Hydrology and Earth System Sciences

A multi-source satellite data approach for modelling Lake Turkana water level: calibration and validation using satellite altimetry data

N. M. Velpuri[1], G. B. Senay[1,2], and K. O. Asante[3]

[1]GISc Center of Excellence, South Dakota State University, Brookings, SD, USA
[2]USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD, USA
[3]Climatus LLC, 800 W El Camino 180, Mountain View, CA, USA

Correspondence to: N. M. Velpuri (velpuri@gmail.com)

Abstract. Lake Turkana is one of the largest desert lakes in the world and is characterized by high degrees of inter- and intra-annual fluctuations. The hydrology and water balance of this lake have not been well understood due to its remote location and unavailability of reliable ground truth datasets. Managing surface water resources is a great challenge in areas where in-situ data are either limited or unavailable. In this study, multi-source satellite-driven data such as satellite-based rainfall estimates, modelled runoff, evapotranspiration, and a digital elevation dataset were used to model Lake Turkana water levels from 1998 to 2009. Due to the unavailability of reliable lake level data, an approach is presented to calibrate and validate the water balance model of Lake Turkana using a composite lake level product of TOPEX/Poseidon, Jason-1, and ENVISAT satellite altimetry data. Model validation results showed that the satellite-driven water balance model can satisfactorily capture the patterns and seasonal variations of the Lake Turkana water level fluctuations with a Pearson's correlation coefficient of 0.90 and a Nash-Sutcliffe Coefficient of Efficiency (NSCE) of 0.80 during the validation period (2004–2009). Model error estimates were within 10 % of the natural variability of the lake. Our analysis indicated that fluctuations in Lake Turkana water levels are mainly driven by lake inflows and over-the-lake evaporation. Over-the-lake rainfall contributes only up to 30 % of lake evaporative demand. During the modelling time period, Lake Turkana showed seasonal variations of 1–2 m. The lake level fluctuated in the range up to 4 m between the years 1998–2009. This study demonstrated the usefulness of satellite altimetry data to calibrate and validate the satellite-driven hydrological model for Lake Turkana without using any in-situ data. Furthermore, for Lake Turkana, we identified and outlined opportunities and

challenges of using a calibrated satellite-driven water balance model for (i) quantitative assessment of the impact of basin developmental activities on lake levels and for (ii) forecasting lake level changes and their impact on fisheries. From this study, we suggest that globally available satellite altimetry data provide a unique opportunity for calibration and validation of hydrologic models in ungauged basins.

1 Introduction

The Intergovernmental Panel on Climate Change (IPCC) Technical Paper on Climate Change and Water stressed the fact that increased demand and reduced availability of fresh water under global climate change will significantly affect agriculture and food security in the 21st century (Bates et al., 2008). Due to increases in population, industrialization, and irrigated agriculture, several surface water resources are rapidly depleting (Vörösmarty et al., 2010). Because of these consequences, it has become increasingly important to accurately identify, quantify, and monitor freshwater resources. In most regions of the world, inland lakes provide important sources of fresh water and influence the local hydrological budget. Furthermore, monitoring changes in lake water levels is essential because they reflect changes in the seasonal distribution of river inflows, precipitation, and evapotranspiration (ET); in some cases integrated over many years (Bates et al., 2008). According to Alsdorf et al. (2007), the measurements required on the variability of surface water are (a) surface water area, A, (b) the elevation of the water surface, h, (c) temporal change, $\partial h/\partial t$, and (d) slope of the water surface, $\partial h/\partial x$. However, such measurements over rivers and lakes/reservoirs are missing in the terrestrial water budget

Table 1. Challenges and opportunities of using satellite data for estimating components of lake water balance in data limited regions.

SL No.	Water balance component	Satellite data/model	Opportunities	Challenges	References
1	Over-the-lake rainfall (Q_{rain})	TRMM, CMORPH, RFE, PERSIANN, NexRAD, or Meteosat, Geostationary weather satellites, etc.	Rainfall data with high spatial and temporal coverage is readily available globally	Raw data under-estimates up to 50 %; calibration required to improve accuracies	Huffman (1995); Joyce et al., (2004); Crow and Ryu et al. (2009)
2	Lake inflows (Q_{in})	Rainfall-runoff models (VIC, SWAT, VegET, USGS GeoSFM, TOPMODEL, etc.)	Established open source models available for direct application and estimation	Accuracy depends on data and model used. Calibration required to improve accuracies	Abbott et al. (1986); Asante et al. (2007); Rostamian et al. (2008); Senay (2008)
3	Over-the-lake evaporation (Q_{evap})	GDAS, GLDAS, NLDAS, MODIS, Energy balance methods (SEBAL, METRIC, SSEB)	Usually accuracies are high; with around 15–30 % relative errors	Accuracy depends on data scale and resolution of the dataset used	Mu et al. (2007); Kalma et al. (2008); Senay et al. (2008)
4	Groundwater storage (GS) – (Q_{gwin} and Q_{gwout})	GRACE (to estimate change in groundwater storage from Total Water Storage (TWS) estimates)	Can be estimated using calibration; recently, GRACE TWS has been used to estimate GS over small lake basins	No direct method to estimate GS using satellite data	Wahr et al. (2004); Becker et al. (2010)
5	Lake outflows (Q_{out})	To some extent (irrigation water use estimation from optical/thermal imagery)	Can be estimated using calibration; RS data used only in lakes where irrigation water use dominates lake outflows	No direct method to estimate using satellite data	Allen et al. (2005); Senay et al. (2007)
6	Lake heights (D_i)	Satellite altimeter; water balance models	Very high accuracies on the order of 3–5 cm	Data available over large rivers, lakes, and reservoirs globally	Birkett (1995); Cretaux and Birkett (2006)

(NASA Science Plan, 2007). Furthermore, while monitoring of surface water variability is a challenging task in ungauged basins, many of the greatest human impacts occur in basins that have no or very limited data (Sivapalan, 2003). Moreover, several forecasts on probability distribution of seasonal rainfall are now becoming available from Regional Climate Outlook Forums (RCOF) (Ogallo et al., 2008). However, process-based models that translate forecasts into variations in lake levels are not yet available. The assessment of lake water balance could provide improved knowledge of regional and global climate change and a quantification of the human impacts on water resources (Cretaux and Birkett, 2006), including capacity to transfer the impacts of climate forecasts on lake levels.

In this study, we use a water balance approach to model lake water levels using multi-source satellite-driven data. Water balance modelling has been widely used in the past for several lake studies (Tate et al., 2006; Kebede et al., 2006; Gibson et al., 2006; Li et al., 2007). Lake levels can be modelled using a water balance approach as

$$D_{(t)} = D_{(t-1)} + Q_{rain} + Q_{in} + Q_{gwin} - Q_{evap}$$
$$- Q_{gwout} - Q_{outflow} - Q_{hw} \qquad (1)$$

where $D_{(t)}$ and $D_{(t-1)}$ are lake depths for current and previous time steps and Q represents the fluxes of the variables

[L] for the current time step; "rain" is direct rainfall over the lake; "in" is incoming runoff contribution into the lake; "gwin" and "gwout" are groundwater contribution to/from the lake; "evap" is over-the-lake ET; "outflow" is surface outflow from the lake; and "hw" is the component of human water withdrawal from the lake. The precision of modelled lake levels using this approach depends on the accuracy of each parameter considered in Eq. (1). Ground truth data are either limited or unavailable in most ungauged basins, but remote sensing satellites offer reliable estimates of hydrologic variables required for water balance modelling at shorter time scales. The opportunities and challenges of using satellite data to derive parameters in Eq. (1) are described in Table 1.

The parameters Q_{rain}, Q_{in}, and Q_{evap} can be modelled or estimated from remotely sensed data; however, it is challenging to estimate the parameters Q_{gwin}, Q_{gwout}, $Q_{outflow}$ and Q_{hw} from satellite data sources. Estimates of rainfall, including over-the-lake rainfall Q_{rain}, can be reliably obtained from satellite-based rainfall. Because of the increased accuracy and availability of satellite-based rainfall products, several studies have recently used satellite-based rainfall estimates for lake level studies (Awange et al., 2008; Swenson and Wahr, 2009; Ricko et al., 2011). However, satellite rainfall estimates often show bias when compared to ground truth observations and require site specific calibration or bias correction for improving model accuracies. Lake inflows or runoff

cannot be directly measured using remote sensing data, but several indirect ways of modelling runoff (Q_{in}) are available. Simple lumped rainfall-based runoff models to complex distributed hydrologic models are available to estimate runoff (Wagner et al., 2004). Satellite-based estimates of evapotranspiration (Q_{evap}) are now becoming increasingly available. It has been found that in arid and semiarid regions, around 90 % or more of the annual precipitation can be evapotranspired; therefore, ET determines the freshwater recharge and discharge from aquifers in these environments (Huxman et al., 2005). Hence, accurate estimation of ET is essential for closing the water balance.

Groundwater fluxes (Q_{gwin} and Q_{gwout}) remain as principle unknowns in most water budget studies. In most regions, information on groundwater does not exist and gauging networks on rivers and lakes have been drastically decreasing since the early 1980s (Kundzewicz et al., 2004). Even though groundwater fluxes into large lakes may typically be smaller than surface water inputs, they may form a significant component of the overall water budget in many lakes (Harvey et al., 2000). Groundwater fluxes are often estimated as a residual of water balance calculations only when other parameters are available to close the water balance. Estimates of groundwater storage have been poorly constrained in most water balance studies using remote sensing data. Recently, the use of GRACE total water storage (TWS) for estimating groundwater storage has been demonstrated by Becker et al. (2010) for small lake basins in Africa. Estimates of lake outflows ($Q_{outflow}$) are mostly unavailable in ungauged basins. While ground truth observations are essential for the quantification of lake outflows, remote sensing offers an indirect estimation of $Q_{outflow}$ with reasonable accuracies. Quantifying human water withdrawal separately from lake outflows is not possible using satellite data. However, when water abstraction for irrigation dominates the natural outflows from the lake, reasonable monthly estimates of outflows can be obtained from remote sensing imagery by quantifying irrigation water use in the downstream irrigated areas (Allen et al., 2005; Senay et al., 2007). Information on lake level heights is required to estimate temporal changes in the lake storage. Recently, lake levels based on satellite altimetry data are becoming globally available for large rivers, lakes, and reservoirs. Most of the data are available from several ungauged basins of the world. Among all satellite data in Table 1, satellite altimetry data are by far the most accurate data available, with errors as low as 3–5 cm. Based on the review of all the parameters, one of the common challenges of using satellite-driven data/models for estimating water balance components is the need for data/model calibration. However, performing calibration is especially difficult in basins where reliable ground truth measurements are unavailable. While ongoing and future research continues to address the challenges of using satellite data and improving the accuracy of satellite estimates, we present an approach that uses satellite altimetry data for calibration and validation of a satellite-driven water

balance model for Lake Turkana, which has no reliable in-situ data.

