A climatological baseline for understanding patterns of seasonal lake dynamics across sub-Sahelian Africa

communications earth & environment

Article

A Nature Portfolio journal

https://doi.org/10.1038/s43247-025-02684-5

A climatological baseline for understanding patterns of seasonal lake dynamics across sub-Sahelian Africa

M. Amadori[1] ✉, A. J. Greife[1,2], L. Carrea[3], M. Pinardi[1], R. Caroni[1], E. Calamita[4,5], L. Serrao[6], R. Maidment[3], S. Bordoni[6], C. Giardino[1], M. Bresciani[1], F. P. Fava[2], M. Schmid[7], M. Ndebele-Murisa[8], T. Nhiwatiwa[9], J-F. Crétaux[10], C. J. Merchant[3,11], X. Liu[12], S. Simis[12], D. Lomeo[12,13], H. Yesou[14], C. Albergel[15] & R. I. Woolway[16]

Lakes in sub-Sahelian Africa are facing growing ecological threats from climate change and human, yet most research has focused on a handful of well-known large lakes. This study analyses 137 lakes, many previously understudied, and identifies consistent seasonal co-variability patterns across meteorological variables, satellite-derived lake physical and biogeochemical variables, and morphological and anthropogenic characteristics. We identify four distinct clusters of lakes, shaped by the atmospheric variability and its synchrony with water temperature seasonality. Within each cluster, we observe three seasonal patterns of chlorophyll-a concentration tied to wet and dry seasons. These patterns align with regional climatic threats in Africa, such as shifts in rainfall seasonality, altered frequency of tropical cyclones and wildfires, thus positioning our study as a framework to assess lake vulnerability across the sub-Sahelian region.

Inland water bodies are complex components of the hydrological, energy and carbon cycles[1] and respond to climate and human activities[2,3]. Natural lakes and artificial reservoirs, further referred to as "lakes", stand out as critical hotspots of biodiversity[4] and provide a diverse range of ecosystem services[5,6] to human communities[7]. This is particularly true for the sub-Sahelian region, where lakes are vital for supporting the livelihoods of local communities[8–10]. However, various processes compromise both the quantity and quality of the lake's waters[11,12] and threaten the long-standing and natural connection between humans and the water environment they rely on. According to the latest Intergovernmental Panel on Climate Change (IPCC) report[13], five out of nine key risks for Africa are related to freshwater resource deterioration, increasing demand and sensitivity to climate-related extreme events. Fast-growing human settlements in the proximity of lakes have resulted in an increase in water demand for a variety of activities, including industrial, agricultural, and energy-related uses[14], which eventually widens the disparity in water access among different social groups[15]. At the same time, a lack of adequate wastewater treatment infrastructure[16,17] and subsequent uncontrolled wastewater inflows have caused a notable decline in water quality in many lakes[18], endangering the food security and health of millions of people[19]. In addition to the direct impact of human populations, sub-Sahelian lakes are threatened by climate change[20,21]. Projections by the IPCC indicate that rising temperatures and altered precipitation patterns will affect lake ecosystems and, consequently, all water-dependent sectors. This is already evidenced by climate-induced ecological transitions, including reduced fish abundance and diversity in many lakes in this region[22,23]. Another example is presented by lake surface warming and increased thermal stability[24] impacting nutrient and oxygen availability[25] and, consequently, phytoplankton dynamics[26]. Tropical lakes might be particularly vulnerable, with the forced signal expected to emerge from natural variability at a faster rate than lakes in other climatic regions[27].

[1]Institute for Electromagnetic Sensing of the Environment, National Research Council, Milan, Italy. [2]Department of Environmental Science and Policy, Università degli Studi di Milano, Milan, Italy. [3]University of Reading, Meteorology Department, Reading, UK. [4]Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland. [5]Department of Geosciences, Eberhard Karls University of Tübingen, Tübingen, Germany. [6]Department of Civil, Environmental and Mechanical Engineering (DICAM), University of Trento, Trento, Italy. [7]Eawag, Swiss Federal Institute of Aquatic Science and Technology, Surface Waters - Research and Management, Kastanienbaum, Switzerland. [8]START International, Harare, Zimbabwe. [9]University of Namibia, Dept. of Fisheries and Ocean Sciences, Henties Bay, Namibia. [10]Laboratoire d'Études en Géophysique et Océanographie Spatiales (LEGOS), Université de Toulouse, CNES-IRD-CNRS-UT3, Toulouse, France. [11]National Centre for Earth Observation, Reading, UK. [12]Plymouth Marine Laboratory, Prospect Place, PL1 3DH, Plymouth, UK. [13]Department of Geography, King's College London, London, UK. [14]ICUBE-SERTIT, Université de Strasbourg, Strasbourg, France. [15]European Space Agency Climate Office, ECSAT, Harwell Campus, Didcot, UK. [16]School of Ocean Sciences, Bangor University, Menai Bridge, Wales, UK. ✉e-mail: amadori.m@irea.cnr.it

These challenges are further worsened by a severe disparity in research funding. Less than 4% of the global budget for climate change research has been allocated to Africa, with 78% going to institutions in North America and Europe, versus a mere 14.5% for African institutions[28]. Moreover, according to the IPCC AR6[29], climate finance commitments in Africa disproportionately support mitigation (61.4%) over adaptation (33.1%), with few exceptions such as Chad, Sudan and Gabon, where adaptation is targeted more strongly.

For most of the last century, traditional limnological research in Africa has heavily relied on fieldwork, with a relatively small fraction of lakes subjected to extensive, in-depth investigations (e.g., Lake Victoria and other African Great Lakes[30,31]). Aside from peer-reviewed science, technical reports produced by local governmental and non-governmental authorities provide a significant source of knowledge, which, however, is frequently inaccessible and limited in terms of both temporal and spatial coverage[32].

Recent work has demonstrated the potential for upscaling limnological investigations from single-lake to global scale by exploiting the synoptic-scale coverage and long-term availability of remote sensing products[33,34]. Continental-scale monitoring of water quantity (i.e., lake level, volume, surface extent) in Africa has already yielded relevant insights on current trends in water availability[35], especially in relation to clusters of rainfall patterns[36]. However, water quality monitoring has so far only been approached at the scale of individual lakes[37] or across relatively small regions[38], where the use of remote sensing has proven vital for prioritising freshwater systems for conservation[8]. This represents a major constraint for understanding emerging risks and implementing effective response strategies.

Identifying the precise controls of lake behaviour in terms of water quality and quantity is one of the most challenging and urgent issues in water-related research. In most cases, the answer does not lie in a single factor, but in an array of conditions or events favourably nested in either a short or long temporal chain[39]. If the level of complexity is reduced, lakes can be used as sentinels of climate change[40] and environmental risk. It then becomes crucial to gain a synoptic understanding of how anthropogenic pressures and climate drivers affect lakes' biophysical and biochemical properties and whether emerging relationships exist among geographical regions or clusters of lakes. In this study, we provide the first synoptic assessment of 137 natural lakes and reservoirs in the sub-Sahelian region. Given the relevance of reservoirs to sub-Sahelian water resources, our study includes them in the analysis and compares seasonal patterns across similar climatic zones, exploring the role of morphology and catchment size. We utilise satellite-derived lake variables from the global "lakes_cci" database provided by the European Space Agency (ESA) under the Climate Change Initiative framework (CCI)[34]. We integrate the CCI dataset on lake water quality and quantity with multiple datasets on key atmospheric variables and proxies of anthropogenic pressures such as land use, land cover and population density. We classify lakes based on the seasonal co-variability between atmospheric drivers and lake responses and deliver analogy maps along with a catalogue of per-lake results. By providing a broader understanding of the typical seasonal synchrony between lake variables and external forcing, we facilitate the identification of otherwise undetected patterns. This knowledge is critical to guide future policy planning and operational monitoring efforts, while at the same time providing data that will support future sub-continental scale research.

Results and discussion

Lakes of sub-Sahelian Africa

The 137 lakes in our sub-Sahelian sample examined in this study exhibit diverse morphological characteristics (Fig. 1a–c). The majority (75%) are under 500 km² in size, with 32% smaller than 50 km². The remaining 25% range from 500 to 67,000 km², with 16% between 500–3000 km², such as lakes Natron and Kivu. Only the largest lakes, Malawi, Tanganyika, and Victoria, exceed 10,000 km². Most large lakes are located between 5°N and 5°S, within the East African Rift Valley, except for West African reservoirs like Lake Volta and Lake Kainji. Among the largest, lakes Malawi and Tanganyika are also the deepest, with average depths of 261 m and 577 m, respectively. Shallow lakes (0.1–7 m, e.g., Lake Natron) make up 45%, medium-depth lakes (7–15 m, e.g., Lake Tana) 30%, and deep lakes (>15 m) 25%, with only lakes Tanganyika, Malawi, and Kivu exceeding 100 m in average depth. The majority of lakes (69%) are at elevations below 1000 m, including two below sea level (lakes Afrera and Assal, at −112 m and −154 m), and 17 lakes (11%) are located above 1400 m, mainly within Ethiopia, followed by Rwanda and Kenya.

The Köppen-Geiger[41] climate zonation classifies sub-Sahelian lakes as 62% tropical, 25% arid and 13% temperate climates, with a clear alignment with land cover class (see Fig. 1d, f). Tropical climates span nearly the entire continent. The "tropical, savannah" subclass is the most dominant, found for 77 out of 85 lakes. Tropical lakes are characterized by the surrounding dense vegetation (23% of all lakes, 22.5% tropical lakes), mostly deciduous (17% of all lakes, 27% of tropical lakes). Croplands are present in the lake watersheds in all climate zones (39% of all lakes) and are found in 37 out of 85 tropical lakes. Arid climates host 34 of the lakes mainly in the "arid, steppe, hot" subclass. These lakes are located in the countries surrounding the Horn of Africa as well as in the southwestern part of Madagascar and in the countries poleward of 16°S latitude. The watershed area of lakes located in arid climate zones are characterised by sparse shrublands (4.3% of all lakes, 18% of arid lakes), grasslands (3% of all lakes, 12% of arid lakes), and croplands (9.4% of all lakes, 38% of arid lakes) reflecting their dry conditions and need for irrigation. Temperate climates characterise the lakes in the eastern parts of South Africa, central Madagascar, Angola, Zambia, Zimbabwe and Mozambique. There, croplands, tree cover, grass- and shrublands are the most common land cover classes. Croplands, making up 22% of the watershed of temperate lakes (3% of all lakes) in the temperate climates, concentrate around the north-western and northern boundaries of the region (i.e., Sierra Leone, Côte d'Ivoire, Nigeria, Burkina Faso, Chad), as well as around the north-eastern African Great Lakes (Victoria, Albert, Edward, Kivu). In the south, croplands are most prominent in Zimbabwe, south of Lake Kariba. Tropical and temperate climate areas are also the most densely populated in sub-Sahelian Africa (Fig. 1g), containing the largest natural lakes and 32% of reservoirs (e.g., Lake Volta, see Fig.1e). The remaining 68% of artificial reservoirs are located in arid and dry climatic zones characterised by limited vegetation classes, like bare ground, shrublands, grasslands and deciduous tree cover. Thirty percent of lakes in our dataset are artificial and were built between the 1960's and 1990's[41] for water abstraction, irrigation, flood control and hydropower purposes.

Lakes in the most densely populated areas have been the target of most frequent research (Fig. 1g, h). Based on our literature screening, there are either no studies in the scientific literature (23%) or less than 10 scientific publications (27%) for half of the lakes considered here. Most unstudied lakes are of natural origin and show a population density of less than 60 persons per km² in the upstream catchment. Forty percent of lakes are the direct subject of 10-200 research articles spanning a wide range of topics. Among these, reservoirs are primarily the target of climate- and water-related research (e.g., Chomen, Lagdo, Tiga reservoirs), often supported by remote sensing technology, while natural lakes (e.g., Mweru, Bangweulu) are preferentially case studies for biological, paleolimnological and geological investigations.

With more exposure in the literature (correlated with the lake size and population density), the range of disciplines studied in connection to a water body also widens to social and political sciences, epidemiology, and economics. Twelve lakes (9%) have been the subject of intensive research (with more than 200 studies each). They are the East African Great Lakes, the West African lakes Chad and Volta, and the Southern African lakes St. Lucia and Kariba. All of them are of natural origin except for the two large reservoirs, Kariba and Volta.

