_{t-1} are the lake levels at time t and $t - 1$, respectively. The difference series (Fig. 6), which were *iid*, identified two significant change points in 1915 (detected in 1937) and 1937 (detected in 1946). These two generally coincide with the period of extremely low lake levels, when the flows in the Shire River stopped. For the two time slices from 1899 to 1946 and 1947 to 2017, the F test under the null hypothesis of equal variances ($\sigma_1^2 = \sigma_2^2$) of the lake levels could not be accepted as $F = \sigma_1^2 / \sigma_2^2 > F_c$, for the critical value F_c . From the foregoing, the lake levels were found to be non-stationary in both the first (mean) and second moments (variance). Therefore, the second change point detected (1946) was the basis of a decision for the study to use the period from 1958 to 2012 for the lake level extremes analysis.

On the other hand, annual rainfall over the lake from 1958 to 2012 (Table 5) had a mean of 1236.1 mm (Stdev = 200 mm) and a relatively low coefficient of variation (CV) of 16.0%. The CV is actually slightly higher than that reported by others (e.g., Lyons et al. 2011, Sene et al. 2017). At the individual stations, the rainfall peaks around Nkhata Bay and Nkhotakota area whereas Karonga and Monkey Bay to the extreme north and south respectively had the lowest mean annual rainfall. In addition, the CV indicated that the rainfall pattern varies more inter-annually at the individual stations as compared with the basin-wide mean rainfall. The mean annual rainfall pattern across the catchment also shows a generally decreasing trend during the period (Fig. 7),

Table 5 Summary of rainfall in Lake Malawi basin during 1958 to 2012

Station	Mean (mm)	Stdv (mm)	CV (%)
Karonga	969.86	178.56	0.18
Mangochi	974.36	201.64	0.21
Monke Bay	963.40	180.57	0.19
Nkhata Bay	1766.15	394.97	0.22
Nkhotakota	1601.53	307.62	0.19
Salima	1140.98	200.48	0.18
Average	1236.05	201.55	0.16

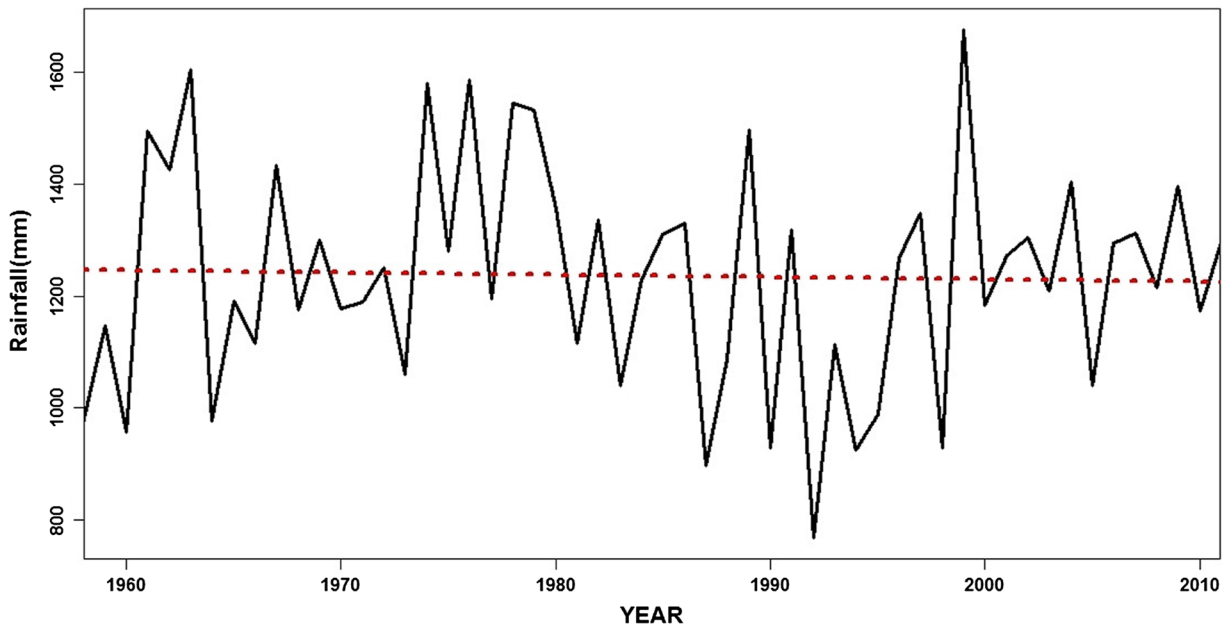


Fig. 7 Temporal pattern of mean annual rainfall in the Lake Malawi basin during 1958 to 2012

although no change point was detected by the Mann-Whitney-Pettit test. However, the MK trends were not statistically significant at $\alpha = 0.05$ level, for both the spatial rainfall as well as the individual stations.

Furthermore, the temperature regime during 1961 to 2004, the period of available record, shows that the basin has undergone considerable temperature increases (Fig. 8). The most recent decade from 1990 to 2001 was

the warmest. The mean annual temperature during the period had a positive MK statistic that was significant at $\alpha = 0.05$ level. The year 1979 was identified as the change point in the mean annual temperature regime by the Mann-Whitney test. The warning due to increased temperature has potential to increase the open water evaporative demand over the lake. Both temperature increases and rainfall decreases have also been

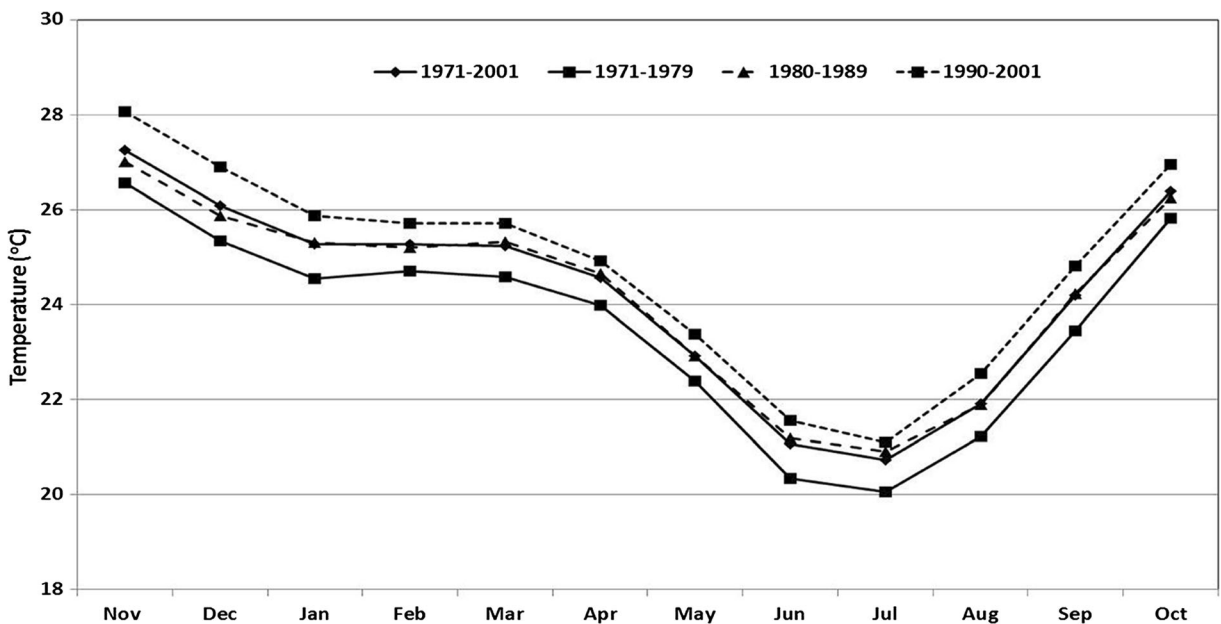


Fig. 8 Inter-decadal monthly mean temperature changes in the Lake Malawi between 1971 and 2001