Justification and objectives of this study

Lake Turkana is one of the largest desert lakes in the world and is characterized by high degrees of inter- and intra-annual fluctuations (Rickett and Johnson, 1996). The hydrology and water balance of the lake has not been well understood due to its remote location and unavailability of in-situ datasets. The most recent study on the hydrology of the lake was carried out by Kallqvist et al. (1988). However, due to the increase in population and agriculture expansion over the last two decades, the Lake Turkana basin has been rapidly undergoing changes with several basin developmental activities taking place in the upstream river basin. The tallest dam in Africa (Gibe-III) is currently under construction on the Omo River, which contributes more than 80 % of the lake inflows. The impact of such changes on the water balance of Lake Turkana is not well understood. Hence, there is an immediate need to understand the relationship between lake levels and upstream processes occurring in the watershed. Since availability of reliable in-situ data is a major problem for the Lake Turkana basin, we present a satellite-driven water balance model to study the impact of upstream basin developmental activities on the Lake Turkana water levels.

The objectives of this study are (a) to demonstrate the use of satellite altimetry data for model calibration and validation when reliable in-situ data for calibration is unavailable, and (b) to establish a calibrated satellite-driven water balance model for Lake Turkana to better understand the interactions between the lake and its watershed. We present a hydrologic modelling approach that integrates digital elevation data, satellite-based rainfall estimates, modelled ET, runoff, and satellite altimetry data to produce information on variations in Lake Turkana levels without relying on in-situ data sources. Potential applications of the calibrated models are also identified.

2   Study area and data used

2.1   Description of the study area

This study is conducted over Lake Turkana, one of the lakes in the Great Rift Valley of East Africa (Fig. 1). The lake is about 250 km long and 15–30 km wide, with an average surface area of nearly 6750 km^2. The lake catchment is 145 500 km^2 and extends over Ethiopia in the north, Kenya in the south, and Sudan and Uganda in the west. The lake has a maximum depth of nearly 110 m and an average depth of 30 m. Three rivers, the Omo, Turkwel, and Kerio, constitute the lake inflows. The Omo River is perennial and meanders nearly 1000 km before emptying into the northern tip of the lake. It accounts for more than 80 % of the lake inflows (Ricketts and Johnson, 1996). In contrast, the Turkwel and
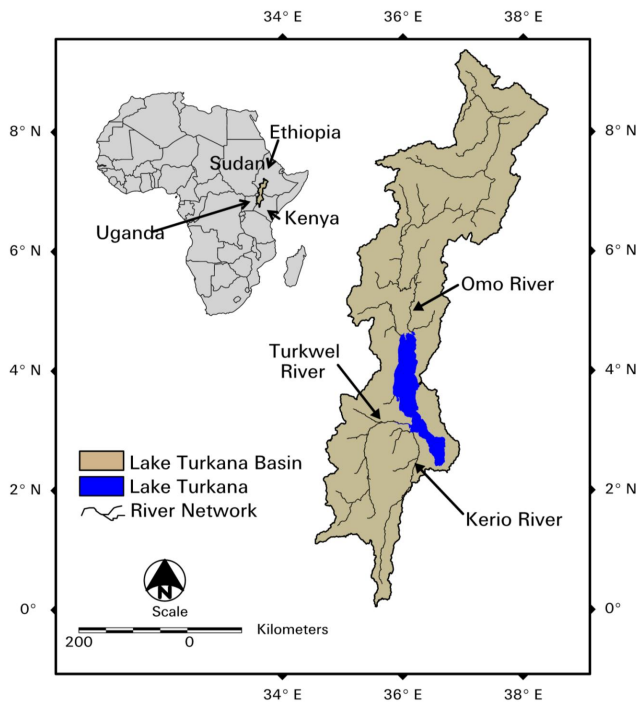
Fig. 1. Map of the study area showing spatial extent of the Lake Turkana basin in East Africa.

Kerio Rivers are intermittent and contribute little to the total volume of the lake (Carr, 1998). Lake Turkana basin has four distinct seasons with two dry periods (December–February and July–August) and two rainy seasons (March–June and September–November). Lake Turkana is considered an endorheic lake with no surface outlet and insignificant seepage (Ricketts and Johnson, 1996). The outflow is dominated only by evaporation.

2.2 Data used

The data used in this study are summarized in Table 2. The National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) produces satellite-based rainfall estimates (RFE) for the Famine Early Warning System (FEWS) project of the U.S. Agency for International Development (USAID). The data have been produced daily with a spatial resolution of 0.1° since June 1995 and are available to the public in near-real time. The product covers the entire African continent and a few surrounding regions. RFE data from June 1995 to 31 December 2000, were produced using the RFE 1.0 algorithm (Herman et al., 1997), and since 1 January 2001, RFE data have been produced using the version 2.0 algorithm (Xie and Arkin, 1996). RFE data from January 1998 to December 2009 are used in this study. The reference evapotranspiration (ET$_o$) data used in this study are produced at the USGS Earth Resources Observation and Science Center from 6-hourly Global Data

Assimilation System (GDAS) climate parameters using the standardized Penman-Monteith equation, then downscaled to 0.1° for this study (Senay et al., 2008). Historical average dekadal (10-day) Normalized Difference Vegetation Index (NDVI) datasets (1982–2006) described by Tucker et al. (2005) from the Advanced Very High Resolution Radiometer (AVHRR) are used to characterize the land surface phenology (LSP) and to estimate actual evapotranspiration (ET$_a$) on a pixel-by-pixel basis at 0.1° resolution. The canopy interception parameter is estimated using the global percent tree cover product produced from MODIS Vegetation Continuous Field (Hansen et al., 2003). The Digital Soil Map of the World (FAO, 1995) is used to estimate water holding capacity (WHC) for the dominant soil type for each grid cell. Shuttle Radar Topography Mission (SRTM) 90-m digital elevation model (DEM) data are obtained from the Consultative Group on International Agricultural Research (CGIAR) Consortium for Spatial Information (CSI) website. These void-filled DEM data are used to derive hydrological derivatives such as (a) streams and river networks and (b) sub-basins and basins. The DEM are also used to estimate lake surface area at various depths. Lake Turkana water level obtained from TOPEX/Posiedon (T/P), Jason-1, 2 and ENVISAT satellite altimetry data is used for calibration and validation of the modelled results. T/P is a joint space mission conducted by the United States and France, primarily designed to measure sea-surface heights since 1992 (Fu et al., 1994). Jason-1 is the T/P follow-on mission and has been measuring ocean surface topography since December 2001. Both T/P and Jason-1 data have also been widely used to study inland lake level variations (Birkett, 1995). Moreover, lake levels derived from satellite altimetry data are highly reliable, with errors on the order of a few centimeters (Birkett, 1995; Alsdorf et al., 2001). Hence, satellite altimetry data are considered as proxy to in-situ lake level measurements and used for model calibration and validation. Satellite-based lake levels have a 10-day temporal resolution (Birkett et al., 1999).

3 Methods

3.1 Deriving lake depth-surface area-volume relationships

Lake Turkana depth-surface area relationship is developed from seamless lake topo-bathymetry (LTB) data. The Lake Turkana bathymetry data obtained from Kallqvist et al. (1988) were draped on the SRTM elevation model to develop seamless LTB data with 90-m resolution. A simple GIS-based model was used to extract surface areas at every 0.5-m interval between the ranges of recent natural fluctuations of the lake (350 to 366 m). Thus, a relationship that explains the variations in Lake Turkana surface area with respect to lake level change was obtained. Similarly, using seamless LTB data, changes in lake volumes were

Table 2. Satellite data and products used in the lake level modelling (LLM) approach.

No	Data	Satellite sensor/ source	Frequency	Resolution/ Scale	References
1	Rainfall estimate for Africa	SSM/I, AMSU	Daily	$0.1° \times 0.1°$	Herman et al. (1997); Xie and Arkin (1996)
2	Global GDAS reference ET	Model assimilated satellite data	Daily	$0.1° \times 0.1°$	Senay et al. (2008)
3	Climatological NDVI	NOAA AVHRR	Dekadal	8 km	Tucker et al. (2005)
4	Digital soil map of the world	National statistics	Single date	1:5 000 000	FAO (1995)
5	Global percent tree cover map	MODIS VCF	Single date	500 m	Hansen et al. (2003)
6	Digital Elevation Model	SRTM	Single date	90 m	Farr and Kobrick (2000)
7	Lake Turkana water levels	TOPEX/Poseidon, Jason-1, ENVISAT	Daily	>200 m	Birkett (1995)

derived. First, Lake Turkana was divided into different water columns. The depth of each column is derived as

$$CD_i = LD - LEB_i \qquad (2)$$

where CD_i is the depth for water column [L] i, LD is the lake depth or lake level [L], and LEB_i is the height [L] obtained from the LTB data. Then, volume of each column of water (CV_i) is obtained as

$$CV_i = CD_i \times CA_i \qquad (3)$$

where CA_i is the area of the column of water $[L^2]$ obtained from the pixel area of the LTB data. Finally, the total volume of the lake (TLV) $[L^3]$ is obtained by the summation of volume of water (CV_i) from the total number of columns (N) as

$$TLV = \sum_i^N CV. \qquad (4)$$

Using Eqs. (2), (3), and (4), lake volumes at every 0.5-m interval between the ranges of natural fluctuations (350 to 366 m) are extracted and the relationship between lake elevation and volume is derived.

3.2 Lake Level Modelling (LLM) approach

Based on the principle of water balance, a multi-sensor physically based hydrologic modelling approach, hereafter called Lake Level Modelling (LLM), is developed to estimate Lake Turkana water levels. The LLM approach (Fig. 2) used in this study can be illustrated in four steps.

3.2.1 Modelling runoff and ET

First, weather data (RFE and GDAS ET_o) are used to estimate runoff [L] on a pixel-by-pixel basis using the phenology-based ET model called VegET (Senay, 2008; Senay et al., 2009). The VegET model is based on standard water balance principles comparable to those outlined in Allen et al. (1998) and Senay and Verdin (2003). The unique

aspect of the VegET model is the use of remotely sensed land surface phenology (LSP) to parameterize the spatial and temporal dynamics of ET and runoff on a grid-cell basis. The modelling approaches in the VegET model can be explained by Eqs. (5) and (6):

$$ET_a = K_{cp} \times K_s \times ET_o \qquad (5)$$

$$SW_{(t)} = SW_{(t-1)} + ((1 - ILC_i) \times RFE_i) - ET_{a_i} \qquad (6)$$

where ET_a is the actual ET; K_{cp} is the LSP-based crop coefficient; K_s is the soil water stress coefficient; ET_o is the global GDAS reference ET; RFE is the satellite-based rainfall estimate; and SW represents soil water content. ILC_i accounts for canopy interception losses interception coefficient, subscript t represents the current modelling time step, and subscript $t - 1$ represents the previous time step. As interception losses depend on vegetation cover and rainfall, an area weighted average interception loss coefficient (ILC) was estimated for each modelling pixel based on the vegetation cover distribution obtained (bare, herbaceous, and tree cover percentage) from the MODIS VCF, which provides the percentage of bare, herbaceous, and tree cover for each pixel (Hansen et al., 2003). ILC varied from a minimum of zero in bare cover types to a high of 35 % in areas with a dense forest cover. The ILC used in this study was found to be within the range of ILC published in the literature (Kelliher et al., 1993; Tate, 1995; Hörmann et al., 1996). The VegET model estimates runoff for each time step based on the principle of soil saturation excess, where soil water content in excess of the WHC of the soil is considered runoff. Variables ET_a, ET_o, RFE, and SW all are in units [L/T]. Further description of the VegET modelling approach is found in Senay (2008) and Senay et al. (2009).

