Surface water dynamics of Africa: Analysing continental trends and identifying drivers for major lakes and reservoirs

Patrick Sogno, Igor Klein, Soner Uereyen, Felix Bachofer & Claudia Kuenzer

➕ View supplementary material ↗

📅 Published online: 18 Oct 2024.

✎ Submit your article to this journal ↗

🔍 View related articles ↗

▣ View Crossmark data ↗

Taylor & Francis
Taylor & Francis Group

Check for updates

Surface water dynamics of Africa: Analysing continental trends and identifying drivers for major lakes and reservoirs

Patrick Sogno [ID][a], Igor Klein [ID][a], Soner Uereyen [ID][a], Felix Bachofer [ID][a] and Claudia Kuenzer [ID][a,b]

aGerman Remote Sensing Data Center, DFD, of the Earth Observation Center, EOC, of the German Aerospace Center, DLR, Oberpfaffenhofen, Germany; bChair of Remote Sensing, Institute for Geography and Geology, University of Wuerzburg, Wuerzburg, Germany

ABSTRACT

Africa has the most dynamic demographic development world-wide. Current projections predict a population of > 3 billion people by the end of the century. Sub-Saharan Africa alone will likely see a 40% increase in population between 2020 and 2050. Although it is well known that large parts of Africa are in a constant state of water stress, its surface water resources remain understudied. This study analyses long-term trends of surface water in Africa. It identifies causal impacts on major lakes and reservoirs for the timeframe 2003–2020, as well as dynamic and causal similarities between the various lakes. For this, a set of daily time series based on Earth observation is employed. Global WaterPack data is used for a daily uninterrupted time series of the continent's surface water area. Additionally, an array of relevant independent variables, namely precipitation, total evapotranspiration, groundwater, soil moisture, and gross primary productivity (GPP) in different land use areas is analysed. For causal identification, the Peter and Clark Momentary Conditional Independence algorithm is used. Findings show that > 42% of African countries and > 34% of African ecoregions experience shrinking surface water area. Over 80% of investigated surface water bodies are driven by the surface water in their upstream subbasins and GPP in agriculturally used areas. About 85% of investigated lakes are significantly driven by agricultural usage, often in the form of water abstraction, as referenced regional studies confirm. Our analysis demonstrates the feasibility of conducting causal analyses of surface water dynamics using Earth observation data. Dynamically similar lakes are often impacted by the same drivers, forming regional lake clusters. Considering the causes identified may greatly help adapt strategies for sustainable development. A causality analysis to identify drivers of surface water dynamics has, to our knowledge, never been performed before on this scale and at such high temporal resolution.

CONTACT Patrick Sogno ✉ patrick.sogno@dlr.de 🖴 German Remote Sensing Data Center, DFD, of the Earth Observation Center, EOC, of the German Aerospace Center, DLR, Muenchener Strasse 20, Oberpfaffenhofen 82234, Germany

Supplemental data for this article can be accessed online at https://doi.org/10.1080/01431161.2024.2412802

1. Introduction

Surface water bodies are an integral part of the terrestrial landscape. They play a role in the hydrosphere, biosphere, and anthroposphere, acting as an above-ground storage of liquid water. Especially in water-scarce regions, the availability of surface water defines its ecological composition and whether humans can live there. However, climate and land use change can challenge the balance of feeding and draining hydrological parameters that define the dynamics of surface water bodies. Ongoing climate change leads to more variable and extreme events (Trisos et al. 2022). This directly affects the inflow of surface water bodies. On the other hand, increasing temperatures may force evapotranspiration and water withdrawal rates to rise (Trisos et al. 2022). Such changes in the natural hydrological dynamic can make traditionally used water resources unreliable or unavailable, which increases water stress (FAO 2011; Yimere and Assefa 2022, Rolle, Tamea, and Claps 2022, Kone et al. 2024). This has a direct impact on the stability of local ecosystems and the livelihood of the local population (Trisos et al. 2022). The African population is projected to surpass 3 billion people by the end of the century, with current demographic trends indicating one of the most dynamic growth trajectories globally (Gu, Andreev, and Dupre 2021), it is essential to have a better understanding of the hydrological processes on the continent (Papa et al. 2023). So far, the renewable water resource potential has not been developed to its full potential (Ahmed et al. 2022; Cofie and Amede 2015; Owusu et al. 2022; Saruchera and Lautze 2019). Especially in areas with scarce or unreliable surface water resources, groundwater may be a reliable alternative to meet water demands. However, access to groundwater is restricted by the investment and infrastructure needed for its extraction. Additionally, in some regions, such as around Lake Chad, the groundwater is too saline to be used for drinking or irrigation (Luxereau, Genthon, and Ambouta Karimou 2012). As a consequence, irrigation often relies on surface water (Yimere and Assefa 2022). Of the total irrigated agricultural area in sub-Saharan Africa, only 6% is fed by groundwater (Ahmed et al. 2022). Figure A1 provides a thematic outline of the manifold impacts on African surface water availability. Due to its short replenishing cycle, surface water can be a sustainable resource as long as it is not overused (Sogno, Klein, and Kuenzer 2022). In light of this, sustainable water planning schemes that go beyond the national scale are paramount (UN Economic Commission for Africa 2000). For this, an accurate understanding of surface water distribution, its dynamics, and its driving factors is critical.

Over the last decades, Earth observation (EO) and in particular remote sensing (RS) has seen increasing popularity for the investigation of hydrological systems on all spatial scales (Sogno, Klein, and Kuenzer 2022). It can provide the observations for long-term monitoring that is essential for informed decision-making, adaptation to climate change, and the reduction of adverse human impact (Phiri et al. 2023). Particularly in Africa, where ground measurements and long time series based on in-situ observations are largely unavailable, satellite RS is a valuable tool for gaining insights into the processes taking place on the ground. Within the thematic context of surface water, Africa-oriented RS studies are often performed for individual lakes or constellations of multiple neighbouring lakes, at a regional, or national scale. The motivation for these studies is often decreasing water availability, unreliable water resources, and increasing pressure on available water resources (Sogno, Klein, and Kuenzer 2022). These studies frequently focus on assessing

how the surface water body (or bodies) changed, identifying drivers of these changes, and evaluating impacts on local communities (Bekele et al. 2018; Deus and Gloaguen 2013). Impacts on surface water are approximated through comparison of land cover at inter-annual scales (e.g. Zekarias et al. 2021), trend comparisons (e.g. Adegbehin et al. 2021; Belete, Diekkrüger, and Roehrig 2016; Ndehedehe et al. 2017), or model-based predictions (e.g., Goshime, Haile, Rientjes, et al. 2021). However, as Uereyen et al. (2022a) show, understanding the influence that climate change and human intervention have on the land surface remains a challenging task that necessitates joint analyses of multivariate time series of all involved spheres. The identification of causal impacts requires high temporal resolution data to be accurate. So far, causal identification approaches in geoscience often rely on bivariate causality analysis frameworks, such as Granger causality (Feldman et al. 2020; Jiang, Liang, and Yuan 2015; Philippon et al. 2005). Due to their complexity, driver analysis studies are still rare. One of the few global investigations is conducted by Yao et al. (2023). Notably, however, this study used multiple linear regression models, which provide limited information on cause-effect relationships and are prone to including spurious and unreliable links (Nowack et al. 2020). In contrast, causality analysis assesses whether there is a directional effect between independent and dependent variables that enhances the predictability of the dependent variable. Yet, within the traditional Granger causality framework, approaches are limited to bivariate analyses (Krich et al. 2020). Multivariate analyses therefore require alternative approaches. Peter and Clark Momentary Conditional Independence (PCMCI) is one of these alternatives and was specifically developed for causal identification in complex systems, such as the ones regularly observed in EO (Runge et al. 2019, 2023). Uereyen et al. (2022b), for example, use this approach to investigate multivariate impacts on land surface dynamics within the Indo-Gangetic river basins.

This study analyses all major standing water bodies in Africa with regard to their dynamics and potential drivers. For this, a multivariate system of directly and indirectly linked climatological and hydrological variables is investigated. Similarities in the dynamic interaction of surface water body areas with investigated independent variables based on cross-correlation analysis are identified. Additionally, causal impacts of subbasin-wide surface water as well as evapotranspiration, precipitation, gross primary productivity (GPP) in natural, irrigated, and unirrigated agricultural areas, soil moisture, and groundwater (each on subbasin and lake scale) on surface water area dynamics are analysed based on the PCMCI algorithm. We consider these variables for the PCMCI analysis because they can have a direct impact on surface water dynamics. The objectives are to 1) identify general trends of surface water in Africa 2) analyse potential drivers of all major lakes and reservoirs on the continent, 3) assess the impact of these drivers on investigated surface water bodies, and 4) identify patterns of dynamic and causal similarity amongst investigated water bodies.

2. Materials & methods

2.1. Study area

Covering the entirety of mainland Africa, Madagascar, as well as São Tomé and Príncipe, the Comoros, and the Canary Islands, the study area covers ca. 29953,000

km². However, due to a general lack of large standing water bodies on the smaller islands, Madagascar is the only island included in this study. For Egypt, only the part west of the Suez Canal is considered in the analysis. Africa spans latitudes between ~ 35.5° North and ~ 35° South, placing it entirely within the subtropical and tropical regions. Elevations are between ca. −100 m.a.s.l. in the Qattara Depression in Egypt and ~ 5,900 m.a.s.l. at the peak of Mount Kilimanjaro. Apart from the mountainous ridge of the Ethiopian Highlands, the East African Rift, and the South African Plateau, most of the continent is at an elevation of < 1,000 m.a.s.l. The orographic features split the continent's surface hydrology into several basins that partially drain into oceans via major rivers (Figure A2 (a)). A considerable portion of surface water drains into endorheic basins, the largest being the Chad basin. The tropical region is mainly comprised of inner-tropical rainforest, monsoon, and savannah climates (Köppen-Geiger classes Af, Am, and Aw, respectively). North of the equator, the transition from tropical to hot desert climate (BWh) via arid hot steppe climate (BSh) is largely homogeneous (Figure A2 (b)). This area is known as the Sahel. At the northern shore of the continent, impacts of the Mediterranean Sea and the west wind drift lead to a wetter and more temperate climate with hot, dry summers (Csa). In the East, the mountainous areas lead to a shift towards cooler climates. In the South, the western area is predominated by hot desert and steppe, while the eastern and especially south-eastern areas are cold steppes and temperate regions.

As visible in Figure A2 (c), population concentration is generally low. Based on the constrained population density estimate for 2020 (Bondarenko et al. 2020), the average population density in Africa is ~ 10 people per km². However, especially within informal settlement structures in urban cores, population density can spike beyond 10,000 people per km². Although the majority of the African population currently lives in rural areas, many areas are rapidly transitioning towards urbanization. It is projected that by the end of the 2030s, the majority of the African population will live in urban areas (Trisos et al. 2022). The population of Africa is concentrated in several regions, especially around the African Great Lakes, in the Ethiopian Highland, along the eastern ridge of the South African Plateau, in the Niger and Volta basins stretching into the Sahel, and in the lower Nile area (Figure A2 (c)). Dense human population visibly overlaps with major lake and reservoir areas and their respective inflowing subbasins (Figure 1). Many lakes are clustered along the East African Rift, along the Niger and Volta, or along the Yobe River, which flows into Lake Chad from the west. Additional clusters exist in eastern South Africa, southern Mozambique, as well as at the Egyptian-Sudanese border. All of these major lakes are of great importance to the local flora and fauna, as well as to the people living there. Most lakes in Africa are exorheic, while only about 27% are endorheic (Figure 1(a)). Endorheic lakes are often brackish or salty, with salinity depending on the seasonal inflow of water (Figure 1(b)). Although such lakes are mostly not used for drinking water, they are still vital for the local population due to the local ecosystems' dependence on these water bodies. Of the investigated major lakes in Africa, ~36% are managed or dammed (Figure 1(c)). These reservoirs are often multi-use. While these dams are most often important for satisfying the water needs of livestock, they are also used for hydroelectricity and irrigation (Zhang and Gu 2023).
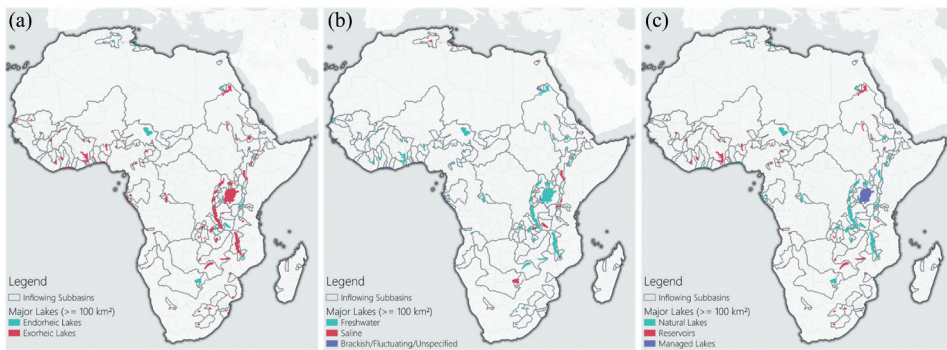
Figure 1. Overview of the major lakes investigated in this study with their respective inflowing subbasins portraying (a) endorheic and exorheic lakes, (b) whether lakes are freshwater, saline, or have a fluctuating salinity, and (c) lake type. Used datasets: basins (Lehner and Grill 2013), lakes (Messager et al. 2016). Background map: OpenStreetMap © OSM contributors.