Table 6 Candidate distributions under SFA

Serial	Series	EVD	Method	$\hat{\mu}(CI)$	$\sigma(CI)$	$\xi(CI)$	AIC	BIC	NLL	DIC
1	AMS	<i>Gumbel</i>	ML	475.16 (474.93, 475.16)	0.73 (0.58, 0.89)	NA	108.93	112.50	52.46	NA
2	AMS	<i>GEV</i>	ML	475.28 (475.04, 475.52)	0.74 (0.58, 0.91)	- 0.29 (- 0.47, 0.11)	103.56	108.92	48.78	NA
3	POT	<i>GPA</i>	ML	NA	0.63(0.38,0.88)	- 0.48 (- 0.80, - 0.17)	9.31	13.18	2.66	NA
4	POT	PP3	ML	477.07 (476.92, 477.21)	0.11 (0.026, 0.20)	- 0.48 (- 0.80, - 0.17)	- 250.14	- 244.35	- 128.07	NA
5	POT	<i>Exp</i>	ML	NA	0.41 (0.30, 0.52)	NA	13.35	15.28	5.67	NA
6	AMS	<i>Gumbel</i>	GMLE	475.16 (474.93, 475.39)	0.73 (0.58, 0.89)	NA	108.927	112.5	52.46	NA
7	AMS	GEV	GMLE	475.29 (475.04, 475.53)	0.75 (0.58, 0.92)	- 0.32 (- 0.48, 0.15)	2.00E+16	2.00E+16	1.00E+16	NA
8	POT	GPA	GMLE	NA	0.52 (0.27, 0.77)	- 0.26 (- 0.78, 0.26)	2.00E+16	2.00E+16	1.00E+16	NA
9	POT	PP3	GMLE	477.20 (476.10, 478.30)	0.20 (- 0.45, 0.85)	- 0.26 (- 1.39, 0.86)	2.00E+16	2.00E+16	1.00E+16	NA
10	POT	Exp	GMLE	NA	0.41 (0.30, 0.524)	NA	13.35	1.53E+01	5.67	NA
11	AMS	<i>GEV</i>	<i>Bayesian</i>	475.26 (474.96, 475.56)	0.78 (0.61, 1.03)	- 0.26 (- 0.47, 0.02)	NA	NA	NA	302.0
12	POT	GPA	Bayesian		0.59 (0.37, 0.86)	- 0.37 (- 0.71, 0.096)	NA	NA	NA	21.8
13	POT	PP3	Bayesian	477.2 (476.97, 477.77)	0.18 (0.055, 0.51)	- 0.37 (- 0.77, 0.023)	NA	NA	NA	- 758.8
14	AMS	GEV	L-Moments	475.3 (475.02, 475.51)	0.75 (0.58, 0.91)	- 0.32 (- 0.54, - 0.12)	NA	NA	NA	NA
15	POT	GPA	L-Moments		0.52 (0.34, 0.78)	- 0.26 (- 0.66, 0.026)	NA	NA	NA	NA

Italicized are acceptable pdfs, their parameter fitting methods, AIC, BIC, NLL, and DIC for modeling AMS or POT lake levels

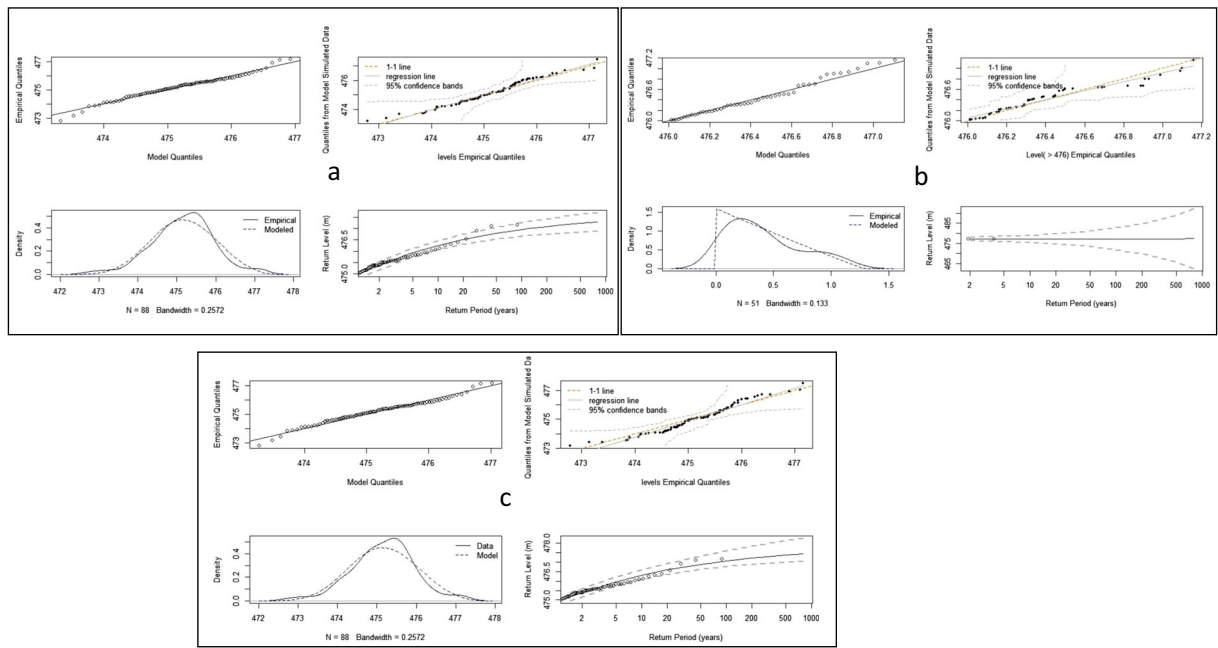


Fig. 9 Diagnostic plots for (a) AMS (GEV, MLE); (b) POT (GPA, MLE); and (c) AMS (GEV, Bayesian)

reported by other studies (e.g., McSweeney et al. 2010, Ngongondo et al. 2015). The mean annual open water evaporation over the lake was 1290 mm (CV = 7.44%), with a statistically significant positive MK trend at $\alpha = 0.05$ level for all stations as well as the average. However, the annual mean evaporation is lower than the 1872 mm that was reported by Crossley et al. (1990) for 1954 to 1980. The evaporation from Thornthwaite’s model was compared with the annual lake evaporation derived from pan measurements for the same stations from 1971 to 1996. It should be noted that these were scanty for which annual values for some years only could be savaged. A pan conversion factor of 1.3 was applied (Linacre, 1994). The pan-derived evaporation were slightly higher (about 2%) than those from Thornthwaite’s model, but the two were strongly correlated according to Pearson’s correlation ($\rho > 88\%$, $p < 0.05$), and both having a significant positive MK trend $\alpha = 0.05$). The Thornthwaite model-derived evaporation was therefore considered to be a reasonable estimate of lake evaporation. Furthermore, the weighted mean annual river discharge for all inflowing rivers into the lake considered was 209 mm and did not have a statistically significant MK trend, although Bua and Lufira Rivers had significant positive and negative trends respectively at $\alpha = 0.05$. The mean is lower than the 300 mm found by Owen et al. (1990) during 1954

and 1980. The Mann-Whitney test identified a change points in river discharge around 2001 and 1977 and 1992 for evaporation, with increases in both variables after the change points.