Lake water colour reflects the climatic region, anthropogenic pressure, and morphological characteristics

Land use and cover in the upstream catchment are tightly connected to lake water colour, as inflows from the surrounding territory impact optically

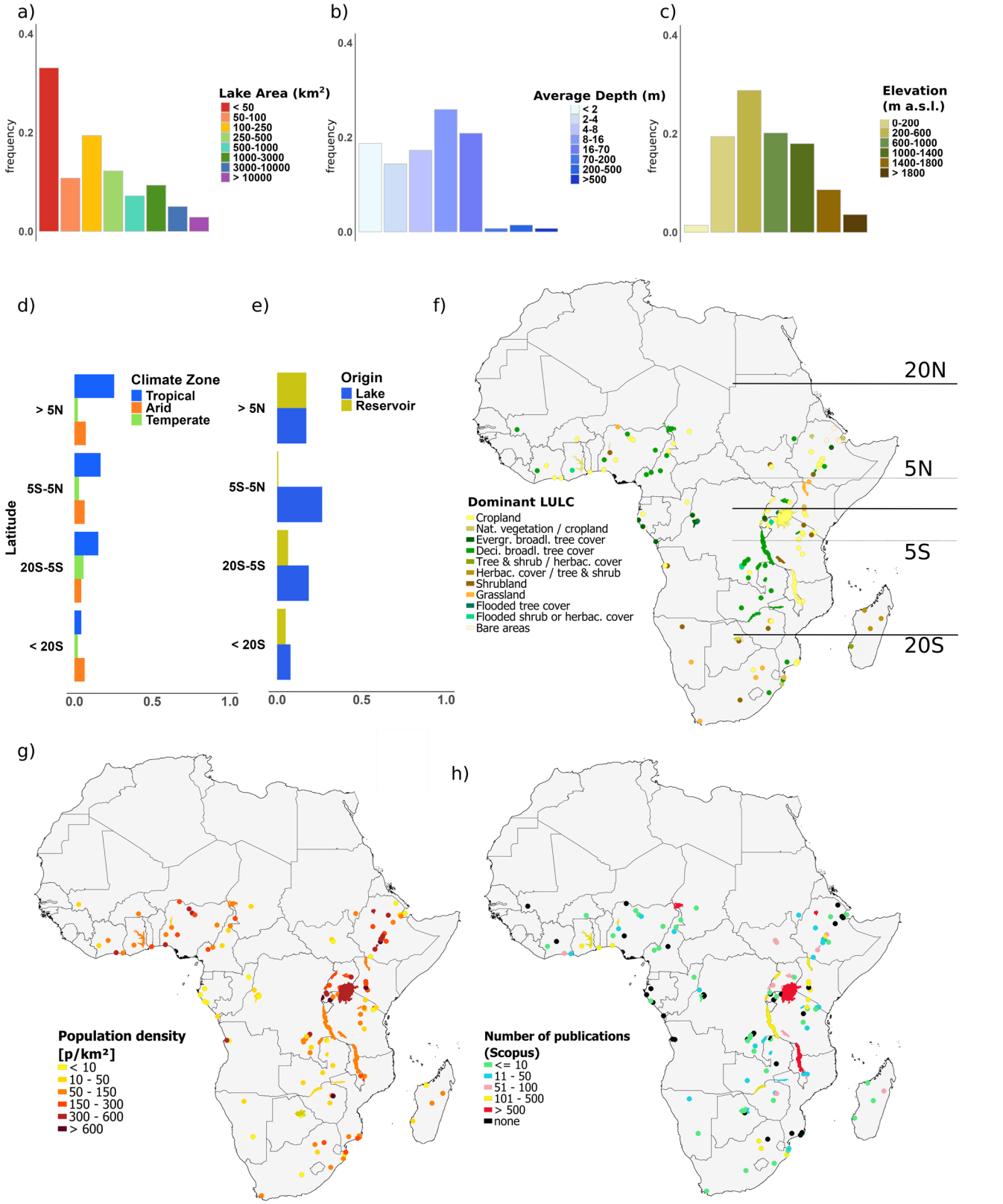

Fig. 1 | Sub-Sahelian lakes considered in this study clustered based on natural and human features. Distribution of lake a surface area[81] in km²; **b** average depth[81] in m; c elevation[81] in m a.s.l.; **d** climatic zones after the modified Köppen-Geiger classification[41]; **e** origin of water bodies; clusters of lakes based on **f** land use and land cover[82] in the upstream catchment; g population density[94] in the upstream catchment; **h** number of publications per lake across all subjects from a Scopus[95] search performed in this study. Maps of all considered features and tables containing per-lake information are provided Table S1 in the Supplementary Materials.

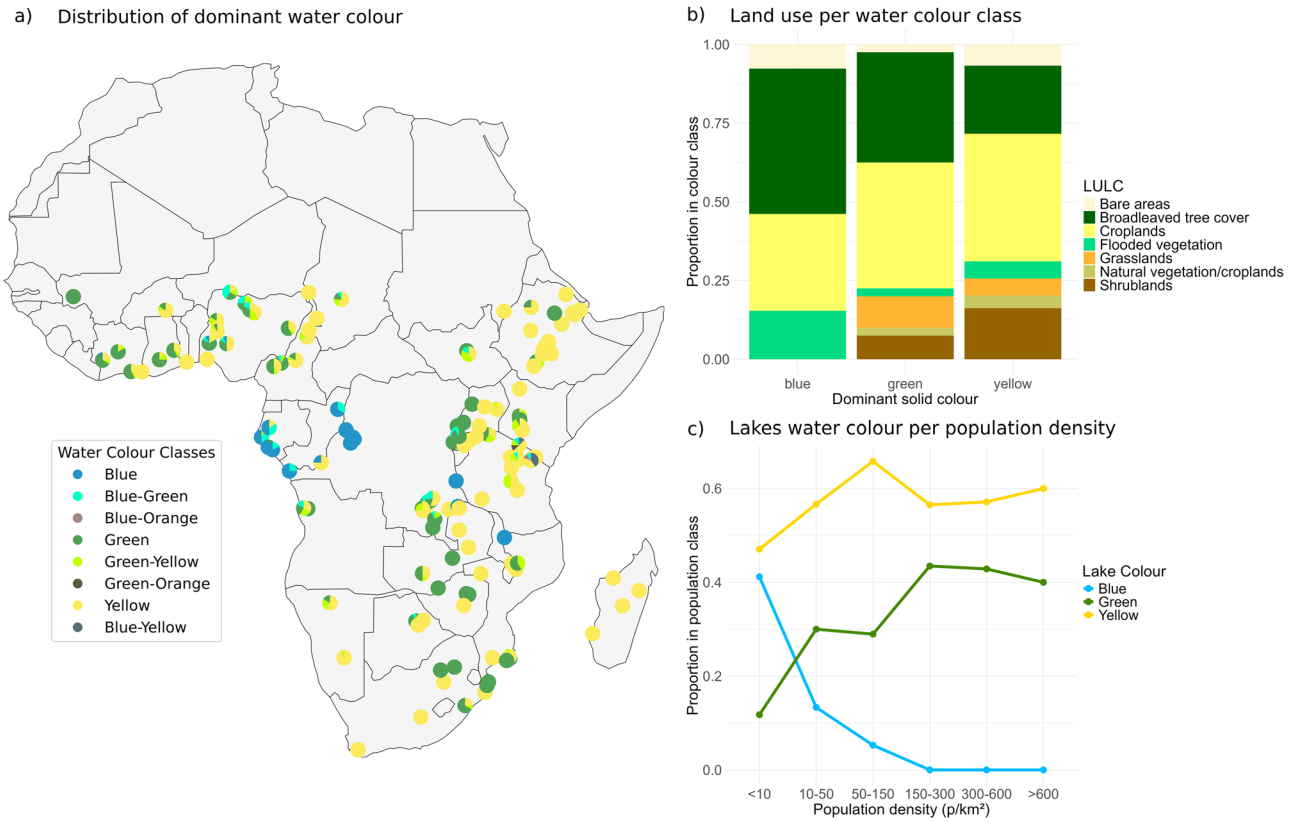

Fig. 2 | Lake water colour and catchment features. a Overview of the unique colours expressed throughout the year by each lake. **b** Distribution of predominant land cover in upstream catchment classes per each lake water colour class (only solid colour classes are displayed, with respect to the total number of lakes per class). **c** Abundance of lake colour classes per population density classes (normalised by the total number of lakes per population density class).

active constituents such as dissolved organic matter, suspended particles and algae. These constituents interact with sunlight to produce a dominant wavelength detectable by satellites and attributable to colour, which is therefore an aggregated indicator of water quality.

The most common water colour across sub-Sahelian African lakes (Fig. 2a) is yellow (54% e.g., Lake Turkana, Lake T'ana), followed by green (29%, Lake Kariba, Lake Kivu) and blue (9%, e.g., Lake Malawi, Lake Tanganyika). A few lakes show orange observations on a daily basis (e.g., Lake Zimbambo and Ambussel). Some lakes display transient colour classes, e.g., green-yellow (4%, as Tiga Dam and Lake Langano), blue-green (2%, Goronyo Dam, Lake Zimbambo), blue-yellow (Lake Mweru Wantipa), or yellow-orange (Lake Ambussel). Blue lakes mainly reside in tropical regions (monsoon, forest and savannah), typically between 10°S and 5°N latitude and west of 10° longitude in Gabon, Congo and the Democratic Republic of Congo. Green and yellow lakes are more evenly distributed across latitudinal and longitudinal gradients. Green lakes are found in arid, temperate and tropical climate zones, while yellow lakes are mostly in tropical and arid regions. Croplands (rainfed and irrigated) characterise the catchments of lakes of each colour classification with the highest contribution in the yellow category (blue: 31% over 13 lakes, green: 33% over 40, yellow: 39% over 74, see Fig. 2b). In contrast, tree cover (broadleaved evergreen or deciduous) is distributed mostly in the tropical zones, leading to the highest contributions to blue lakes' catchments (blue: 38%, green: 35%, yellow: 20%). The catchments of blue lakes are further constituted by 15% of flooded tree cover (fresh or brackish). Shrublands, typical of arid climate zones, comprise 11% of yellow lakes' catchments, while grasslands contribute 10% to the catchments of green lakes. Blue lakes are found exclusively in areas of low population density and are entirely absent once population increases over 50–150 p/km^2 (Fig. 2c). Green lakes are found in all population density brackets but are most persistent in medium densities. Yellow lakes are present in every population density bracket and peak in number at the threshold density of 50–150 p/km^2. This indicates that once an anthropogenic pressure point is reached in terms of population density, clear, blue oligotrophic systems can no longer be sustained as eutrophication increases. Including observations of lake area and average depth into these considerations it appears that the blue lakes Tanganyika and Malawi, persisting at moderate population densities, not only cover an extensive area (>29,000 km^2) but also show high average depths (577 and 261 m respectively). These blue lakes also have long hydraulic residence times and respond slowly to nutrient inputs. Thus, their maintained clarity may reflect a delayed response to high anthropogenic pressure rather than equilibrium with nutrient loads and true resilience. Lakes Victoria and Kivu are examples of persistent green lakes in areas of high population density (>600 p/km^2). While Lake Victoria is comparatively shallow (41 m), it covers an area of 67,000 km^2. Lake Kivu, while 12 times smaller than Lake Victoria (2400 km^2), resembles Lake Malawi in average depth (237 m). The yellow lakes in this dataset do not exceed average depths of 35 m and, on average, comprise smaller areas. The shallowest and smallest lakes can appear yellow even at low population densities. Lake size and depth can therefore be considered important features buffering potential anthropogenic stressors.

Large-scale climate variability shapes the seasonality of atmospheric variables

Sub-Sahelian lakes are influenced by overlapping climate patterns. We consider the climatological year of spatially averaged atmospheric variables from reanalysis (ERA5-Land[42]) and satellite-derived precipitation within the immediate catchment from CHIRPS[43] and TAMSAT[44] datasets (only TAMSAT is shown). The seasonal variability of most climatic factors depends strongly on geographic location, especially latitude (Fig. 3a–f), while longitudinal patterns mainly impact rainfall and relative humidity (not shown). The magnitude and timing of seasonal maxima and minima vary across different climatic variables and latitudes. The Intertropical

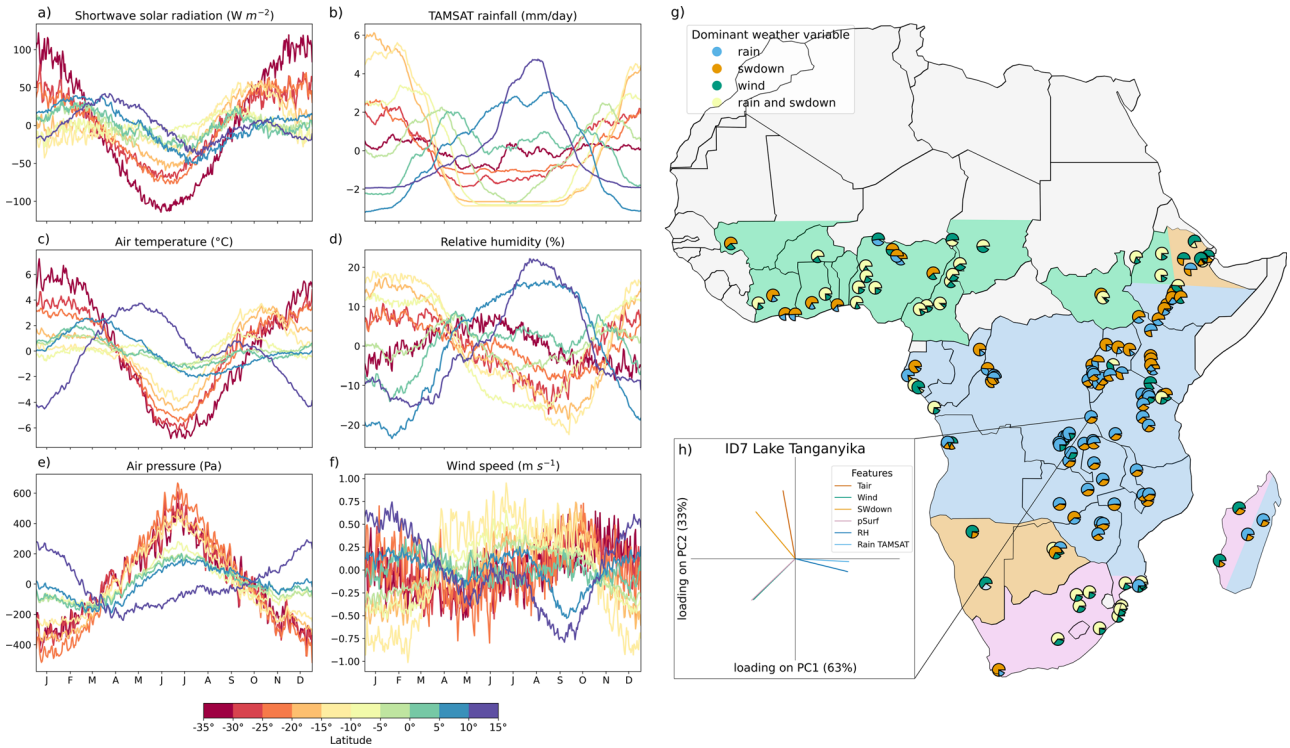

Fig. 3 | Seasonality and Principal Components of atmospheric variables. **a–f** Climatological year showing the average seasonal evolution of atmospheric variables in lakes across latitudes (colorbar). Each line represents the climatological data, with the annual mean removed, and averaged across lakes within the same latitude range. Climatologies for all variables considered in this study for each lake are available in Table S.2 in the Supplementary Materials. **g** Lake clusters based on climate variables predominantly explained by the first two principal components (PCs). The colour fraction in each circle indicates the proportion of variance explained by each PC (i.e., PC1 corresponds to the larger, PC2 to the smaller fraction). Background colour blue, green, pink and orange identify countries with common main climatic factors. Inset h displays an example from Lake Tanganyika of per-lake PCA results that were used to create the map in **g**). Similar plots for all lakes are available in Table S.3 in the Supplementary Materials.