2.2. Used datasets

As the focus of this study is on the investigation of trends, patterns, and causal relationships that impact surface water dynamics, it is entirely based on datasets that are fully published and have undergone a validation or accuracy assessment. Each utilized dataset is presented below and an overview table of the input data to the analysis is provided in the supplementary material (Table A1). As mentioned in the previous chapter, causal impacts can only be observed up to the temporal scale of the used datasets. We therefore focused on high temporal resolution datasets, although higher spatial resolution alternatives (especially for surface water extent) may be available.

2.2.1. Surface water extent

The DLR Global WaterPack (GWP) is a dynamic surface water extent product based on MODIS Terra and Aqua observations. It covers the timeframe of 2003 to 2023. Within this study, the observed timeframe was from 2003 up to and including 2020. The GWP has a spatial resolution of ~250 m and a temporal resolution of one day. It provides binary information on surface water coverage per pixel (0 – no water, 1 – water) (Klein et al. 2017). The dataset has already been used in various EO studies (e.g. Klein et al. (2021); Uereyen et al. (2022b)) and is available here: https://doi.org/10.15489/vcalr2s1qv66. GWP time series for analysed lakes was used as the dependent variable in this study. To account for the causal impact of upstream surface water on lake area dynamics, the surface water area was spatially aggregated over the entire inflowing subbasin as one of the independent variables.

2.2.2. Surface water level

Lake water level time series based on satellite altimetry were included in the dynamic similarity analysis. We utilized analysis-ready data from two sources: HYDROWEB and DAHITI. Both datasets are calculated from multi-mission altimetry data. HYDROWEB's lake database was presented in detail by (Crétaux et al. 2011). The DAHITI approach and its database were presented by (Schwatke et al. 2015). Both datasets use similar input

data, partially from the same sensors. We utilized both datasets to maximize the number of lakes for which water level time series can be obtained. For lakes where both HYDROWEB and DAHITI have a time series, there was nearly no deviation between the two datasets in our study area. This allowed us to use both datasets in conjunction to increase the temporal resolution. In the dynamic similarity analysis, we considered water level data for all lakes in Africa that are observed for the entirety of our time frame of interest (2003–2020). Table A2 provides an overview of the water bodies with water level data and the dataset it stems from. Both databases are openly accessible (HYDROWEB: https://hydroweb.next.theia-land.fr/.; DAHITI: https://dahiti.dgfi.tum.de/en/products/water-level-altimetry/).

2.2.3. Meteorological variables
ERA5-Land is a dataset developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). It is based on the ERA5 reanalysis but has some improvements that make it more accurate for land applications. The dataset has a spatial resolution of 9 km and a temporal resolution of 1 hour (Muñoz-Sabater et al. 2021). Depending on the parameter, the measurement unit varies. Temperature (T) and dewpoint temperature (D) are measured in kelvin. Potential evaporation (PET) and total evaporation (ET) are measured in metres of water equivalent. PET is open water evaporation (pan evaporation), while ET is all evaporation plus an approximation of transpiration from vegetation. As per ECMWF Integrated Forecasting System convention, downward fluxes are positive, meaning that negative values indicate evaporation and positive values indicate condensation. Surface solar radiation downwards (SSRD) is provided in joules per square metre and characterizes the incoming short-wave solar radiation. Total precipitation (P) is given in metres of water equivalent. It covers liquid and solid water but excludes fog, dew, and precipitation that evaporates in the atmosphere before it arrives at the surface of the Earth. The dataset is available here: https://climate.copernicus.eu/climate-reanalysis.

2.2.4. Vegetation
Gross primary production or gross primary productivity (GPP) is the amount of CO_2 that is assimilated through photosynthesis. The measurement unit is kg C m^{-2}. In this study, GPP was used to approximate vegetation activity and biomass. The dataset used was developed by Joiner and Yoshida (2020). This dataset provides global gridded GPP estimates based on MODIS Terra and Aqua data that was used to globally upscale GPP estimates from FLUXNET 2015 eddy covariance tower sites (Joiner and Yoshida 2020). This dataset has a spatial resolution of ~5 km and a nominal temporal resolution of 1 day and is available here: https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1835. The MODIS product that the GPP product is based on is produced with a rolling 16-day window. It covers the timeframe from March 2000 until and including July 2020, which spans the majority of the study period. For this study, pixels were filled with NA values if the quality flag of the dataset indicated missing data.

2.2.5. Underground water storage
The Global Land Water Storage dataset (GLWS2.0), developed by Gerdener et al. (2023), was used for assessing soil moisture (SM) and groundwater (GW). This dataset is based on an assimilation of monthly GRACE/-FO mass change maps into the WaterGap hydrological

model using the ensemble Kalman filter. In essence, observed total water storage anomaly (TWSA) is distributed into the different water storages based on WaterGap. This process considers uncertainties both in the model and in the observations by attributing corresponding weights in the assimilation equation. The dataset has a spatial resolution of ~50 km and a monthly temporal resolution and is available here: https://doi.pangaea.de/10.1594/PANGAEA.954742. It covers the timeframe 2000–2019, which spans the majority of the study period.

2.2.6. Climate modes

Two indices were used in this analysis that characterize large-scale oscillations associated with extreme weather events, leading to droughts and floods in Africa (Jarugula and McPhaden 2023; Lüdecke et al. 2021): the Indian Ocean Dipole (IOD) index and the North Atlantic Oscillation (NAO) index. The IOD is based on a dipole in sea surface temperature between the Arabian Sea and the eastern Indian Ocean, south of Indonesia (Jarugula and McPhaden 2023). The index is based on satellite altimetry observations from TOPEX/Poseidon and Jason I-III, which measure changes in sea level. The data is available in a weekly format here: https://sealevel.jpl.nasa.gov/data/vital-signs/indian-ocean-dipole. The NAO index is calculated based on a north-south dipole in atmospheric pressure at sea level between the Icelandic subpolar region (typically a low) and the subtropical Azores (typically a high) (Lüdecke et al. 2021). Data on the NAO index is available as a daily time series here: https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml.

2.2.7. HydroLAKES

The HydroLAKES dataset includes standing inland surface water bodies across the globe. In total > 1.4 million polygons of natural and human-made lakes with a minimum surface area of 10 ha are included (Messager et al. 2016). The most recent version of the dataset, which is from 2016, was used in the study. The dataset is available here: https://www.hydrosheds.org/products/hydrolakes. All lakes with a surface water area of >100 km^2 are included in the analysis.

2.2.8. HydroBASINS

The subbasins considered in this analysis are based on the HydroBASINS dataset. The HydroBASINS layers are derived from HydroSHEDS and were developed by Lehner and Grill (2013). The most recent version of this dataset is from 2014 and is available here: https://www.hydrosheds.org/products/hydrobasins. Of the two formats of this dataset available, this study used the standard format, which does not cut out lake areas. The dataset holds subbasins on various levels of aggregation, which are in line with the Pfafstetter coding system. For this analysis, upstream subbasins were aggregated for every lake for a Pfafstetter level of 6.

2.2.9. Agriculturally used areas

To estimate the impact of human intervention on surface water dynamics, we employed two datasets that characterize agriculturally used areas. Huising, Mwangi, and Buyengo (2020) described agriculturally used areas in general (including farmland, plantations and pastures) as part of the Soils4Africa project. Meier, Zabel, and Mauser (2018) showed

irrigated agricultural fields globally. We used both static datasets as masks with the GPP time series introduced earlier to differentiate between natural primary productivity and human-used primary productivity. Here, we differentiated again between non-irrigated and irrigated areas. Both the agriculturally used areas and the irrigated areas dataset are freely available here https://www.soils4africa-h2020.eu/s4a-maps-agricultural-land-in-africa., and here https://doi.pangaea.de/10.1594/PANGAEA.884744, respectively.

2.3. Methodology

This study analyses potential drivers of surface water dynamics in Africa using a causal identification. The main advantage of this over purely correlation-based approaches is that spurious findings are minimized (Nowack et al. 2020). However, several preprocessing steps have to be taken to ensure that the framing assumptions of the causality analysis are not violated. Figure 2 shows the full workflow. The main processing steps are explained in the following subsections.

2.3.1. Preprocessing

Nearly all variables that were included in this analysis are natively in a raster format. Only water level and the indices for NAO and IOD are time series in tabular format. All major African water bodies (i.e. those with an area of >100 km^2) in the HydroLAKES dataset were included in the study. This threshold has been set due to spatial resolution constraints. For convenience, these water bodies will be commonly referred to as lakes from here on out. In total, the lakes that were included in this analysis make up > 90% of the African lake surface area (243,212 km^2 of 265,342 km^2), according to HydroLAKES. Since HydroLAKES is a static dataset from 2016, a buffer of up to 10 km for each lake was applied to account for possible lake growth beyond the HydroLAKES polygon. The buffer creation respects neighbouring basins and buffer growth does not cross into the ocean. Buffers of neighbouring lakes were not allowed to overlap. In instances where buffers overlapped, the contested area was given to the lake with the closer centre point. In areas where a buffer extends into another lake, the contested area is assigned to the lake whose boundary was crossed.

For each lake and its respective subbasin, the investigated variables were aggregated spatially for each day in the observed timeframe. In the case of GPP, we utilized the agriculturally used areas mask and the irrigated croplands mask to differentiate between natural GPP, agricultural GPP at large, and GPP in irrigated areas, specifically. In the following, these three variables will be abbreviated to nGPP, aGPP, and iGPP.

The variables that are originally at a lower temporal resolution were upscaled to a daily resolution using a third-order polynomial interpolation, thereby creating $n = 27$ daily time series for the $n = 15$ included variables. As indicated in Figure 2, from this point forward, all variables (water extent, water level, T, D, PET, ET, SSRD, P, nGPP, aGPP, iGPP, SM, GW, NAO index, and IOD index) were handled together as one 'set' for each of the lakes. Each set contained all raster-based variables on a lake and an upstream subbasin level, lake water level where available, as well as the indices for NAO and IOD. Throughout the analysis, the water extent time series on a lake level was the dependent variable. All independent variables (i.e. all other variables) were
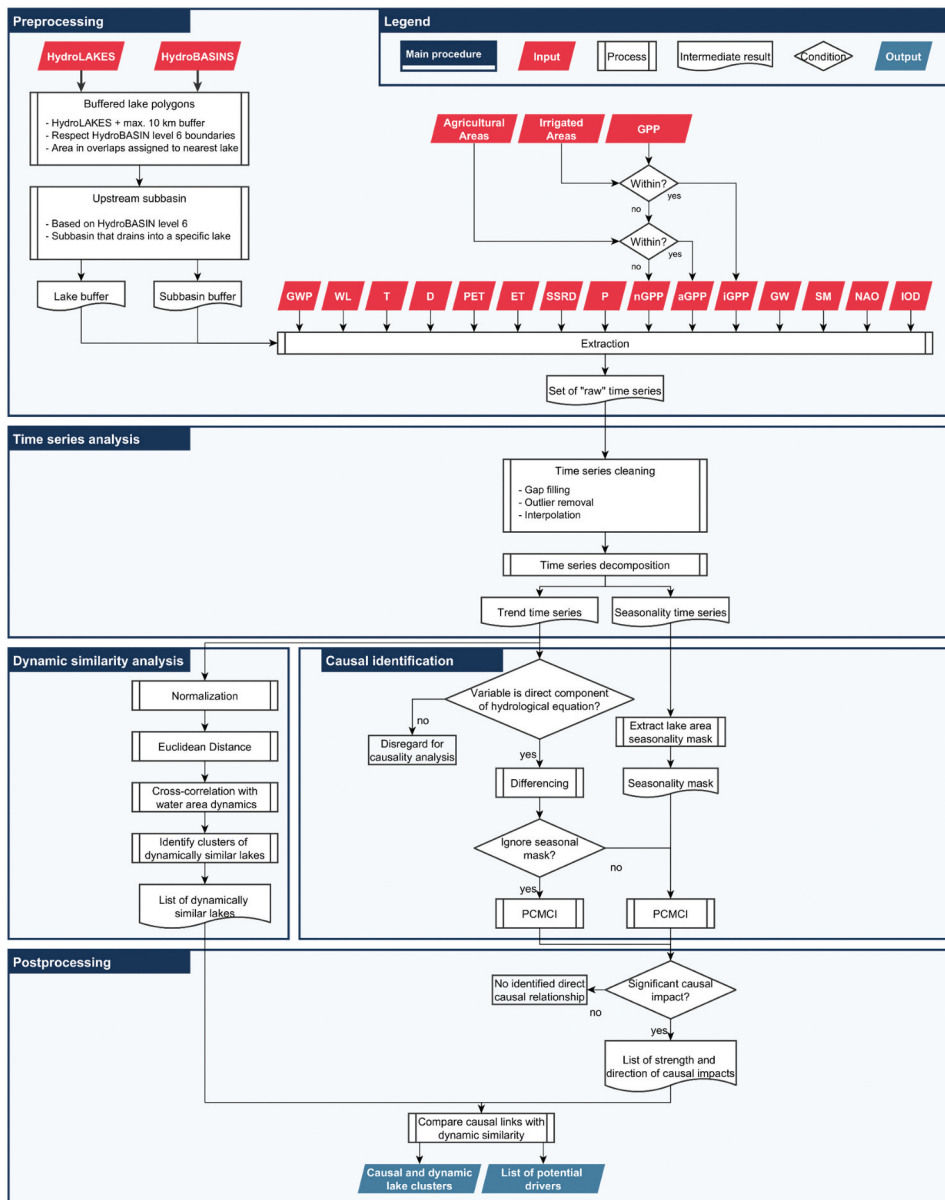
Figure 2. Workflow of the driver analysis. The processing is split into five main steps. From left to right the input variables are: HydroLAKES shapefile and HydroBASINS shapefile, global WaterPack, water level, air temperature, dew point temperature, potential evapotranspiration, total evapotranspiration, solar irradiation downwards, precipitation, natural, agricultural GPP, GPP in irrigated areas, groundwater, soil moisture, North Atlantic Oscillation index, Indian Ocean Dipole index. Note that input variable names are abbreviated in the figure.

