Long-term algal blooms mapping and crucial driving factors analysis based on Landsat series imagery in Lake Victoria (2001-2021)

Daiqi Zhong[1,2], Yi Lin[1,2], Jie Yu[1,2,*], Chen Gao[1,2], Lin He[1,2], Yufei Song[1,2], Yuxuan Yang[1,2], Xin Chen[1,2]

[1] College of Surveying and Geo-Informatics, Tongji University, 200092 Shanghai, China – (2410271, linyi, 2011_jieyu, gaochen, 2010964, 2333591, yyx2000, chenxin)@tongji.edu.cn
[2] Research Centre for Remote Sensing Technology and Application, Tongji University, 200092 Shanghai, China

Keywords: Algal blooms, Spatiotemporal distribution, Long-term trend, Driving factor analysis, Lake Victoria

Abstract

Algal blooms constitute an emerging threat to global inland water quality. As one of the biggest lake and important water resource in the world, Lake Victoria is facing recurrent proliferation of water hyacinth and cyanobacteria. To better manage and improve water resources, the spatiotemporal distribution and long-term trends of algal blooms must be understood, as well as the driving factors. In this study, we used more than 20 years of Landsat series images to extract and map algal bloom occurrences based on a neural network model in the Transform framework and to revel the long-term changes law and trends of cyanobacterial. Results showed that both the bloom occurrence frequency (BOF) and the affected areas exhibited an increasing trend with the rates of 2.87%·yr^{-1} and 981 $km^2 \cdot yr^{-1}$, respectively. Some crucial driving factors were selected to analyze climatic and anthropogenic impact on bloom occurrences. For meteorological factors, lake water volume has shown positive correlation with BOF (r equals 0.58), while precipitation is positively correlated with BOF variations (r equals 0.57), with 1-year lag. The annual precipitation range that contributes to BOF increase in Lake Victoria was estimated to be approximately 98-118 $km^3 \cdot yr^{-1}$. For anthropogenic factors, socio-economic development, expansion of built-up areas and croplands, and decrease in ecological land areas, such as wetland and grassland, largely contributed to the BOF increase in Lake Victoria. Based on the above results, the degree of influence of the factors was analyzed using grey relational analysis (GRA), with lake water volume and socio-economic being the predominant driving factors. This study provides valuable insights into the long-term algal bloom occurrence dynamics in Lake Victoria, and could provide important data support for the ecological safety and sustainable use of the lake.

1. Introduction

Water is the fundamental in sustaining life and achieving sustainable society, as mentioned in United Nations' Sustainable Development Goals (UN, 2015). Inland lakes, as the key component of water resources, play a vital role in ecological balance (Lin et al., 2024). However, over 60% of global lakes have been facing the eutrophication situation with the algal blooms emerging persistently (Wang et al., 2018; Hou et al., 2022). The persistent occurrence of algal blooms can lead to oxygen depletion in the water (Paerl et al., 2016), deteriorating water quality (Amorim et al., 2021), and thus negatively affecting fisheries and the safety of drinking water for residents (Treuer et al., 2021).

Excessive inputs of nitrogen and phosphorus are considered the main causes of algal blooms (Lin et al., 2024), with these nutrients primarily originating from human activities such as agricultural runoff, untreated sewage discharge, and industrial wastewater (Khan and Mohammad, 2014). Some studies indicated that climate change may also directly or indirectly promote the occurrence of algal blooms (Paerl et al., 2016; Ojok et al., 2017; Ho et al., 2019). Additionally, changes in precipitation patterns and the increasing frequency of extreme weather events contribute to the further eutrophication of the lake (Amanullah et al., 2020; Ojok et al., 2017). Researches have demonstrated that surface runoff caused by precipitation contributes to an increased input of nutrients into water bodies, providing essential nutrients for algal growth (Michalak et al. 2013). Additionally, the precipitation induces vertical mixing between the upper and lower water layers of lakes, enhancing dissolved oxygen levels and thereby promoting algal proliferation (Liu et al., 2020). The combined effects of climate change and human activities are accelerating the degradation of the lake's ecosystem (Akurut et al., 2014; Baltodano et al., 2022), leading to more frequent and intense algal blooms. Therefore, understanding the impact of climate change and human activities on the frequency and intensity of algal blooms is crucial for formulating effective environmental protection measures.