The Shire River is the only outlet from Lake Malawi at Mangochi, with a mean daily discharge of 473.0 m^3s^{-1} between 1957 and 2001. The daily discharge had a statistically significant positive trend the period from 1957 to 2001, whereas those for monthly and annual timescales were not significant at $\alpha = 0.05$ level. However, it should be noted that the Shire River outflow was not considered as a covariate in this study, as the river is regulated with barrage operation rules about

Table 7 Return levels for AMS with GEV and MLE

T-Year	Level (95% CI)	
	AMS	POT
2-year	475.41 (475.20, 475.61)	477.11 (475.97, 478.25)
5-year	476.03 (475.80, 476.27)	477.18 (475.41, 478.94)
10-year	476.37 (476.11, 476.63)	477.21 (474.75, 479.67)
20-year	476.65 (476.33, 476.96)	477.24 (473.80, 480.67)
50-year	476.95 (476.54, 477.35)	477.26 (471.92, 482.60)
100-year	477.13 (476.63, 477.63)	477.27 (469.82, 484.73)

Table 8 Test results for models with covariates in the location parameter ($\hat{\mu}$)

Model#	Covariate(s)	AIC	BIC	NLL	LR (<i>p</i> value)	Nested model
AMS ₁	Rainfall	98.87	106.01	45.43	0.0097	AMS ₀
AMS ₂	Discharge	99.31	106.45	45.66	0.0125	AMS ₀
<i>AMS₃</i>	<i>Evaporation*</i>	<i>92.47</i>	<i>99.61</i>	<i>42.23</i>	<i>0.0003</i>	AMS ₀
AMS ₄	IOD	102.43	109.57	47.22	0.0771	AMS ₀
AMS ₅	SOI	103.82	110.96	47.91	0.1879	AMS ₀
AMS ₆	Evap+Rainfall	94.30	103.13	43.10	0.6067	AMS ₃
AMS ₇	Evap+Discharge	91.37	100.29	40.69	0.07844	AMS ₃
AMS ₈	Rainfall+Discharge	100.28	109.20	45.14	0.4441	AMS ₁
AMS ₉	Evap+Rainfall+Discharge	92.52	103.22	40.26	0.0018	AMS ₈

In italics are the best non-stationary model

72 km downstream from the lake for purposes of lake level stabilization (Shela 2000).

Stationary lake level frequency analysis

This section presents results of candidate models for lake level frequency analysis under stationary conditions (SFA). Table 6 shows members of the EVD family that were evaluated, the parameter fitting methods, and the evaluation criteria results. Based on the evaluation criteria (lowest AIC/BIC/NLL and highest DIC), the results suggest that the best models for extreme lake levels were the GEV and Gumbel distributions (serials 2 and 3 in Table 6) when fitted to the AMS series and the GPA and exponential distributions when fitted to the POT series with a lake level threshold (*u*) of 476.1 m ASVD (serial 3). The MLE method was also found to be the best parameter fitting method for the models. However, the results of the GMLE for the GEV and Gumbel

were comparable with those from MLE and the MLE were found to be the better method by the Quantile-Quantile plots (figures not shown). In addition, the GEV distribution when fitted to the AMS with Bayesian parameter fitting also had acceptable results, according to the DIC, which was the highest. However, the Gumbel distribution was not acceptable for modeling the AMS when compared with the GEV according to the likelihood ratio (LR) test at $\alpha = 0.05$ level ($p < 0.05$). This is also evident from the negative shape factor (Table 6), which suggests that the AMS series of the lake levels have an upper bound, which is not typical of a Gumbel distribution (Gillend and Katz 2016).

Figure 9a, b, and c shows the diagnostic plots of the remaining three candidate models: AMS (GEV, MLE), POT (GPA, MLE), and the AMS (GEV, Bayesian). A visual assessment of the QQ (Fig. 8, upper left panels), density (lower left panels), and return level plots (lower right panels) show a better performance from AMS

Table 9 Estimated fitted models of the location parameter ($\hat{\mu}$) with covariate (*C*)

Model	$\hat{\mu} = \mu_0 + \mu_1 * C$	$\hat{\sigma}(se)$	$\xi(se)$
AMS ₁	474.57 + 1.33 * <i>P</i>	0.73 (0.09)	- 0.42 (0.13)
AMS ₂	474.68 + 1.38 * Disch	0.67 (0.08)	- 0.26 (0.12)
<i>AMS₃</i>	<i>475.90 - 1.42 * Evap</i>	<i>0.59 (0.07)</i>	<i>- 0.17 (0.13)</i>
AMS ₄	475.68 - 0.88 * IOD	0.73 (0.08)	- 0.32 (0.08)
AMS ₅	475.59 - 0.62 * SOI	0.71 (0.08)	- 0.26 (0.11)
AMS ₆	475.71 - 1.31 * Evap + 0.26 * <i>P</i>	0.60 (0.07)	- 0.20 (0.14)
AMS ₇	475.39 - 1.20 * Evap + 0.93 * Disch	0.60 (0.07)	- 0.14 (0.14)
AMS ₈	474.55 + 0.78 * <i>P</i> + 0.72 * Disch	0.70 (0.10)	- 0.34 (0.16)
AMS ₉	475.56 - 1.31 * Evap - 0.52 * <i>P</i> + 1.32 * Disch	0.54 (0.07)	- 0.09 (0.14)

(GEV; MLE and Bayesian) as compared with POT (GPA, MLE). Furthermore, the density and the fitted against empirical quantile plots suggest a better fit by AMS (GEV, MLE) than AMS (GEV, Bayesian). In addition, the normal AMS (GEV, MLE) model was more acceptable than a similar type with a log link function to the scale parameter ($\sigma(\varnothing = \log \sigma)$), according to the LR test where $p > 0.05$. From the foregoing, the AMS (GEV, MLE) was selected for the analysis of floods both under stationary and non-stationary conditions. The modeled T-year lake levels of the AMS (GEV, MLE) with parameters ($\mu = 475.28 \pm 0.24$, $\sigma = 0.74 \pm 0.16$, and $\xi = -0.29 \pm 0.2$) are shown in Table 7. The results in Table 7 suggest that more frequent T = 2-year lake levels of 475.41 ± 0.21 m ASVD, and the lake levels may reach 477.13 ± 0.5 m ASVD once in every T = 100 years. The T = 100 year level is comparable with the lake level peak of May 1980, suggesting a 100-year cycle in the lake levels of this magnitude. However, the extremely low lake levels between 1915 and 1935 (about 469 m ASVD) were not captured by the models, suggesting that they are a very rare event. The stationary model with the GEV is henceforth referred to as AMS₀.

Non-Stationary Frequency Analysis

This section presents the results of NFA of AMS series of Lake Malawi levels. The GEV (MLE) was fitted to the AMS series with consideration of each of the covariates (rainfall, river discharge, evaporation, IOD, and SOI) either in μ , σ , or both. In addition, various combinations of the covariates μ were investigated.

Single covariate modeling of the location parameter

Table 8 shows the results of the non-stationary models that were developed. Models AMS₁ to AMS₅ involved constraining the scale parameter (σ) in the stationary model (AMS₀) as each of these is nested within the stationary model (AMS₀), such that setting the location parameter $\mu_1 = 0$ should yield the AMS₀ model (Prosdocimi et al. 2015; Su and Chen 2019). Based on the values of the AIC, BIC, and NLL, which should be a minimum, it can be seen in Table 8 that the model (AMS₃) with evaporation as a covariate in the location parameter has the best performance, followed by the rainfall-based (AMS₁) and the inflowing river mean discharge-based model (AMS₂). According to the LR tests with the nested model AMS₀, the models with

evaporation (AMS₃), rainfall (AMS₁), or discharge (AMS₂) as a covariate in the location parameter were also all acceptable with better performances as compared with the stationary model case (AMS₀) without covariates ($p < 0.05$). This indicates that adding each of the covariates (rainfall, discharge, and evaporation) improves the performance of the AMS₀ in the frequency analysis of AMS lake levels. On the other hand, the results show that the covariates IOD (AMS₄) and SOI (AMS₅) were not immediate key factors in the frequency analysis of AMS series of the lake levels, as both had higher values of AIC, BIC, and NLL. In addition, their LR tests with AMS₀ resulted in the acceptance of the null hypothesis of no significant influence of these covariates in modeling the location parameter (μ) with all $p > 0.05$. Since there is no specified cutoff point in the AIC, DIC, and DIC values as criteria for the evaluation of models, taking the lowest p values of the LR test is an indicator of better model performance. The LR test between each of the three candidate models (AMS₁, AMS₂, and AMS₃) with the stationary model (AMS₀) was instead used at $\alpha = 0.05$, since all of them are nested within this model if $\mu_1 = 0$. The LR test results in Table 8 show that the evaporation-based model ($p = 0.0003$) still had the better performance, as compared with the rainfall- ($p = 0.0097$) and discharge ($p = 0.0125$)-based models.