used in dynamic similarity analysis to analyse behavioural similarities between the lakes. All independent variables with a direct impact on the lake area (i.e. surface water of the upstream subbasin, ET, P, nGPP, aGPP, iGPP, SM, GW) were considered in the causal identification.

2.3.2. Time series analyses

The time series in each set were gap-free. All time series were resampled to a daily resolution using a third-order polynomial interpolation. In observation-based time series, artefacts may still be present. To address this, outliers were removed using an additive time series decomposition and an interquartile range of 3.0 to classify outlier values. Any gaps that stem from outlier detection were closed using a third-order polynomial interpolation.

These initial processing steps produced sets of 'clean' time series for all investigated lakes (Figure 3(a)). All time series were split into trend, seasonality, and residual using an additive decomposition. The trend and seasonality time series, as well as the original full time series (Figure 3(b)), were used in further processing.

For the linear trend analysis, monthly aggregations of the daily GWP time series were produced for all African countries and ecoregions. For these, the minimum, mean, and maximum water area for each month for all areas of interest were derived. Then, linear regressions were fit for the entire study period (2003–2020) using the mean water area time series. The slopes of these regressions indicate the long-term water area trend. Trend significance is represented by p-values.

2.3.3. Dynamic similarity analysis

To evaluate if some lake extents respond similarly to the studied independent variables, their dynamic similarity was investigated. This analysis was performed on the trend time series only. As shown in Figure 3, the trend time series after decomposition were stripped of noise and seasonality. This allowed us to find similar behaviour (i.e. similar lake extent changes and similar responses to independent variables) in lakes at different times. All trend time series of all sets were normalized using a z-score using the following Equation (1).

$$Z_t = \frac{X_t - \mu}{\sigma} \tag{1}$$
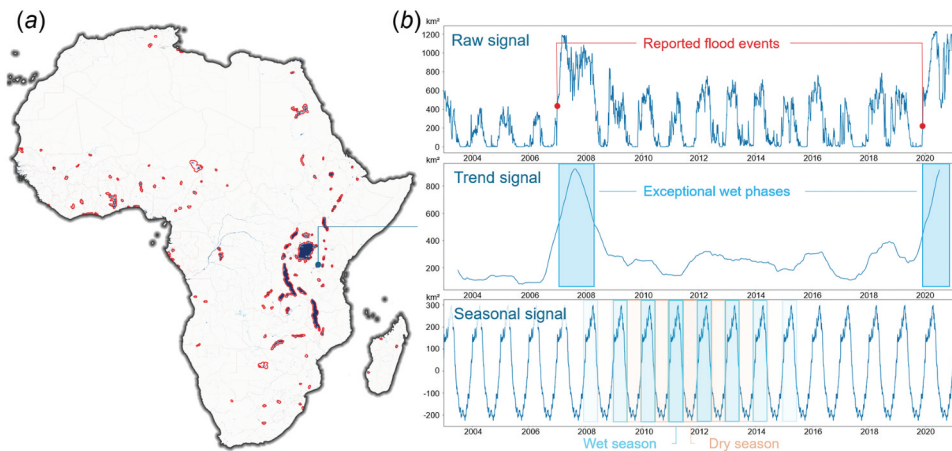
Figure 3. Major lakes in Africa (a) shown with buffered extent in red and example of a time series decomposition (b). An example time series decomposition for lake eyasi, an endorheic salt lake in the East African Rift. Reported flood events: OCHA (2023).

Here, Z_t is the z-score normalized value at time t, X_t is the original value at time t, μ is the mean of the time series, and σ is the standard deviation. The normalized trend time series of the independent variables were compared to the normalized trend time series of the dependent variable via the Euclidean distance between each observation of a first-order differencing of the two curves as shown in the following Equation (2).

$$D_t = \sqrt{((A_t - A_{t-1}) - (B_t - B_{t-1}))} \tag{2}$$

Where the Euclidean distance D at a time t was calculated as the absolute value of the difference between the changes in time series **A** (lake surface water area) and **B** (independent variable), which were expressed through a first-order differencing. By comparing the Euclidean distance for each observation, distance curves were produced for all independent variables for all lakes. Then the cross-correlation between the distance curves for all lakes was calculated as presented in the following Equation (3).

$$R_\tau = \frac{\sum_{t=1}^{n-\tau}(x_t - \bar{x})(y_{t+\tau} - \bar{y})}{\sqrt{\sum_{t=1}^{n-\tau}(x_t - \bar{x})^2 \sum_{t=1}^{n}(y_t - \bar{y})^2}} \tag{3}$$

Where **x** and **y** are distance curves for two lakes that are to be compared to each other, and $\bar{X}$ and $\bar{Y}$ are their respective means. τ is the time lag. To account for similar behaviour that takes place at very different times, $\tau \in [-3650, 3650]$. The cross-correlation R was calculated for each τ. Subsequently, the maximum cross-correlation was determined for all distance curves that were compared between two lakes and it was analysed for how many independent variables the maximum cross-correlation exceeded 0.5. This method assumes that lakes with the same causal links identified in PCMCI analysis and high dynamic similarity are subject to the same process because they respond similarly to the same drivers. The results of the dynamic similarity analysis were further used in postprocessing.

2.3.4. Causal identification

For the causal identification, the decomposed trend signal of the variables that are directly part of the hydrological cycle was used. The following drivers were used: evapotranspiration, precipitation, natural GPP, GPP in irrigated areas, GPP in unirrigated agricultural areas, soil moisture, and groundwater (each for the upstream subbasin and the buffered lake area), as well as surface water area (for the upstream subbasin. The PCMCI algorithm is utilized to identify possible causal impacts of the investigated independent variables on the lake area. This approach was developed by Runge et al. (2019) and consists of two main steps, a PC_1 condition selection and an MCI test. In the first step, potential causal parents are identified. These causal parents are correlations of the independent variables and the dependent variable with a time lag ensuring the anteriority of the independent variable. Iteratively, the potential causal parents are reduced to only a few links for which the partial correlation under consideration of all other independent variables at all considered time lags remains significant. The PC_1 portion of the analysis ultimately results in a convergence, reducing the field to only a few potential causal parents. This set contains actual causal links, as well as potential false positives. Therefore, the MCI step tests for conditional independence, removing false positives.

As explained by Runge et al. (2019), the causal identification framework rests on several assumptions. For this analysis, it was assumed that all common drivers are covered in the observed variables and that all observed conditional independencies stem from the causal graphical structure. Further, it was assumed that no contemporaneous causal effects exist and that the analysed time series are stationary. To ensure accurate outcomes from the causal discovery, most of the auto-correlation within the time series was filtered out by disregarding the seasonality and residual signal. The remaining trend time series was first-order differentiated so that stationarity can be assumed.

The causal impact of the potential drivers is investigated based on the differentiated trend time series. This was done for the entire study period, but also for seasonally explicit subsets that cover the dry season and wet season of a lake. Within the causal identification, these seasonality subsets were enforced using a seasonality mask. This seasonality mask is a temporal mask that covers part of the investigated time series in causal identification using PCMCI. It masks a temporal subset of the dependent variable for which causal links are searched. In this analysis, two seasonality masks were employed: one to only consider periods when a lake's seasonality curve is decreasing (= dry season) and one to only consider periods when a lake's seasonality curve is increasing (= wet season). A maximum time lag of 90 days was allowed. This number was selected as a compromise between considering causal impacts with long lags and the computational effort that comes with identifying potential causal relationships in a multivariate setting. Variables with short lags should not be impacted by the high number of allowed lags, as PCMCI is robust to potential links with long lags that can be explained via links with shorter lags. In the season-specific runs, this time lag may extend into the masked part of the dependent variable's time series. The results of the three causality identification runs undergo postprocessing steps.

Causally related independent variables that have been identified in the PCMCI analysis were considered if they had a high enough significance level ($p < 0.05$). Causal links that cross this threshold were considered relevant, and their causal impact was investigated further. This impact is defined by the strength value of an identified PCMCI connection (Krich et al. 2020). As indicated in Section 2.3.4., it was assumed that lakes with similar causal relationships and high dynamic similarity are subject to the same processes and by extension similar causal impact. In this processing step these causally and dynamically similar links were identified and analysed.

3. Results

The results of the time series dynamics and causality analyses are presented below, starting with the visualization of linear trends on administrative and ecoregion levels and ending with graphics that communicate the dynamic and causal similarity of analysed lakes.

3.1. Linear trends

As illustrated in Figure A2 (*b*), Africa spans Köppen-Geiger climate zones ranging from Af to EF, with various meteorological systems exerting influence over regional weather patterns. Naturally, there are differences in the impacts on and interactions with the

surface water hydrology. Apart from meteorological differences, there are additional differences in soil permeability, biosphere interaction, and human interaction. The trend of surface water area therefore varied for different regions on the continent. Figure 4 visualizes the general long-term trends of surface water availability based on GWP data. Figure 4(a) presents the trends in surface water per country. Trends are shown for the entire investigated timeframe (2003–2020). Of the 54 countries in Africa, 28 had an overall negative trend, while 24 showed an overall positive trend. However, for most countries the significance level of long-term trends was low. In total, significant surface water area trends could be identified in 16 countries. For six of these, trends were negative, while the
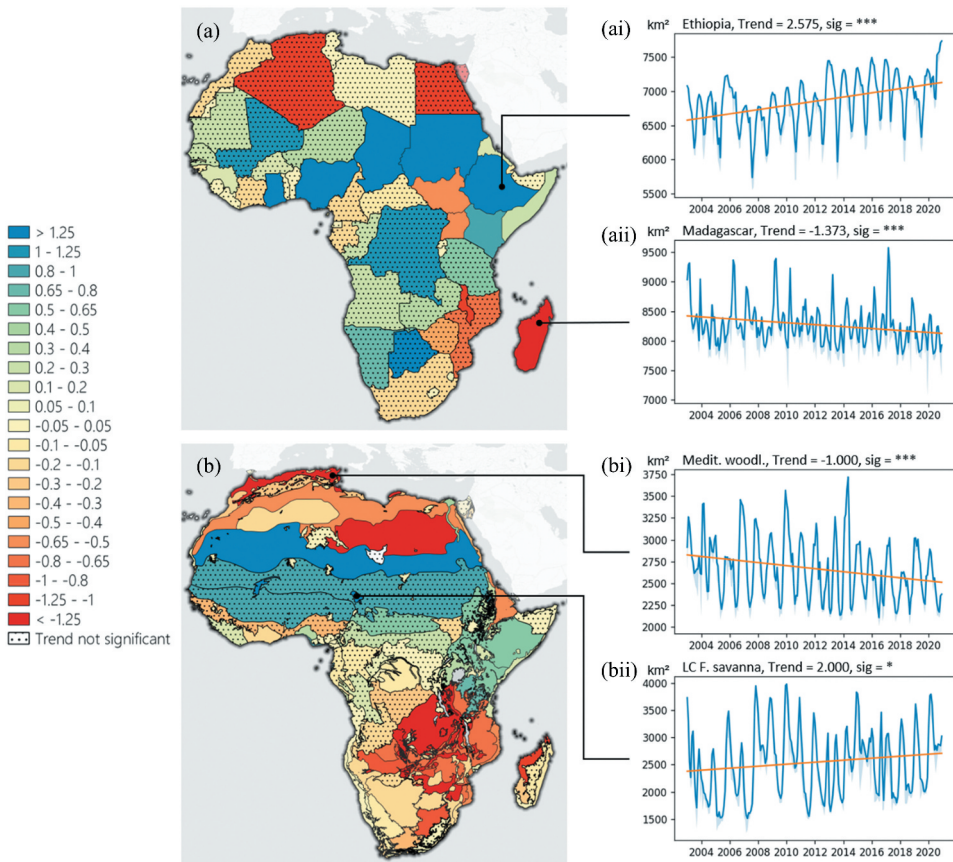
Figure 4. Long-term linear trends (2003–2020) for Africa on a country (a) and ecoregion (b) scale. Areas, where the trend significance was low (p-value > 0.05), are dotted in the maps. Trends are shown for two examples of (a) and (b), respectively: (a1 – Ethiopia, a2 – Madagascar, b1 – Mediterranean woodlands and forests, and b2 – Lake Chad flooded savanna. Here, the significance of the trend is communicated as '*', '**', and '***' corresponding to p-values of ≤ 0.05, ≤ 0.01, and ≤ 0.001, respectively. The blue line in the trend plots shows the respective mean surface water area per month based on daily observations. The blue area in the trend plots covers the values between the minimum and maximum observed water area per month. The orange line is the linear regression. Trends are given in km^2 per month. Ecoregions: Dinerstein et al. (2017).

other 10 exhibited positive trends. Trends in surface water area spanned from -1.373 km^2 per month in Madagascar to $+5.814$ km^2 per month in Ghana.