Lake Victoria, as the largest freshwater lake in Africa and the second-largest freshwater lake in the world, provides various ecosystem services to countries like Tanzania, Uganda, and Kenya, including drinking water, fisheries resources, agricultural irrigation, and transportation (Downing et al., 2014). Due to its unique geographical location straddling the equator and its vast water surface area, Lake Victoria not only has a significant impact on the regional ecological environment but has also become a key focus for global water resource conservation and research (Tungaraza et al., 2012). Like many other inland lakes, Lake Victoria has been affected by eutrophication over a prolonged period (Gidudu et al., 2021), facing increasingly severe algal blooms (Olokotum et al., 2020). However, the issue of algal blooms has been poorly investigated on the African continent, so as Lake Victoria (Svirčev et al., 2019). Most of the reviewable researches have focused on the bays & gulfs in Kenya as a result of the high frequency of blooms in this region (Olokotum et al., 2020), while few studies have been conducted on lake-wide bloom spatial and temporal distributions and patterns.

Compared to traditional ground-based water quality monitoring, remote sensing has the advantages of wide spatial coverage, high temporal resolution, and access to long-term time series data (Lin et al., 2024). This allows researchers to conduct large-scale, long-term monitoring of the lakes, capturing the spatiotemporal distribution and trends of algal blooms. To effectively monitor and manage the algal blooms in Lake Victoria, remote sensing technology, particularly using Landsat-series satellite imagery, offers a powerful tool. With a spatial resolution of 30 meters and a data record spanning over 40 years, the Landsat series provides an ideal data source for studying long-term algal bloom and their environmental drivers.

This study aims to extract the spatiotemporal distribution changes of algal blooms in Lake Victoria over a 21-year period from 2001 to 2021 using long-term Landsat satellite imagery.

Furthermore, we analyzed the factors impact on algal blooms by integrating meteorological data and human activity data. We tried to examine two questions in this study: (1) what is the spatiotemporal frequency of the algal blooms occurred in Lake Victoria? (2) What is the extent of the meteorological and anthropogenic impact on algal blooms?

2. Study Area and Materials

2.1 Study Area

Lake Victoria is located in East Africa (31°39′E–34°53′E, 0°20′N–3°S), stretching approximately 400 km from north to south and about 320 km from east to west, crossing the equator (Figure 1). With a total shoreline length of around 3500 km, it is the second-largest freshwater lake in the world and the largest freshwater lake situated in developing countries and tropical regions. The lake is shared by Kenya (6%), Tanzania (51%) and Uganda (43%). Lake Victoria not only serves as a crucial water body for the survival of several African nations but is also a key subject of global water resource monitoring efforts (Nyamweya et al., 2023).

The climate of the Lake Victoria basin varies depending on the region. In general, the area experiences two distinct rainy seasons, the long rains from March to May and the short rains from October to December. Precipitation levels vary significantly, from 870 to 1,561 mm in Uganda and from 400 to 2,736 mm in Tanzania (Kizza et al., 2009). Water temperatures in Lake Victoria typically range from 23°C to 29°C throughout the year (Muggidde et al., 2005), and are less affected by seasonal variations, with the warm waters creating favorable conditions for algal growth. The population around Lake Victoria primarily engages in agriculture and fishing. The expansion of agricultural activities, coupled with rapid urbanization, has increased the load of organic pollutants and nutrients entering the lake, contributing to its eutrophication (Osoro et al., 2016; Olokotum et al., 2020). Overfishing and aquaculture have also disrupted the ecological balance of the lake, indirectly affecting the proliferation of algae (Egessa et al., 2018; Nyamweya et al., 2023).

In this study, 1 degree buffer zone is extracted for attribution analysis of climate change and human activities (Lin et al., 2023).

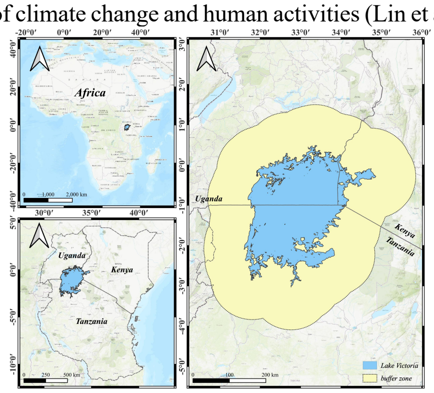

Figure 1. Lake Victoria and 1 degree buffer zone.

2.2 Research Data

Landsat-series satellite images (TM/ETM+/OLI) with a total of 8657 between the observing period 2001-2021 were used to extract algal blooms in Lake Victoria, and is available from the US Geological Survey (http://earthexplorer.usgs.gov) and through Google Earth Engine (https://earthengine.google.com). De-cloud operation was performed using Fmask algorithm (Zhu & Woodcock, 2012) after the image surface reflections were obtained to remove the cloud interference.