3.2.2 Source-to-sink routing algorithm

Modelled runoff is routed using a source-to-sink routing algorithm (Asante, 2000; Olivera et al., 2000). This method is a simplification of St. Venant equations and incorporates
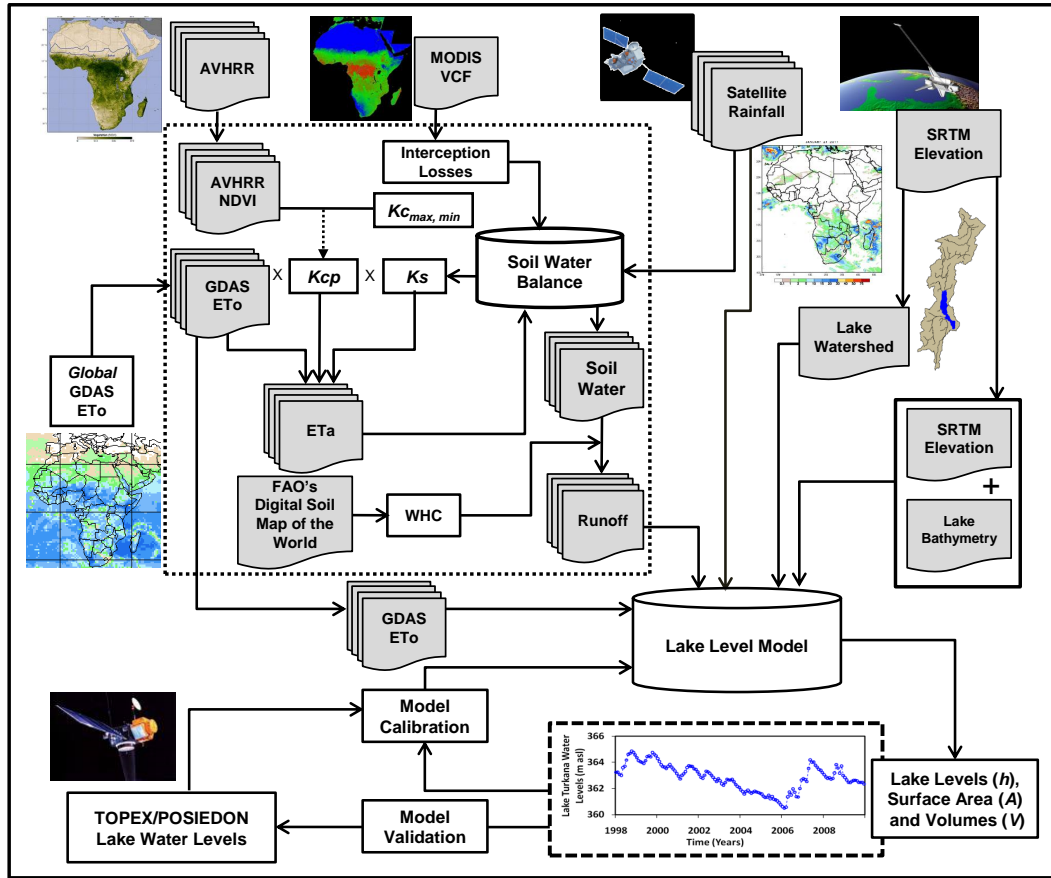
Fig. 2. Flowchart showing the lake level modelling (LLM) approach using multi-source satellite data.

translation (advection) and redistribution (dispersion) processes in the flow path. We chose this method because storage-based routing algorithms are computationally intensive, and Gong et al. (2009) reported that when the reproduction of discharge dynamics at a basin outlet is an important objective, cell-to-cell methods can be replaced by source-to-sink methods. First, the Lake Turkana basin was divided into 39 time-area zones using flow length and flow velocity information such that the value for each zone represents time spent [in days] by the runoff generated in different time-area zones before it reaches the lake. The response function or the first-passage-time distribution for time-area zone j is then estimated based on the diffusion equation model (Lettenmaier and Wood, 1992; Naden et al., 1999; Asante, 2000; Olivera et al., 2000), which can be written mathematically as

$$u_j = \frac{1}{2\,(\Pi\,D)^{0.5}}\,\frac{x}{t^{1.5}}\,\exp\left[-\frac{(c\,t-x)^2}{4\,D\,t}\right] \qquad (7)$$

where u_j [1/T] is the response function of time-area zone j at the lake, x [L] is the mean distance for each time-area zone to the lake, t is the time interval [T], c [L/T] is the celerity or advective velocity of the river, and D [L^2/T] is the diffusion coefficient.

Snell and Sivapalan (1994) showed that the dispersion coefficient depends on the first two moments of the flow path lengths, with the assumption of a constant flow velocity c, and longitudinal dispersion D throughout the catchment. Parameters c and D in Eq. (7) for the Lake Turkana basin were derived as

$$c = 3\,v/2 \quad \text{and} \quad D = R_j\big/2\,S\,B \qquad (8)$$

where v [L/T] is the average velocity of the river (0.3 m s^{-1}) obtained as a function of slope and land cover distribution using the simplified Manning's equation. R_j [L/T] is the time series of modelled runoff generated for time-area zone j at the lake; S [L/L] is the average riverbed slope along the flow path (0.003 m m^{-1}); and B [L] is the average width of the river. High resolution Google Earth imagery was used to measure the width of the river over 100 transects along the Omo, Turkwel, and Kerio Rivers. The width of the rivers ranged from 1 m near the head of the rivers to about 210 m near the Omo river delta, with an average width of 100 m.

By convoluting the time series runoff generated in time-area zone j with the response function u_j, we get the runoff

volume contribution of each time-area zone, RV_j, expressed as

$$RV_j = A_j \, R_j \, \Theta \, u_j \qquad (9)$$

where A_j [L^2] is the area of the time-area zone j, and the Θ symbol stands for the convolution integral.

Finally, the hydrograph for the lake is calculated as the sum of contributions of all the time-area zones that drain into the lake, which is represented as

$$TRV_{(t)} = \sum_{j=1}^{j=n} RV_j \qquad (10)$$

where $TRV_{(t)}$ [$L^3\,T^{-1}$] is the total runoff volume hydrograph from all the time-area zones j for the modelling time step i, RV_j [$L^3\,T^{-1}$] is the contribution of time area zone j, and the sum applies to all 39 (n) sources or time-area zones that drain into the lake.

3.2.3 Simulation of Lake Turkana water levels

A water balance model is applied to derive Lake Turkana water level variations. Total monthly over-the-lake rainfall Q_{rain} was derived from the RFE data, and the total monthly runoff contribution to the lake Q_{runoff} was obtained using Eq. (10). Total monthly over-the-lake ET was obtained from GDAS ET_o data. Becker et al. (2010) analyzed total water storage data from GRACE gravimetry over Lake Turkana basin and indicated that the TWS is mainly influenced by the surface water, and that groundwater contribution in the basin is insignificant. Further, several researchers have concluded that Lake Turkana's groundwater inflows and outflows are considered minimal or insignificant (Yuretich and Cerling, 1983; Cerling, 1986; Avery, 2010), and because Lake Turkana is a saline lake and cannot be used directly for drinking or irrigation, Q_{hw} is considered negligible for Lake Turkana. Hence, for Lake Turkana, Eq. (1) is simplified as

$$D_{(t)} = D_{(t-1)} + Q_{rain} + Q_{runoff} - Q_{evap} \pm \varepsilon. \qquad (11)$$

The lake level model is formulated to handle Eqs. (5) to (11). The lake water balance model was run at a monthly time scale. The model was run from January 1998 to December 2009. Since GDAS ET_o data are available from 2001, long-term monthly mean ET_o values were used from 1998 to 2000. GDAS ET_o is the sum of evaporation from the soil surface and transpiration from a standardized reference clipped grass surface (Allen et al., 1998). However, evaporation from open water bodies like lakes and rivers is lower than the pan evaporation or reference ET_o (Allen and Tasumi, 2005) and can be represented by ET_f. We used an ET_f of 0.75 to produce comparable over-the-lake ET losses obtained from the literature (Yuretich and Cerling, 1983; Cerling, 1986; Avery, 2010). Initial Lake Turkana water level information for January 1998 was obtained from the French Space Agency

website. An error term (ε) is introduced in Eq. (11) to compensate for the data and modelling errors. Parameter ε is estimated through model calibration.

3.2.4 Estimation of variations in lake volumes

The lake volume at each time step (t) is computed as

$$LV_{(t)} = LV_{(t-1)} + \Delta S_{(t)} \qquad (12)$$

where $LV_{(t)}$ is the volume of the lake at time step t; $LV_{(t-1)}$ is the volume of the lake at previous time steps; and $\Delta S_{(t)}$ is the change in storage, which is obtained as

$$\Delta S_{(t)} = TRV_{(t)} + V_{rain_{(t)}} - VET_{(t)} \pm V\varepsilon_{(t)}. \qquad (13)$$

Here, TRV_i is the total runoff volume contribution obtained from Eq. (10); $V_{rain_{(t)}}$ is the volume of rainfall received over the lake obtained from RFE rainfall; $VET_{(t)}$ is the volume of ET losses over the lake; and $V\varepsilon)_{(t)}$ is the volume contribution of the error for the time step (t). The initial volume of the lake is obtained using depth-volume relationship derived using Eq. (4).

3.3 Uncertainties in LLM approach

In physically based modelling, it is important to distinguish between the predictive performance of a model and its ability to explain environmental phenomena (Beven, 2001). Uncertainty in hydrologic models includes uncertainties in (a) the structure of the model, (b) the model parameters/input data, and (c) the solution of the model (Addiscott et al., 1995). In most hydrologic models, when using satellite data such as the LLM approach, major uncertainties in the model outputs can be attributed to the model parameters or input data. To understand the uncertainty in the LLM model, the impact of the bias in the input data is to be understood. But it is neither possible nor desirable to evaluate and eliminate all of the uncertainties associated with data and models because resources are always limited and must be used effectively (Van Rompaey and Govers, 2002). Hence, parameters that are likely to contribute most to the uncertainties associated with the model results were evaluated. In the case of the LLM model, parameters such as WHC, interception losses, groundwater fluxes, NDVI, and DEM are static across years and hence would result in minimal random errors. Errors in other parameters such as rainfall, runoff, and ET would critically affect modelling results. Validation of RFE rainfall over the Ethiopian highlands using gauge data suggested that RFE can be reliably used for early warning systems to empower the decision making (Dinku et al., 2008). However, RFE underestimates rainfall during peak rainy seasons and overestimates in other seasons (Laws et al., 2004), with an average bias of $-0.15\,\mathrm{mm\,day^{-1}}$ (NOAA/CPC, 2002). Few validation studies indicate estimates of errors for different locations in Africa and over different time periods (Laws et al., 2004; Dinku et al., 2008). However, we cannot extrapolate

these errors for the Lake Turkana basin due to its complexity. Further studies in this direction are required to determine the true errors in RFE data over the period 1998–2009 for the Lake Turkana basin. Even though the relationship between rainfall and runoff is not linear from individual storms, we assume a constant rainfall-runoff coefficient to evaluate the propagation of mean bias in RFE into modelled runoff. Other important input in the LLM approach is the ET data. Currently, there is no published evidence of validation of GDAS reference ET in Africa. However, modelled ET data validated in the United States showed a mean underestimation up to 5 % over 16 sites (Senay et al., 2008). Finally, the relationships between the change in the lake levels in relation to basin rainfall, runoff, and ET were derived and the impact of these errors on the modelled lake water levels was understood.