Generally, countries on the northern coast of Africa all showed decreasing, albeit non-significant, trends ranging from Morocco's -0.112 km^2 per month to Egypt's -1.217 km^2 per month. Countries in the Sahel saw mostly stable to positive trends, the strongest being recorded in Sudan ($+4.449$ km^2 per month). Significant positive trends in surface water area were particularly evident in the eastern Sahel. Conversely, significant negative trends were observed for Madagascar, Malawi, Uganda, South Sudan, Djibouti, and Burundi. In East African countries north of and including Tanzania, surface water was increasing, with the strongest trend visible in Ethiopia (Figure 4(ai)). Countries in Central Africa saw stable water areas or slightly decreasing trends that are not significant. Lastly, most countries in the African Southeast see negative trends, the strongest being in Malawi (-1.165 km^2 per month), and Madagascar (-1.373 km^2 per month). Madagascar was faced with a particularly strongly declining trend (Figure 4(aii)) that indicates a decrease in surface water area of nearly 16.5 km^2 per year.

Ecoregion-scale surface water trends (Figure 4(b)), provide a more nuanced understanding of the surface water area changes on a country scale. In total, 66 of the 132 African ecoregions experienced decreasing surface water area trends. In 45 of those, trends were significant. While increasing surface water area trends could be identified in 45 ecoregions, trends were significant in only 19 of Along the African Northern coast, the Mediterranean Forests, Woodlands & Scrub biomes exhibited significant decreasing surface water trends (Figure 4(bi)) In contrast, the Sahelian biomes, particularly the Sahelian Acacia savanna, the Inner Niger Delta flooded savanna, and the South Sahara desert exhibited strong increasing trends. Slight increases in surface water trends were visible for much of Eastern Africa. For Madagascar, much of the negative trend on a country scale stemmed from decreasing surface water area in the Madagascar dry deciduous forest on the western side of the island. Significant negative surface water developments for much of the Southeast African ecoregions, such as the Central Zambezian wet miombo woodlands, have also become clear. Here, the overall trend was -11.347 km^2 per month, amounting to a water area loss of ca. 136 km^2 per year. In contrast, although with significant inter-annual variability, the surface water trend for the Lake Chad region (Figure 4(bii)) was positive and significant for the investigated time-frame ($+1.541$ km^2 per month).

3.2. *Driver analysis*

While the investigation of linear trends helps visualize large-scale changes, it cannot explain the underlying reasons for these changes. To this end, a driver analysis using the PCMCI algorithm was performed. Causal relationships over the entire investigated timeframe, as well as for temporal subsets that portray the lakes' dry and wet seasons were identified. In the following, the most dominant causal links are analysed for each lake. Additionally, the total causal impact of external drivers is identified. Causal relationships with all independent variables that are directly part of the surface water hydrology could be found. This includes drivers that characterize gain (e.g. precipitation), storage (e.g. soil moisture), and loss (e.g. evapotranspiration) on a subbasin and lake scale. An

overview of lake characteristics, trends, and main drivers is also provided in Table A3 in the supplementary material.

3.2.1. Main causal impact on African lakes and reservoirs

Figure 5 presents the main driver for all major lakes in Africa over the entire observed study period, dry season, and wet season (5(*a*), 5(*b*), and 5(*c*), respectively). Regarding the entire study period, the most frequent dominant causal link was with surface water on a subbasin scale. This link was the most dominant in 67 of the 124 investigated African major lakes. Clear geographical aggregations were not identifiable. The second most frequent main driver was aGPP on a subbasin scale (17 of 124). Again, no clear aggregation to a specific geographic region was visible. GPP was the main driver for 48 lakes (~39% of investigated lakes). This accounts for all types of GPP within the subbasin as well as GPP within the direct lake vicinity. Within the GPP group, aGPP was the dominant driver for most lakes. It was the most dominant driver for 17 lakes at the subbasin level and 13 lakes in direct proximity to the lake. Following aGPP, nGPP was the dominant driver for 9 lakes at the subbasin level and 5 lakes in proximity to the lake, while iGPP was the dominant driver for 2 lakes at the subbasin level and 2 lakes in direct proximity to the lake. The remaining ~ 7% of lakes that were not mainly impacted by upstream surface water or GPP of any kind were mainly impacted by ET within the subbasin ($n = 3$), P in the direct lake vicinity ($n = 2$), GW in the upstream subbasin ($n = 1$), or SM in the direct lake vicinity ($n = 1$). For two surface water bodies, Sebkha Adhilbate/Sebkha Tadet and Mellahet el Brega, the main driving impact over the entire timeframe could not be identified.

During the dry season (Figure 5(b)), ~53% of lakes were impacted mostly by upstream surface water. The second most frequent main driver was nGPP on a subbasin scale ($n = 13$). Without distinguishing between land use or variable impact on subbasin level and lake level, 41 lakes were mostly impacted by GPP. Within the GPP group, aGPP was the dominant driver for most lakes (subbasin: 12, lake: 8), nGPP was the second most frequent dominant driver (subbasin: 13, lake: 4), and iGPP was the least frequent dominant driver (subbasin: 3, lake: 1). Together, upstream surface water and GPP were the most important drivers for ~ 86% of major African lakes during the dry seasons. The remaining lakes were dominantly driven by P on a subbasin scale ($n = 2$), SM on a subbasin scale ($n = 2$), ET on a subbasin scale ($n = 2$), or GW on a subbasin scale ($n = 1$). For 10 lakes, no main driving impact could be identified during the dry seasons. These lakes are almost exclusively concentrated around the equator (between 7° N and 4° S). The only surface water body outside this area, for which no drivers could be found is the Hakskeen Pan in South Africa, a periodically flooded salt pan. There was no apparent spatial aggregation for main impactors.

For the wet season (Figure 5(c)), 54 of the 124 investigated lakes (~44%) were mainly driven by surface water in the upstream subbasin. nGPP in the upstream subbasin was the dominant causal link for 14 lakes and thereby the second most frequent dominant causal link. Without distinguishing between land use, and taking into account both subbasin level and lake level, GPP was the most dominant driver for 51 lakes (~41%). Within the GPP group, aGPP was the most frequent dominant driver (subbasin: $n = 12$, lake: $n = 12$). The second most frequent dominant driver was nGPP (subbasin: $n = 14$, lake: $n = 9$). iGPP was the least frequent dominant driver (subbasin:
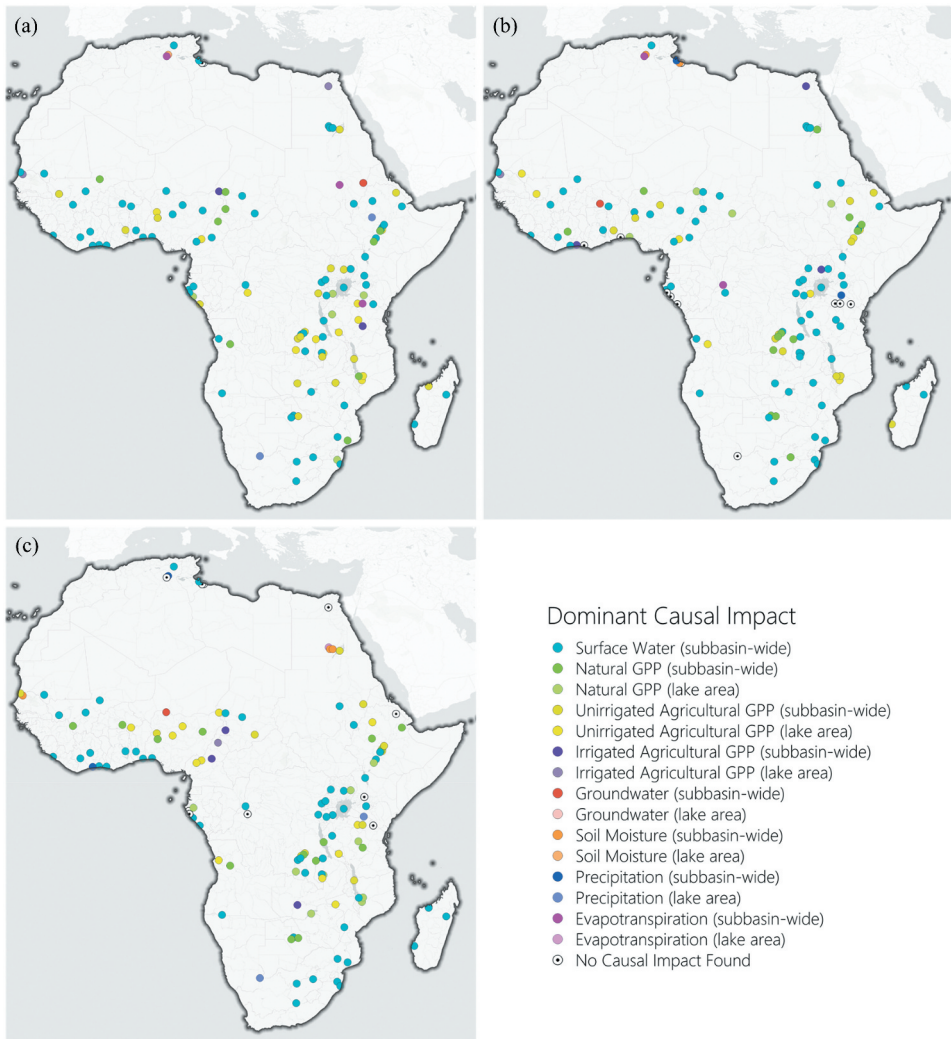
Figure 5. Dominant causal impact on major lakes for (a) the entire time frame; (b) the lakes' dry seasons; and (c) the lakes' wet seasons.

$n = 3$, lake: $n = 1$). In total, ~85% of major African lakes were mainly impacted by upstream surface water or GPP during the wet season. The remaining ~ 15% of lakes were dominantly linked to P (subbasin: $n = 2$, lake: $n = 2$), SM on a subbasin level ($n = 3$), GW on a subbasin level ($n = 1$), or ET on a lake level ($n = 1$). For 10 lakes, the main driver during the wet season could not be identified. These water bodies are concentrated into two groups. One exists close to the equator (between 1° N and 4° S), while the other is confined to hot desert climates but more geographically extensive (Köppen-Geiger class BWh, see Figure A2b). As with the full study period observations and the dry-season-specific run, no clear spatial patterns were recognizable for the most prominent causal impact during the wet season.

Most shifts of the most dominant causal impact do not follow a clear spatial pattern. For ~ 17% of investigated lakes, the most dominant causal impact did not change, whether the entire time series, only the dry season, or only the wet season was considered. Of these lakes, ~91% were always mainly impacted by upstream surface water. Shifts in the main driver between different observed temporal windows often occurred between different GPP land uses in the upstream subbasin, and the lake vicinity. This shift occurred for ~ 73% of lakes within this group.