Two types of potential impact drivers were selected in this study, including meteorological and anthropogenic factors (Table 1). For the meteorological factors, the lake surface water temperature (LSWT) (Carrea et al., 2024) and precipitation (https://daac.gsfc.nasa.gov/) supporting the findings of this study are publicly available, and the lake water volume (LWV) data are derived from Lin et al.'s (2020) study. Anthropogenic factors include land use/cover changes (LUCC) and socio-economic indicators which reflects the variation and development of human activities. LUCC data can be accessed from Climate Change Initiative (CCI, https://climate.esa.int/en/). Socio-economic indicators include Gross Domestic Product (GDP) which is available from the World Bank (https://www.worldbank.org/), and population (POP) which can be acquired from the WorldPop (Open Spatial Demographic Data and Research - WorldPop).

Source	Time	Temporal Resolution	Use
Landsat Series	2001-2021	Monthly Annually	Algal bloom detection
LSWT precipitation LWV	2001-2021 2001-2019 2001-2019	Monthly Annually Annually	Meteorological driver analysis
LUCC GDP POP	2001-2020 2001-2021 2001-2021	Annually Annually Annually	Anthropogenic driver analysis

Table 1. Research data list.

3. Methodology

3.1 Algal Bloom Extraction

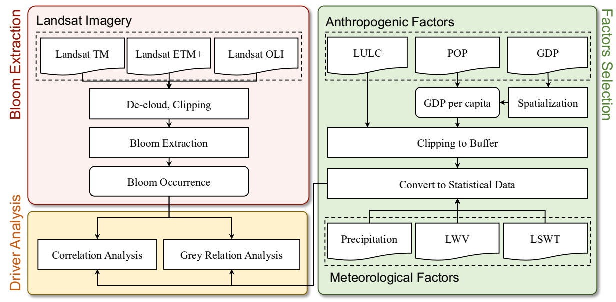

Figure 2. Flowchart of this study.

1) Feature space construction: To effectively extract algal bloom areas, in addition to the original six multispectral bands, three additional features were calculated, including NDVI, FAI, and the greenness component of the K-T transformation, resulting in a total of nine feature maps to enhance algal bloom.

2) Extraction model construction: We designed an algal bloom extraction model based on Vision Transformer, consisting with an encoder and decoder. The encoder is used to extract hidden features, which converts the input data into a higher dimensional feature space. The decoder is used to convert the high-dimensional features back into the reference data space to predict the output data.

The encoder is consisted of a four-layer progressively down-sampling and pyramid-structured Transformer block. Each block contains multi-head self-attention (MSA) and feedforward neural network (FFN), where FFN is essentially a two-layer multi-layer perceptron (MLP). To better distinguish between water bodies and algal blooms, the model inputs the high-response features of both algal blooms x_a' and water bodies x_w' separately into the block. The output is the difference between the two sets of features learned by the encoder, which can be expressed as following formulas:

$$x' = \left[x_a^{(i)\prime}; x_w^{(i)\prime}\right] = \mathrm{MSA}\left(\left[x_a^{(i)}; x_n^{(i)}\right]\right) \quad (1)$$

$$x_a^{(i+1)} = \mathrm{FFN}_1\left(x_a^{(i)\prime}\right) \quad x_w^{(i+1)} = \mathrm{FFN}_2\left(x_w^{(i)\prime}\right) \quad (2)$$

$$out^{(i)} = x_a^{(i+1)} - x_n^{(i+1)} \quad (3)$$

where, $x_a^{(i)\prime}$ and $x_w^{(i)\prime}$ represent the ith layer's feature maps for algal blooms and water bodies separately; $out^{(i)}$ represents the output feature maps of the ith layer.

The decoder is consisted of several MLP layers, inspired by Xie et al. (2021). Firstly, the output from the encoder $out^{(i)}$ is upsampled to a size of $\frac{H}{4}\times\frac{W}{4}$ with given feature channel dimension C and concatenates to F. Then, F is passed through two successive linear transformation and upsampling layers to gradually restore it to the original image size and the number of target classes. The formula are as follows:

$$F_i = US_{\left(\frac{H}{4}\times\frac{W}{4}\right)}\left(MLP\left(out^{(i)}\right)\right) \quad (4)$$

$$F = Concat([F_1; \dots; F_4]) \quad (5)$$

$$out = US_{(H\times W)}\left(W_1 \times Upsample_{\left(\frac{H}{2}\times\frac{W}{2}\right)}(W_2 \times F)\right) \quad (6)$$

where, $US(\cdot)$ represents the upsampling operation; $Concat(\cdot)$ represents the concatenation operation; W_1 and W_2 are the weight matrix of the two linear layers.