Convergence Zone (ITCZ) is the main mode of intra-annual climate variations across the continent and follows the seasonal path of the sun[45]. However, regional factors, such as topography and oceanic modes of variability (such as ENSO and the Indian Ocean Dipole) also play an important role[46–48].

Consistent with recent advances in the theoretical understanding of monsoon systems[49], lakes located in the proximity of the ITCZ (20°S-5°N) show a clear seasonal pattern. In these lakes increases in surface air temperature, which follows the solar radiation cycles, (Fig. 3a, c) precede the development of convection and precipitating storm systems (Fig. 3b). As a result, warmer air temperatures are associated with higher precipitation and lower surface air pressure (Fig. 3e). Conversely, away from the ITCZ (north of 5°N, south of 20°S), air masses tend to sink, inhibiting cloud formation and increasing surface air pressure. Lakes in those regions experience low-to-no precipitation and higher surface air pressure when air temperatures are cooler. Seasonal patterns of relative humidity (Fig. 3d) closely follow those of precipitation. The seasonality of wind speed (Fig. 3f) is less apparent and instead associated with the Northeast and Southeast trade winds across Northern and Southern Africa, respectively (as well as local features, such as the West Africa Monsoon winds).

Temperate and arid lakes in the Northern regions (10°N to 0°S, e.g., Lake Chad) show a single rainy season, while wind climatology shows a bimodal distribution. A long, windy period coincides with the dry season due to North-Easterly winds from the Sahara Desert, and a shorter peak is observed during the wet summer season (from July to September). Tropical equatorial lakes (0 to 10°S, e.g., Lake Victoria) exhibit two rainy seasons. Among these, in some East African lakes, wind speed clearly reflects the Somali jet, which develops as part of the Indian Summer Monsoon, with the highest peak from June to August[50]. Southward, lakes at 10 to 20°S (e.g., Lake Kariba) show a unimodal rainy season, while South-African lakes (20 to 30°S) show a flat rainfall distribution, signifying persistent arid conditions. Here the wind peak shifts towards September-November and coincides with the driest conditions associated with South-Easterly winds caused by the Atlantic Ocean high-pressure system.

The interplay between wet, dry and windy season defines regions of analogous climate drivers

As large-scale atmospheric processes regulate the climatological annual cycle of different atmospheric variables, many of them exhibit similar seasonal evolution. These groupings depend mostly on the geographical location of the lakes, with factors such as lake surface area and altitude also influencing the variations. To identify clusters of climate regulators (Fig. 3g), we apply a Principal Component Analysis (PCA) to reduce six atmospheric variables (Fig. 3a–f) into two Principal Components (PCs) that capture the most variance.

In East and Central African lakes (countries coloured in blue), the first and the second PC resemble the seasonality of either the hydrological cycle (represented by rainfall and humidity) or insolation (expressed as shortwave solar radiation, "swdown" for brevity). In 51 lakes (37%) PC1 associates with the hydrological cycle, while PC2 with solar radiation (40 lakes) or wind/air pressure (11 lakes). PCA results for these lakes are similar to those of Lake Tanganyika shown in Fig. 3h. These lakes are primarily tropical (e.g., Lake Victoria, Tanganyika, Malawi, Kivu), while fewer classify as arid (e.g., lakes Turkana, Eyasi, Sua-Pan) or temperate (e.g., Alaotra, Mutanda, Burera), with a clear distinction between the dry and wet seasons, driven by the ITCZ migration and monsoonal rains. In a smaller but considerable number of lakes (32, i.e., 22%), shortwave radiation holds most of the explained variance of PC1 (from 47% to 73%), and rainfall is well aligned with PC2. These lakes are primarily tropical lakes located at the Equator and spanning all longitudes from Côte D'Ivoire (e.g., Kossour Reservoir, lakes Buyo and Aby,

Ebrie Lagoon) to Kenya and Ethiopia (e.g., lakes Baringo, Bogoria, Awasa). Further arid (e.g., Kariba) and temperate (e.g., Naivasha) lakes spread in extratropical latitudes but mostly along a longitudinal slice between 26 to 36°E. All these lakes are characterised by net wet and dry seasons, but they are located at the edges of the seasonal migration of the ITCZ in Africa[51].

In lakes located outside 5°N and 20°S, rainfall and solar radiation both contribute to PC1, while PC2 generally follows wind and/or air pressure seasonality. The fact that rainfall and solar radiation contribute to the same PC means that they are either correlated or anti-correlated, and can therefore be explained by the same seasonal process. 23 of these lakes are located at the northernmost latitudes (e.g., lakes Volta, Kainji, Lagdo, Bankim and Chomen, countries coloured in green), marking the transition between the dry Sahel's climate and the humid tropical region, and the development of North-Easterly trade winds. In these regions, the rainy season (June-September) coincides with the arrival of the West African Monsoon, which brings persistent cloud cover blocking shortwave radiation. For this reason, rainfall peaks exactly when shortwave radiation is at its minimum, e.g., in August. The remaining lakes are located in the temperate Eastern areas of South-Africa (countries coloured in pink), where the interplay between the ITCZ migration and subtropical high-pressure systems regulates local climate. In these lakes (e.g., Gariep dam, Lake Vaaldam), shortwave and rainfall both peak during the austral summer (November to March), which coincides with the rainy season. The reason is that precipitation in this area is mainly of convective origin, thus it is caused by surface heating, which is at its maximum in summer.

Rainfall does not contribute to any principal component in 16 lakes (12%, in countries coloured in orange). In these lakes, rainfall indeed shows minimal seasonal variation compared to other atmospheric variables like solar radiation and wind speed. The lakes classified in this group include Lakes Hayala, Melka Wakena, Abayata, Abbe (Ethiopia), Assal (Djibuti), Natron, Jipe (at the border between Tanzania and Kenya), Kinkony, Ihotri (western Madagascar), Hardap, Etosha Pan (Namibia), Nkokwane Pan (Botswana), Quiminha (Angola), Kiri, Zobe (Nigeria), Mantanali (Mali). These lakes are located in arid or semi-arid climatic zones or in areas subject to pronounced dry seasons, leading to very extreme conditions. In fact, many of these lakes are endorheic, like Etosha Pan, Assal, Natron, Abbe, Kinkony, with high evaporation rates leading to saline or alkaline water. Lakes Ihotry and Kinkony also fall in this class, reflecting the significantly drier climate of western Madagascar compared to the east, where heavy rainfalls from moist Indian Ocean air lead to a different classification for lakes Alaotra and Itasy.

Climate drivers regulate lake water quality and quantity on a seasonal scale

The response of lakes to climate drivers varies by latitude, morphology, hydrology and ecology. Here we examine whether satellite-derived lake variables, namely Lake Surface Water Temperature (LSWT), Lake Water Level (LWL), Chlorophyll-a concentration (chlorophyll-a) and turbidity (TURB), correlate with principal climate components. The Pearson correlation coefficient r between their climatologies and the PCs is considered for single lakes (see Fig. 4a for Lake Tanganyika), while the absolute value is considered to cluster multiple lakes (Fig. 4b, d–f). Full results from these analyses are shown in the supplementary material (Table S.1–S.3).

LSWT strongly correlates with major climatic drivers ($|r|> 0.9$ and p-value < 0.01 in 52% of cases). Using $|r|> 0.5$ as a "strong" correlation threshold, 66% of lakes strongly correlate with PC1 (with a left-skewed distribution across lakes and median value 0.86), and 30% with PC2 (median value 0.55), while 4% show negligible correlation. Tropical lakes like Tanganyika display positive LSWT-rainfall correlations (Fig. 4d) with almost no lag (Fig. 4c), as the wet season coincides with the predominantly cloud-covered warmest period of the year. These lakes indeed experience wind-driven cooling during the dry season[52]. Strong positive correlation between LSWT and dry season is instead found in 28 lakes (20%) where temperate (e.g., Lake Naivasha) or arid (e.g., Lake Kariba) climates enhance surface warming through the shortwave radiation penetration[53].

Chlorophyll-a and turbidity correlate less strongly than LSWT to the primary climatic drivers, with higher median correlations to PC1 (median $|r| = 0.80$ and 0.44 for turbidity and chlorophyll-a, respectively, and a left-skewed distribution) than PC2 (0.46 and 0.48, with a uniform distribution). In 18% of lakes chlorophyll-a and turbidity are aligned with a correlation $|r|>0.2$ to the same PCs. These lakes show a clear signature of rainfall season in both chlorophyll-a and turbidity climatology (Fig. 4e, f), such as lakes in the tropical and subtropical region (e.g., Victoria, Kyoga, Kivu, Kariba), a few coastal West African lakes (e.g., Buyo, Togo) and two lakes in Madagascar (Alaotra, Itasy). Turbidity often responds quickly to rainfall (less than one month), while chlorophyll-a lags by 40–55 days (Fig. 4c), probably reflecting nutrient-driven phytoplankton growth. In 60% of lakes (e.g., Tanganyika, Chad, Volta, Mweru), turbidity correlates with rainfall, while chlorophyll-a reflects radiative and wind-driven hydrodynamic processes (e.g., Tanganyika and Volta).

The primary difference between natural and artificial lakes revealed by our analysis is the correlation between rainfall and LWL variations. Of the 54 lakes for which LWL is available, 68% exhibit a positive correlation (i.e., $r > 0.2$) with rainfall-driven PCs. Notably, 29 of these water bodies show an optimal lag of over 3 months (over 90 days), 11 of which (38%) are reservoirs, and the remaining 18 (62%) are natural lakes. Within the natural lakes exhibiting this delay, seven are very large (>1000 km²), such as Tanganyika, Malawi, Turkana, Kivu, Bangweulu and Rukwa. The other 11 are still substantial in size (100–800 km²) and have upstream catchments up to 100 times larger than their surface area. For instance, Lake Kisale is 300 km² in size with an upstream catchment of 36000 km². Additionally, two of these are wetlands (i.e., Alaotra and Bangweulu) characterized by very slow flow in their upstream large floodplains and relatively minor climatological water level variations.

These findings indicate that reservoirs show a weaker or delayed response to rainfall, likely due to artificial regulation. Similarly, the natural lakes with delayed responses tend to share morphological characteristics - particularly large surface areas and extensive catchments - which influence the timing of their response to rainfall. For what relates to water quality variables, in lakes and reservoirs located in close proximity or within the same latitudinal range, the principal components of atmospheric variability and LSWT seasonality are often similar. However, the seasonal dynamics of water quality variables, particularly chlorophyll-a, can vary. For example, in the arid regions of South-Eastern Africa (latitude >20°S), Lake Bambene (Mozambique) and Loskop reservoir (South Africa) exhibit similar PCs, as well as comparable LSWT, chlorophyll-a and turbidity seasonality. At latitudes between 5°S and 20°S in East Africa (Tanzania), Lake Sulunga and Mtera Reservoir share similar PCs and seasonality in LSWT and turbidity, yet chlorophyll-a dynamics slightly differ. The shallow Lake Sulunga reaches its first chlorophyll-a peak less than one month (21 days) after the end of the rainy season, whereas in the deeper Mtera reservoir, the peak occurs during the dry season. A similar pattern is observed in West-Central Africa (Angola) at comparable latitudes, where the shallow Lake Quilunda responds more quickly and synchronously to rainfall with earlier chlorophyll-a peaks compared to the deeper Quiminha reservoir. These variations are not limited to a specific region or lake type, but occur across latitudes and between natural lakes and reservoirs alike. Generally, shallow water bodies tend to exhibit more frequent chlorophyll-a peaks and to react more rapidly to rainfall than deeper lakes or reservoirs. This suggests that basin morphology, rather than lake origin, plays a more significant role in shaping ecological responses when other factors (i.e., local climate, upstream catchment characteristics) are the same.