3.2.2. Identified external impact groups of African lakes

To understand the total causal impact from external impacts, meaning potential drivers other than upstream surface water, the individual causal relationships of investigated independent variables were aggregated into three external impact groups: The atmosphere group (consisting of P and ET), the biosphere group (consisting of nGPP, aGPP, and iGPP), and the soil group (consisting of soil moisture and groundwater). In this regard, external impactors were all independent variables included in the PCMCI analysis apart from surface water on a subbasin level. Going forward, these external impact groups will be referred to as EIGs. As explained in Section 2.3.3, causal impacts can be seen as largely independent from each other, allowing us to aggregate them into EIGs by adding the individual causal impacts together. This masks the causal impact of the individual variables within one EIG but allows us to understand the total impact of that EIG on a surface water body. Overall, a significant causal impact from GPP irrespective of land use was observed for 109 lakes. GPP in agriculturally used areas (aGPP, and iGPP) drove 105 of the investigated lakes. Most of the lakes unaffected by GPP are seasonal salt lakes. Atmospheric parameters had a significant impact on 66 lakes, and GW and SM had a significant impact on 52 lakes. For easier visualization positive causal impacts were aggregated separately from negative causal impacts (Figure 6a, c, e vs. b, d, f), thereby showing external positive and negative driving impacts separately. The colouration of the lakes was based on the normalized impact of the positive/negative driver groups.

Considering the full time series (Figure 6a, b), the positive driving EIGs could be identified for ~ 71% of all investigated lakes (Figure 6(a)). There was no clear spatial pattern to the driving impacts of the individual variable groups. Of the lakes with an identified positive driving EIG, about ~ 55% were driven by biosphere interaction. These lakes are distributed all over the continent with no clear spatial agglomeration. Over 43% of lakes were at least partly driven by biosphere interaction that is subject to human intervention.

For ~39% of lakes, atmospheric interaction was part of the positive driving EIGs. These lakes are dispersed all over the continent. About 71% of these lakes are natural. The remainder are reservoirs or dammed lakes. Of the natural lakes, ~55% are endorheic and – with the exception of Lake Abayata – they all are situated in hot desert ('BWh') and steppe ('BSh') climates. Negative driving impacts were similarly heterogeneous (Figure 6(b)). No clear spatial patterns of individual impact groups were identifiable. Of the lakes for which negative driving impacts could be observed ($n = 86$), ~94% were negatively impacted by biosphere interaction, ~43% were negatively impacted by atmosphere interaction, and ~ 33% were negatively impacted by soil interaction. Human intervention, in the form of GPP in agriculturally used areas (aGPP, and iGPP) negatively impacted 57% of all lakes. Irrigated agriculture negatively impacted 29% of all lakes.
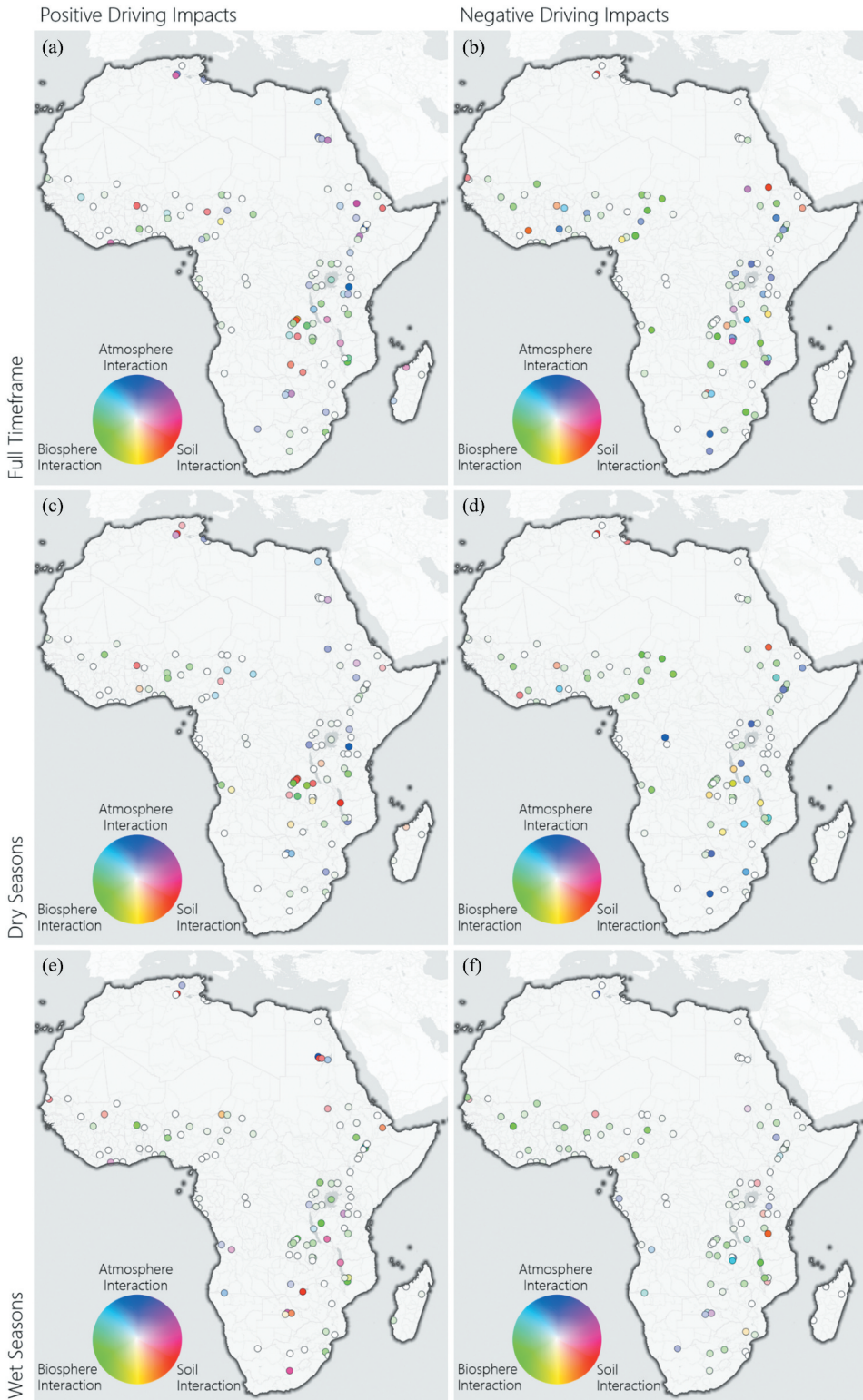
Figure 6. Composition of positive and negative driving impacts on the dynamics of all major lakes over the entire observed time frame (a), (b), the dry seasons (c), (d), and the wet seasons (e), (f).

For the dry season (Figure 6(c)), a pronounced positive driving impact from biosphere interaction was observed for 38 lakes. About 39% of the lakes that were driven by biosphere interaction are reservoirs or dammed lakes. Fewer lakes (21), were positively driven by soil interaction. Lastly, 18 lakes were positively driven by atmospheric interaction. Most of these lakes are concentrated in Eastern Africa Negative driving impacts during the dry season (Figure 6(d)) could be identified for 79 lakes. About 71% of these lakes were negatively driven by biosphere interaction during the dry season. As is visible in Figure 6(d), this negative impact was most prominent in Western and Central Africa. Additionally, 12 lakes were negatively driven by atmosphere interaction, and six were negatively driven by soil interaction.

For the wet season of investigated surface water bodies, biosphere interaction was a prominent positive driver, especially for lakes in the East and Southeast of Africa. In total, 40 lakes were positively driven by biosphere interaction. Soil interaction positively affected 18 lakes throughout the continent, especially seasonal water bodies such as the Ntwetwe Pan and the Sua Pan in Botswana. Lastly, atmospheric interaction was a positive driver in 17 cases throughout the continent (Figure 6(e)). Lake Nasser in Egypt along with the nearby Toshka lakes, Lake Kivu in Rwanda, Lake Tanganyika in the Democratic Republic of the Congo, Guerrah Et Tarf in Algeria, and several reservoirs in the South of the continent were primarily impacted by this driver group. For 81 lakes, a negative driver could be identified that impacts the lakes' wet seasons. 54 lakes were negatively driven by biosphere interaction. In 70% of these cases, a negative causal relationship existed with aGPP on a subbasin level. In ca. 39% of these cases, a negative causal relationship existed with nGPP. To a lesser degree, atmospheric interaction could be a negative driver. It was identified as such for 12 lakes, most of which are situated in the eastern part of the continent. Additionally, soil interaction was a negative driver in some cases. A majority of those are reservoirs, such as Mtera Reservoir in Tanzania, for which this negative driving relationship was especially apparent (Figure 6(f)).

3.2.3. Dynamic similarities and causal similarities

Based on the dynamic similarity of African lakes, they were aggregated into clusters (Figure 7(a)). In total, 21 clusters could be identified. As explained in section 2.3.3, dynamic similarity is defined as an alikeness in which lakes and reservoirs behave in response to any of the independent variables (listed in Table A1). These clusters each consist of two to 14 lakes. Lakes without a dynamically similar counterpart were not assigned to any clusters. Dynamic similarity clusters were often confined to one basin and were defined by spatial proximity. However, there were several remote connections between clusters, such as the connection between the Chad cluster (H) and the Zambezi cluster (R). Several isolated clusters without any remote connections could be identified as well. One such example is the Nasser-Toshka cluster (C). As with the other isolated clusters (Chott cluster (A), Sebkha cluster (B), Senegal cluster (D), Ivory Coast cluster (F), Lagos-Nokoue Cluster (G), Sanaga cluster (I), Gabon cluster (J), Congo cluster (K), Edward cluster (O), and the Salt Pan cluster (S)), the isolation was due to the exceptional surface water hydrology of the lakes included. In general, more isolated clusters existed in Western Africa than in the East. In the western part of the continent, even the two identified major clusters (E, H) were largely independent of each other. Only two remote connections between the Volta-Niger
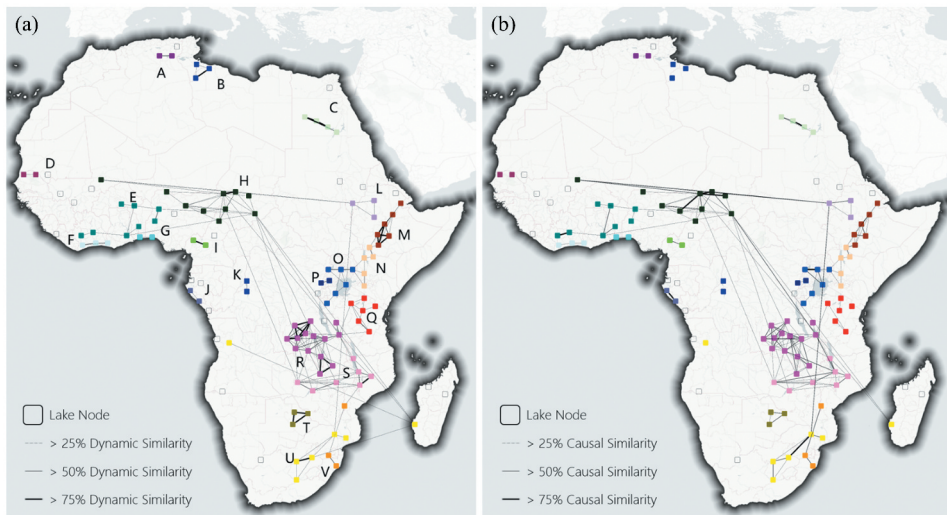
Figure 7. Dynamic similarity (a) and causal similarity (b). Lakes are attributed to clusters based on their dynamic similarity. Identified clusters are color-coded and enumerated with letters A-V. While all variables are considered for dynamic similarity, causal similarity is based only on causally related independent variables. Therefore, causal similarity links show connections between lakes that have similar dynamic interactions with the same causally related variables.

cluster (E) and the Chad cluster (H) could be identified. In contrast, remote connections between the Chad cluster (H) and the Zambezi cluster (R) were much more numerous.

Along the East African Rift, south of the Nile cluster (L), dynamic similarities were more abundant. Still, the degree of similarity varied, which allowed us to identify seven distinct dynamic similarity clusters from the Ethiopian Rift Valley to Mozambique. Specifically, the Ethiopian Rift cluster (M) was identified, extending from Abhe Bid in the Northeast to Lake Shala in the Southwest. Although the lakes Abaya and Chamo are geographically situated within the Ethiopian Main Rift as well, they were identified as part of the Lake Turkana cluster (N) further south along the Gregory Rift. Along the western branch of the East African Rift, the Albertine Rift the Victoria cluster (O) was identified. It consists of Lake Victoria and most of its surrounding lakes that drain into the Nile. West of the Victoria cluster the Edward cluster was identified, which only contains the lakes George and Edward. The Eyasi cluster (Q), southeast of the Victoria cluster, contains Lake Eyasi as well as all endorheic lakes between Lake Natron and Lake Sulunga. The Nyumba ya Mungu and Mteru reservoirs also belong to this cluster. Southwest of the Eyasi cluster is the Mweru cluster (R). This cluster consists of two cores with very high dynamic similarity. The first core consists of the lakes Kabamba, Zimbambo, Kisale, Upemba, and Kabele, which all lie in the Upemba Depression. The second core comprises the lakes Bangweulu, Kampolombo, and Chifunabuli, which all belong to the Bangweulu wetlands. Lastly, along the southern stretches of the East African Rift, the Zambezi (S) cluster could be identified, consisting of the lakes Malawi, Malombe, Chilwa, and Chiuta, the Itezhi-Tezhi Dam, and the Cahora Bassa, and Kariba reservoirs. All of the lakes and reservoirs in the Zambezi cluster drain into the Zambezi or its tributaries. Lastly, adjacent to the previously

mentioned Salt Pan cluster (T), two additional clusters are present in Southern Africa. One is the Southern cluster (U), which contains multiple South African reservoirs, but also the remotely connected Lake Ihotry, in Madagascar and Capanda Dam in Angola. The other one is the Southeastern cluster (V), which is almost exclusively made up of reservoirs.