The estimated models for the location parameter μ with single and combined covariates, the estimated constrained scale (σ) and shape parameters (ξ), and their standard errors (se) are shown in Table 9. Physically, models AMS₁ and AMS₂ show the additive effects of rainfall and discharge to the location parameter, whereas AMS₃ to AMS₅ suggests the respective subtractive effects of evaporation, IOD, and SOI on the location parameter. This means that higher rainfall and river inflows may lead to increases in lake levels, whereas higher evaporation and negative phases of IOD and SOI results in lower lake levels.

Joint covariate modeling of the location parameter

Having eliminated the IOD and SOI as important covariates in the location parameter for modeling the lake level AMS, the study then investigated the combined influences of evaporation, rainfall, and discharge on the location parameter (AMS₆ to AMS₉ in Table 8). The results have shown that AMS₃ (evaporation-based model) was still the best model, and the other two

Table 10 Test results for models with covariates in the scale parameter ($\hat{\sigma}$)

Model#	Covariate	AIC	BIC	NLL	LR (p value)	Nested Model
AMS ₁₀	Rainfall	102.25	109.39	47.12	0.0689	AMS ₀
AMS ₁₁	Discharge	105.40	112.54	48.70	0.6930	AMS ₀
AMS ₁₂	Evaporation	100.71	107.85	46.36	0.0277	AMS ₀
AMS ₁₃	IOD	104.93	112.07	48.47	0.4298	AMS ₀
AMS ₁₄	SOI	101.78	108.92	46.89	0.0520	AMS ₀

covariates (rainfall and discharge) may therefore add some value to the location parameter of this model. Consequently, AMS₃ was considered to be nested in AMS₆ and AMS₇. In addition, AMS₁ was considered to be nested in AMS₈, since the based model was the second best. The results in Table 8 show that AMS₇ (combining evaporation and discharge) was the best performer of the three, with the lowest AIC, BIC, and NLL values. However, none of these demonstrated through the LR test that the addition of a covariate would improve the performance of the nested model as all $p > 0$. Finally, AMS₉ (combing all three covariates) fails short of being the best model according to the AIC, BIC, and NLL values shown in Table 10. The LR test was therefore used with AMS₉ and against AMS₆, AMS₇, and AMS₈. The best (lowest) p value of 0.0018 for LR test was only found when AMS₈ was nested within AMS₉. The implication is that only the addition of evaporation would be of some value in the performance of each of the combined models. The estimated models for the location parameter with combined covariates, the estimated constrained scale (σ) and shape parameters (ξ), and their standard errors (se) are also shown in Table 9. Physically, models AMS₁ and AMS₂ show the additive effects of rainfall and discharge to the location parameter, whereas AMS₃ to AMS₅ suggests the respective subtractive effects of evaporation, IOD,

and SOI on the location parameter. In addition, the study also found that there were no significant improvements in the performance of each of these 9 NSFA models when log link functions were used in the scale parameter ($\emptyset = \log(\sigma)$).

Single covariate models of the scale parameter

This section presents results on the effect of adding a covariate in modeling the scale parameter (σ) only while constraining the location parameter (μ). The stationary model (AMS₀) served as the nested model for the LR tests in the evaluation of all five models (AMS₁₀ to AMS₁₄). The results of the test results are shown in Table 10 whereas the actual model parameters are shown in Table 11. Overall, the results suggest that varying σ while constraining the location did not result in significant improvements in the model performance of AMS₀ according to the AIC, BIC, and NLL values (Table 10). It can be noted that all of these values are relatively higher than those in models for a varying location parameter. However, the model with evaporation (AMS₁₂) as a covariate in σ still outperformed the other four models (AMS₁, AMS₂, AMS₄, and AMS₅), with lowest values of AIC (100.71), BIC (107.85), and NLL (46.36), and lowest value of $p = 0.028$ for the LR test. It can also be noted that AMS₈ was the only model among the five that rejected the null hypothesis of no significant influence in the scale parameter as a covariate according to the LR test.

Table 11 Estimated fitted models of the scale parameter ($\hat{\sigma}$) with covariate (C)

Model#	$\hat{\mu}$ (se)	$\hat{\sigma} = \sigma_0 + \sigma_1 * C$	ξ (se)
AMS ₁₀	475.21 (0.13)	0.35 + 0.87 * P	- 0.49 (0.14)
AMS ₁₁	475.30 (0.13)	0.83 - 0.25 * Disch	- 0.25 (0.14)
AMS ₁₂	475.48 (0.10)	0.23 + 1.05 * Evap	- 0.08 (0.13)
AMS ₁₃	475.25 (0.13)	0.85 - 0.22 * IOD	- 0.31 (0.09)
AMS ₁₄	475.32 (0.10)	0.30 + 0.78 * SOI	- 0.20 (0.09)

Covariate modeling of both μ and σ

For varying both $\hat{\mu}$ and $\hat{\sigma}$, only the three covariates (rainfall, discharge, and evaporation) whose influence was found in the location parameter were considered. The stationary model (AMS₀) and each of the models incorporating variations in μ only served as the nested models for the respective covariates. The results

Table 12 Test results for models with covariates in both $\hat{\mu}$ and $\hat{\sigma}$

Model	Covariate(s)	AIC	BIC	NLL	Nested model	LR (<i>p</i> value)	Nested model	LR (<i>p</i> value)
AMS ₁₅	Rainfall, Rainfall	99.25	108.17	44.63	AMS ₀	0.016	AMS ₁	0.203
AMS ₁₆	Discharge, Discharge	100.35	109.27	45.18	AMS ₀	0.027	AMS ₂	0.3272
AMS ₁₇	Evap, Evap	92.76	101.68	41.38	AMS ₀	0.0006	AMS ₃	0.1915
AMS ₁₈	IOD, IOD	103.04	111.96	46.52	AMS ₀	0.104	AMS ₄	0.238
AMS ₁₉	SOI, SOI	102.81	111.73	46.41	AMS ₀	0.0933	AMS ₅	0.0831

(Tables 12 and 13) show that AMS₁₇, with evaporation in $\hat{\mu}$ and $\hat{\sigma}$, has the best performance, as it has lowest values of AIC, BIC, and NLL. The LR test however suggests that having a covariate in both $\hat{\mu}$ and $\hat{\sigma}$ does significantly improve the performance of the AMS₀ for the rainfall, evaporation, and discharge-based models (AMS₁₅, AMS₁₆, AMS₁₇, respectively) with $p < 0.05$. AMS₁₇ is also the best model among the three as $p = 0.0006$ is the lowest of the LR ratio results. However, this LR ratio is larger than the LR ratio results of AMS₃, the model with evaporation as a covariate in $\hat{\mu}$. This aspect is validated by the LR test results in Table 12, where all 5 models with covariates $\hat{\mu}$ accepted the null hypothesis of no significant improvements in their performance with covariates in $\hat{\sigma}$ ($p > 0.05$).