3.4   Model calibration

Model calibration is an essential step in developing a reliable and useful hydrologic model. This becomes a necessity especially when parameters from satellite-driven data are used for estimating water balance components. We were unable to obtain reliable in-situ data over the Lake Turkana basin. The Omo River discharge data obtained from Ethiopia had numerous data gaps. Discharge data for the Turkwel and Kerio Rivers were also not available. Unavailability of reliable data is a common problem in most basins. Since ground truth data were not available, we used lake levels derived from satellite altimetry data as proxy to ground truth for calibrating the model. Data from the years 1998 to 2003 were used for calibration. In this study, magnitude differences in lake levels were minimized by estimating parameter ε to account for errors. Not accounting for ε could lead to errors in magnitude while performing water balance of the lake. Estimate of ε was assumed to be equal to zero in the initial model run. Based on the initial model estimates, ε values were varied for each model run. Lake levels modelled using different combinations of varying ε were compared with satellite altimetry data, and the ε parameter that provided the minimum value of mean absolute error (MAE) was selected.

3.5   Validation of modelled lake water levels using satellite altimetry data

Ideally, in-situ observations of lake levels are required for validating modelled estimates. But for Lake Turkana, such in-situ observations are not available. This is true in most ungauged basins. Therefore, modelled Lake Turkana water levels were validated using satellite altimetry data estimated from TOPEX/Poseidon (T/P), Jason-1, and ENVISAT. Data from 2004–2009 were used for validation of model results. Since a mass balance approach is used to derive Lake Turkana water levels and volumes, validation was performed using both lake levels and lake volumes. However,

satellite altimetry produces only lake level information; thus, lake volume information was derived from satellite altimetry data using Eq. (12), where change in storage ($\Delta S_{(t)}$) was derived as

$$\Delta S_{(t)} = \left[\left(D_{(t)} - D_{(t-1)} \right) \times \mathrm{LA}_{(t)} \right]. \tag{14}$$

Here, $D_{(t)}$ and $D_{(t-1)}$ are the depths for current and previous time steps, and $\mathrm{LA}_{(t)}$ is the lake surface area at time step t.

3.6   Model accuracy

Model results (both lake levels and volumes) were compared with the altimetry data to evaluate the model performance. Pearson's correlation coefficient (r) is estimated to observe the degree of relationship between satellite altimetry data and the modelled lake level data for calibrated, validation, and combined time periods. Improvements in r for each dataset are tested for significance using Fisher's z-test. Further, to derive statistical goodness of fit of the modelled lake water levels, several statistical estimates were computed. First, root mean square error (RMSE) was computed using the following equation:

$$\mathrm{RMSE} = \sqrt{\frac{\sum\limits_{i=1}^{n} (P_i - O_i)^2}{N}} \tag{15}$$

where P is the modelled lake water level, O is the altimetry lake water level, N is the total number of observations, and i represents time step.

Willmott and Matsuura (2005) reported that MAE is more appropriate over RMSE in assessing average model performance because MAE is not influenced by large errors. MAE was computed using the following equation:

$$\mathrm{MAE} = \frac{\sum\limits_{i=1}^{n} |P_i - O_i|}{N}. \tag{16}$$

Also, a widely used measure in hydrology, the Nash-Sutcliffe Coefficient of Efficiency (NSCE), was used to compute the model efficiency. The advantage of NSCE is that it accounts for the model errors in estimating the mean of the observed datasets. The NSCE is an indicator of the model's ability to predict the 1:1 line (Nash and Sutcliffe, 1970). A value of 1 represents a perfect match and a value of 0 or less is no more accurate than predicting the mean value. NSCE was computed using the following equation:

$$\mathrm{NSCE} = 1 - \frac{\sum\limits_{i=1}^{n} (P_i - O_i)^2}{\sum\limits_{i=1}^{n} \left(O_i - \bar{O} \right)^2} \tag{17}$$

where $\bar{O}$ is the mean value of the observed variable. Finally, mean bias error (MBE) between the modelled lake water

levels and satellite measurements is computed using the following equation:

$$\text{MBE} = \frac{1}{N} \sum_{i=1}^{N} P_i - O_i. \tag{18}$$

To understand the significance of each estimate of error statistic, percent error with respect to the long-term natural variability of Lake Turkana water levels was computed. Modelled lake level data from January 2003 to December 2009 were used for the accuracy assessment.

4 Results and discussion

4.1 Lake level modelling

First, lake level, surface area, and volume relationship are derived for Lake Turkana (Fig. 3). Modelled Lake Turkana water levels from January 1998 to December 2009 are shown in Fig. 4. Visual analyses of patterns observed in modelled lake levels show that seasonal variations and patterns in lake water levels are captured reasonably well. Since the end of 1999, lake water levels gradually declined until mid-2006. However, after mid-2006, the model showed a steep increase in the lake water levels by the end of 2007 and then a gradual decrease by the end of 2009. In this section, the patterns observed in the modelled lake levels are compared with rainfall and climatic patterns observed in the region. For comparison, the lake water level variations for 1998–2009 are divided into five time periods. The trends observed in each time period are compared with general rainfall trends and are supported by citations from the literature.

- *Period 1 (1998):*
  The model results show an increase in the lake water level up to 1.5 m until the end of 1998. The 1997–1998 El Niño caused heavy rains over East Africa (Galvin et al., 2001; Behera et al., 2005). Anyamba et al. (2001) reported that during this period, East Africa had above normal NDVI due to excess rainfall, and southern Africa had below normal NDVI due to a rainfall deficit. This trend is captured by the model (Fig. 4). This increase in the trend up to 1.5 m of lake level shown by the model is corroborated by Birkett et al. (1999), who reported a ~2 m increase in Lake Turkana water levels during this time period.

- *Period 2 (1999–2002):*
  After the heavy El Niño rains, there was a prolonged period of below average rainfall for four consecutive years until 2003. WFP (2000) reported that drought in 2000 was estimated as the worst on record for East Africa. Furthermore, Anyamba et al. (2002) reported that most of the Horn of Africa had NDVI deficits on the order of 30 % to 80 % below normal. The model results show that the lake water levels decreased gradually
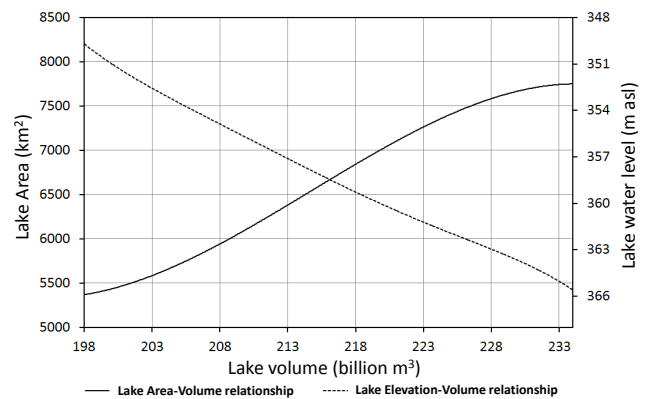

Fig. 3. Relationship between Lake Turkana depth, surface area, and volume derived from seamless lake elevation bathymetry dataset and GIS modelling.

until 2003. As a result of dry weather and possibly very high ET losses, the lake water levels dropped 2 to 3 m during this period.

- *Period 3 (2003–2006):*
  As the Lake Turkana basin experienced below normal rainfall during this period, modelled lake water levels show that water levels declined, and by early 2006, lake levels reached the minimum observed level during the 1998–2009 time period.

- *Period 4 (2006–2007):*
  The model results for this period show a steep increase in the lake water levels. This increase is caused by high runoff generated by heavy rains that occurred in 2006 and 2007. The model shows that the lake water levels increased up to 3 m because of the heavy rainfall and subsequent floods in Ethiopia in 2006 and early 2007. The Disaster Preparedness and Prevention Agency (DPPA) of Ethiopia confirmed that the floods in the southern Omo River Valley caused a few hundreds of casualties and displaced nearly a thousand people (IFRC, 2006). Moreover, Moges et al. (2010) reported that the 2006 floods in Ethiopia were exceptionally severe in intensity, water volume, geographical coverage, and damage.

- *Period 5 (2008–2009):*
  During this period, modelled lake levels show a gradual decrease, with lake levels dropping up to 2 m because of normal to below normal rainfall in 2008 and 2009.