Clusters of lake functioning match regions of analogous vulnerability to climate change

Climatology similarities among atmospheric and lake variables highlight regional patterns in sub-Sahelian lake responses to climatic factors. In Fig. 5 we define four main clusters (A, B, C, D) of lakes based on similar main climatic drivers and on the co-variability of LSWT and turbidity. Each cluster is subdivided into three subcategories (1,2,3), which assign the

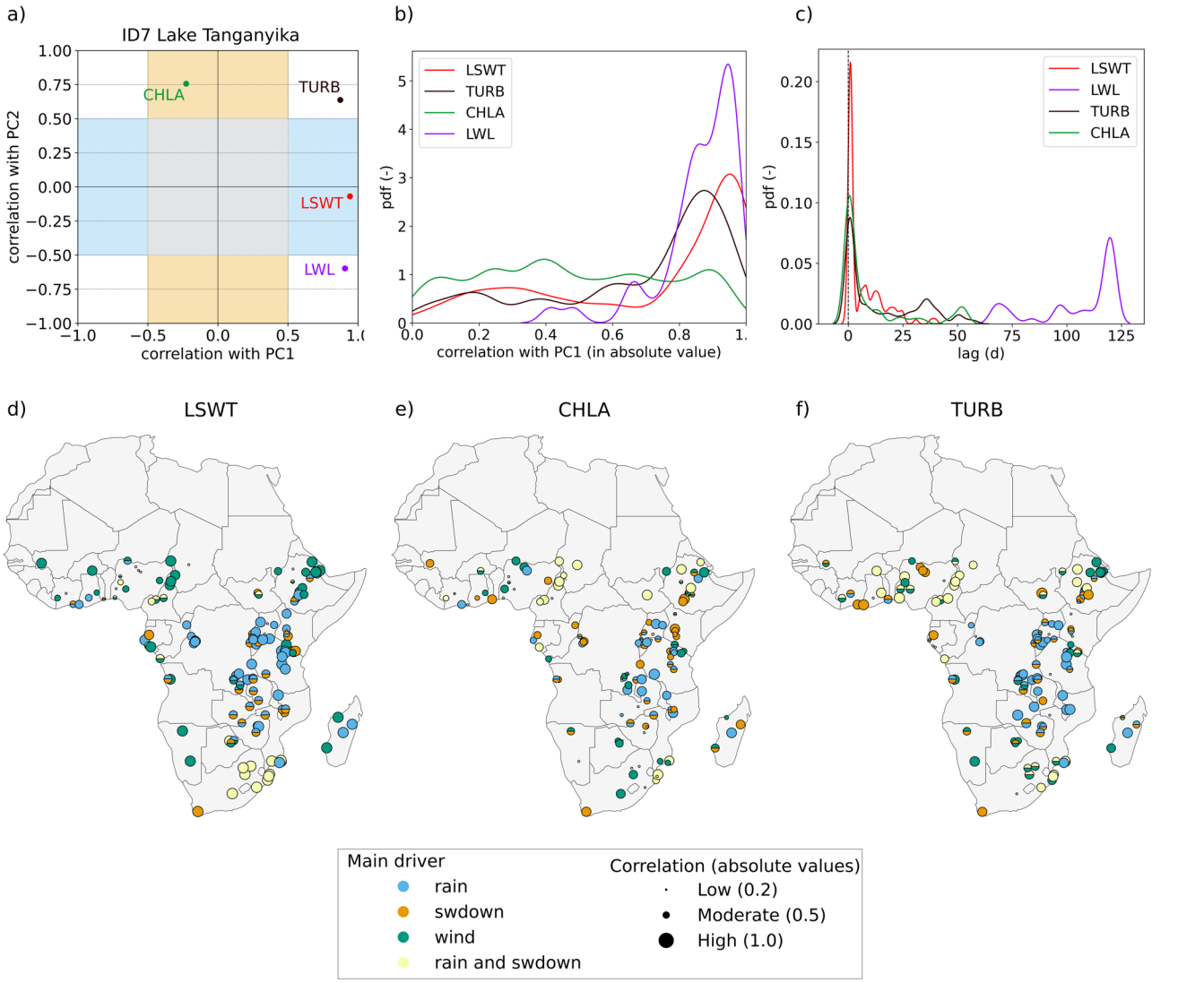

Fig. 4 | Correlations between atmospheric PCs and lake variables. a Example of correlation plot for Lake Tanganyika (similar plots for all lakes are available in Table S.3 in the Supplementary Materials). The coloured shadowed bands represent the areas where the lake variables are weakly or not correlated with PC1 (vertical band) or PC2 (horizontal band). In the example for Lake Tanganyika, if a dot falls in the blue band, the lake variable is well correlated (or anti-correlated) with PC1 (associated with rainfall in this lake) and weakly or not correlated with PC2 (associated with shortwave radiation); if in the orange band, it is well correlated (or anti-correlated) with PC2 and weakly or not correlated with PC1; if in white corners, it is correlated or anticorrelated with both PCs; if in the central grey box, it is not correlated or weakly correlated with any PCs. **b** distribution of correlations between lake variables (in absolute values) and PC1; **c** distribution of optimal lags maximising correlation between lake variables and rainfall PCs; **d**–**f** lake clusters representing the PCs showing the largest correlation (either positive or negative) with the variable displayed in each panel: **d** Lake Surface Water Temperature (LSWT), **e** Chlorophyll-a (CHLA), **f** Turbidity (TURB). The dots are coloured depending on the climate variable explained by the PC with whom lake variables show the largest significant correlation (p-value > 0.05). The radius of each single-coloured dot equals the correlation. Double-coloured dots are reported for those lakes where lake variables show significant correlation (or anticorrelation) with both PC1 and PC2. The radius of double-coloured dots is given as the mean of the two correlations (in absolute values). Small white circles indicate that the p-value is larger than 0.05 or that the correlation is lower than 0.2 in absolute values.

seasonality of algal blooms for each cluster either in the dry, wet, or both seasons, respectively. These clusters partially follow a regional distribution and can be directly related to known heterogeneity of climate vulnerability across Africa as reported by the latest IPCC report[29] and by the Joint Research Commission (JRC) World Atlas of Desertification[54].

Cluster A comprises 14 lakes where rainfall contributes little to the atmospheric variance, with shortwave radiation and wind representing the main seasonal signatures visible in LSWT. These lakes are predominantly located along the eastern margin of North East Africa (NEAF) in tropical/savanna climate regions and in the arid/desert zones of West South Africa (WSAF), regions characterized by consistently low annual rainfall. Despite limited precipitation, the rainfall season is still visible in the turbidity and chlorophyll-a seasonality of these lakes. In class A1, three lakes show algal blooms typically occurring during the dry season, e.g., Koka Reservoir, where cyanobacteria dominance peaks during the dry and minor rain season[55]. Class A2 comprises five lakes in which chlorophyll-a concentrations generally peak during the wet season, e.g., lake Abayata, which is phosphorus-limited in the dry season and was found to be an exception in Ethiopian lakes, where the rainy season normally causes reduced transparency and inhibits the development of algal blooms[56]. We classify the remaining three lakes as A3. In these lakes, chlorophyll-a peaks occur throughout the year, like in Etosha Pan and the Hardap reservoir in Namibia. According to the latest IPCC report[29], lakes of class A located in NEAF are experiencing an increase in temperature extremes and heavy precipitation events. Changes in rainfall seasonality are projected for the short rain season, which is expected to become longer, and confidence is

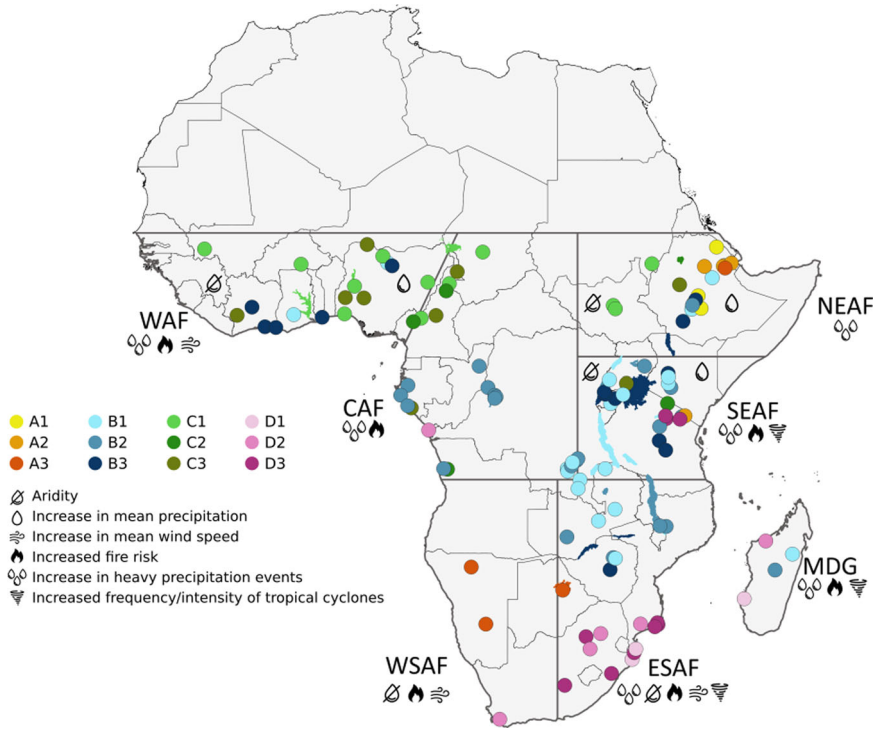

Fig. 5 | Clusters of lake functioning based on PCA results and correlation analysis with lake variables. Acronyms indicate regions defined in the IPCC[29] report (Chapter 9, Figure 9.14): North East Africa (NEAF), South East Africa (SEAF), Madagascar (MDG), East South Africa (ESAF), West South Africa (WSAF), Central Africa (CAF) and West Africa (WAF). Icons depict the climate change vulnerability in the African regions as projected by the IPCC report and by the JRC World Atlas of Desertification[54].

high on decreasing aridity in the eastern areas of NEAF. This might lead to prolonged blooming periods in lakes of classes A2 and A3, and potential shifts in phytoplankton productivity in lakes of class A1 due to the mismatch in nutrient-timing. For the two lakes in WSAF, projections predict decreasing rainfall and rising frequency of fire-conducive weather. Such a combination, in the context of lake vulnerability, can be translated to a change in the origin of nutrients from runoff to wind resuspension and nutrient deposition, which might affect phytoplankton composition and the timing of algal blooms.

Cluster B groups 75 lakes where atmospheric variance is mostly associated with rainfall, relative humidity, and shortwave radiation. In these lakes, the rain season signature is clearly visible in LSWT and turbidity. Almost all tropical lakes of South East Africa (SEAF) reside in this class, together with tropical and temperate lakes in the northernmost ESAF (East South Africa) and NEAF (North East Africa). Tropical lakes in Ethiopia (e.g., Langano), Kenya (e.g., Bogoria) and Zambia (e.g., Mweru) show greater influence of shortwave radiation on seasonal chlorophyll-a peaks (B1, 29 lakes, 39%). A similar proportion of lakes in this class show that chlorophyll-a peaks during the wet season (27, e.g., Malawi, Kyoga, 36%) or during both wet and dry season (B3, e.g., Victoria, Cabora Bassa, Kivu, 19 lakes, 25%). The large, deep and oligotrophic lakes Tanganyika (B1), Malawi (B2) and Kivu (B3) are interesting representatives of the three sub-clusters. Literature for these lakes reports reduced transparency during the wet season (Lake Malawi[57]; Kivu[58]; Tanganyika[59]) and phytoplankton maxima associated with diatom blooms during the dry season (June–September) induced by wind-driven mixing of the water column[60]. During the wet climate characterising the rest of the year, thermal stratification limits nutrient exchange across the epilimnion and pico-cyanobacteria dominate with lower concentrations. It seems that our climatology for lakes Malawi and Kivu captures the chlorophyll-a peak concurrent with rainfall[58,61] and underestimates the peak during the dry season in Lake Kivu[26], albeit detecting its existence. In Tanganyika, our chlorophyll-a results align with established lake seasonality[62], where wind-driven hydrodynamics regulate nutrient availability and chlorophyll-a[63] during the dry season. According to the latest IPCC report[29], the regions where cluster B lakes are more abundant are experiencing an increase in temperature extremes and heavy precipitation events. Similar to cluster A, the IPCC predicts changes in rainfall seasonality and decreasing aridity in the eastern areas for NEAF and SEAF (mainly Ethiopia, Kenya and Uganda). As a result, lakes in this region are vulnerable to longer periods of reduced water clarity, increased nutrient loading, stronger thermal stratification, and shifts in phytoplankton dynamics.

Clusters C and D include the remaining lakes where atmospheric variance is strongly impacted by wind seasonality and its interaction with the wet and dry seasons. Cluster C includes 29 lakes located in western extra-tropical regions close to the border with Sahel (WAF), along the equatorial western coast (CAF), and in the western countries of NEAF (Sudan, South Sudan), where rainfall and shortwave radiation are anti-correlated. Here, the alternation of North-East trade winds and the West African monsoon determines the alternation of wet and dry seasons. In this cluster wind speed anticorrelates with LSWT, such that cooler water temperatures occur during windy seasons. Differences related to the timing of algal blooms and reduced clarity lead to a further distinction into clusters C1 (15 lakes, 47%, e.g., Chad), C2 (5, 25%, e.g., T'ana, Lagdo, Natron) and C3 (9, 28%, e.g., Buyo, Chomen). The majority of lakes in cluster C show visible correlation with wind seasonality in the climatology of chlorophyll-a and turbidity. This is the case, for example, for most West African lakes (from Ghana to Gabon). In Lake Volta, two main peaks in LSWT and chlorophyll-a directly correlate with the two windy seasons, the wet southerly monsoon (peaking in June to August) and the dry northerly "Harmattan" winds (November to March)[64]. Turbidity instead peaks with rainfall in August, when LSWT and chlorophyll-a decline slightly, consistent with existing knowledge in terms of timing[65] and range of concentrations[66] of algal blooms. The regions south of the Sahel belt, particularly West Africa, are expected to experience an increase in mean wind speed and in the frequency of heavy rainfalls, while precipitation trends remain uncertain[29]. These areas have undergone alternating multidecadal wet and dry periods, with severe droughts causing dramatic reductions in water bodies, such as Lake Chad[67], followed by increased annual precipitation partially compensating for the water loss (since 1980s). Additionally, frequent burning in these regions has been correlated to higher turbidity and algal blooms in certain lakes, including Chad and Volta[68]. The IPCC predicts a rising frequency of fire-conducive weather, which, combined with stronger winds and heavy rainfall,

enhances the risk of deteriorating water quality due to terrestrial nutrient inputs in these lakes.