Discussion

Table 14 summarizes the ranks of the various models for $\hat{\mu}$ and $\hat{\sigma}$ based on the AIC, BIC, NLL, and LR test results. From the table, it is apparent that evaporation-based models dominate. AMS₁₇, AMS₉, and AMS₁₂ can be removed from the list, as they are all evaporation-based and did not significantly improve the performance of AMS₃, according to the LR tests. From the foregoing, it is apparent that the lake open water evaporative demand has the most significant influence on the AMS series, followed by rainfall over the

lake and inflowing discharge. The lake evaporative demand is implicitly a function of the temperature regime in an unlimited moisture supply environment like Lake Malawi. Many studies have reported on increasing temperatures in Malawi (Warnatzsch and Reay 2015 and references within, Ngongondo et al. 2015) and over Southern Africa (Engelbrecht et al. 2015, Ngongondo et al. 2015, Maure et al. 2018). In the “Temporal hydro-climatic pattern of Lake Malawi basin (1899 to 2017)” section, this study also reported on statistically significant increases in the temperature and the evaporation regimes of Lake Malawi. In their modeling studies, Crossely et al. (1990) established that reduced rainfall by 30% in the Lake Malawi basin resulted in increased evaporation over the lake by up to 4%, with mean inflowing river discharge decreasing by up to 48%. Awange et al. (2008) also reported a significant influence of rainfall in the recession of Lake Victoria during (1976–1999). However, that study only considered rainfall and discharge from some inflowing rivers.

Figure 10 shows the lake level duration curves for estimated 2 years for the three acceptable non-stationary models (AMS₁, AMS₂, and AMS₃). Based on the Weibull plotting position (Vogel and Fennessey, 1994 & 1995), these are compared with the corresponding stationary model (AMS₀) ($p = 0.5$ non-exceedance or the median and 100-year quantiles $p = 0.99$ non-exceedance) and the actual AMS series. In the figures, high

Table 13 Estimated fitted models for $\hat{\mu}$ and $\hat{\sigma}$ with covariate (C)

Model	$\hat{\mu} = \mu_0 + \mu_1 * C$	$\hat{\sigma} = \sigma_0 + \sigma_1 * C$	$\hat{\xi}$ (se)
AMS ₁₅	474.76 + 1.03 * P	0.48 + 0.47 * P	- 0.47 (0.13)
AMS ₁₆	474.72 + 1.21 * Disch	0.91 - 0.62 * Disch	- 0.15 (0.16)
AMS ₁₇	475.78 - 1.20 * Evap	0.35 + 0.48 * Evap	- 0.73 (- 0.07)
AMS ₁₈	475.68 - 0.88 * IOD	0.85 - 0.31 * IOD	- 0.30 (0.09)
AMS ₁₉	475.46 - 0.39 * SOI	0.29 + 0.78 * SOI	- 0.18 (0.10)

Table 14 Rankings of the models for $\hat{\mu}$ and $\hat{\sigma}$

Model#	Covariate(s)	AIC	BIC	NLL	LR (<i>p</i> value)	Nested model
AMS ₃	Evaporation*	92.47	99.61	42.23	0.0003	AMS ₀
AMS ₁₇	Evap, Evap	92.76	101.68	41.38	0.0006	AMS ₀
AMS ₉	Evap+Rainfall+Discharge	92.52	103.22	40.26	0.0018	AMS ₈
AMS ₁	Rainfall	98.87	106.01	45.43	0.0097	AMS ₀
AMS ₂	Discharge	99.31	106.45	45.66	0.0125	AMS ₀
AMS ₁₂	Evaporation	100.71	107.85	46.36	0.0277	AMS ₀

lake levels have low exceedance probabilities in a particular year whereas low lake levels have high exceedance probabilities. Under stationary conditions, the T = 2-year and 100-year lake levels are estimated at 475.41 ± 0.21 m ASVD and 477.13 ± 0.50 m ASVD, respectively, as shown by the horizontal dashed lines in Fig. 10a and b. These have constant values with fixed exceedance probabilities (*p* = 0.5, *T* = 2 – Year, and *p* = 0.01, *T* = 100 – Year). Under non-stationary conditions, considerable magnitude changes can be seen in the variations of both modeled T = 2-year and 100-year lake levels. All three models agree in quantile changes from the rare high lake levels, to the more common lower lake levels for both return periods, with magnitude changes from a maximum of 476.3 m ASVD (AMS₂) to a minimum of 474.7 m ASVD (AMS₃). The minimum is approximately equal to the annual mean lake level.

The results in Fig. 10a further show that the non-stationary T = 2-year lake level exceeds the stationary 2-year lake level approximately 60% of the time for all models. Subsequently, the 2-year non-stationary lake levels will be equal or lower than the stationary 2-year lake level for the remaining 40% of the time. In the high lake level section of Fig. 10a, which are the rare events with low exceedance probabilities in a year, it can be seen that all three non-stationary models are in approximate agreement in terms of the modeled magnitudes. However, the low lake level section of the curves shows that AMS₃, with evaporation in μ (Fig. 10a), has the most change in the 2-year lake level quantiles, while AMS₁ and AMS₂ are in approximate agreement. In addition, the results suggest that in any year, the T = 2-year lake level has changed in magnitude and is becoming equal to the more frequently observed lake level with magnitudes similar to the normal AMS.

For the T = 100 year lake levels in Fig. 10b, considerable variations in the magnitudes of the lake levels can

also be seen in all three models. These range from a maximum of 477.9 m ASVD (AMS₂) to a minimum of 476.1 m ASVD (AMS₁). In addition, AMS₁ has quantiles above the stationary 100-year levels for approximately 18% of the time, whereas AMS₂ and AMS₃ account for 50% and 60% of the time, respectively. It can also be noted that AMS₃ quantiles exhibit the most changes for the T = 100-year lake levels.

Conclusion

Variations in the levels of Lake Malawi are already having considerable impacts to the livelihoods of the riparian communities and the socio-economic development of Malawi. Since the peak of 1979/1980, a downward pattern of the levels is evident from the empirical record. Consequently, the lake levels cannot be assumed to be stationary. However, there is no documented study in the literature that has attempted to account for this trend. In this study, new FFA approaches incorporating non-stationary have therefore been applied to the lake levels to understand and compared with those from traditional stationary methods. A multi-model framework ranging from the following: change point and trends analysis, series (AMS or PDS); pdfs (Gumbel, GEV, PE3, GPA, and exponential); parameter fitting (MLE; GMLE, Bayesian, and L-Moments); and covariates (rainfall, discharge, evaporation, IOD, and SOI) was used to identify the best approach for modeling the frequency of the lake levels. Based on commonly used model evaluation criteria (AIC, BIC, DIC, and LR tests), the model with the GEV when fitted to the AMS using the MLE was the best performing among the candidate distributions as opposed to the GPA on POT series. In addition, open water evaporation was found to be the most dominant covariate to the location