The dynamic similarity clusters are distinguishable in Figure 7(b), attesting that the similarities also exist in causality and driver impacts. Some of the remote connections are resolved in Figure 7(b). However, within the individual clusters, many lakes shared causal similarities. Such was the case for example in the Chad cluster and the Ethiopian Rift cluster. Within the Mweru cluster, lakes in the Upemba Depression had very high degrees of causal similarity. Since causal similarities occur in instances where similar drivers and similar dynamic reactions of the water area to the investigated independent variables could be identified, these patterns appear to hold a thematic meaning, which is discussed in the following section.

4. Discussion

4.1. Implications of dynamic and causal similarities

As presented, the main causal impact on most lakes was related to upstream surface water. Additionally, several EIGs were identified. The causal impacts of individual EIGs did not adhere to neighbourhoods or basin borders. Of course, every observed lake is unique, but despite this individuality, clusters of dynamic and causal similarity could be identified. In the following, the findings for some of these clusters are interpreted and an assessment of the identified potential drivers is given. Generally, most clusters were impacted by human intervention, foremost water withdrawal for agricultural use, as could be seen from the driving impacts of aGPP and iGPP. This effect was less pronounced for clusters that primarily consist of endorheic salt lakes; here, atmosphere interaction, especially water loss to evapotranspiration, could be the main impactor. Each of the analysed lakes is important for their local ecosystem and population. However, in the context of this paper, the discussion is focused on two exemplary clusters that exist in water-scarce regions with relatively dense populations for which a satisfying amount of scientific literature to compare to this study's findings is available.

4.1.1. Cluster E

Cluster E lies in a relatively densely populated area with much of the unurbanised land being used for agriculture (Bondarenko et al. 2020; Huising, Mwangi, and Buyengo 2020). At the same time, natural water availability from rainfall is deemed unreliable, which poses problems for agriculture. Consequently, crops in the area are watered using various techniques, such as diverting water from streams and rivers to dugouts and reservoirs of varying sizes (Ofosu 2014). To increase the reliability of surface water, several multipurpose dams exist. In fact, all water bodies within the Volta-Niger cluster are reservoirs. Along tributaries to the White Volta, which ultimately drains into Lake Volta, lie Bagré Dam and Lake Kompienga. The reservoir of Bagré Dam was overall mainly driven by aGPP on a lake scale. This impact was negative. As Bagré Dam is a multipurpose dam, one of its uses is for irrigation of local agricultural fields. As such, on a subbasin level, iGPP had a negative causal impact on the lake area trend.

During the wet season, the lake was mainly driven by upstream surface water from the upstream subbasin. This observed impact was positive and hints towards a lake recharge from incoming surface water. Additionally, aGPP and nGPP – both on the lake and the subbasin level – had a positive impact on the lake area. On the other hand, iGPP had a negative impact during this time of lake area increase. Lake Kompienga was positively driven by upstream surface water. On the other hand, aGPP and nGPP on a lake and subbasin scale had a negative impact. Further, evapotranspiration was a negative driver of lake area dynamics. We hypothesize that this lake is impacted by human intervention, just as Bagré Dam, but that the magnitude of this impact in relation to other causal factors, such as upstream surface water and precipitation, is lower than for Bagré Dam. For Lake Volta, Ndehedehe et al. (2017) report that due to upstream human interventions, understanding the natural hydrological variability and the influence of climate variations on the basin's freshwater systems is complicated. They summarize that declining trends in the water level of Lake Volta (e.g. 2011–2015) are consistent with declines in rainfall and net precipitation, and limited stream flow from the Volta River system (Ndehedehe et al. 2017). This aligns with this study's findings, which identified upstream surface water as one of the main driving factors for lake area dynamics over the entire time series. Further, evapotranspiration in the lake's vicinity was found to have a negative causal impact on lake dynamics. This loss of water may have led to a decrease in lake area in phases of decreasing precipitation, as was the case for the phase from 2002 to 2014 (Ndehedehe et al. 2017). Upstream interventions, such as holding back water for agricultural use (Ndehedehe et al. 2017), were also identified in the causality analysis. Specifically, aGPP within the inflowing Volta Basin had a negative causal impact on the lake surface area during the lake's wet season, which coincides with the main growing season. Additionally, iGPP had a negative impact on Lake Volta. This impact was most prominent during the lake's dry season, which may hint towards water abstraction for off-season farming in the upstream basin.

For Nangbeto Dam, which is situated east of Lake Volta at the Mono River, very similar causal impact patterns were visible. Here, upstream surface water was the main driver throughout the time series, but all types of GPP had a significant negative causal impact overall and especially during the lake's wet season (i.e. the main growing season). The negative impact of GPP in agricultural areas, both irrigated and unirrigated, was the most prominent. Studying water balance components of the Mono River Basin, Houteta et al. (2023) draw the same conclusions and attribute the water loss during the phase of May to October (which coincides with the growing season for vegetation in the area) to increased vegetation activity.

For Kainji Reservoir and Jebba Reservoir, which lie along the Niger, the causality analysis unveiled different driver impacts. For both reservoirs, over the entire study period, the main positive driver is GPP. Kainji Reservoir, which drains into Jebba Reservoir, was primarily driven by GPP from agricultural areas – both irrigated and unirrigated. Jebba Reservoir was mostly driven by natural and unirrigated agricultural GPP in the upstream subbasin. For both lakes the causal impact of GPP was complex. Considering the seasonal cycles of GPP and reservoir area, there was a seemingly clear driver-follower relationship that is led by subbasin-wide GPP. The results regarding subbasin-wide aGPP and iGPP impact align with findings provided by Aich et al. (2016) who concluded that land use/land cover changes in the Niger River Basin from natural

vegetation to intensely used croplands and pastures may lead to increased flooding (i.e. increasing water area). During the lakes' dry seasons, the main driver for Kainji Reservoir was upstream surface water. This causal impact was positive, hinting towards an underlying dependency of the lake on the upstream surface water availability. For the lakes' wet seasons and over the entire time series, the impact of upstream surface water was in part positive, and in part negative. Both reservoirs are not only used for irrigation of agricultural areas nearby but also for flood control (Oyerinde et al. 2015). Dam management may explain the unclear causal link to upstream surface water.

In the far West of the cluster, two more reservoirs, Kossou Reservoir and Buyo Reservoir, are situated. Both of them lie within Côte d'Ivoire. Kossou Reservoir was positively driven by upstream surface water and negatively driven by GPP in the inflowing subbasin, irrespective of the land use. That being said, during the lake's wet season, the negative impact of iGPP was far more dominant than that of aGPP and nGPP. Water abstraction and land use change, as presented by (Kouame et al. 2019) could explain this behaviour. Buyo Reservoir was also positively impacted by upstream surface water. Over the entire timeframe, a negative link with iGPP was identified. In addition, both reservoirs seem to be impacted by groundwater infiltration, as a link with GW suggests (Kouame et al. 2019) and (Obahoundje et al. 2022), also proposed a combination of human intervention and groundwater infiltration as potential drivers, at least for Kossou Reservoir.

4.1.2. Cluster M

Lake Ziway was negatively impacted by GPP over the entire observed timeframe. Especially GPP in agricultural land use areas in the lake's immediate vicinity had a negative impact. Goshime et al. (2021a) report that there is extensive water withdrawal from the lake for irrigation, which seems to confirm a strong negative driving impact of iGPP in the direct lake vicinity, for the entire time frame, as well as during the lake's wet season. Jha et al. (2021) state that climate change will likely increase temperatures in the region and decrease precipitation, which may lead to decreasing runoff and lake surface area. The causal impact of this interaction on the surface water trend could not yet be identified for Lake Ziway but may become apparent in the future. For Lake Abayata, however, a strong positive causal connection with precipitation could be identified. This falls in line with the suggestion of Goshime et al. (2021a) that the same issues that affect Lake Ziway, also affect Lake Langano and Lake Abayata. All three lakes are very close to each other and share high dynamic and causal similarities (Figure 7). The characteristic causal relationship with aGPP and iGPP that would hint towards water abstraction for agricultural use could be identified for Lake Langano as well. However, as Goshime et al. (2021b) confirm, the water abstraction from Lake Langano is much smaller compared to Lake Ziway, along with its causal impact. Additionally, the water abstraction from Lake Langano is done by diverting incoming streams. For Lake Abayata, the characteristic causal connections with GPP that would hint at agricultural use could only be conclusively identified on the subbasin scale. As Belete, Diekkrüger, and Roehrig (2016) point out, the water abstraction from Lake Abayata is mostly for industrial use. The continued abstraction of incoming water is expected to lead to severe water scarcity (Goshime, Haile, Rientjes, et al. 2021). Further south in the Ethiopian Main Rift cluster is Lake Shala, which is neighbouring Lake Abayata, and only separated from it by a volcanic caldera rim. Shala is not directly affected by water abstraction, since its waters cannot be used due

to high alkalinity (Goshime, Haile, Rientjes, et al. 2021). This is in accord with the lack of a clear long-term trend for the lake's water area. Apart from a clear positive causal effect of inflowing surface water, a complex, relationship with GPP was identified. Agricultural GPP from unirrigated areas in the inflowing subbasin had an overall negative effect on the lake and is especially inhibiting during the lake's wet season. An impact from irrigated agricultural fields could not be identified. A positive causal connection to natural GPP existed during the dry season and considering the entire timeframe. Literature on Lake Shala in general (Elias et al. 2019), and on the background of this interaction specifically, is sparse. We hypothesize that while the observed positive link with nGPP describes a causal relationship, it may not be a driving relationship. Instead, the subbasin-wide nGPP could be driven by the same underlying factor, the availability of water in the subbasin. More in-depth analyses should be undertaken to clarify whether GPP drives lake surface water area in this example. Further, Mengistu, Demlie, and Abiye (2019) suggest a possible interaction with groundwater. While this could not be identified within our analysis, a possible pathway for this interaction is given by Baumann, Förstner, and Rohde (1975). They state that groundwater reservoirs are recharged by 'meteoric waters', i.e. incoming precipitation and ultimately infiltration of surface water into the groundwater reservoir.