Cluster D includes 19 lakes where rainfall and shortwave radiation are well correlated. These lakes are typically located in East South Africa (ESAF), western Madagascar (MDG), and two lakes located in SEAF (Manyara and Ambussel, in Tanzania). These lakes are sensitive to South-Easterly trade winds that carry moisture from the ocean during the austral summer from November to March. Temperate inland and tropical coastal lakes show LSWT following shortwave radiation seasonality, which shows the same shape as rainfall and impacts chlorophyll-a and turbidity. Lakes further distribute across clusters D1 (3,16%, i.e., lakes Ihotry and Kosi), D2 (6, 31%, e.g., Kinkony, Vaaldam, Bambene), D3 (10,53%, e.g., Manyara, Hartbeespoort dam). In Lake Vaaldam, both turbidity and chlorophyll-a are correlated with the rainfall and shortwave radiation driver, and anticorrelated with wind speed. Wind peaks during the dry season (between September and November), while turbid and nutrient-rich inputs are normally associated with phytoplankton summer blooms during storms occurring between January and March. These regions are classified as vulnerable to hydrological droughts and to tropical cyclones[69], with a projected increase of mean wind speed and a decrease in mean precipitation[29]. These risks associated with climate change can be translated into lake-specific impacts such as enhanced wind-driven mixing and prolonged residence times, which, when combined with nutrient pulses from extreme events, may amplify turbidity and alter phytoplankton dynamics.

We note that the clusters we observe are well aligned with the findings of Sogno et al.[36], who evaluated dynamic and causal similarities among the African lakes based on climatic variables, water availability and vegetation. If our Fig.5 is compared with the results by Sogno et al., it is straightforward to note that their cluster E and M closely replicate the groups of West African lakes (e.g., Volta, Chad, Buyo and Kainji, our class C) and Ethiopian rift valley lakes (e.g., Abhe-Bid, Afrera, our class A). We also see an interesting connection between South African lakes (mainly reservoirs) and the two lakes along the western coast of Madagascar, largely through the influence of wind and ocean moisture climatic drivers and the lower impact of rainfall on the overall water quality in south-western Madagascar (class C1). Our findings expand on Sogno et al. by showing that climate patterns shaping water availability across spatial and temporal scales also resemble lake physical and ecological behaviour. As such, lakes located in regions with similar climate change vulnerability face comparable risks to their water quality and availability.

Challenges and opportunities of an up-scaled approach over sub-Sahelian African lakes

Expanding research from a single well-studied lake to numerous, largely unexplored lakes at a sub-continental scale offers the opportunity to identify broader patterns but requires streamlined approaches. Despite clear advantages, there are three main limitations of global lake-climate studies like ours: (i) use of multiple global datasets for statistical analysis, which requires evaluation of their uncertainties when applied at individual-lake scales; (ii) reliance on spatially averaged atmospheric and satellite-derived data for test sites where spatial heterogeneity might be relevant; (iii) evaluation of a climatological year, with potential issues such as gaps in time-series, bias-correction needs, and variable complexities like chlorophyll-a concentration and turbidity.

In relation to (i), the uncertainty in global datasets arises from multiple sources: random errors in the assimilated observations (ERA5-Land and CHIRPS), inaccuracies in model parameterizations (ERA5-Land), retrieval algorithm errors (CHIRPS, TAMAT, ESA CCI), limitations in spatial and temporal resolution, and inhomogeneous data availability. In the African context, the lack of in-situ validation data is a structural limitation that cannot be resolved in this work. However, algorithmic uncertainties, such as random errors, can be mitigated through averaging, as these errors tend to cancel out when aggregated across time and/or space. In contrast, systematic biases may persist. For physical variables, such as water temperature, level and atmospheric data, the uncertainty is small compared to the natural interannual variability and has a slight seasonal dependence. In contrast, the uncertainty of biogeochemical variables does not exhibit a strong seasonal variability, suggesting that seasonal pattern detection remains reliable despite potential biases in the absolute value, which can exceed 50% in some lakes. We note, however, that in data-scarce environments, the quantification of uncertainty itself can be rather uncertain. Observations of surface water temperature and atmospheric variables are more abundant for the assessment of ERA5 and LSWT products than measurements of bio-optical water properties. Therefore, we deemed it more robust to use the best available datasets, acknowledging their inherent limitations, rather than discarding valuable signals potentially obscured by high and at times exaggerated, uncertainty. Overall, our results are only marginally affected by the uncertainty in the used datasets, as they are primarily based on spatial and temporal patterns rather than on absolute values.

As for (ii), we examined lakes of various sizes and shapes, where spatial variability in water quality variables, such as LSWT, chlorophyll-a, or turbidity is high. In large lakes, this variability is influenced by factors such as proximity to settlements or agricultural areas (e.g., in Lake Volta[66]), tributary inflows (e.g., Koka reservoir[70]), or three-dimensional flow dynamics affecting vertical and horizontal nutrient transport (e.g., in Lake Tanganyika[71]). Spatial averaging often dilutes this heterogeneity or can generate outliers under cloud cover. Many lakes in our sample feature multiple sub-basins or dendritic shapes, especially in reservoirs like Lake Kariba, where upstream areas exhibit riverine characteristics and downstream areas lacustrine characteristics[72]. This issue also applies to semi-enclosed regions, such as the Winam Gulf in Lake Victoria, which shows dramatically different water colour dynamics and correlations with precipitation compared to the rest of the lake[73]. While these complexities pose challenges for spatial averaging, they also present opportunities. The large-scale approach applied in this study is adequate to detect regionally representative patterns. Complementary analyses of spatial variability at daily or climatological scales could complement our findings, highlighting where spatial variability alters a lake's response to climatic and anthropogenic drivers.

Concerning (iii), the use of climatological means is crucial for studying climate-driven shifts or extremes in lake ecosystems[74]. The first step in such analyses is to define the average seasonal behaviour of the relevant geophysical or geochemical variables. These long-term averages are a necessary baseline against which trends, anomalies, and shifts can be identified relative to what is considered "normal" or representative of past conditions. A well-defined climatology is also critical for distinguishing between natural variability from true extremes. We applied bias correction where evidently necessary (e.g., LSWT, LWL), while keeping most other time series unbiased, assuming that a seasonal pattern exists in all lake variables. However, interconnected variables, shifting lake surface extent[67], phytoplankton abundance changes[26] and altered mixing regimes[23] complicate this analysis. Some lakes, like Turkana and Naivasha, exhibit fluctuating water level seasonality, or no yearly cycle at all (e.g., Lake Bogoria), or too sparse data to detect one (e.g., Lake Baringo). These outliers point to larger-scale changes that warrant further investigation.

Identifying a climatology for chlorophyll-a and turbidity is uncommon in limnology, which typically focuses on direct trends. Using logarithmic scales for these variables, a climatology for turbidity is relatively easy to detect. This is because turbidity variations are primarily driven by physical processes such as sediment resuspension and inputs from the catchment, with only occasional contributions from biological activity. By contrast, chlorophyll-a peaks are favoured by light availability, nutrients, temperature, water mixing, and biotic factors, complicating attribution to climatic drivers. For most lakes, we provide annual patterns for turbidity and chlorophyll-a as a reference to identify dynamic lakes or deviations indicating sensitivity to extremes.

Conclusions

Understanding a lake's typical behaviour is essential for tracking and interpreting its changes over time.

In this study, we identify coherent seasonal patterns among atmospheric variables and examine how these patterns are reflected in lake responses. This provides a first-order understanding of how large-scale climatic regimes influence lake behaviour on a seasonal scale.

Using principal component analysis (PCA), we cluster lakes based on the seasonal co-variation of atmospheric variables. While these clusters show some correspondence with established climate zones, they go beyond that traditional classification by capturing the synchronicity of seasonal atmospheric dynamics. By linking these atmospheric clusters to lake-specific variables, such as water temperature, biogeochemistry, and water level, we explore how broad-scale climatic forcing is imprinted on lake systems. Although this approach does not establish causality, it highlights meaningful temporal synchrony between atmospheric patterns and lake responses.

Our regional framework groups lakes that experience similar climatic and anthropogenic pressures, which manifest in consistent physical (i.e., water temperature and level) and geochemical (i.e., chlorophyll-a concentration, turbidity, colour) responses and potential vulnerabilities. These climate and ecological analogues can critically inform future monitoring strategies. Insights from a well-monitored and extensively studied lake can offer predictive guidance for other, less-studied lakes within the same cluster, and help to prioritize key variables for monitoring, especially when direct measurements are unavailable. This is especially crucial in sub-Sahelian Africa, where significant knowledge gaps remain, and there is an urgent need for harmonized monitoring data, continued long-term remote sensing efforts, large-scale coordination at both political and scientific levels[62], and research questions grounded in existing local knowledge.

Where long-term in-situ records are unavailable, the climatologies developed in this study offer a foundational reference. These baselines, validated where possible through local and scientific knowledge, are now publicly available as supporting materials. They represent a significant step toward addressing long-standing data gaps for many understudied lakes. Ultimately, this study establishes a solid base upon which future research and monitoring strategies in such vulnerable regions can build and expand.

Methods

Global open-source datasets

Lake-specific variables. Data from the European Space Agency (ESA) Lakes Climate Change Initiative (CCI) project for the "Lakes" Essential Climate Variable (ECV) are exploited to obtain time series of lake variables derived from satellite imagery. Lakes in this dataset were selected to be globally representative of the largest inland waters and to cover a wide range of ecological settings and characteristics[34]. The criteria for lake selection were based on lake area (starting from the largest) and on distance from the nearest land (stretch of open water at least 2 km from land). Additional lakes were then considered because of their in-situ data availability or because they were investigated under Copernicus activities. Among the 2024 lakes included in the dataset (version 2.1), we first selected all lakes located in sub-Sahelian Africa (including those in Madagascar), for a total of 140 study cases. We discarded Lake Ist'ifanos (Ethiopia), Pool Malebo (Republic of the Congo and the Democratic Republic of the Congo), and the West Ngalabalab swamp (Sudan) as they displayed almost no water occurrence in the period observed[75].

For the selected 137 lakes, we extract the full time series of maps of Lake Surface Water Temperature (LSWT, in K), chlorophyll-a concentration (in mg/m^3), turbidity (in NTU), Lake water Leaving Reflectances (LWLR, normalised), with a pixel size resolution of ~1 km. In addition, the time series of Lake Water Level (LWL, in m a.s.l.) is obtained for 54 lakes. The temporal coverage of the different thematic variables is the following: LSWT from 1995 to 2022; chlorophyll-a, turbidity and LWLR from 2002 to 2022 (with a gap between 2012 and 2016 for all lakes but Lake Victoria, Cahora Bassa Reservoir and Lake Kivu); LWL from 1992 to 2022.

Meteorological variables from ERA5-Land reanalysis. Atmospheric forcing is obtained from model reanalysis products provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). We extract data from the ERA5-Land[42] reanalysis model, which combines the ERA5 reanalysis forcing with a higher resolution land model and provides data at hourly resolution with a grid cell size of (0.1°x0.1°). We note, however, that 43% of the lakes considered in this study have surface extensions smaller than the model grid size (approximately 100 km^2). ERA5-Land variables are downloaded from the Google Earth Engine (GEE) daily aggregated collection (ECMWF_ERA5_LAND_DAILY_AGGR) for each lake by considering the lake mask available in the Lakes_cci dataset[76]). We extracted data from 1992 to 2022 to cover the full range of availability of CCI products. We consider air temperature at 2 m above ground (t2m, in K), shortwave downward solar radiation (ssrd, in J/m^2), dew point temperature at 2 m a.s.l. (d2m, in K), wind speed at 10 m above ground (w, in m/s), precipitation (tp, in m), air pressure (sp, in Pa), Relative humidity (RH) is computed from d2m following Lawrence 2005[77].

Satellite-derived precipitation. For a better characterization of precipitation, the TAMSAT[44] v3.1 and CHIRPS[43] v2.0 satellite-derived rainfall datasets are included, alongside rainfall estimates from ERA5-Land. The use of multiple datasets can help quantify observational uncertainty, which can be substantially large in places over Africa with scarce in situ observations[78]. For example, ERA5 reanalysis has been shown to allow for a better closure of the water balance of Lake Tanganyika than CHIRPS, IMERG or other rainfall datasets[79]. The TAMSAT and CHIRPS products, which are widely used in both operational and research applications across Africa, are based on thermal-infrared (TIR) satellite imagery to identify the cold cloud tops of precipitating convective storm systems. In turn, the satellite imagery is used to derive cold cloud duration (CCD) maps, which acts as a proxy for rainfall. CCD maps are then calibrated using rainfall from either rain gauges (in the case of TAMSAT) or microwave-based satellite rainfall estimates (in the case of CHIRPS). CHIRPS includes an additional step of merging contemporaneous rain gauge measurements with the TIR-based satellite estimate to improve skill within the vicinity of the rain gauge. TAMSAT and CHIRPS both provide rainfall estimates since the early 1980s-present at the daily timestep.