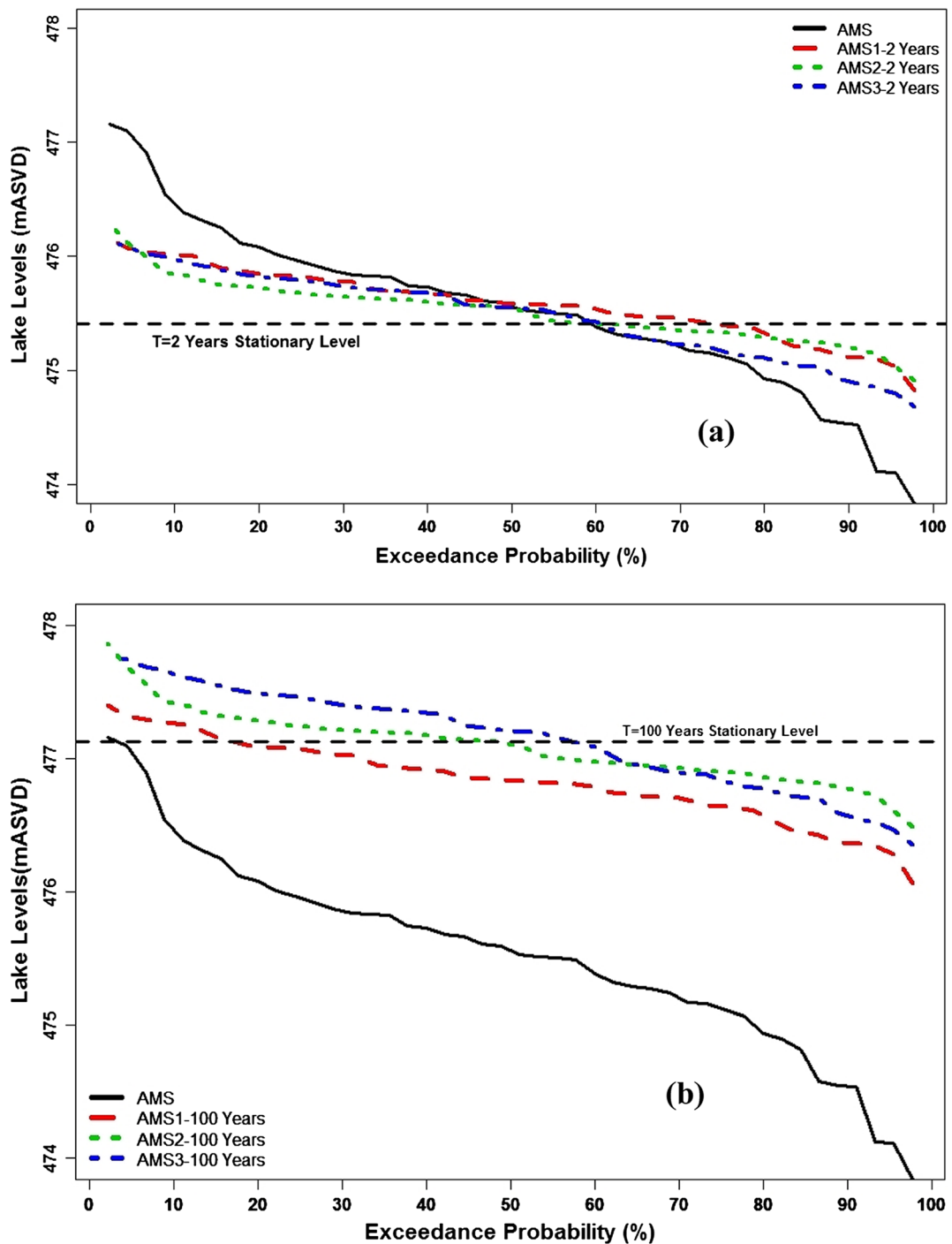


Fig. 10 Lake level duration curves of AMS series, AMS0, and the best non-stationary models. **a** With $T = 2$ years return period and **b** $T = 100$ years return period

parameter (μ) of the non-stationary GEV distribution, followed by rainfall over the lake and river discharge. All covariates were found to have no significant influence on the scale parameter (σ). However, the IOD and

SOI, large-scale low climate variability indices in the area, had non-significant impact on the variation of the lake levels. The results suggest that non-stationary models are more ideal in the evolution of lake quantiles

and should be incorporated in infrastructural design and flood zone planning among others.

Acknowledgments The Malawi Ministry of Agriculture, Irrigation and Water Development, Water Resources Department, and the Climate Justice Project provide the lake levels and river discharge data. The temperature and pan evaporation data were kindly sourced from the Malawi Department of Climate Change and Meteorological Services for.

Funding information Open Access funding provided by University of Oslo (incl Oslo University Hospital). We acknowledge the financial support from the Research Council of Norway through the project “Advancing Frequency Analysis of Nonstationary hydrological extremes for reducing flood risk in a changing climate” (Forskerprosjekt–FRINATEK) at the University of Oslo, Norway.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ajami, H., Sharma, A., Band, L., Evans, J., Tuteja, N., Amirthanathan, G., & Bari, M. (2017). On the non-stationarity of hydrological response in anthropogenically unaffected catchments: an Australian perspective'. *Hydrology and Earth System Sciences*, *21*, 281–294. <https://doi.org/10.5194/hess-2016-353>.
- Awange, J. L., Ogalo, B., J.-H., Were, P., Omondi, P., Omutte, P., & Omullo, M. (2008). Falling Lake Victoria water levels: is climate a contributing factor? *Climatic Change*, *89*, 281–297. <https://doi.org/10.1007/s10584-008-9409-x>.
- Bayazit, M. (2015). Nonstationarity of hydrological records and recent trends in trend analysis: a state-of-the-art review. *Environmental Processes*, *2*, 527–542.
- Beguieria, S., & Vicente-Serrano, S.M. (2017). Package "SPEI". Calculation of the standardised precipitation-evapotranspiration index. <https://cran.r-project.org/web/packages/SPEI/SPEI.pdf>.
- Bootsma, H. A., & Hecky, R. E. (1993). Conservation of the African Great Lakes: a limnological perspective. *Conserv. Biol*, *7*, 644–656.
- Botero, B. A., & Frances, F. (2010). Estimation of high return period flood quantiles using additional non-systematic information with upper bounded statistical models. *Hydrol. Earth Syst. Sci.*, *14*, 2617–2628.
- Calder, I. R., Hall, R. L., Bastable, H. G., Gunston, H. M., Shela, O., Chirwa, A., & Kafundu, R. (1994). The impacts of land use change on water resources in sub-Saharan Africa: a modelling study of Lake Malawi. *J. Hydrol*, *170*, 123–135.
- Coles, S. (2001). *An introduction to statistical modelling of extreme values*. London: Springer. <https://doi.org/10.1007/978-1-4471-3675-0>.
- Core Team, R. (2013). R: A language and environment for statistical computing. In *R Foundation for statistical computing*. Vienna: Austria. URL <http://www.R-project.org/>.
- Cui, B.-L., & Li, X.-Y. (2016). The impact of climate changes on water level of Qinghai Lake in China over the past 50 years. *Hydrology Research*, *47*(2), 532–542.
- De Luca, D. L., & Galasso, L. (2018). Stationary and non-stationary frameworks for extreme rainfall time series in southern Italy. *Water*, *10*, 147. <https://doi.org/10.3390/w10101477>.
- De Paola, F., Giugni, M., Pugliese, F., Annis, A., & Nardi, F. (2018). GEV parameter estimation and stationary vs. non-stationary analysis of extreme rainfall in African test cities. *Hydrology*, *5*, 28. <https://doi.org/10.3390/hydrology5020028>.
- Debele, S. E., Bogdanowicz, E., & Strupczewski, W. G. (2017). Around and about an application of the GAMLSS package to non-stationary flood frequency analysis. *Acta Geophys*, *65*, 885–892. <https://doi.org/10.1007/s11600-017-0072-3>.
- Drayton, R. S. (1984). Variations in the level of Lake Malawi. *Hydrological Sciences Journal*, *2*(1), 1–12.
- Dubois, N., Saulnier-Talbot, E., Mills, K., Gell, P., Battarbee, R., Bennion, H., Chawchai, S., Dong, X., Francus, P., Flower, R., Gomes, D. F., Gregory-Eaves, L., Humane, S., Kattel, G., Jenny, J.-P., Langdon, P., Massaferro, J., McGowan, S., Mikomägi, A., Thi Minh Ngoc, N., Sandaruwan Ratnayake, A., Reid, M., Rose, N., Saros, J., Schillereff, D., Tolotti, M., & Valero-Garcés, B. (2018). First human impacts and responses of aquatic systems: a review of palaeolimnological records from around the world. *The Anthropocene Review*, *(1)*, 28–68.
- Engelbrecht, F., Adegoke, J., Bopape, M.-J., Naidoo, M., Garland, R., Thatcher, M., McGregor, J., Katzfey, J., Werner, M., Ichoku, C., & Gatebe, C. (2015). *Environ. Res. Lett.*, *10*, 085004. <https://doi.org/10.1088/1748-9326/10/8/085004>.
- Fathian, F., Modarres, R., & Dehghan, Z. (2016). Urmia Lake water-level change detection and modeling. *Model. Earth Syst. Environ*, *2*, 203.
- François, B., Schlef, K. E., Wi, S., & Brown, C. M. (2019). Design considerations for riverine floods in a changing climate – a review. *Journal of Hydrology*, *574*, 557–573.
- Garousi, V., Najafi, A., Samadi, A., Rasouli, K., & Khanaliloo, B. (2013). Environmental crisis in Lake Urmia, Iran: a systematic review of causes, negative consequences and possible solutions. In *Proceedings of the 6th International Perspective on Water Resources & the Environment (IPWE)*. Izmir: Environmental and water resources institute of ASCE.
- Ghosh, S., & Resnick, S. (2010). A discussion on mean excess plots. *Stoch Process Appl*, *120*, 1492–1517.
- Gilleland, E. (2018). *Extreme value Analysis, exRemes R Package*.
- Gillend, E., & Katz, R. W. (2016). Extremes: an extreme value analysis package in R. *J of Statistical Software*, *72*(8), 1–39.
- Gownaris, N. J., Routntos, K. J., Kaufman, L., Kolding, J., Lwiza, K. M. M., & Pikitch, E. K. (2018). Water level fluctuations