4.2. Limitations and uncertainty

All variables investigated are sourced from validated and analysis-ready datasets. The investigated time series are thus largely reliable. To minimize the impact of potential observation errors even further, outliers were removed and the remaining gaps were filled in all time series. In the driver analysis, PCMCI was employed for causal identification. Thereby, this study overcomes the shortcomings of conventional correlation-based driver analyses. For this, however, the investigated time series had to be preprocessed so that the premises that exist for causal identification are fulfilled: Preprocessing the data was done to such a degree that stationarity of the time series can be assumed. Despite this, it is still possible that some of the analysed time series may not be stationary after preprocessing. PCMCI is comparatively robust to nonstationarity, but spurious links may occur if a causal dependency on a common unknown nonstationarity exists. Seasonality was excluded as far as possible. However, because a standardized time series decomposition technique was used, some degree of autocorrelation may still be present in the cleaned time series. While this remaining autocorrelation may still introduce uncertainty into the identified causal relationship, the PCMCI algorithm is particularly suited for such cases (Nowack et al. 2020). However, the limitations of PCMCI also apply to this study. In particular, detection rates may be impacted negatively in highly deterministic systems. True causal links between two variables may be seen as independent because of a third variable that fully determines either of them. Some algorithms can handle causal identification in such scenarios, but at the cost of higher false positive detection rates (Runge et al. 2019). While the amount of random noise in the investigated time series makes such a scenario unlikely, it is necessary to conduct additional investigations to confirm the causalities that have been found. It is also important to note that all of the independent variables that were studied are also affected by drivers. The causal relationships that drive water dynamics in upstream subbasins, for example, warrant their own analysis. We also

note that all datasets included in this study have their limitations which may affect the obtained results. In the case of GLWS2.0, inaccuracies in the WaterGAP hydrological model, which is being used in the assimilation process of GRACE/-FO may be inherited (Gerdener et al. 2023). Further, the limited spatial and temporal resolution of the soil moisture and groundwater time series may introduce uncertainty. By resampling the time series to a higher resolution, we assume that the unobserved points between two observations can be described through interpolation. Thereby small regional changes or highly dynamic behaviour will be overlooked. That being said, GLWS2.0 already exceeds other GRACE/-FO-based products significantly in its spatial resolution (Gerdener et al. 2023). High temporal resolution may not be as essential for below-ground water, as the reaction time of these factors is normally quite slow (Becker et al. 2022). Most importantly, parts of the impact of human intervention, e.g. in the form of water withdrawal, could not be considered in this analysis. To this date, there is no reliable dataset on this factor for the entire continent of Africa, especially not at a viable temporal resolution for this analysis. The presented analysis seeks to estimate the impact of human intervention by splitting the GPP signal into natural, irrigated, and other agriculturally used areas. Through this, surface water abstraction for irrigation and the impact of non-irrigated agriculture on surface water dynamics may be presented. However, the datasets used for this land-use-based differentiation are static. Increases of agricultural areas, crop rotations, or renaturations therefore remain unobserved. There is potential for future enhancement for example in the consideration of human impact as a dedicated driving variable. Proxies for human land use beyond the distinction of GPP into natural, irrigated, and other agriculturally used areas could prove valuable, though their inclusion warrants careful consideration of the uncertainty that they might introduce to the analysis. Where available, runoff data could be used to accurately show how dependent lakes and reservoirs are on their inflows. It is important to acknowledge that while comparing surface water changes in the upstream subbasin with a target lake may provide some insight into potential relationships, this analysis should be considered a preliminary estimate rather than a definitive conclusion. Further research and analysis are required to fully understand the relationship between the investigated variables. Lastly, it can be argued that the dynamics of water area may not fully represent the changes in water volume of lakes and reservoirs, which is the more relevant term for hydrology and resource planning. The derivation of lake volume based on remote sensing data is possible, as multiple studies have shown for various geographic regions and at various spatial scales. However, such estimations are most often based on static bathymetric data, and changes in water volume are derived from changing lake water area or level (Sogno, Klein, and Kuenzer 2022). Thus, the drivers of lake area or lake height dynamics are likely very similar to the drivers of lake volume dynamics. In this study, the focus was water area time series based on the GWP as it currently is the only available dataset with daily observations for the investigated timeframe for the entire continent.

Despite these limitations, the findings of this analysis are thematically relevant. To the best of our knowledge, continent-wide causal identification and similarity analysis of standing surface water bodies at this temporal extent or resolution has never been conducted before. The implications of these findings for future research and decision-makers are discussed in the following.

4.3. Potential directions for the future

The presented findings reveal that despite the unique nature of each lake shaped by complex hydrological settings and human interventions, there are still identifiable similarities. These similarities are mostly found within the same basin, but may also be in entirely different regions. In complex hydrological scenarios, such as the Chad Basin or the Ethiopian Main Rift cluster, where lakes are heavily impacted by human intervention, it is especially important to understand the causal relationships that drive surface water bodies. The methodology presented for the first time allows for such thematically profound analyses on a continental scale. As shown, the causal relationships found in this analysis generally align with the proposed drivers in literature. However, especially in light of the dynamically evolving situation in Africa with regard to ecology, socio-economic development, and climate change adaption, we stress the need for a transferable and continentally applicable approach for causal analysis. Considering the identified causal impacts on lakes in individual regions can contribute to more sustainable planning of water usage schemes or conservation projects. We specifically see added value for interested parties in international development aiming to support adapted sustainable planning at a continental scale. The developed dynamic and causal similarities bring together the identification of local dynamics and the grouping to larger clusters, relevant in decision-making beyond the national scale. However, there remains an acute need for more research at the regional level to provide the necessary context for large-scale causality analyses such as the one presented.

5. Conclusion

Reliable water access is crucial for sustainable development in Africa. In many regions of the continent, surface water access is already unreliable. Climate change and the predicted strong population increase will likely exacerbate this situation unless sustainable water planning schemes are implemented. To enhance the current understanding of the continent's surface water distribution, dynamics, and driving factors, all major standing water bodies in Africa were analysed regarding their dynamics and potential drivers for the timeframe 2003–2020. This study has shown that for many regions in Africa, clear linear trends are visible that warrant careful consideration of current surface water usage. The results show that:

- Of the 54 African countries, 28 show an overall negative trend in surface water area, while 24 have an overall positive trend. For most countries the significance level of long-term trends is low. Only six countries have significant decreasing trends, while 10 have significant increasing trends.
- Out of the 132 African ecoregions studied, 66 exhibit decreasing trends in surface water area, with 45 of these trends being significant. Increasing surface water area trends could be identified in 45 ecoregions. However, only in 19 of those areas are the trends significant.
- For a majority of investigated water bodies (~54%) a dominant causal link with upstream surface water could be observed for the entire timeframe. Second to this, ~27% of water bodies are most dominantly driven by GPP in agriculturally used areas.

- GPP within a subbasin or on a lake level has a significant impact on almost all lakes (n = 109). This impact can be largely linked to intense agricultural usage and water abstraction, as referenced regional studies confirm. For ~ 85% of investigated lakes, GPP in agriculturally used areas is a significant driver.
- Over the observed timeframe, human intervention, in the form of GPP in agriculturally used areas, negatively impacted 57% of all investigated lakes. Irrigated agriculture negatively impacted 29% of all investigated lakes.
- Our analysis demonstrates the feasibility of conducting causal analyses of surface water dynamics using Earth observation data.
- Clusters of lakes and reservoirs exhibiting similar dynamic interactions and causal connections with the independent variables can be identified. In most cases, these clusters are spatially autocorrelated, although remote connections exist.

Accounting for the drivers of individual lakes may support the development of appropriate usage and conservation strategies. Further, it enables large-scale comparisons, such as the identification of causally similar lakes, which may in turn hold potential for sustainable planning beyond the national scale. For over 20 years it has been well known that this is a necessary next step, now even more so considering the dynamic climatic, demographic, and socioeconomic developments that are taking place in Africa. The threats to water security that arise from human intervention and climate change will likely increase in severity with increasing population pressure in the coming years. Only sustainable planning that considers the complex causal relationships that underlie water availability can circumvent this disastrous scenario.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Patrick Sogno http://orcid.org/0000-0001-7474-4035
Igor Klein http://orcid.org/0000-0003-0113-8637
Soner Uereyen http://orcid.org/0000-0003-3733-0049
Felix Bachofer http://orcid.org/0000-0001-6181-0187
Claudia Kuenzer http://orcid.org/0009-0007-4933-5898

Data availability statement

Data will be made available upon reasonable request.

References

Adegbehin, A. B., E. O. Iguisi, Y. O. Yusuf, and C. K. Dauda. 2021. "Climate Variability and Impact Vulnerability Status of IrrigationWater Resources on Rice and Tomato Production Downstream of Tiga Dam, Nigeria." *FUDMA Journal of Sciences* 5 (1): 126–134. https://doi.org/10.33003/fjs-2021-0501-545.

Ahmed, Z., D. Gui, Z. Qi, and Y. Liu. 2022. "Poverty Reduction Through Water Interventions: A Review of Approaches in Sub-Saharan Africa and South Asia." *Irrigation and Drainage* 71 (3): 539–558. https://doi.org/10.1002/ird.2680.

Aich, V., S. Liersch, T. Vetter, S. Fournet, J. C. M. Andersson, S. Calmanti, F. H. A. van Weert, F. F. Hattermann, and E. N. Paton. 2016. "Flood Projections within the Niger River Basin Under Future Land Use and Climate Change." *Science of the Total Environment* 562:666–677. https://doi.org/10.1016/j.scitotenv.2016.04.021.

Baumann, A., U. Förstner, and R. Rohde. 1975. "Lake Shala: Water Chemistry, Mineralogy and Geochemistry of Sediments in an Ethiopian Rift Lake." *Geologische Rundschau* 64 (1): 593–609. https://doi.org/10.1007/BF01820685.

Becker, B., F. Reichel, D. Bachmann, and R. Schinke. 2022. "High Groundwater Levels: Processes, Consequences, and Management." *WIREs Water* 9 (5): e1605. https://doi.org/10.1002/wat2.1605.

Bekele, B., W. Wu, A. Legesse, H. Temesgen, and E. Yirsaw. 2018. "Socio-Environmental Impacts of Land Use/Land Cover Change in Ethiopian Central Rift Valley Lakes Region, East Africa." *Applied Ecology and Environmental Research* 16 (5): 6607–6632. https://doi.org/10.15666/aeer/1605_66076632.

Belete, M. D., B. Diekkrüger, and J. Roehrig. 2016. "Characterization of Water Level Variability of the Main Ethiopian Rift Valley Lakes." *Hydrology* 3 (1): 1. https://doi.org/10.3390/hydrology3010001.

Bondarenko, M., D. Kerr, A. Sorichetta, and A. Tatem. 2020. "Census/projection-Disaggregated Gridded Population Datasets, Adjusted to Match the Corresponding UNPD 2020 Estimates, for 183 Countries in 2020 Using Built-Settlement Growth Model (BSGM) Outputs." [Internet]. Accessed by October 25, 2023. https://doi.org/10.5258/SOTON/WP00685.

Cofie, O., and T. Amede. 2015. "Water Management for Sustainable Agricultural Intensification and Smallholder Resilience in Sub-Saharan Africa." *Water Resources and Rural Development* 6:3–11. https://doi.org/10.1016/j.wrr.2015.10.001.

Crétaux, J.-F., A. Arsen, S. Calmant, A. Kouraev, V. Vuglinski, M. Bergé-Nguyen, M.-C. Gennero, et al. 2011. "SOLS: A Lake Database to Monitor in the Near Real Time Water Level and Storage Variations from Remote Sensing Data." *Advances in Space Research* 47 (9): 1497–1507. https://doi.org/10.1016/j.asr.2011.01.004.

Deus, D., and R. Gloaguen. 2013. "Remote Sensing Analysis of Lake Dynamics in Semi-Arid Regions: Implication for Water Resource Management. Lake Manyara, East African Rift, Northern Tanzania." *Water* 5 (2): 698–727. https://doi.org/10.3390/w5020698.

Dinerstein, E., D. Olson, A. Joshi, C. Vynne, N. D. Burgess, E. Wikramanayake, N. Hahn, S. Palminteri, P. Hedao, R. Noss, et al. 2017. "An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm." *BioScience* 67 (6): 534–545. https://doi.org/10.1093/biosci/bix014.

Elias, E., W. Seifu, B. Tesfaye, W. Girmay, and M. Tejada Moral. 2019. "Impact of Land Use/Cover Changes on Lake Ecosystem of Ethiopia Central Rift Valley.Tejada Moral M, Editor." *Cogent Food & Agriculture* 5 (1): 1595876. https://doi.org/10.1080/23311932.2019.1595876.

FAO. 2011. *Water for Agriculture and Energy in Africa: The Challenges of Climate Change*. Rome: FAO.

Feldman, A. F., D. J. Short Gianotti, I. F. Trigo, G. D. Salvucci, and D. Entekhabi. 2020. "Land-Atmosphere Drivers of Landscape-Scale Plant Water Content Loss." *Geophysical Research Letters* 47 (22): e2020GL090331. https://doi.org/10.1029/2020GL090331.

Gerdener, H., J. Kusche, K. Schulze, P. Döll, and A. Klos. 2023. "The Global Land Water Storage Data Set Release 2 (GLWS2.0) Derived via Assimilating GRACE and GRACE-FO Data into a Global Hydrological Model." *Journal of Geodesy* 97 (7): 73. https://doi.org/10.1007/s00190-023-01763-9.

Goshime, D. W., A. T. Haile, R. Absi, and B. Ledésert. 2021a. "Impact of Water Resource Development Plan on Water Abstraction and Water Balance of Lake Ziway, Ethiopia." *Sustainable Water Resource Management* 7 (3): 36. https://doi.org/10.1007/s40899-021-00516-w.

Goshime, D. W., A. T. Haile, T. Rientjes, R. Absi, B. Ledésert, and T. Siegfried. 2021b. "Implications of Water Abstraction on the Interconnected Central Rift Valley Lakes Sub-Basin of Ethiopia Using WEAP." *Journal of Hydrology: Regional Studies* 38:100969. https://doi.org/10.1016/j.ejrh.2021.100969.

Gu, D., K. Andreev, and M. E. Dupre. 2021. "Major Trends in Population Growth Around the World." *China CDC Weekly* 3 (28): 604–613. https://doi.org/10.46234/ccdcw2021.160.

Houteta, D. K., K. Atchonouglo, J. G. Adounkpe, B. Diwediga, Y. Lombo, K. E. Kpemoua, and K. Agboka. 2023. "Changes in Land Use/Cover and Water Balance Components During 1964–2010 Period in the Mono River Basin, Togo-Benin." *Geography, Environment, Sustainability* 15 (4): 171–180. https://doi.org/10.24057/2071-9388-2021-098.