For each lake, the precipitation from each dataset was extracted using lake catchment shapefiles provided by the Lake-TopoCat[80] dataset - a global lake drainage topology and catchment database. Precipitation can be sorted in three categories: (i) direct rainfall into the lake; (ii) rainfall that falls over the immediate catchment and drains directly into the lake; (iii) rainfall that has fallen over upstream lakes and catchments that then eventually flows downstream into the lake via rivers or streams, and/or groundwater flow (e.g., aquifers). While all important, these precipitation sources operate on different time-scales, which vary for each lake. For this study, we extracted only the precipitation that fell over the immediate catchment of each lake, as this provides an indication of the precipitation variability on seasonal timescales relevant for this study. Precipitation from upstream areas (for lakes within such networks) typically influences timescales longer (i.e., over 12 months) than intra-annual or intra-seasonal patterns, making it less critical for our focus.

Ancillary information. Lake morphological characteristics are extracted from further open datasets. Lake area, average depth, and elevation are obtained from HydroLAKES[81]. The climatic zones were extracted from high-resolution (1 km) Köppen-Geiger maps for 1991-2020[41]. Land use land cover (LULC) data from a reanalysis based on the CCI_LC maps ranging from 1992-2015 was obtained from Digital Earth Africa[82], and extracted on a catchment level leveraging the TopoCat[80] dataset. Information on the origin of waterbodies is obtained from the Lakes_cci metadata. Population counts were obtained from the constrained population density estimate for 2020[83], at 3 arc second resolution (approx. 100 m at the equator), from which population density was calculated per catchment. Additionally, the Joint Research Commission

(JRC) Water Occurrence tool for Africa[75] is used to verify the existence and extent of water surface in the lakes of our sample.

A literature search on Scopus is performed to obtain an overview of the volume of scientific research carried out in the lakes across various disciplines. Papers mentioning the lake name and the country in the title, keywords and abstract are searched with the following query: `(Lake OR Reservoir OR Lac OR Lagoon OR Lagune OR Dam OR Pool OR Pond OR Wetland OR Lagoa OR Crique) Lakename OR AND (Country)`. This allows avoiding the exclusion of results for those lakes in East Africa whose names are homonymous to lakes in Australia, the United States, and Canada (e.g., Lake Albert, George, Victoria, respectively). The query also comprises additional documented names and countries (for transboundary lakes) by adding an "OR" after Lakename and after Country keys. For example, the query for Lake Malawi is: `(Lake OR Reservoir OR Lac OR Lagoon OR Lagune OR Dam OR Pool OR Pond OR Wetland OR Lagoa OR Crique) Malawi OR Niassa OR Nyasa AND (Malawi OR Mozambique OR Tanzania)`. Such search terms therefore exclude those scientific contributions that do not explicitly mention one of the lakes in our sample. This applies for example to those studies that mentioned the name of the lakes in the materials and methods section. We are also aware that the method might also inadvertently assign studies on different lakes within the same country, such as Chad or Malawi, to lakes named after the country, leading to potential overestimation of the scientific studies associated with those lakes. However, as this might apply to a little number of lakes, we believe the method provides an indication of how well each lake has been studied, without demanding the level of detail and accuracy required for a literature review.

Multivariate analysis

Estimation of climatological year. For each variable, we evaluate the climatological year to assess its average seasonal evolution and the standard deviation from 20 years of observations. From now on, "climatology" refers to the annual time series obtained by considering the climatological mean of each day of the year (doy). The climatology is computed based on the statistical properties and temporal availability of the variables involved in the multivariate analysis.

The climatology for the TAMSAT and CHIRPS rainfall is computed for each doy after logarithmic transformation of the rainfall data, which can have a large variance. Moreover, the distribution of rainfall is typically skewed. The log-transformation makes the distribution as 'normal' as possible so that the calculation of the mean and the standard deviation are statistically valid. The mean and the standard deviation are calculated on the doy using a smoothing window of ±7 days and the maximum likelihood estimator is found more suitable than a simple mean: $\underline{x} = \exp(m_x + 0.5s_x^2)$ where m_x and s_x^2 are the sample mean and variance of the log transformed data[46]. The climatologies are computed over the period 2002-2022. Despite considerable uncertainties, rainfall estimated from ERA5-Land, CHIRPS and TAMSAT show consistent annual variability both in terms of order of magnitude and timing of the rainy season(s) (as reported in the climatology plots provided in Table 2 of Supplementary materials).

The climatology of the ESA CCI Lakes v2.1 LSWT is computed for each day of the year (doy) as the mean of the available LSWTs of quality levels >=3 between 1995 and 2022 on the doy ± 7 days in order to obtain a smooth curve. The standard deviation is also computed to capture the variability of the LSWT within the period considered.

The climatology of the ESA CCI Lakes v2.1 chlorophyll-a and turbidity is computed for each day of the year with three different methods: (1) as the mean of the available values, (2) as the mean of the log-transformed values, (3) maximum likelihood estimator (as for the rainfall) of values. The climatology has been computed between 2002 and 2022 on the doy ± 15 days to obtain a smooth curve. The mean of the log-transformed data has been chosen for the analysis as the most representative of the log-normal distribution of chlorophyll-a and turbidity. We point out that, for the majority of the lakes, chlorophyll-a and turbidity have a gap between 2012 and 2016 (see the section on the lake-specific data) which can influence the climatology.

The climatology of the ESA CCI Lakes v2.1 LWL is computed for doy as the mean of the available LWL after removing the annual mean from the timeseries. Such detrending is performed in order to avoid that significant changes in the average annual LWL (as registered in many lakes and reservoirs in our sample, e.g., Lake Chad) impact the shape of climatology. We point out that for a few lakes the climatology could still be affected by long term changes impacting the overall seasonality of LWL variations (e.g., Lake Naivasha and Baringo, see also Supplementary Materials).

The climatology of atmospheric variables from ERA5-Land is computed from the daily time series of the spatially averaged values downloaded from GEE. The mean and the standard deviation therefore refer to the value that the single variable has had in each doy from 1992 to 2022.

Estimation of water colour. Water colour, as the dominant wavelength, is derived from daily spatially averaged LWLR using the hue angle algorithm by Van der Woerd and Wernand, 2018[84], supplemented with coefficients for Sentinel-3 by Ye and Sun, 2022[85]. First, daily time series of dominant wavelengths are evaluated for each lake. Frequency and distribution of these daily observations across all lakes are evaluated using a histogram (see Fig. S16). We define 10 colour classes following Lehmann et al., [86] with a subtle adaptation of the threshold values to our dataset, which presents sparser observations in green and blue and a few orange observations. First, daily dominant wavelength (λ_{dom}) observations are classified as solid colours, with threshold values as blue ($\lambda_{dom} < 495$ nm); green ($495\ \text{nm} \le \lambda_{dom} < 560$ nm); yellow ($560\ \text{nm} \le \lambda_{dom} < 590$ nm); and orange ($\lambda_{dom} \ge 590$ nm). Solid colours (blue, green, yellow, orange) are assigned to a month when ≥ 60% of observations in that month are classified as the respective solid colour. Transient colours are then assigned based on the percentage of observations between solid colour classes in a month. For example, the class green-yellow was assigned when in a month ≥ 30% of observations were green and ≥ 40% of observations were yellow. All binning criteria are listed in Table S.4 in the Supplementary Information. We refer to "colour climatology" as the result of such monthly grouping based on the prevailing colour. Lakes are then clustered based on the predominant colour and the number of unique colours expressed throughout the year.

Estimation of uncertainty for reanalysis and remote sensing global datasets. The global datasets used for estimating the climatological year of atmospheric variables, rainfall and lake-specific variables provide an estimation of the uncertainty at the daily scale.

For the ERA5-Land reanalysis products, ECMWF recommends evaluating the Ensemble of Data Assimilations (EDA) system products for ERA5[42], based on the ten-member ensemble data assimilation system[87] at 3-hourly resolution and at half the ERA5 horizontal resolution. From this, the ensemble mean and spread can be calculated and used to estimate the reanalysis uncertainty. The ensemble spread includes the random uncertainties in assimilated observations and uncertainties in the physical parameterizations. We therefore downloaded the EDA ensemble members for t2m, ssrd, d2m, w, tp, sp from the ERA5 daily aggregated reanalysis (derived-era5-single-levels-daily-statistics) from 1992 to 2022. For each variable we calculated the daily ensemble mean and standard deviation and aggregated them on a monthly basis. Maps of monthly ensemble standard deviation[88] are provided in Supplementary Materials. A direct use of uncertainty values in each pixel is not recommended by ECMWF, however the obtained maps efficiently deliver an idea of where reanalysis products are more or less accurate and how uncertainties follow large-scale atmospheric dynamics across Africa on a climatological basis.

In satellite-based rainfall estimates such as those from TAMSAT and CHIRPS the main source of uncertainty is from the indirect relationship between cloud top temperature and rainfall at the surface, which can be further exacerbated by sparse and unevenly distributed rain gauge data used for calibration and bias correction, as well as challenges in capturing localised rainfall climates that can complicate the cold cloud-rainfall relationship. However, while both TAMSAT and CHIRPS do not

operationally provide uncertainty estimates, both datasets have been extensively evaluated across Africa and demonstrate typically good to very good skill across Africa, indicating they can be used as a reliable indicator of rainfall variability on multiple time and spatial scales. Such studies include Dinku et al. [89] (East Africa), Houngnibo et al. [90] (West Africa), Bamweyana et al. [91] (Uganda), Bagiliko et al. [92] (Ghana and Zambia) and Maidment et al. [44] (Mozambique, Niger, Nigeria, Uganda, and Zambia). Consequently, such datasets are widely used in many operational applications across multiple sectors across Africa (for example, climate, agriculture and finance).

The ESA CCI Lakes v2.1 provides detailed documentation[93] on the uncertainty associated with the variables included in the dataset, which have been thoroughly validated through comparison with in-situ measurements[34]. Uncertainties come along with lake variables as pixel-by-pixel quantities for each lake at up to a daily scale. The uncertainty in LSWT is estimated by accounting for propagated, retrieval, and sampling uncertainties and it includes all sources of uncertainty (sensor errors, modelling errors, prior errors, retrieval indeterminacy and data sampling for gridding) expected for a valid retrieval of LSWT. The propagated uncertainty reflects how errors in satellite observations are amplified during the retrieval process. Retrieval uncertainty represents the range of LSWT values that could reasonably match the satellite observations, due to ambiguity introduced by the atmosphere between the lake surface and the satellite sensor. This component also includes the effect of the prior value used in the retrieval process (optimal estimation). Sampling uncertainty, relevant during the gridding process, occurs when only a portion of the lake within a grid cell is visible or measurable. Further, quality levels are associated with per-pixel LSWT and they are an indication of the confidence of the retrieved value of LSWT and its uncertainty estimation.

Uncertainty for chlorophyll-a and turbidity is determined for each algorithm underlying the dynamic weighted blending of algorithms based on optical water types pre-classification. The characterisation of uncertainty of each algorithm and optical water type is carried out against in situ bio-optical data sets as part of Lakes_cci product development. The per-pixel uncertainty estimates included with the Lakes_cci dataset extrapolate from these uncertainty models and are, therefore, not specific to the observation conditions or a specific lake - they only describe the anticipated performance for the specific combination of algorithm weights determined for each observation. It is possible for lake-specific observation bias to be under- or over-stated in this estimate. Such biases are expected to be relatively systematic within the confines of a single water body, thus supporting trend analysis even if uncertainty estimates are of the same order of magnitude as the trend. Typically, algorithm selections perform worse in the vicinity of lake shores and these may be observed to have higher per-pixel uncertainties based on the abovementioned global uncertainty patterns, linked with specific optical water types and associated algorithm performance. This should not be confused with observation-specific uncertainties, as these are not presently reported. Algorithms associated with different concentration ranges may also be associated with different degrees of uncertainty. Therefore, to avoid introducing sampling bias (e.g., seasonal), data filtering based on uncertainty estimates is not recommended, whereas several quality flags associated with observation conditions and algorithms operating outside of their training scope, should be observed.

LWL uncertainty is quantified as the standard deviation of the data used to derive LWL for each satellite overpass, reflecting the variability associated with the altimetric retrieval. This uncertainty largely depends on the size of the lake and the transect of the satellite pass and typically ranges from a few centimetres to a few decimetres, with higher values observed for narrow or complex-shaped lakes. Intercomparison with some available in situ lake level in North and South America and in Europe, for small to large lakes, have shown that the uncertainty provided in the CCI dataset are in good agreement with the uncertainty of ground measurements, with no identified seasonal variability. Although uncertainty can vary on a lake-by-lake basis, the validation of the dataset as provided and commented in Carrea et al. [34] makes us relatively confident that uncertainty on this variable generally remains within negligible to the purpose of this study.

We extracted uncertainty values for sub-Sahelian lakes on a monthly basis for LSWT, Chlorophyll-a and Turbidity. Maps of monthly median values (reported in the Supplementary Materials, Figs. S11-S13) show that the uncertainties are approximately 0.5-1 K for LSWT, 60–70% for Turbidity, 40-50% for Chlorophyll-a. LSWT uncertainty clearly depends on the season, with larger values during the rainy season due to cloud cover, while uncertainty of water quality variables shows limited seasonal variation. When compared with the interannual variability of each parameter, quantified as the standard deviation around monthly means (Fig. S14), these uncertainties generally fall within or below the expected natural variability range.