3) Bloom occurrence calculation: Monthly bloom occurrences are calculated after algal bloom detection and extraction by summing the average daily occurrences. To mitigate the randomness and uncertainty caused by cloud interference in satellite imagery, bloom occurrence frequency (BOF) in each year is further calculated as follows:

$$BOF = \frac{1}{S} \times \frac{1}{\sum_{m=1}^{p} N} \times \sum_{m=1}^{p}\sum_{n=1}^{N} O_{n,m} \quad (7)$$

where, N is the number of images with no or few clouds, m stands for the month applied (for monthly $p = 1$; for annually $p = 12$), $O_{n,m}$ denotes the number of bloom occurrence in one month, S denotes the total lake pixels.

Slope trend analysis was employed to evaluate long-term trends of bloom occurrence, which is a linear regression model based on the least squares technique. The formula can be expressed as:

$$slope = \frac{T\sum_{t=1}^{T} t \cdot O_t - \sum_{t=1}^{T} t \sum_{t=1}^{T} O_t}{T\sum_{t=1}^{T} t^2 - \left(\sum_{t=1}^{T} t\right)^2} \quad (8)$$

where, T denotes the study period (equals 21), O_t is the annual algal bloom occurrences in the t-th year.

3.2 Driver Analysis

3.2.1 Meteorological Factors

1) Precipitation: Monthly average precipitation rate was acquired from TRMM 3B43 data (unit in $mm \cdot hr^{-1}$), which records global precipitation from 1998 to 2019. The monthly average precipitation rate of Lake Victoria was calculated by average the pixel value within the region. The annual precipitation was then obtained by summing the monthly precipitation, and can be expressed as follows:

$$P = \sum_{i=1}^{12} \bar{p} \times 10^{-6} \times \bar{A} \times 24 \times days_i \quad (9)$$

where, P is the annual precipitation of Lake Victoria (unit in $km^3 \cdot yr^{-1}$), $\bar{p}$ is the monthly average precipitation rate of Lake Victoria (unit in $mm \cdot hr^{-1}$), $\bar{A}$ is the average area of Lake Victoria in 20 years ($\bar{A}$=66228.74 km^2) and $days_i$ is the number of days in each month.

2) Lake Surface Water Temperature (LSWT): LSWT is derived from ESA CCI dataset, which contains daily lake hydrological information. In this study, monthly average LSWT of Lake Victoria was calculated by average the water temperature in each year.

3) Lake Water Volume (LWV): LWV was derived from a "water level – water area – water volume" model proposed by Lin et al. (2020), where the relative water volume of Lake Victoria was estimated by modelling the effective relationship between water level, water area and water volume variations.

3.2.2 Anthropogenic Factors

1) Land Use/Cover Changes (LUCC): The LUCC data was acquired from ESA CCI landcover products, which is an annual global landcover dataset and has been updated to 2020. Considered the similarity of some landcover types, we aggregated the original 38 types and reclassified to 7 main types within the 1° buffer zone of Lake Victoria, including cropland, forest, grassland, wetland, built-up, water body and other dominated by unused land and sparse vegetation.

2) POP and GDP: POP and GDP are collected to reflect the level of socio-economic development, which are clipped using the 1° buffer shape file. To estimate the GDP value in the buffer zone, spatialization was performed by applying the nighttime light (NTL) product which is proved to be consistent with GDP (Lin et al., 2022). The spatialization process is carried out using the following formula (Lin et al., 2022):

$$GDP_i = \frac{DN_i}{\sum_{country} NTL_i} \times GDP_{country} \quad (10)$$

where, GDP_i represents the estimated GDP value for the ith pixel, NTL_i represents the DN value for the ith pixel in the NTL data, and $GDP_{country}$ represents the statistical GDP value for each country.

3.2.3 Grey Relational Analysis

Impact factors often exhibit complexity and non-linear constraints, characterized by a wide range of uncertainty or "greyness" (Allen et al., 1998). In this study, Grey Relational Analysis (GRA) is employed to quantitatively describe the impact of various factors on BOF. GRA measures the correlation between two sequences, X and Y, to describe the degree of their relationship. The calculation formula is shown as follows:

$$\Delta_i = |Y_i^* - X_i^*| \quad (11)$$

$$R(X,Y) = \frac{1}{N}\sum_{i=1}^{N} \frac{\Delta_{min} + \rho\Delta_{max}}{\Delta_i + \rho\Delta_{max}} \quad (12)$$

where, X^* and Y^* are the normalized X and Y, Δ_i represents the absolute difference between X_i^* and Y_i^*, Δ_{min} and Δ_{max} are the minimum and maximum differences across the series, ρ is the distinguishing coefficient (usually equals 0.5), and N is the number of observations.