4.2 Uncertainties in LLM approach

The relationships between rainfall, runoff, and ET on changes in Lake Turkana water levels are shown in Fig. 5. The monthly data are classified into wet and dry months with
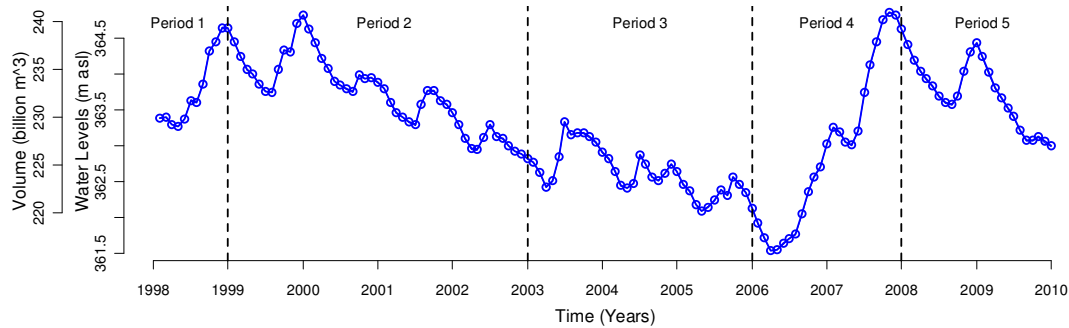
Fig. 4. Lake Turkana water levels modelled using the lake level modelling (LLM) approach and multi-source satellite data. Estimated errors with respect to the modelled runoff and ET data are too small to be visible with respect to the data points.
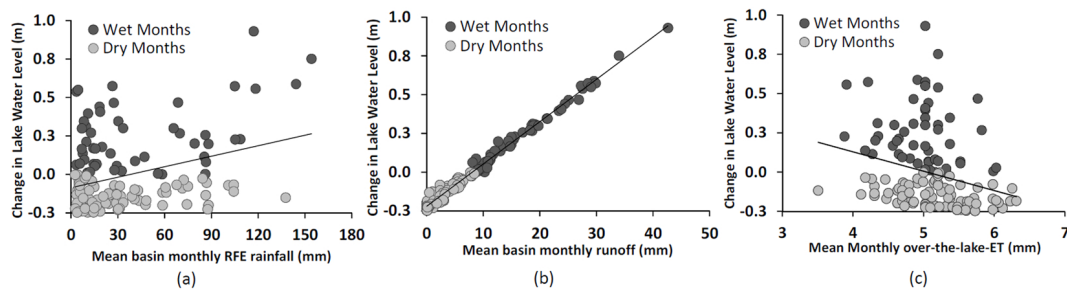

Fig. 5. Relationship between key model parameters and change in Lake Turkana water levels: **(a)** relationship between mean basin monthly RFE rainfall and change in lake water levels ($Y = 0.0023\,X - 0.0863$); **(b)** relationship between mean monthly runoff and change in lake water levels ($Y = 0.0273\,X - 0.2205$); and **(c)** effect of ET on change in lake water ($Y = -0.1236\,X + 0.6231$). All the trend lines and equations are derived using combined data (wet + dry months). Monthly data from 1998–2009 are used in this analysis.

respect to the lake, where wet and dry months correspond to the months when the lake level increased and decreased, respectively. In Fig. 5a, the relationship between rainfall and lake level changes is not always linear, as rainfall has to meet the soil moisture and other storage demands in the basin before generating runoff. On the other hand, once the runoff is generated and reaches the lake, it shows a linear relationship with the lake level changes (Fig. 5b). However, basin runoff/inflows have to be more than the evaporative demand of the lake to cause a net increase in lake levels. Figure 5c shows that ET over wet months does not show any relation. However, it shows a strong relation over the dry months, when the effect of ET on the lake level changes is substantial.

Using relationships derived in Fig. 5, the impact of the errors on the lake water levels is estimated. It is found that the bias in RFE data (-0.15 mm/day) would translate to up to 1 cm month^{-1} of error in the modelled lake levels during peak rainy seasons (March to June and September to November). The runoff coefficient of 0.21 is obtained from monthly analysis between rainfall and modelled runoff data. Using this coefficient, the error in monthly runoff data is estimated to be up to 0.3 mm month^{-1}, which would further introduce up to 2.5 cm month^{-1} of error in the modelled lake level data over peak rainy seasons. Together, rainfall and runoff would result in an error of up to 3.5 cm month^{-1}. The magnitude

of error during other months would be less as the number of days of rainfall would be low. Assuming consistent errors globally, errors in the ET data (up to 5 % underestimation) are introduced in the model, and their impact on the modelled lake levels is estimated using the relationship obtained in Fig. 5c. Our results indicate that errors in ET data would translate to up to 4 to 5 cm month^{-1} of error in lake levels. More evaluation is needed to understand the impact of these errors on lake level dynamics. Total errors in the modelled lake levels would be compensated by the constant parameter ε in Eq. (11), which is estimated by the calibration process.

4.3 Model calibration

Figure 6a shows the comparison of un-calibrated modelled lake levels and volumes with altimetry-based lake levels. Figure 6a shows that the patterns and seasonal variations in water level fluctuations are captured reasonably by the model. However, the un-calibrated model shows an overestimation with a difference in magnitude when compared to the altimetry data, with an MAE of 0.96 m. During the calibration process, and since un-calibrated lake levels were showing overestimation when compared to satellite altimetry data, the value for ε was considered negative and varied from -0.1 to 0 m per month (with an increment of 0.01 m). The
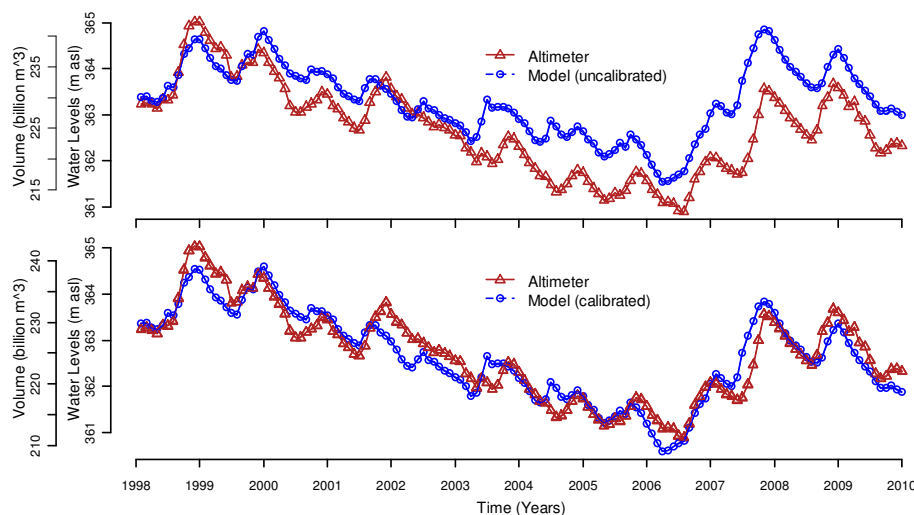
Fig. 6. Comparison of Lake Turkana water levels modelled using the LLM approach and satellite altimetry data: **(a)** un-calibrated lake levels $[\varepsilon = 0]$; **(b)** calibrated lake levels for magnitude $[\varepsilon = 0.06\,\mathrm{m\,day^{-1}}]$. The calibration is performed on data from 1998 to 2003 and the model is validated using data from 2004 to 2009.

MAE for different combinations of varying ε ranged from 0.31 to 3.4 m. An $\mathrm{ET_f}$ value of 0.75 and an ε value of 0.06 m per month provided the least MAE (0.31 m) when compared to the satellite altimetry data. To validate the $\mathrm{ET_f}$ fraction, total actual ET losses over the lake were estimated and compared with the published ET losses for Lake Turkana. Using $\mathrm{ET_f}$ of 0.75, an average rate of ET from Lake Turkana is estimated as $2.3\,\mathrm{m\,yr^{-1}}$, and this estimate is found to be within the range of ET losses reported in the literature (Yuretich and Cerling, 1983; Cerling, 1986, Avery, 2010). Hence, these parameter estimates of $\mathrm{ET_f}$ and ε obtained by calibration are considered for further modelling.

4.4 Model validation using satellite altimetry data

Modelled lake levels are validated using lake levels estimated from satellite altimetry data (Fig. 6a and b). After considering ε of $0.06\,\mathrm{m\,month^{-1}}$, the error in modelled vs. altimetry was reduced with MAE of 0.31, 0.27, and 0.29 m over the calibration, validation, and combined periods. Total monthly over-the-lake rainfall, over-the-lake ET, and total monthly runoff into the lake are shown in Fig. 7a–c. Modelled monthly lake water levels from January 1998 to December 2009 are illustrated against altimetry data in Fig. 7d. Possible reasons for the errors observed between the model and altimetry-based lake level estimates are listed here.

In the LLM approach, the model-based Lake Turkana water levels are primarily driven by runoff and ET. The increase in the lake water levels is driven by the runoff derived from the rainfall estimates. The differences seen while the lake water levels are increasing could be attributed to inaccuracies in the satellite rainfall estimates or the modelling errors. On the other hand, the decline in the lake water levels is mostly

dependent on the over-the-lake ET and ε. The slope of the declining trend as seen in modelled lake levels matches reasonably well with the altimetry data, which means that the error contributed from ET could be minimal.

The wetland complex located in the Omo River Delta could act as a temporary reservoir and possibly reduce the flow rate, which could result in the errors in the modelled estimates. Another reason for the difference could be caused by a small percentage of subsurface groundwater drainage occurring in the basin. Information on the subsurface drainage occurring in the upper Lake Turkana basin is not available. Other sources of discrepancy in modelled lake levels could also be due to (a) changes in lake surface pressure, (b) wind-driven events or tides, or (c) fluctuations in the volume of the column due to an alternating temperature or composition, which could also influence lake water levels (Mercier et al., 2002).

During 2002, 2003, and 2004, the peaks of modelled lake levels tend to show some discrepancy with the peaks seen in satellite altimetry estimates. This could be due to the occurrence of low flows, as the basin received below average rainfall during these years. Further investigation is required to understand the differences in peak flows during a low flow year.

The use of the constant ε could result in differences between the modelled and the satellite altimetry data. In reality, ε would vary with time of the year. However, accurate estimates of ε for unit time step is a challenging task unless the uncertainty in the data and model is clearly understood.

Minor discrepancies seen after 2003 can be also explained by the Gilgel Gibe hydroelectric dam-I on the lower Omo River, commissioned in 2004. The impact of the dam on the
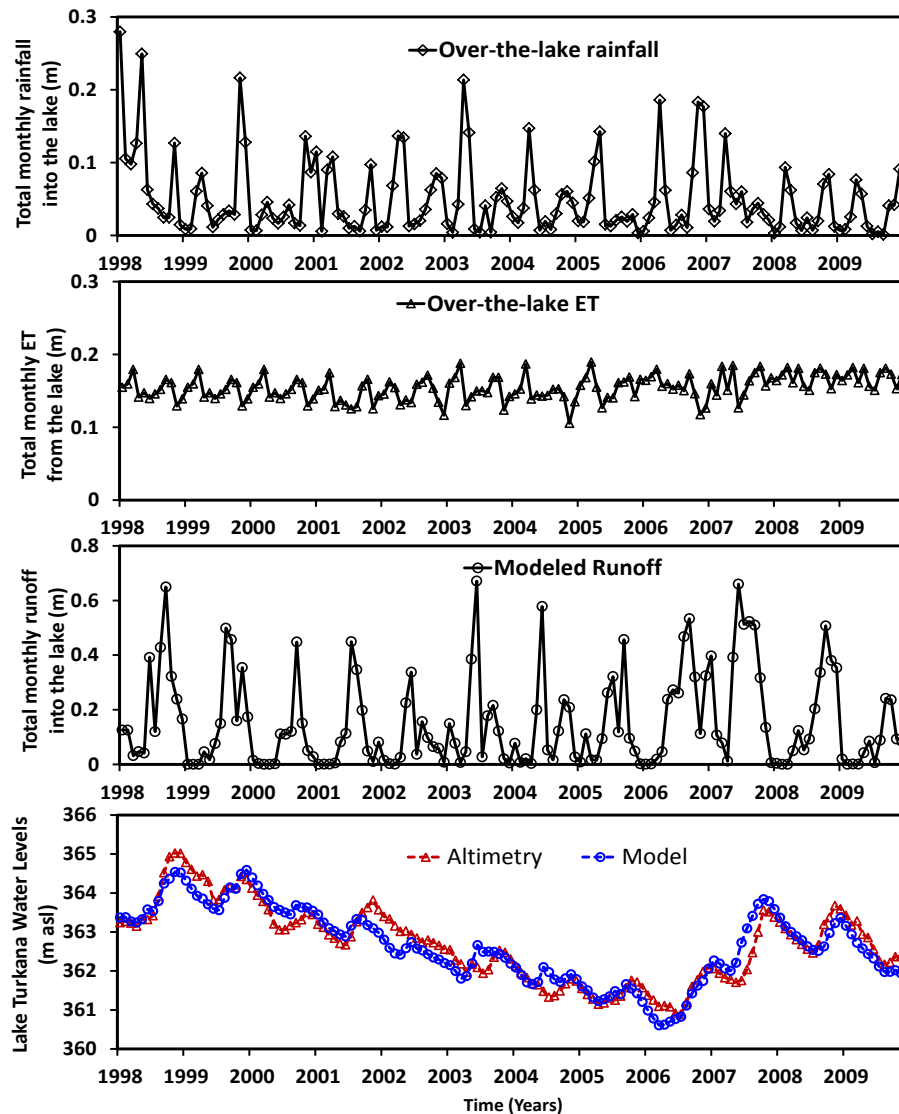
Fig. 7. (a) Total average basin rainfall (m month^{-1}) modelled using satellite rainfall estimates; **(b)** total over-the-lake ET$_o$ (m month^{-1}) estimated from GDAS ET$_o$; **(c)** total average basin runoff (m month^{-1}) modelled for the Lake Turkana basin; **(d)** Lake Turkana water levels for 1998–2009. Modelled daily lake levels (in blue) and lake level estimates from TOPEX/Poseidon (T/P), Jason-1, and ENVISAT satellite altimetry data (in dark red).