Huising, J., E. Mwangi, and J. Buyengo. 2020. *D4.1: Map of Agricultural Land of Continental Africa*. Soils4Africa_D4.1_v01. Wageningen, The Netherlands: Soils4Africa. Supported by the Horizon 2020 programme of the EU. Project no. 869200.

Jarugula, S., and M. J. McPhaden. 2023. "Indian Ocean Dipole Affects Eastern Tropical Atlantic Salinity Through Congo River Basin Hydrology." *Communications Earth & Environment* 4 (1): 1–9. https://doi.org/10.1038/s43247-023-01027-6.

Jha, R., V. P. Singh, V. Singh, and L. B. Roy, Thendiyath R, editors. 2021. *Climate Change Impacts on Water Resources: Hydraulics, Water Resources and Coastal Engineering*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-64202-0.

Jiang, B., S. Liang, and W. Yuan. 2015. "Observational Evidence for Impacts of Vegetation Change on Local Surface Climate Over Northern China Using the Granger Causality Test." *Journal of Geophysical Research Biogeosciences* 120 (1): 1–12. https://doi.org/10.1002/2014JG002741.

Joiner, J., and Y. Yoshida. 2020. "Satellite-Based Reflectances Capture Large Fraction of Variability in Global Gross Primary Production (GPP) at Weekly Time Scales." *Agricultural and Forest Meteorology* 291:108092. https://doi.org/10.1016/j.agrformet.2020.108092.

Klein, I., U. Gessner, A. J. Dietz, and C. Kuenzer. 2017. "Global WaterPack - a 250 M Resolution Dataset Revealing the Daily Dynamics of Global Inland Water Bodies." *Remote Sensing of Environment* 198:345–362. https://doi.org/10.1016/j.rse.2017.06.045.

Klein, I., S. Mayr, U. Gessner, A. Hirner, and C. Kuenzer. 2021. "Water and Hydropower Reservoirs: High Temporal Resolution Time Series Derived from MODIS Data to Characterize Seasonality and Variability." *Remote Sensing of Environment* 253:112207. https://doi.org/10.1016/j.rse.2020.112207.

Kone, S., A. Balde, P. Zahonogo, and S. Sanfo. 2024. "A Systematic Review of Recent Estimations of Climate Change Impact on Agriculture and Adaptation Strategies Perspectives in Africa." *Mitig Adapt Strateg Glob Change* 29 (2): 18. https://doi.org/10.1007/s11027-024-10115-7.

Kouame, Y. M., S. Obahoundje, A. Diedhiou, B. François, E. Amoussou, S. Anquetin, R. S. Didi, et al. 2019. "Climate, Land Use and Land Cover Changes in the Bandama Basin (Côte D'Ivoire, West Africa) and Incidences on Hydropower Production of the Kossou Dam." *The Land* 8 (7): 103. https://doi.org/10.3390/land8070103.

Krich, C., J. Runge, D. G. Miralles, M. Migliavacca, O. Perez-Priego, T. El-Madany, A. Carrara, and M. D. Mahecha. 2020. "Estimating Causal Networks in Biosphere–Atmosphere Interaction with the PCMCI Approach." *Biogeosciences* 17 (4): 1033–1061. https://doi.org/10.5194/bg-17-1033-2020.

Lehner, B., and G. Grill. 2013. "Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the world's Large River Systems." *Hydrological Processes* 27 (15): 2171–2186. https://doi.org/10.1002/hyp.9740.

Lüdecke, H.-J., G. Müller-Plath, M. G. Wallace, and S. Lüning. 2021. "Decadal and Multidecadal Natural Variability of African Rainfall." *Journal of Hydrology: Regional Studies* 34:100795. https://doi.org/10.1016/j.ejrh.2021.100795.

Luxereau, A., P. Genthon, and J.-M. Ambouta Karimou. 2012. "Fluctuations in the Size of Lake Chad: Consequences on the Livelihoods of the Riverain Peoples in Eastern Niger." *Regional Environmental Change* 12 (3): 507–521. https://doi.org/10.1007/s10113-011-0267-0.

Meier, J., F. Zabel, and W. Mauser. 2018. "A Global Approach to Estimate Irrigated Areas – a Comparison Between Different Data and Statistics." *Hydrology and Earth System Sciences* 22 (2): 1119–1133. https://doi.org/10.5194/hess-22-1119-2018.

Mengistu, H. A., M. B. Demlie, and T. A. Abiye. 2019. "Revisão: Potencial dos recursos hídricos subterrâneos e status do desenvolvimento dos recursos hídricos subterrâneos na Etiópia." *Hydrogeology Journal* 27 (3): 1051–1065. https://doi.org/10.1007/s10040-019-01928-x.

Messager, M. L., B. Lehner, G. Grill, I. Nedeva, and O. Schmitt. 2016. "Estimating the Volume and Age of Water Stored in Global Lakes Using a Geo-Statistical Approach." *Nature Communications* 7 (1): 13603. https://doi.org/10.1038/ncomms13603.

Muñoz-Sabater, J., E. Dutra, A. Agustí-Panareda, C. Albergel, G. Arduini, G. Balsamo, S. Boussetta, M. Choulga, S. Harrigan, H. Hersbach, et al. 2021. "ERA5-Land: A State-Of-The-Art Global Reanalysis Dataset for Land Applications." *Earth System Science Data* 13 (9): 4349–4383. https://doi.org/10.5194/essd-13-4349-2021.

Ndehedehe, C. E., J. L. Awange, M. Kuhn, N. O. Agutu, and Y. Fukuda. 2017. "Analysis of Hydrological Variability Over the Volta River Basin Using in-Situ Data and Satellite Observations." *Journal of Hydrology: Regional Studies* 12:88–110. https://doi.org/10.1016/j.ejrh.2017.04.005.

Nowack, P., J. Runge, V. Eyring, and J. D. Haigh. 2020. "Causal Networks for Climate Model Evaluation and Constrained Projections." *Nature Communications* 11 (1): 1415. https://doi.org/10.1038/s41467-020-15195-y.

Obahoundje, S., A. Diedhiou, K. L. Kouassi, M. Y. Ta, E. M. Mortey, P. Roudier, and D. G. M. Kouame. 2022. "Analysis of Hydroclimatic Trends and Variability and Their Impacts on Hydropower Generation in Two River Basins in Côte d'Ivoire (West Africa) During 1981–2017." *Environmental Research Communications* 4 (6): 065001. https://doi.org/10.1088/2515-7620/ac71fa.

OCHA. 2023. *Disasters | ReliefWeb* [Internet]. Accessed by November 22, 2023. https://reliefweb.int/disasters.

Ofosu, E. A. 2014. *Flood Based Irrigation in the White Volta Sub Basin: Status and Potential* 14. Wageningen, The Netherlands: Spate Irrigation Network.

Owusu, S., O. Cofie, M. Mul, and J. Barron. 2022. "The Significance of Small Reservoirs in Sustaining Agricultural Landscapes in Dry Areas of West Africa: A Review." *Water* 14 (9): 1440. https://doi.org/10.3390/w14091440.

Oyerinde, G. T., F. C. C. Hountondji, D. Wisser, B. Diekkrüger, A. E. Lawin, A. J. Odofin, and A. Afouda. 2015. "Hydro-Climatic Changes in the Niger Basin and Consistency of Local Perceptions." *Regional Environmental Change* 15 (8): 1627–1637. https://doi.org/10.1007/s10113-014-0716-7.

Papa, F., J.-F. Crétaux, M. Grippa, E. Robert, M. Trigg, R. M. Tshimanga, B. Kitambo, A. Paris, A. Carr, A. S. Fleischmann, and M.de Fleury. 2023. "Water Resources in Africa Under Global Change: Monitoring Surface Waters from Space." *Surveys in Geophysics* 44 (1): 43–93. https://doi.org/10.1007/s10712-022-09700-9.

Philippon, N., E. Mougin, L. Jarlan, and P.-L. Frison. 2005. "Analysis of the Linkages Between Rainfall and Land Surface Conditions in the West African Monsoon Through CMAP, ERS-WSC, and NOAA-AVHRR Data." *Journal of Geophysical Research Atmospheres* 110 (D24): 110(D24. https://doi.org/10.1029/2005JD006394.

Phiri, H., D. Mushagalusa, C. Katongo, C. Sibomana, M. Z. Ajode, N. Muderhwa, S. Smith, G. Ntakimazi, E. L. R. De Keyzer, D. Nahimana, et al. 2023. "Lake Tanganyika: Status, Challenges, and Opportunities for Research Collaborations." *Journal of Great Lakes Research* 49 (6): S0380133023001946. https://doi.org/10.1016/j.jglr.2023.07.009.

Rolle, M., S. Tamea, and P. Claps. 2022. "Climate-Driven Trends in Agricultural Water Requirement: An ERA5-Based Assessment at Daily Scale Over 50 Years." *Environmental Research Letters* 17 (4): 044017. https://doi.org/10.1088/1748-9326/ac57e4.

Runge, J., A. Gerhardus, G. Varando, V. Eyring, and G. Camps-Valls. 2023. "Causal Inference for Time Series." *Nature Reviews Earth and Environment* 4 (7): 487–505. https://doi.org/10.1038/s43017-023-00431-y.

Runge, J., P. Nowack, M. Kretschmer, S. Flaxman, and D. Sejdinovic. 2019. "Detecting and Quantifying Causal Associations in Large Nonlinear Time Series Datasets." *Science Advances* 5 (11): eaau4996. https://doi.org/10.1126/sciadv.aau4996.

Saruchera, D., and J. Lautze. 2019. *Small Reservoirs in Africa: A Review and Synthesis to Strengthen Future Investment*. Colombo, Sri Lanka: International Water Management Institute (IWMI). https://doi.org/10.5337/2019.209.

Schwatke, C., D. Dettmering, W. Bosch, and F. Seitz. 2015. "DAHITI – an Innovative Approach for Estimating Water Level Time Series Over Inland Waters Using Multi-Mission Satellite Altimetry." *Hydrology and Earth System Sciences* 19 (10): 4345–4364. https://doi.org/10.5194/hess-19-4345-2015.

Sogno, P., I. Klein, and C. Kuenzer. 2022. "Remote Sensing of Surface Water Dynamics in the Context of Global Change—A Review." *Remote Sensing* 14 (10): 2475. https://doi.org/10.3390/rs14102475.

Trisos, C. H., I. O. Adelekan, E. Totin, A. Ayanlade, J. Efitre, A. Gemeda, K. Kalaba, C. Lennard, C. Masao, Y. Mgaya. 2022. "Africa". In *Climate Change 2022: Impacts, Adaptation and Vulnerability*, edited by H. -O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, and B. Rama, 1285–1455. Cambridge, UK and New York, NY, USA: Cambridge University Press. https://doi.org/10.1017/9781009325844.011.

Uereyen, S., F. Bachofer, I. Klein, and C. Kuenzer. 2022a. "Multi-Faceted Analyses of Seasonal Trends and Drivers of Land Surface Variables in Indo-Gangetic River Basins." *Science of the Total Environment* 847:157515. https://doi.org/10.1016/j.scitotenv.2022.157515.

Uereyen, S., F. Bachofer, and C. Kuenzer. 2022b. "A Framework for Multivariate Analysis of Land Surface Dynamics and Driving Variables-A Case Study for Indo-Gangetic River Basins." *Remote Sensing* 14 (1): 24. https://doi.org/10.3390/rs14010197.

UN Economic Commission for Africa. 2000. *The Africa Water Vision for 2025: Equitable and Sustainable Use of Water for Socioeconomic Development*. Addis Ababa, Ethiopia: United Nations Economic Commission for Africa (ECA).

Yao, F., B. Livneh, B. Rajagopalan, J. Wang, J.-F. Crétaux, Y. Wada, and M. Berge-Nguyen. 2023. "Satellites Reveal Widespread Decline in Global Lake Water Storage." *Science* 380 (6646): 743–749. https://doi.org/10.1126/science.abo2812.

Yimere, A., and E. Assefa. 2022. "Current and Future Irrigation Water Requirement and Potential in the Abbay River Basin, Ethiopia." *Air, Soil and Water Research* 15:11786221221097929. https://doi.org/10.1177/11786221221097929.

Zekarias, T., V. Govindu, Y. Kebede, and A. Gelaw. 2021. "Geospatial Analysis of Wetland Dynamics on Lake Abaya-Chamo, the Main Rift Valley of Ethiopia." *Heliyon* 7 (9): e07943. https://doi.org/10.1016/j.heliyon.2021.e07943.

Zhang, A. T., and V. X. Gu. 2023. "Global Dam Tracker: A Database of More Than 35,000 Dams with Location, Catchment, and Attribute Information." *Scientific Data* 10 (1): 111. https://doi.org/10.1038/s41597-023-02008-2.