- and the ecosystem functioning. *Journal of Great Lakes Research*, 44, 1154–1163.
- Han, Q., Zhang, S., Huang, G., & Zhang, R. (2016). Analysis of long-term water level variation in Dongting Lake, China. *Water*, 8(7), 306. <https://doi.org/10.3390/w8070306>.
- Hisdal, H., Stahl, K., Tallaksen, L. M., & Demuth, S. (2001). Have streamflow droughts in Europe become more severe or frequent? *Int. J. Climatol*, 21, 317–333. <https://doi.org/10.1002/joc.619,2001a>.
- Hounkpè, J., Diekkrüger, B., Badou, D. F., & Afouda, A. A. (2014). Non-stationary flood frequency analysis in the Ouémé River Basin, Benin Republic. *Hydrology*, 2(4), 210–222. <https://doi.org/10.3390/hydrology2040210>.
- Iacobellis, V., Gioial, A., Manfreda, S., & Fiorentino, M. (2011). Flood quantiles estimation based on theoretically derived distributions: regional analysis in Southern Italy. *Nat. Hazards Earth Syst. Sci*, 11, 673–695.
- Jaiswal, R. K., Lohani, A. K., & Tiwari, H. L. (2015). Statistical analysis for change detection and trend assessment in climatological parameters. *Environmental Processes*, 2, 729–749. <https://doi.org/10.1007/s40710-015-0105-3>.
- Jury, M. J. (2014). Malawi's Shire river fluctuations and climate. *Journal of Hydrometeorology*, 15, 2039–2049. <https://doi.org/10.1175/JHM-D-13-0195.1>.
- Kafumbata, D., Jamu, D., & Chiotha, S. (2014). Riparian ecosystem resilience and livelihood strategies under test: lessons from Lake Chilwa in Malawi and other lakes in Africa. *Philosophical Transaction of Royal Society B, Biological Sciences*, 69, 20130052. <https://doi.org/10.1098/rstb.2013.0052>.
- Karlsson, I. B., Sonnenborg, T. O., Jensen, K. H., & Refsgaard, J. C. (2014). Historical trends in precipitation and stream discharge at the Skjern River catchment, Denmark. *Hydrol. Earth Syst. Sci*, 18, 595–610.
- Kendall, M. G. (1975). *Rank correlation methods* (4th ed.). London: Charles Griffin.
- Kundzewicz, Z. W., & Robson, A. (2004). Change detection in hydrological records—A review of the methodology. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, 49(1), 7–19. <https://doi.org/10.1623/hysj.49.1.7.53993>.
- Kumambala, P. N., & Ervine, A. (2010a). Water balance model of Lake Malawi and its sensitivity to climate change. *Open Hydrology J.*, 4, 152–162.
- Kumambala, P., & Ervine, A. (2010b). Water balance model of lake malawi and its sensitivity to climate change. *The Open Hydrology Journal*, 4(1), 152–162. <https://doi.org/10.2174/1874378101004010152>.
- Li, L., Ngongondo, C., Gong, L., & Xu, C.-Y. (2013). Comparison of the global TRMM and WFD precipitation datasets on large-scale hydrological modelling in southern Africa. *Hydrology Research*, 45(5), 770–788. <https://doi.org/10.2166/nh.2012.175>.
- Linacre, E. T. (1994). Estimating U.S. Class A Pan evaporation from few climate data. *Water International*, 19(1), 5–14. <https://doi.org/10.1080/02508069408686189>.
- Lopez, J., & Frances, F. (2013). Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates, *Hydrol. Earth Syst. Sci.*, 17, 3189–3203. <https://doi.org/10.5194/hess-17-3189-2013>.
- Lyons, R. P., Kroll, C. N., & Scholz, C. A. (2011). An energy-balance hydrologic model for the Lake Malawi Rift Basin, East Africa. *Global and Planetary Change*, 75, 83–97.
- Machado, M. J., Botero, B. A., López, J., Francés, F., Díez-Herrero, A., & Benito, G. (2015). Flood frequency analysis of historical flood data under stationary and non-stationary modelling. *Hydrol. Earth Syst. Sci.*, 19, 2561–2576.
- Madsen, H., Rasmussen, P. F., & Rosbjerg, D. (1997). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 1: At-site modeling. *Water Research*, 33(4), 737–757.
- Madsen, H., Lawrence, D., Lang, M., Martinkova, M., & Kjeldsen, T. R. (2014). Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. *J. Hydrol.*, 519, 3634–3650. <https://doi.org/10.1016/j.jhydrol.2014.11.003>.
- Makwinja, R., Phiri, T., Kosamu, I. B. M., & Kaonga, C. C. (2017). Application of stochastic models in predicting Lake Malawi water levels. *International Journal of Water Resources and Environmental Engineering*, 9(9), 191–2000. <https://doi.org/10.5897/IJWREE2017.0740>.
- Mann, H. B. (1945). Non-parametric test against trend. *Econometrica*, 13, 245–259.
- Marchant, R., Mumbi, C., Behera, S., & Yamagata, T. (2006). The indian ocean dipole – The unsung driver of climatic variability in East Africa. *African Journal of Ecology*, 45, 4–16. <https://doi.org/10.1111/j.1365-2028.2006.00707.x>.
- Martins, E. S., & Stedinger, J. R. (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, 36, 737–744.
- Maure, G., Pinto, I., Ndebele-Murisa, M., Muthige, M., Lennard, C., Nikulin, G., Dosio, A., & Meque. (2018). The southern African climate under 1.5 °C and 2 °C of global warming as simulated by CORDEX regional climate models. *Environ. Res. Lett.*, 13, 065002. <https://doi.org/10.1088/1748-9326/aa190>.
- McSweeney, C., New, M. & Lizcano G., 2010. *UNDP climate change country profiles*, Malawi, The UNDP Climate Change Country Profiles.
- Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, F. M., Kundzewicz, Z. W., Lettenmaier, D. R., & Stouffer, R. J. (2008). Stationarity is dead: whither water management? *Science*, 319, 573–574.
- Neuland, H. (1984). Abnormal high water levels of Lake Malawi?—An attempt to assess the future behaviour of the lake water levels. *Geojournal*, 9(4), 323–334.
- Ngongondo, C., Xu, C.-Y., Gottschalk, L., & Alemaw, B. (2011). Evaluation of spatial and temporal characteristics in Malawi: a case of data scarce region. *Theoretical and Applied Climatology*, 106, 79–93.
- Ngongondo, C., Xu, C. Y., Tallaksen, L. M., & Alemaw, B. (2015). Observed and simulated changes in the water balance components over Malawi during 1971–2001. *Quaternary International*. <https://doi.org/10.1016/j.quaint.2014.06.028>.
- Nicholson, S. E. (1998). Historical fluctuations of Lake Victoria and other lakes in the Northern Rift Valley of East Africa. In J. T. Lehman (Ed.), *Environmental change and response in East African Lakes. Monographiae Biologicae* (Vol. 79). Dordrecht: Springer.
- Nicholson, S. E. (2001). A semi-quantitative, regional precipitation dataset for studying African climates of the nineteenth