lake water levels is not clearly understood. Further, ET losses from the reservoir would also decrease the total volume of water that would end up in the lake and could subsequently lead to the delay in the lake level hydrograph. The effect of the Gibe-I dam on the lake levels is not modelled, as information on the operational strategies for the dam is unavailable.

4.5 Model accuracy

Accuracy assessment is performed by comparing both modelled lake water levels and lake volumes with the estimates from satellite altimetry data. The Pearson's correlation coefficients and percentage of errors were similar for both cases (lake levels and lake volumes). Hence, we only presented

accuracy results of lake levels in Table 3. The un-calibrated modelled lake water levels and the satellite measurements yielded a high degree of correlation with Pearson's correlation coefficient (r) values of 0.87, 0.92, and 0.86 for calibration, validation, and combined periods. Although the accuracy of un-calibrated modelled estimates are high, this method can only be used to study the long-term trends in lake level variations when ground truth data are not available. On the other hand, model accuracy was significantly improved when calibrated with limited ground truth data. Calibrated lake levels showed a higher degree of correlation with correlation coefficient (r) values of 0.89, 0.90, and 0.93 for calibration, validation, and combined time periods, respectively. The improvement in r value for the combined period was

Table 3. Results of calibration and accuracy assessment of modelled lake water levels using satellite altimetry data.

Dataset	Calibration period (1998–2003)	Validation period (2004–2009)	Combined period (1998–2009)
Pearson's correlation coefficient (r):			
Un-calibrated	0.87*	0.92*	0.86*
Calibrated	0.89*	0.90*	0.93*
Error statistic** [units]			
RMSE [m]	0.36 (7 %)	0.35 (10 %)	0.36 (9 %)
MAE [m]	0.31 (6 %)	0.27 (7 %)	0.29(7 %)
Bias [m]	−0.08 (−2 %)	0.00 (0 %)	−0.04 (−1 %)
NCSE [no units]	0.87	0.80	0.87

* p-value $\leq$ 0.001; ** error statistic estimates are made using calibrated data.

Note: Value next to each error statistic estimate denotes percent error with respect to the natural lake level variability of 4.8 m.

Fig. 8. Scatterplot of modelled water levels vs. satellite altimetry-based lake level data. Also shown is the scatterplot of modelled lake volumes vs. altimetry-based lake volumes. Each point represents mean monthly modelled (calibrated) lake water levels or corresponding lake volume over the combined period (January 1998 through December 2009). 1:1 line is shown as dotted line.

found to be significant with the Fisher's statistic at 95 % significance level. Error statistics in Table 3 were estimated using calibrated lake levels for calibration, validation, and combined time periods. The model efficiency estimated using NSCE is found to be 0.87, 0.80, and 0.87 for calibration, validation, and combined periods, respectively. For the validation period, the RMSE and MAE were found to be 0.35 m and 0.27 m, respectively, and the model showed no mean bias error. The MAE and RMSE are found to be 10 %, respectively, and 7 % of the long-term natural variability observed for Lake Turkana (4.8 m). As a result, the LLM approach can be used to model lake levels with confidence. Figure 8 illustrates a scatterplot between the modelled and the satellite altimetry measurements. The modelled versus satellite altimetry data lie reasonably on the 1:1 line.

4.6 Variations in Lake Turkana water levels

Lake Turkana shows a high degree of seasonal variability. Based on our modelling results, we found that the annual ET losses from the lake were between 2.1 and 2.3 m. Over-the-lake rainfall contributes only up to 30 % of the lake evaporative demand. Lake inflows and evaporation losses are the two key factors affecting lake water levels. Since over-the-lake precipitation amounts to only up to one-tenth of the evaporation losses, the increase in lake level is mainly caused by inflows from the Omo River. A decline in lake level is highly influenced by the ET losses from the lake. On average, Lake Turkana water levels would start to rise from July and reach a peak level by October–November. Thereafter, due to the reduction in the inflows from the Omo River, the lake would decline gradually until the end of summer. During the modelling time period, Lake Turkana showed seasonal variations of 1–2 m. The lake level fluctuated up to 4 m between the years 1998–2009.
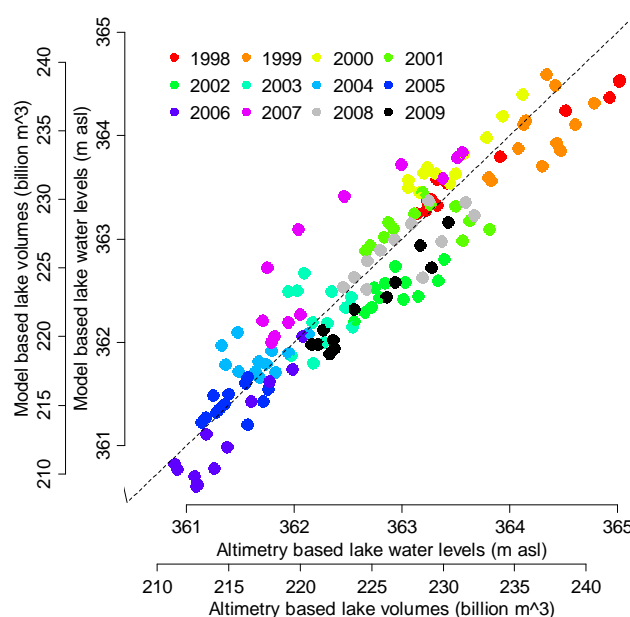
4.7 Lake Turkana water level modelling: opportunities and challenges

4.7.1 Use of satellite altimetry data for model calibration

Satellite data/models for lake level studies are subject to high errors. Hence, one of the challenges of using satellite-driven data/models for estimating water balance components is the need for data/model calibration. But calibration is especially difficult in areas where reliable gauge measurements are unavailable. However, accurate and consistent satellite altimetry-based lake level data are available for over 150 large lakes and reservoirs globally (Cretaux et al., 2011). Altimetry data on river height are also available for large river basins around the world. So far, satellite altimetry-based lake level data have not been used for calibration or validation of hydrologic models, especially over lakes in data scarce regions. Recently, Getirana et al. (2010) used altimetry data on river height to validate a hydrologic model for the Negro River basin. In this study, we demonstrate an approach using satellite altimetry data for model calibration and validation. Enhanced accuracy due to calibration and validation of hydrologic modelling enables the use of satellite-driven data for understanding the interaction between lakes and watersheds. However, calibration of the model could be a challenging task where satellite altimetry data or gauge data are unavailable.

4.7.2 Towards quantification of upstream impacts on lake water levels

The assessment of lake water balance would provide improved knowledge of regional and global climate change and a quantification of the human impacts on water resources (Cretaux and Birkett, 2006). The use of a multi-source satellite data approach offers a unique advantage to understand the upstream impacts on Lake Turkana water levels. Upstream basin processes such as changes in LCLU, groundwater abstraction, irrigation water use, construction of dams, or any newly imposed water regulations influence the Lake Turkana inflows. Gathering such data could be a challenging task using remote sensing datasets. However, in this study, the use of phenology information based on climatological NDVI for runoff modelling (Senay et al., 2009) has provided satisfactory results for Lake Turkana. However, future research should focus on the use of current NDVI for runoff modelling to capture the upstream impacts on lake water levels. Furthermore, the Ethiopian government is currently building a series of dams on the Omo River. Setting up a well calibrated and validated water balance model is a first step towards understanding the interaction and potential impacts of the dams on the hydrology of Lake Turkana.

4.7.3 Use of satellite rainfall data for Lake Turkana water level modelling

Recently, satellite-based rainfall estimates are being created by combining data from a combination of sensors and from different sources to improve accuracy, coverage, and resolution. The usability of RFE-assimilated products demonstrated in this study would enhance the efforts towards monitoring surface water bodies, especially in ungauged basins. The RFE dataset has been successfully used as an early warning tool in several parts of the world by the USA's Famine Early Warning Systems (FEWS) Network. However, considering the accuracies of RFE, modelling results from the LLM approach should be thoroughly calibrated and validated. As pointed out by Artan et al. (2007), our study also demonstrated that calibrated and validated model results could only be used for monitoring lake levels, whereas un-calibrated results could only be used to infer relative year-to-year lake level changes.