Principal component analysis. A Principal Component Analysis (PCA) is performed across all climatic variables to gather a standardised description of the seasons and dominant weather features in each lake. The PCA is performed on each lake singularly. The features x_i are the climatologies of four ERA5-Land variables (air temperature and pressure, shortwave radiation and relative humidity, see the above section for a detailed description of the variables) and rainfall from either ERA5-Land, TAMSAT or CHIRPS datasets. The contribution of rainfall from, TAMSAT and CHIRPS are evaluated in three different PCAs, such that the number of features $p = 5$ for each test. Only results including TAMSAT are reported in the results section.

Each feature x_i is a time-series composed by $n = 365$ samples (29th of February is removed from the computation) and is normalised by removing its mean value μ_i and dividing by its standard deviation σ_i over the n values as in Eq. 1.

$$x_i' = \frac{x_i - \mu_i}{\sigma_i} \tag{1}$$

The data matrix X for the PCA is therefore composed by appending the normalised features in a matrix with $i = 1, \ldots, p$ columns (features) and $t = 1, \ldots, n$ rows (samples), as in Eq. (2).

$$X = \left[x'_1, \ldots, x'_i, \ldots, x'_p\right] \tag{2}$$

The covariance matrix C is then computed as in Eq. (3) as a matrix p x p. Its eigenvectors w_i are p vectors with p values representing the direction of the i-th principal component, while its eigenvalues λ_i are scalars quantifying the amount of variance explained by the i-th principal component.

$$\begin{aligned} C &= \tfrac{1}{n-1} X^T X \\ Cw_i &= \lambda_i w_i \end{aligned} \tag{3}$$

A matrix W is constructed by appending the eigenvectors w_i in descending order according to the value of their respective eigenvalue λ_i (Eq. 4). As for almost all lakes the variance explained by the first two components $(\lambda_1 + \lambda_2)$ exceeds 80% of the total variance, only two principal components are retained.

$$W = \left[w_{i,1}, w_{i,2}\right] \tag{4}$$

From now on, we will refer to the principal components with the index $j = 1, 2$.

The amount of variance of the i-th feature explained by the j-th principal components (e.g., displayed for all involved features in panel h in Fig. 3) is carried by the terms $w_{i,j}$ stored in the eigenvector's matrix W (also known with the term loadings).

The time series of PC1 and PC2 are obtained by projecting the original data matrix X on the principal component space W as in Eq. (5). Thus, PC_1 and PC_2 represent the transformed data into the directions of maximum

variance.

$$PC = X \cdot W \tag{5}$$

To obtain comparable results across the different lake PCAs, we first identify what PC_j holds the most variance of each feature. To this aim, we compute a score $a_{i,j}$ based on the loading of the i-th feature on the j-th principal components as in Eq. (6):

$$a_{i,j} = \frac{\left|w_{i,j}\right|}{\left|w_{i,1}\right| + \left|w_{i,2}\right|} (\text{with } j = 1, 2) \tag{6}$$

where $w_{i,1} + w_{i,2}$ is the total variance of the feature explained by PC_1 and PC_2. The module is applied to the loadings in order to account for their arbitrary sign on PC_1 and PC_2.

We consider one variable to be well described by the j-th principal component when more than 55% of its total explained variance is captured by PC_j ($a_{i,j} > 0.55$). We then list the variables satisfying the above criterion and classify the main meteorological variable to be associated to each PC. Since in each lake features can group differently along PC_1 and PC_2, our classification is based on a first distinction between variables associated with the rainy season (rainfall, relative humidity) and those associated with the dry season (shortwave radiation mostly, representative of the radiative energy coming from the sun). A third group of variables (air pressure, temperature, wind speed) can both align with one of the two seasons-defining PCs or be completely independent on that and define their own PC. The class assignment logic is therefore based on the definition of four main classes: (1) rain, (2) swdown (abbreviation for shortwave solar radiation), (3) rain and swdown, (4) wind, where the:

1. "rain" class is attributed to PC_j when the list of variables with score $a_{i,j} > 0.55$ does not include shortwave radiation and includes rainfall or relative humidity;
2. "swdown" class is attributed to PC_j when the list of variables with score $a_{i,j} > 0.55$ does not include rainfall but includes shortwave radiation or air temperature or air pressure;
3. "rain and swdown" class is attributed to PC_j when the list of variables with score $a_{i,j} > 0.55$ includes both rainfall and shortwave radiation.
4. "wind" class is attributed to PC_j when the list of variables with score $a_{i,j} > 0.55$ does not include rain nor shortwave radiation and includes wind or atmospheric pressure.

Correlation analysis between water quality variables and PCs

Once the complexity of weather variables interacting with one-another is reduced to two main climatic drivers, we compute the linear correlation between the obtained PCs and the climatologies of the lake variables (i.e., LSWT, chlorophyll-a, turbidity, LWL). The value of the Pearson correlation coefficient r is plotted in the vector space of PC1 and PC2 to visually identify which of the two eigenvectors shows the largest similarity with the lake variables (see e.g., Fig. 4a). We search for the optimal lag that maximises $|r|$ between lake variables and the PCs and allow a maximum lag of 120 days for LWL, 60 days for all other variables. The p-value is also computed, and results with p-value > 0.05 are ignored.

We cluster the lakes based on what PC shows the largest correlation with each lake variable. When $|r_{j,k}|>0.5$, we consider PC_j as correlated with the lake variable k. We set 0.2 as the minimum threshold for significant (albeit weak) correlation. Correlation with both PC_1 and PC_2 is allowed and displayed in Fig. 4d–f) as bi-colour dots.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data used in this work have been downloaded from publicly available datasets. Lakes_cci: https://catalogue.ceda.ac.uk/uuid/7fc9df8070d34cacab8092e45ef276f1/ ERA5-LAND: https://developers.google.com/earth-engine/datasets/catalogue/ECMWF_ERA5_LAND_DAILY_AGGR?hl=it Hydrolakes: https://www.hydrosheds.org/products/hydrolakes TAMSAT: https://research.reading.ac.uk/tamsat/data-access/ CHIRPS: https://www.chc.ucsb.edu/data Lake-TopoCat: https://doi.org/10.5281/zenodo.7916729 Population density: https://hub.worldpop.org/geodata/listing?id=78 JRC Africa Knowledge Platform: https://africa-knowledge-platform.ec.europa.eu/explore_maps JRC World Atlas of Desertification: https://wad.jrc.ec.europa.eu/ Per-lake results obtained via post-processing of public data can be found at the following links: https://gws-access.jasmin.ac.uk/public/cds_c3s_lakes/CCI_African_lakes_paper/htmltab1/African_Lakes_List_Tab1.html https://gws-access.jasmin.ac.uk/public/cds_c3s_lakes/CCI_African_lakes_paper/htmltab2_climatologies/African_Lakes_List_Tab2.html https://gws-access.jasmin.ac.uk/public/cds_c3s_lakes/CCI_African_lakes_paper/htmltab3_results/African_Lakes_List_Tab3.html.

Received: 3 February 2025; Accepted: 7 August 2025;
Published online: 20 August 2025

References

1. Tranvik, L. J. et al. Lakes and reservoirs as regulators of carbon cycling and climate. *Limnol. Oceanogr.* **54**, 2298–2314 (2009).
2. Williamson, C. E., Saros, J. E., Vincent, W. F. & Smol, J. P. Lakes and reservoirs as sentinels, integrators, and regulators of climate change. *Limnol. Oceanogr.* **54**, 2273–2282 (2009).
3. Woolway, R. I. et al. Global lake responses to climate change. *Nat. Rev. Earth Environ.* **1**, 388–403 (2020).
4. Heino, J. et al. Lakes in the era of global change: moving beyond single-lake thinking in maintaining biodiversity and ecosystem services. *Biol. Rev.* **96**, 89–106 (2021).
5. Sterner, R. W. et al. Ecosystem services of Earth's largest freshwater lakes. *Ecosyst. Serv.* **41**, 101046 (2020).
6. Inácio, M., Barceló, D., Zhao, W. & Pereira, P. Mapping lake ecosystem services: A systematic review. *Sci. Total Environ.* **847**, 157561 (2022).
7. Kummu, M., Moel, Hde, Ward, P. J. & Varis, O. How close do we live to water? A global analysis of population distance to freshwater bodies. *PLOS ONE* **6**, e20578 (2011).
8. Danaher, C., Newbold, T., Cardille, J. & Chapman, A. S. A. Prioritizing conservation in sub-Saharan African lakes based on freshwater biodiversity and algal bloom metrics. *Conserv. Biol.* **36**, e13914 (2022).
9. Kafumbata, D., Jamu, D. & Chiotha, S. Riparian ecosystem resilience and livelihood strategies under test: lessons from Lake Chilwa in Malawi and other lakes in Africa. *Philos. Trans. R. Soc. B Biol. Sci.* **369**, 20130052 (2014).
10. McClain, M. E. Balancing water resources development and environmental sustainability in Africa: A review of recent research findings and applications. *AMBIO* **42**, 549–565 (2013).
11. Chapman, C. A. et al. The future of sub-Saharan Africa's biodiversity in the face of climate and societal change. *Front. Ecol. Evol.* **10**, 790552 (2022).
12. Santos, S. et al. Urban growth and water access in sub-Saharan Africa: Progress, challenges, and emerging research directions. *Sci. Total Environ.* **607–608**, 497–508 (2017).
13. Niang, I. et al. *Africa. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. (Cambridge University Press, 2014).
14. Nyamweya, C. S. et al. A century of drastic change: Human-induced changes of Lake Victoria fisheries and ecology. *Fish. Res.* **230**, 105564 (2020).
15. Savelli, E., Mazzoleni, M., Di Baldassarre, G., Cloke, H. & Rusca, M. Urban water crises driven by elites' unsustainable consumption. *Nat. Sustain.* **6**, 929–940 (2023).