- century, part I. Overview of the data set. *Climatic Change*, 50, 317–353.
- Okonkwo, C., Demoz, B., & Gebremariam, S. (2014). Characteristics of Lake Chad level variability and links to ENSO, precipitation, and river discharge. *Sci. World J.*, 13. <https://doi.org/10.1155/2014/145893>(2014).
- Owen, R.B., Crossley, R., Johnson, T.C., Tweddle, D., Kornfield, I., Davison, S., Eccles, D.H. & Engstrom, D.E. (1990). Major low lake levels of Lake Malawi and their implications for speciation rates in cichlid fishes. Proceedings of the Royal Society, London, Biological Sciences 240:519–553.
- Pettitt, A. N. (1979). A non-parametric approach to the change point problem. *Appl Stat*, 28(2), 126–135.
- Pomposi, C., Funk, C., Shukla, S., Harrison, L., & Magadzire, T. (2018). Distinguishing southern Africa precipitation response by strength of El Niño events and implications for decision-making. *Environ. Res. Lett.*, 13, 074015. <https://doi.org/10.1088/1748-9326/aacc4c>.
- Prosdocimi, I., Kjeldsen, T. R., & Miller, J. D. (2015). Detection and attribution of urbanization effect on flood extremes using nonstationary flood-frequency models. *Water Resources Research*, 4244–4262. <https://doi.org/10.1002/2015WR017065>.
- Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Appl Stat*, 54, 507–554. <https://doi.org/10.1111/j.1467-9876.2005.00510>.
- Rosbjerg, D., & Madsen, H. (2004). Advanced approaches in PDS/Modeling of extremes hydrological events. *Hydrology*, 1, 217–220.
- Ross, R.J. (2015). *Sequential and batch change detection using parametric and nonparametric methods*. R Package CPM.
- Salas, J. D., Obeysekera, J., & Vogel, R. M. (2018). Techniques for assessing water infrastructure for nonstationary extreme events: a review. *Hydrol. Sci. J.*, 63(3), 325–352.
- Sayer, C. A., Máiz-Tomé, L., Akwany, L. O., Kische-Machumu, M. A., Natugonza, V., Whitney, C. W., Omondi, R., Nshutiayesu, S., & Kabuye, C. S. (2018). The importance of freshwater species to livelihoods in the Lake Victoria Basin. In C. Sayer, L. Máiz-Tomé, & W. R. T. Darwall (Eds.), *Freshwater biodiversity in the Lake Victoria Basin*. IUCN Global Species Programme.
- Sene, K. M., Piper, B., Wykeham, D., McSweeney, R. T., Tych, W., & Beven, K. (2017). Long-term variations in the net inflow record for Lake Malawi. *Hydrology Research*, 48.3, 851–866.
- Shela, O. (2000). Naturalisation of Lake Malawi levels and Shire River flows: challenges of water resources research and sustainable utilisation of the Lake Malawi-Shire River System. Sustainable Use of Water Resources Maputo: 1st WARFSA/WaterNet Symposium, 1-12. Available online at <https://floodobservatory.colorado.edu/SiteDisplays/Shire2SHELA.PD>
- Singh, V. P. (1998). *Entropy-based parameter estimation in hydrology* (p. 368). Dordrecht: Springer.
- Song, C., Huang, B., Richards, K., Ke, L., & Hien, V. P. (2014). Accelerated lake expansion on the Tibetan Plateau in the 2000s: induced by glacial melting or other processes? *Water Resour. Res.*, 50, 3170–3186.
- Šraj, M., Viglione, A., Parajka, J., & Blöschl, G. (2016). The influence of non-stationarity in extreme hydrological events on flood frequency estimation. *J. Hydrol. Hydromech.*, 64, 2016, 4, 426–437. <https://doi.org/10.1515/johh-2016-0032426>.
- Stefanidis, K., Kostara, A., & Papastergiadou, E. (2016). Implications of human activities, land use changes and climate variability in Mediterranean lakes of Greece. *Water*, 8, 483. <https://doi.org/10.3390/w8110483>.
- Su, C., & Chen, X. (2019). Assessing the effects of reservoirs on extreme flows using nonstationary flood frequency models with the modified reservoir index as a covariate. *Advances in Water, Resources*, 124, 29–40.
- Úbeda, B., Di Giacomo, A. S., Neiff, J. J., Loisele, S. A., Poi, A. S., Gálvez, J. Á., Casco, S., & Cózar, A. (2013). Potential effects of climate change on the water level, flora and macrofauna of a large neotropical wetland. *PLoS One*, 8(7), e67787. <https://doi.org/10.1371/journal.pone.0067787> Print 2013.
- Vainu, M. & Terasmaa, J. (2014). *Changes in climate, catchment vegetation and hydrogeology as the causes of dramatic lake-level fluctuations in the Kurtne Lake District, NE Estonia*.
- Vasilidiades, L., Galiatsatou, P., & Louka, A. (2015). Nonstationary frequency analysis of annual maximum rainfall using climate covariates. *Water Resour Manage*, 29, 339–358. <https://doi.org/10.1007/s11269-014-0761-5>.
- Vogel, R. M., & Fennessey, N. M. (1994). Flow duration curves I: new interpretation and confidence intervals. *Journal of Water Resources Planning and Management*, 120(4).
- Vogel, R. M., & Fennessey, N. M. (1995). Flow duration curves II: a review of applications in water resources planning. *Water Resources Bulletin*, 31(6).
- Wamatzsch, E. A., & Reay, D. S. (2015). Temperature and precipitation change in Malawi: evaluation of CORDEX-Africa climate simulations for climate change impact assessments and adaptation planning. *Science of The Total Environment*, 65, 378–392.
- Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., O’sterle, H., Adam, J. C., Bellouin, N., Boucher, O., & Best, M. (2011). Creation of the watch forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *J Hydrometeorol*, 12(5), 823–848.
- Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH ForcingData methodology applied to ERA-Interim, reanalysis data. *Water Resources Research*, 50, 7505–7514. <https://doi.org/10.1002/2014WR015638>.
- WMO (World Meteorological Organization) (1988). *Analyzing long time series of hydrological data with respect to climate variability*. WCAP-3, WMO/TD No.224, p 12
- Yang, X. K., & Lu, X. X. (2014). Drastic change in China’s lakes and reservoirs over the past decades. *Scientific Reports*, 4, 6041.
- Ye, Z., Liu, H., Chen, Y., Shu, S., Wu, Q., & Wang, S. (2017). Analysis of water level variation of lakes and reservoirs in Xinjiang, China using ICESat laser altimetry data (2003–2009). *PLoS ONE*, 12(9), e0183800. <https://doi.org/10.1371/journal.pone.0183800>.