4.7.4 Seasonal forecasting of lake level variations

The Intergovernmental Authority on Development (IGAD) Climate Prediction and Applications Center (ICPAC) releases a seasonal climate outlook statement for every three months for the Greater Horn of Africa (GHA) region, which includes the Lake Turkana basin. This seasonal climate outlook provides information on probability of rainfall in terms of percentage of above-normal, near-normal, and below-normal rainfall occurrences summarized from model
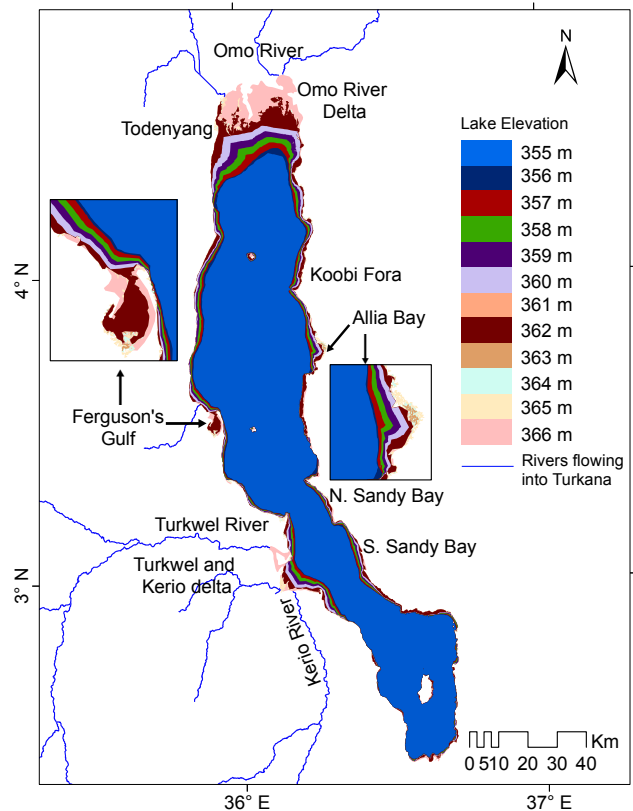

Fig. 9. Lake Turkana water levels at various depths. The shallow regions along the lake shoreline show the most productive fishing zones.

forecasts provided by the Global Producing Centres (GPCs), statistical modelling, and expert analysis and interpretation (Ogallo et al., 2008). Currently, no model is available to operationally translate seasonal rainfall forecast information into useful applications. We suggest that a satellite-driven water balance model can be integrated to translate seasonal rainfall forecast information in Lake Turkana water level forecasts. However, more research and application development is required in this direction.

4.7.5 Impact of lake level change on Lake Turkana fisheries

Lake Turkana water levels are very critical for fisheries production. Small changes in Lake Turkana levels often result in large changes in fish yields for Lake Turkana (Kolding, 1992). The most productive fishing zones in Lake Turkana are found along the shallow areas of the lake, as shown in Fig. 9. The most significant fish producing area, Ferguson's Gulf on the western side of the lake, is vulnerable to lake level changes because it dries when the lake level falls below 362 m (Avery et al., 2010). This gulf remained dry during 2005–2006. During this period, the fish catch records dropped from 9000 to 2500 t (Avery et al., 2010).

Furthermore, fish availability and extent are also dependent on the changes in littoral/inshore habitat distribution, extent of turbidity, and nutrient availability, which are directly linked to lake inflows. Because of this, continuous monitoring of lake inflows, apart from lake levels, should be undertaken (Kolding, 1992). The use of satellite altimetry data for model calibration would enable us to reliably understand and simulate the impact of individual components of lake water balance. This approach can also be used to generate possible changes in the lake levels based on the short-term climate forecasts.

4.7.6 Operational monitoring of Lake Turkana level variations

Application of this approach over other complex basins could be a challenging task, especially due to the complexity and poor constraining of certain water balance components such as Q_{gwin}, Q_{gwout} and $Q_{outflow}$. However, the water balance of Lake Turkana using multi-source satellite data can be satisfactorily used to model lake water level variations. Results indicated that calibrated lake levels captured the observed trends reasonably well (Fig. 6). Therefore, the multi-sensor-driven physical hydrologic model presented here can be used for operational monitoring of Lake Turkana. The satellite data used in this study are available for download from different sources in near-real time. RFE rainfall data, GDAS ET_o, and NDVI data are available for download in near-real time with a few days of lag. Other static datasets, such as SRTM DEM, the Digital Soil Map of the World, and MODIS VCF are also available for download at no cost. Future NASA missions, such as the Global Precipitation Measurement (GPM) and the Visible Infrared Imager Radiometer Suite sensor on board the National Polar Orbiting Operational Environmental Satellite System (NPOESS), will enable reliable estimation of climate variables and improve the accuracy of rainfall and ET products, making the LLM approach more useful.

5 Conclusions

The objectives of this study are (a) to demonstrate the use of satellite altimetry data for model calibration and validation when reliable in-situ data are unavailable and (b) to establish a calibrated satellite data-driven water balance model for Lake Turkana to improve understanding of the interactions between the Lake Turkana and its watershed. Since most satellite-driven data/models require calibration, we presented an approach to calibrate and validate the water balance model for Lake Turkana using a composite lake level product of TOPEX/Poseidon, Jason-1, and ENVISAT satellite altimetry data. The use of satellite altimetry data made it possible to calibrate a satellite-driven hydrologic model without using any in-situ data. The model results showed

that the satellite-driven lake level modelling approach could satisfactorily capture the patterns and seasonal variations of the lake water level fluctuations, including the effect of El Niño/floods in 1998 and 2006, and the effect of drought in 2000. Validation results showed that model-based lake levels are in good agreement with observed satellite altimetry data with a Pearson's correlation coefficient of 0.90 and with model efficiency of 0.80 during the validation period. Further, error estimates were found to be within 10 % of the natural variability of Lake Turkana, giving high confidence on the modelled lake level estimates.

It was found that the lake inflows and over-the-lake ET are the two main driving forces of Lake Turkana water levels. Over-the-lake rainfall contributes only up to 30 % of the lake evaporative demand. During the modelling time period, Lake Turkana showed seasonal variations of 1–2 m. The lake level fluctuated up to 4 m between 1998 and 2009. This study demonstrated the usefulness of satellite altimetry data (a) to calibrate and validate the hydrologic model, especially in ungauged basins, and (b) to establish a reliable water balance model for understanding the interactions between the lake and its watershed. Furthermore, for Lake Turkana, we identified opportunities and challenges of using a calibrated satellite-driven water balance model for (i) quantitative assessment of the impact of upstream basin developmental activities on lake levels and (ii) the use of seasonal rainfall forecasts for assessing lake level changes and their impact on fisheries. From this study, we suggest that globally available satellite altimetry data provide a unique opportunity to study similar ungauged basins in different parts of the world.

References

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., and Rasmussen, J.: An Introduction to the European Hydrological System-Systeme Hydrologique European "SHE" 1: History and Philosophy of a Physical, Distributed Modeling System, J. Hydrol., 87, 45–59, 1986.

Addiscott, T. M., Smith, J., and Bradbury, N.: Critical evaluation of models and their parameters, J. Environ. Qual., 24, 803–807, 1995.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements, FAO Irrigation and Drainage Paper, Rome, 56, 1998.

Allen, R. G. and Tasumi, M.: Evaporation from American Falls Reservoir in Idaho via a combination of Bowen ratio and eddy covariance, Proceedings of EWRI Conference, Anchorage, Alaska, 15–19 May 2005.

Allen, R. G., Tasumi, M., Morse, A., and Trezza, R.: A landsat based energy balance and evapotranspiration model in the Western US water rights regulation and planning, Irrig. Drain. Syst., 19, 251–268, 2005.

Alsdorf, D., Birkett, C., Dunne, T., Melack, J., and Hess, L.: Water level changes in a large Amazon lake measured with space-borne radar interferometry and altimetry, Geophys. Res. Lett., 28, 2671–2674, 2001.

Alsdorf, D. E., Rodriguez, E., and Lettenmaier, D. P.: Measuring surface water from space, Rev. Geophys., 45, RG2002, doi:10.1029/2006RG000197, 2007.

Anyamba, A., Tucker, C. J., and Eastman, J. R.: NDVI anomaly patterns over Africa during 1997/1998 ENSO warm event, Int. J. Remote Sens., 22, 1847–1859, 2001.

Anyamba, A., Tucker, C. J., and Mahoney, R.: From El Niño to La Niña: Vegetation response over east and south Africa during 1997–2000 periods, J. Climate, 15, 3096–3103, 2002.

Artan, G., Gadian, H., Smith, J. L., Asante, K. O., Badaragoda, C. J., and Verdin, J. P.: Adequacy of satellite derived rainfall data for stream flow modeling, Nat. Hazards, 43, 167–185, 2007.

Asante, K. O.: Approaches to continental scale river flow routing, Ph.D. dissertation, University of Texas at Austin, 281, 2000.

Asante, K. O., Artan, G. A., Pervez, S., Bandaragoda, C., and Verdin, J. P.: Technical documentation of the Geospatial Stream flow Model (GeoSFM), USGS EROS Center, 2007.

Avery, S.: Hydrological impacts of Ethiopia's Omo basin on Kenya's lake Turkana water levels and fisheries, African Development Bank, Tunis, 146, 2010.

Awange, J. L., Ogalo, L., Bae, K. H., Were, P., Omondi, P., Omute, P., and Omullo, M.: Falling Lake Victoria water levels: Is climate a contributing factor?, Climatic Change, 89, 281–297, 2008.

Bates, B. C., Kundzewicz, Z. W., Wu, S., and Palutikof, J. P.: Climate Change and Water, Technical Paper of the IPCC, IPCC Secretariat, Geneva, 210, 2008.

Becker, M., Llovel, W., Cazenave, A., Gunter, A., and Cretaux, J. F.: Recent hydrological behavior of the East African great lakes region inferred from GRACE, satellite altimetry and rainfall observations, Comptes Rendus Geoscience, 342, 223–233, 2010.

Behera, S. K., Luo, J. J., Masson, S., Delecluse, P., Gualdi, S., Navarra, A., and Yamagata, T.: Paramount impact of the Indian Ocean dipole on the east African short rains: A CGCM study, J. Climate, 18, 4514–4530, 2005.

Beven, K.: On explanatory depth and predictive power, Hydrol. Process., 15, 3069–3072, 2001.

Birkett, C. M.: The contribution of TOPEX/POSEIDON to the global monitoring of climatically sensitive lakes, J. Geophys. Res., 100, 25179–25904, 1995.

Birkett, C. M. and Mason, I. M.: A new global lakes database for a remote sensing program studying climatically sensitive large lakes, J. Great Lakes Res., 21, 307–318, 1995.

Birkett, C. M., Murtugudde, R., and Allan, T.: Indian Ocean climate even brings floods to East Africa's Lakes and the Sudd Marsh, Geophys. Res. Lett., 26, 1031–1034, 1999.

Carr, C. J.: Patterns of vegetation along the Omo River in the southwest Ethiopia, Plant Ecol., 135, 135–163, 1998.

Cerling, T. E.: A mass-balance approach to basin sedimentation: Constraints on the recent history of the Turkana basin, Palaeogeogr. Palaeocl., 43, 129–151, 1986.

Cretaux, J. and Birkett, C.: Lake studies from satellite radar altimetry, Comptes Rendus Geosciences, 338, 1098–1112, 2006.

Cretaux, J., Jelinski, W., Calmant, S., Kouraev, A., Vuglinski, V., Berge-Nguyen, M., Gennero, M.-C., Nino, F., Abarca Del Rio, R., Cazenave, A., and Maisongrante, P.: SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data, Adv. Space Res., 47, 1497–1507, 2011.

Crow, W. T. and Ryu, D.: A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals, Hydrol. Earth Syst. Sci., 13, 1–16, doi:10.5194/hess-13-1-2009, 2009.

Dinku, T., Chidzambwa, S., Ceccato, P., Connor, S. J., and Ropelewski, C. F.: Validation of high resolution satellite rainfall products over complex terrain, Int. J. Remote Sens., 29, 4097–4110, 2008.