16. Chen, S. S., Kimirei, I. A., Yu, C., Shen, Q. & Gao, Q. Assessment of urban river water pollution with urbanization in East Africa. *Environ. Sci. Pollut. Res. Int.* **29**, 40812–40825 (2022).
17. Sanitation and Wastewater Atlas of Africa | GRID-Arendal. https://www.grida.no/publications/471.
18. Takyi, S. A. et al. Urbanization against ecologically sensitive areas: effects of land use activities on surface water bodies in the Kumasi Metropolis. *Int. J. Urban Sustain. Dev.* **14**, 460–479 (2022).
19. Jones, E. R. et al. Sub-Saharan Africa will increasingly become the dominant hotspot of surface water pollution. *Nat. Water* **1**, 602–613 (2023).
20. Nhamo, G. & Mutanda, G. W. Impact of climate change on Africa's major lakes: a systematic review incorporating pathways of enhancing climate resilience. *Front. Water* **6**, https://doi.org/10.3389/frwa.2024.1443989 (2024).
21. Ogutu-Ohwayo, R., Natugonza, V., Musinguzi, L., Olokotum, M. & Naigaga, S. Implications of climate variability and change for African lake ecosystems, fisheries productivity, and livelihoods. *J. Gt. Lakes Res.* **42**, 498–510 (2016).
22. Muringai, R. T., Mafongoya, P. L. & Lottering, R. Climate change and variability impacts on Sub-Saharan African fisheries: a review. *Rev. Fish. Sci. Aquac.* **29**, 706–720 (2021).
23. O'Reilly, C. M., Alin, S. R., Plisnier, P.-D., Cohen, A. S. & McKee, B. A. Climate change decreases aquatic ecosystem productivity of Lake Tanganyika, Africa. *Nature* **424**, 766–768 (2003).
24. Woolway, R. I. & Merchant, C. J. Worldwide alteration of lake mixing regimes in response to climate change. *Nat. Geosci.* **12**, 271–276 (2019).
25. Hecky, R. E. et al. Deoxygenation of the deep water of Lake Victoria, East Africa. *Limnol. Oceanogr.* **39**, 1476–1481 (1994).
26. Rugema, E. et al. Long-term change of phytoplankton in Lake Kivu: The rise of the greens. *Freshw. Biol.* **64**, 1940–1955 (2019).
27. Huang, L. et al. Emergence of lake conditions that exceed natural temperature variability. *Nat. Geosci.* **17**, 763–769 (2024).
28. Overland, I. et al. Funding flows for climate change research on Africa: where do they come from and where do they go? *Clim. Dev.* **14**, 705–724 (2022).
29. Trisos, C. H. et al. *Africa. In: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. (Cambridge University Press, 2023). https://doi.org/10.1017/9781009325844.011.
30. Bootsma, H. A. & Hecky, R. E. A comparative introduction to the biology and limnology of the African Great Lakes. *J. Gt. Lakes Res.* **29**, 3–18 (2003).
31. Deirmendjian, L. et al. Limnological changes in Lake Victoria since the mid-20th century. *Freshw. Biol.* **66**, 1630–1647 (2021).
32. Anghileri, D. et al. Global water challenges in Sub-Saharan Africa and how to strengthen science-policy dialogues on transboundary governance and cooperation. *J. Environ. Manage.* **365**, 121417 (2024).
33. Calamita, E., Lever, J. J., Albergel, C., Woolway, R. I. & Odermatt, D. Detecting climate-related shifts in lakes: A review of the use of satellite Earth Observation. *Limnol. Oceanogr.* **69**, 723–741 (2024).
34. Carrea, L. et al. Satellite-derived multivariate world-wide lake physical variable timeseries for climate studies. *Sci. Data* **10**, 30 (2023).
35. Papa, F. et al. Water resources in Africa under global change: monitoring surface waters from space. *Surv. Geophys.* **44**, 43–93 (2023).
36. Sogno, P., Klein, I., Uereyen, S., Bachofer, F. & Kuenzer, C. Surface water dynamics of Africa: Analysing continental trends and identifying drivers for major lakes and reservoirs. *Int. J. Remote Sens.* **45**, 9538–9568 (2024).
37. Dube, T., Mutanga, O., Seutloali, K., Adelabu, S. & Shoko, C. Water quality monitoring in sub-Saharan African lakes: a review of remote sensing applications. *Afr. J. Aquat. Sci.* **40**, 1–7 (2015).
38. Gebreegziabher, G. A., Degefa, S. & Furi, W. A review of the shrinking and expanding Eastern Africa rift valley lakes: The case of Ethiopian and Kenyan lakes. *J. Hydrol. Reg. Stud.* **54**, 101909 (2024).
39. Irani Rahaghi, A. et al. Combined Earth observations reveal the sequence of conditions leading to a large algal bloom in Lake Geneva. *Commun. Earth Environ.* **5**, 1–12 (2024).
40. Adrian, R. et al. Lakes as sentinels of climate change. *Limnol. Oceanogr.* **54**, 2283–2297 (2009).
41. Beck, H. E. et al. High-resolution (1 km) Köppen-Geiger maps for 1901–2099 based on constrained CMIP6 projections. *Sci. Data* **10**, 724 (2023).
42. Muñoz-Sabater, J. et al. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* **13**, 4349–4383 (2021).
43. Funk, C. et al. The climate hazards infrared precipitation with stations —a new environmental record for monitoring extremes. *Sci. Data* **2**, 150066 (2015).
44. Maidment, R. I. et al. A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa. *Sci. Data* **4**, 170063 (2017).
45. Nicholson, S. E. The ITCZ and the seasonal cycle over Equatorial Africa. *Bull. Am. Meteorol. Soc.* **99**, 337–348 (2018).
46. Baker, A. et al. Rainfall recharge thresholds in a subtropical climate determined using a regional cave drip water monitoring network. *J. Hydrol.* **587**, 125001 (2020).
47. Scanlon, B. R. et al. Linkages between GRACE water storage, hydrologic extremes, and climate teleconnections in major African aquifers. *Environ. Res. Lett.* **17**, 014046 (2022).
48. Seka, A. M. et al. Spatio-temporal analysis of water storage variation and temporal correlations in the East Africa lake basins. *J. Hydrol. Reg. Stud.* **41**, 101094 (2022).
49. Geen, R., Bordoni, S., Battisti, D. S. & Hui, K. Monsoons, ITCZs, and the concept of the global monsoon. *Rev. Geophys.* **58**, e2020RG000700 (2020).
50. Jain, S., Mishra, S. K., Anand, A., Salunke, P. & Fasullo, J. T. Historical and projected low-frequency variability in the Somali Jet and Indian Summer Monsoon. *Clim. Dyn.* **56**, 749–765 (2021).
51. Suzuki, T. Seasonal variation of the ITCZ and its characteristics over central Africa. *Theor. Appl. Climatol.* **103**, 39–60 (2011).
52. Thiery, W. et al. The impact of the African Great Lakes on the regional climate. *J. Climate* **28**, 4061–4085 (2015).
53. Calamita, E., Vanzo, D., Wehrli, B. & Schmid, M. Lake modeling reveals management opportunities for improving water quality downstream of transboundary tropical dams. *Water Resour. Res.* **57**, e2020WR027465 (2021).
54. Cherlet, M. et al. *World Atlas of Desertification*. (Publications Office of the European Union, Luxembourg, 2018). https://doi.org/10.2760/06292.
55. Major, Y., Kifle, D., Spoof, L. & Meriluoto, J. Cyanobacteria and microcystins in Koka reservoir (Ethiopia). *Environ. Sci. Pollut. Res* **25**, 26861–26873 (2018).
56. Zinabu, G.-M. The effects of wet and dry seasons on concentrations of solutes and phytoplankton biomass in seven Ethiopian rift-valley lakes. *Limnologica* **32**, 169–179 (2002).
57. Ngochera, M. J. & Bootsma, H. A. Temporal trends of phytoplankton and zooplankton stable isotope composition in tropical Lake Malawi. *J. Gt. Lakes Res.* **37**, 45–53 (2011).
58. Sarmento, H., Isumbisho, M. & Descy, J.-P. Phytoplankton ecology of Lake Kivu (eastern Africa). *J. Plankton Res.* **28**, 815–829 (2006).
59. Moshi, H. A. et al. Community monitoring of coliform pollution in Lake Tanganyika. *PLOS ONE* **17**, e0262881 (2022).
60. Ndebele-Murisa, M., Charles, F. M. & Lincoln, R. A review of phytoplankton dynamics in tropical African lakes. *South Afr. J. Sci.* **106**, 13–18 (2010).
61. Chavula, G. M. S. et al. Lake Malawi/Niassa/Nyasa basin: Status, challenges, and research needs. *J. Gt. Lakes Res.* **49**, 102241 (2023).

62. Plisnier, P.-D. et al. Need for harmonized long-term multi-lake monitoring of African Great Lakes. *J. Gt. Lakes Res.* **49**, 101988 (2023).
63. Bergamino, N. et al. Examining the dynamics of phytoplankton biomass in Lake Tanganyika using Empirical Orthogonal Functions. *Ecol. Model.* **204**, 156–162 (2007).
64. Viner, A. B. Hydrobiology of Lake Volta, Ghana. *Hydrobiologia* **35**, 209–229 (1970).
65. Tay, C. K. Integrating water quality indices and multivariate statistical techniques for water pollution assessment of the Volta Lake, Ghana. *Sustain. Water Resour. Manag.* **7**, 71 (2021).
66. Appiah Boamah, L., Nyamekye, C., Gyamfi, C., Ballard, J. Q. & Anornu, G. K. Mapping and estimating water quality parameters in the Volta Lake's Kpong Headpond of Ghana using regression model and Landsat 8. *Cogent Eng* **11**, 2307165 (2024).
67. Pham-Duc, B. et al. The Lake Chad hydrology under current climate change. *Sci. Rep.* **10**, 5498 (2020).
68. Caroni, R. et al. Investigating the impact of wildfires on lake water quality using earth observation satellites. *Appl. Sci.* **14**, 2626 (2024).
69. Mittermaier, D. et al. *The Climate Conflict Vulnerability Index (CCVI) - Technical Documentation v1.1*. (2024).
70. Assegide, E. et al. Spatiotemporal dynamics of water quality indicators in Koka Reservoir, Ethiopia. *Remote Sens* **15**, 1155 (2023).
71. Horion, S. et al. Optimized extraction of daily bio-optical time series derived from MODIS/Aqua imagery for Lake Tanganyika, Africa. *Remote Sens. Environ.* **114**, 781–791 (2010).
72. Calamita, E. et al. Sixty years since the creation of Lake Kariba: Thermal and oxygen dynamics in the riverine and lacustrine sub-basins. *PLOS ONE* **14**, e0224679 (2019).
73. Nakkazi, M. T., Nkwasa, A., Martínez, A. B. & van Griensven, A. Linking land use and precipitation changes to water quality changes in Lake Victoria using earth observation data. *Environ. Monit. Assess.* **196**, 1104 (2024).
74. Woolway, R. I. et al. Multivariate extremes in lakes. *Nat. Commun.* **15**, 4559 (2024).
75. European Commission. Africa Knowledge Platform. *European Commission Africa Knowledge Platform* https://africa-knowledge-platform.ec.europa.eu/ (2024).
76. Carrea, L., Merchant, C. & Simis, S. Lake mask and distance to land dataset of 2024 lakes for the European Space Agency Climate Change Initiative Lakes v2. Zenodo https://doi.org/10.5281/zenodo.6699376 (2022).
77. Lawrence, M. G. *The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications*. (2005) https://doi.org/10.1175/BAMS-86-2-225.
78. Washington, R. et al. African Climate Change: Taking the Shorter Route. *Bull. Am. Meteorol. Soc*. **87**, 1355–1366 (2006).
79. Gbetkom, P. G. et al. Lake Tanganyika basin water storage variations from 2003–2021 for water balance and flood monitoring. *Remote Sens. Appl. Soc. Environ.* **34**, 101182 (2024).
80. Sikder, M. S. et al. Lake-TopoCat: a global lake drainage topology and catchment database. *Earth Syst. Sci. Data* **15**, 3483–3511 (2023).
81. Messager, M. L., Lehner, B., Grill, G., Nedeva, I. & Schmitt, O. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nat. Commun.* **7**, 13603 (2016).
82. Harper, K. L. et al. A 29-year time series of annual 300 m resolution plant-functional-type maps for climate models. *Earth Syst. Sci. Data* **15**, 1465–1499 (2023).
83. Bondarenko, M., Kerr, D., Sorichetta, A. & Tatem, A. & WorldPop,. Census/projection-disaggregated gridded population datasets, adjusted to match the corresponding UNPD 2020 estimates, for 51 countries across sub-Saharan Africa using building footprints. *University of Southampton* https://doi.org/10.5258/SOTON/WP00683 (2020).
84. Van der Woerd, H. J. & Wernand, M. R. Hue-angle product for low to medium spatial resolution optical satellite sensors. *Remote Sens* **10**, 180 (2018).
85. Ye, M. & Sun, Y. Review of the Forel–Ule Index based on in situ and remote sensing methods and application in water quality assessment. *Environ. Sci. Pollut. Res.* **29**, 13024–13041 (2022).
86. Lehmann, M. K., Nguyen, U., Allan, M. & Van der Woerd, H. J. Colour Classification of 1486 Lakes across a Wide Range of Optical Water Types. *Remote Sens* **10**, 1273 (2018).
87. Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
88. ECMWF. Section 8.1.2 ENS Mean and spread - Forecast User Guide - ECMWF Confluence Wiki. https://confluence.ecmwf.int/display/FUG/Section+8.1.2+ENS+Mean+and+Spread.
89. Dinku, T. et al. Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Q. J. R. Meteorol. Soc.* **144**, 292–312 (2018).
90. Houngnibo, M. C. M. et al. Validation of high-resolution satellite precipitation products over West Africa for rainfall monitoring and early warning. *Front. Clim.* **5**, https://doi.org/10.3389/fclim.2023.1185754 (2023).
91. Bamweyana, I., Musinguzi, M. & Kayondo, L. M. Evaluation of CHIRPS satellite gridded dataset as an alternative rainfall estimate for localized modelling over Uganda. *Atmos. Clim. Sci.* **11**, 797–811 (2021).
92. Bagiliko, J., Stern, D., Ndanguza, D. & Torgbor, F. F. Validation of satellite and reanalysis rainfall products against rain gauge observations in Ghana and Zambia. *Theor. Appl. Climatol.* **156**, 241 (2025).
93. Simis, S. et al. *Lakes_CCI+ - Phase 2, D2.3. End-to-End ECV Uncertainty Budget (E3UB)*. (2023).
94. Carioli, A., Schiavina, M., Freire, S. & MacManus, K. GHS-POP R2023A - GHS population grid multitemporal (1975-2030). (2023) https://doi.org/10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE.
95. Scopus. Data retrieved using an automated script via Scopus database. https://www.scopus.com.

Acknowledgements

This work is supported by the Lakes_cci (contract n. 40000125030/18/I-NB) project which is part of the ESA Climate Change Initiative. The authors express their sincere gratitude to Sebastiano Piccolroaz, Marco Toffolon, Michael Matiu, Henk Dijkstra and Alessandro Chesini for insightful discussions on methods and results. Special thanks go to Rafa Marcè and Joachim Jansen for their inspiring feedback on the objectives and challenges of global approaches to limnology, and to the Climate Journal Club of the University of Trento for stimulating cross-disciplinary conversations about climate change. The authors are also grateful to the anonymous reviewers for their constructive feedback, which contributed to the improvement and significance of this work.

Author contributions

Conceptualization: M.A. and L.C. Methodology: M.A., A.J.G., L.C., M.P., R.M., C.G. Validation: M.A., L.C., R.C., E.C., M.N., T.N., S.B., D.L., M.S. Formal analysis: M.A. and A.J.G. Investigation: M.A., A.J.G., L.C., M.P., R.C., S.B. Resources: L.C., J.F.C., S.S., X.L., H.Y., C.A. Visualization: M.A. and A.J.G. Supervision: C.G., M.P., M.B., E.C., RIW Writing—original draft: M.A., A.J.G., L.C., L.S. Writing—review and editing: M.A., A.J.G., L.C., M.P., R.C., E.C., L.S., R.M., S.B., C.G., M.B., F.P.F., M.S., M.N., T.N., J.F.C., C.J.M., X.L., S.S., D.L., H.Y., C.A., R.I.W. All authors contributed critically to the drafts and gave final approval for publication.

Competing interests

The authors declare no competing interests.

Ethics

This research was conducted in collaboration with local researchers. Roles and responsibilities were agreed upon in advance and the study's relevance to local contexts was discussed, ensuring that the research outcomes are meaningful and beneficial. Relevant local and regional research was considered and cited to ensure appropriate acknowledgment of prior work.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-025-02684-5.

Correspondence and requests for materials should be addressed to M. Amadori.

Peer review information *Communications Earth & Environment* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Haihan Zhang and Alireza Bahadori. A peer review file is available.