FAO.: Digital Soil Map of the World Version 3.5 (CD-ROM), FAO, Rome, 1995.

Farr, T. G. and Kobrick, M.: Shuttle Radar Topography Mission produces a wealth of data, EOS Trans. AGU, 81, 583–585, 2000.

Fu, L.-L., Christensen, E. J., Yamarone Jr., C. A., Lefebvre, M., Menard, Y., Dorrer, M., and Escudier, P.: TOPEX/POSEIDON mission overview, J. Geophys. Res., 24, 24369–24381, 1994.

Galvin, K. A., Boone, B., Smith, M., and Lynna, J.: Impacts of climate variability on East African pastoralists: linking social science and remote sensing, Clim. Res., 19, 161–172, 2001.

Getirana, A. C. V., Bonnet, M.-P., Rotunno Filho, O. C., Collischonn, W., Guyot, J.-L., Seyler, F., and Mansur, W. J.: Hydrological modelling and water balance of the Negro River basin: evaluation based on in situ and spatial altimetry data, Hydrol. Process., 24, 3219–3236, 2010.

Gibson, J. J., Prowse, T. D., and Peters, D. L.: Hydroclimatic controls on water balance and water level variability in Great Slave lake, Hydrol. Process., 20, 4155–4172, 2006.

Gong, L., Widen-Nilsson, E., Halldin, S., and Xu, C. Y.: Large-scale runoff routing with an aggregated network response function, 368, 237–250, 2009.

Hansen, M. C., Defries, R., Townshend, J., Carrol, M., Dimiceli, C., and Sohlberg, A.: Global percent tree cover at a spatial resolution of 500 meters: first results of the MODIS VCF algorithm, Earth Interact., 7, 1–15, 2003.

Harvey, F. E., Rudolph, D. L., and Frape, S. K.: Estimating Ground water flux into large lakes: Application in the Hamilton Harbor, Western Lake Ontario, Ground Water, 38, 550–565, 2000.

Herman, A., Kumar, V. B., Arkin, P. A., and Kousky, J. V.: Objectively Determined 10-Day African Rainfall Estimates Created for Famine Early Warning Systems, Int. J. Remote Sens., 18, 2147–2159, 1997.

Hörmann, G., Branding, A., Clemen, T., Herbst, M., Hinrichs, A., and Thamm, F.: Calculation and simulation of wind controlled canopy interception of a beech forest in Germany, Agr. Forest Meterol., 79, 131–148, 1996.

Huffman, G. J., Alder, R. F., Rudolf, B., Schneider, U., and Keehn, P. R.: Global precipitation estimates based on a technique for combining satellite based estimates, rain gauge analysis and NWP model precipitation information, J. Climate, 8, 1284–1295, 1995.

Huxman, T., Wilcox, B., Breshears, D., Scott, R., Snyder, K., Small, E. E., Hultine, K., Pockman, W. T., and Jackson, R. B.: Ecohydrological implications of woody plant encroachment, Ecology, 86, 308–319, 2005.

IFRC.: Ethiopia Floods, Operations Update, International Federation of Red Cross, www.ifrc.org/docs/appeals/06/MDRET00301.pdf (last access: 9 November 2011), 2006.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P.: CMORPH: A method that produces global precipitation estimates from passive microwave and Infrared data at high spatial and temporal resolution, J. Hydrometeorol., 5, 487–503, 2004.

Kallqvist, T., Lien, L., and Liti, D.: Lake Turkana Limnological Study 1985–1988, Norwegian Institute of Water Resources, Report number 0-85313, 98, Norway, 1988.

Kalma, J. D., McVicar, R., and McCabe, F.: Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data, Surv. Geophys., 29, 421–469, 2008.

Kebede, S., Travi, Y., Alamayehu, T., and Marc, V.: Water balance of Lake Tana and its sensitivity to fluctuations in rainfall, Blue Nile basin, J. Hydrol., 316, 233–247, 2006.

Kelliher, F. M., Leuning, R., and Schulze, E. D.: Evaporation and canopy characteristics of coniferous forests and grasslands, Oecologia, 95, 153–163, 1993.

Kolding, J.: A summary of Lake Turkana: An ever-changing mixed environment, Mitt. Internat. Verein. Limnol., Stuttgart, 23, 25–35, May 1992.

Kundzewicz, Z. W., Graczyk, D., Maurer, T., Przymusinska, I., Radziejewski, M., Svensson, C., and Szwed, M.: Detection of change in world-wide hydrological time series of maximum annual flow, GRDC Report 32, 36, 2004.

Laws, K. B., Janowiak, J. E., and Huffman, G. J.: Verification of rainfall estimates over Africa using RFE, NASA MPA-RT and CMORPH, AMS Hydrology Conference, 2004.

Lettenmaier, D. P. and Wood, E. F.: Hydrologic forecasting, in: Handbook of Hydrology, edited by: Maidment, D. R., 26, 26.8, 1992.

Li, X. Y., Xu, H. Y., Sun, Y. L., and Zhang, D. S.: Lake level change and water balance analysis of Lake Qinghai, West China, Water Resour. Manage., 21, 1506–1516, 2007.

Mercier, F., Cazenave, A., and Maheu, C.: Interannual lake level fluctuations (1993–1999) in Africa from Topex/Poseidon: connections with ocean-atmosphere interactions over the Indian Ocean, Global Planet. Change, 32, 141–163, 2002.

Moges, S., Alemu, Y., and Mcfeeters, S.: Flooding in Ethiopia: Recent history and 2006 flood: Implications for the Nile Basin, edited by: Kloos, H. and Legesse, W., Cambria Press, NY, 285–305, 2010.

Mu, Q., Zhao, M., and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorithm, Remote Sens. Environ., 115, 1781–1800, 2011.

Naden, P., Broadhurst, P., Tauveron, N., and Walker, A.: River routing at the continental scale: use of globally-available data and an a priori method of parameter estimation, Hydrol. Earth Syst. Sci., 3, 109–123, doi:10.5194/hess-3-109-1999, 1999.

NASA Science Plan 2007–2016.: Science Plan for NASA's Science Mission Directorate 2007–2106, http://science.nasa.gov/media/medialibrary/2010/03/31/Science_Plan_07.pdf (last access: 9 November 2011), 2007.

Nash, E. J. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – a discussion of principles, J. Hydrol., 10, 282–290, 1970.

NOAA/CPC.: The NOAA Climate Prediction Center, African Rainfall Estimation Algorithm, V2.0, www.cpc.ncep.noaa.gov/products/fews/RFE2.0_tech.pdf (last access: 9 November 2011), 2002.

Ogallo, L., Bessemoulin, P., Ceron, J. P., Mason, S., and Connor, S. J.: Adapting to climate variability and change: The climate outlook forum process, WMO Bulletin, 57, 93–102, 2008.

Olivera, F., Flamiglietti, J., and Asante, K.: Global scale flow routing using a source-to-sink algorithm, Water Resour. Res., 36, 2197–2207, 2000.

Ricketts, R. D. and Johnson, T. C.: Climate change in the Turkana basin as deduced from a 4000 year long δO^{18} record, Earth Planet. Sc. Lett., 142, 7–17, 1996.

Ricko, M., Carton, J. A., and Birkett, C.: Climatic effects on Lake basins, Part I: Modeling Tropical lake levels, Am. Meteorol. Soc., 24, 2983–2999, 2011.

Rostamian, R., Aazam, J. A., Afyuni, M., and Mousavi, S. F.: Application of a SWAT model for estimating runoff and sediment in two mountain basins, Hydrol. Sci., 53, 977–988, 2008.

Senay, G. B.: Modeling landscape evapotranspiration by integrating land surface phenology and water balance algorithm, Algorithms, 1, 52–68, 2008.

Senay, G. B. and Verdin, J. P.: Characterization of yield reduction in Ethiopia using a GIS-based crop water balance model, Can. J. Remote Sens., 6, 687–692, 2003.

Senay, G. B., Budde, M., Verdin, J. P., and Melesse, A. M.: A coupled remote sensing and SSEB approach to estimate actual evapotranspiration from irrigated fields, Sensors, 7, 979–1000, 2007.

Senay, G. B., Verdin, J., Lietzow, R., and Melesse, M.: Global daily reference evapotranspiration modeling and evaluation, J. Am. Water Res. Assoc., 44, 969–979, 2008.

Senay, G. B., Asante, K., and Artan, G.: Water balance dynamics in the Nile basin, Hydrol. Process., 23, 3675–3681, 2009.

Sivapalan, M.: Prediction in ungauged basins: a grand challenge for theoretical hydrology, Hydrol. Process., 17, 3163–3170, 2003.

Snell, J. D. and Sivapalan, M.: On geomorphological dispersion in natural catchments and the geomorphological unit hydrograph, Water Resour. Res., 30, 2311–2323, 1994.

Swenson, S. and Wahr, J.: Monitoring the water balance of Lake Victoria, East Africa, from Space, J. Hydrol., 370, 163–176, 2009.

Tate, E., Sutcliffe, J., Conway, D., and Farquharson, F.: Water balance of Lake Victoria: update to 2000 and climate change modeling to 2100, Hydrol. Sci., 49, 563–574, 2004.

Tate, K. W.: Interception on rangeland watersheds, Rangeland watershed program, U. C. Cooperative Extension and USDA Natural Resources Conservation Service, 36, 1995.

Tucker, C. J., Pinzon, J. E., Brown, M. E., Slayback, D., Pak, E., Mahoney, R., Vermont, E., and Nazmi, El.-S.: An Extended AVHRR 8-km NDVI Data Set Compatible with MODIS and SPOT Vegetation NDVI Data, Int. J. Remote Sens., 14, 4485–4498, 2005.

Van Rompaey, A. J. J. and Govers, G.: Data quality and model complexity for regional scale soil erosion prediction, Int. J. GIS, 16, 663–680, 2002.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Reidy Liermann, C., and Davies, P. M.: Global threats to human water security and river biodiversity, Nature, 467, 555–561, 2010.

Wagener, T., Wheater, H. S., and Gupta, H. V.: Rainfall-Runoff Modeling in Gauged and Ungauged Catchments, Imperial College Press, London, UK, 300, 2004.

Wahr, J., Swenson, S., Zlotnicki, V., and Velicogna, I.: Time-variable gravity from GRACE: First results, Geophys. Res. Lett., 31, 1–4, 2004.

Willmott, C. J. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error in assessing average model performance, Clim. Res., 30, 79–82, 2005.

World Food Program (WFP): Kenya's drought: No sign of any let up, WFP, Rome, Italy, http://reliefweb.int/node/70268 (last access: 9 November 2011), 2000.

Xie, P. and Arkin, P. A.: Analysis of Global Monthly Precipitation Using Gauge Observations, Satellite Estimates, and Numeric al Model Prediction, J. Climate, 9, 840–858, 1996.

Yuretich, R. F. and Cerling, T. E.: Hydrogeochemistry of Lake Turkana, Kenya: mass balance and mineral reactions in an alkaline lake, Geochim. Cosmochim. Acta, 47, 1099–1109, 1983.