4. Results and Discussion

4.1 Algal Bloom Occurrence

4.1.1 Accuracy Validation

The error matrix and accuracy results are listed in Table. 2. The user accuracy and producer accuracy of algal bloom extraction are 93.52% and 98.63%, respectively. The overall accuracy is 97.49% and Kappa equals 0.9271, indicating the capability of the proposed model on algal bloom extraction.

User Accuracy		Producer Accuracy	
Bloom	Non-bloom	Bloom	Non-bloom
93.52%	98.63%	95.12%	98.16%
Kappa	0.9271	Overall Accuracy	97.49%

Table 2. Error matrix.

Several segmentation models were further compared with the proposed model to examine the performance on bloom extraction task. Most models performed well visually except for Segmenter model (Figure. 3). Deeplab v3 and FCN prone to misdetections and omissions in details. Segformer and ours have the best results, but Segformer was prone to misdetection on some small area cyanobacteria extraction. Three evaluation metrics, mIOU, aACC, and mACC, were calculated for the results of different models (Table. 3). In general agreement with the visual results, Segformer and our proposed model have the highest accuracy among all models tested, while our model was basically 1%-2% higher than Segformer.

Model	mIOU	aAcc	mACC
FCN	79.63	90.09	88.28
Deeplab v3	81.51	91.53	88.3
Deeplab v3 plus	84.06	92.83	89.72
SegFormer	87.1	94.13	92.76
Segmenter	78.74	89.71	88.08
Ours	**88.33**	**94.63**	**94.32**

Table 3. Metrics comparison.

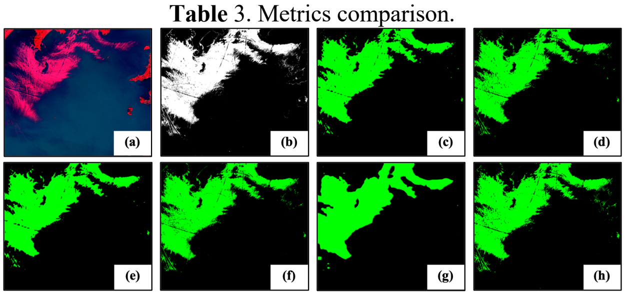

Figure 3. Visual comparison of different models: (a) False color image (composited with NIR, Red, Green); (b) Ground truth by visual interpretation; (c) DeepLab v3; (d) DeepLab v3 plus; (e) FCN; (f)SegFormer; (g) Segmenter; (h) Ours.

4.1.2 Long Trend Analysis

The algal blooms existed persistently in Lake Victoria (Figure 4), especially in bays and gulfs, which agrees with previous studies (Mbonde et al., 2015). Four areas with severe bloom occurrences are zoomed in as shown in Figure 5, including Bukakata, Buluube, Speke Gulf, and Mwanza Gulf. The spatial distribution of the bloom occurrence in Lake Victoria also showed a tendency to develop from the outside inward with increasing years, predominantly on the southeastern side of the lake.

The total bloom-affected area of Lake Victoria expanded by 17,489 km^2 with a growth rate of 1.53 times between 2001 and 2021, which corresponds to an expansion of about 900 km^2 per year (Figure 5). The growth rate of the BOF in Lake Victoria was 45.45% (+2.78% yr^{-1}) during the observation period. Spatially, the area with a significant trend of increasing frequency (slope>0.1) was 36.49% larger than the area with the decreasing trend. The long-term trend analysis suggests that the fluctuations between bloom-affected areas and BOF in Lake Victoria were generally consistent over the period 2001 to 2021. The affected areas and the BOF in three countries have all shown an increasing trend. In Uganda (Figure 5b), the affected areas increased by 6669.20 km² with 1.25 times growth, and the BOF increased by 47.09%. Spatially, the BOF showed an increasing trend on the eastern side, while the northern side exhibited a decline. In Kenya (Figure 5c), Lake Victoria, mainly comprising bays and gulfs, experienced the largest increase in bloom-affected areas among the three countries, expanding by 1101.55 km², with 8.93 times growth. The BOF rose by 37.25%. The overall spatial distribution in Kenya showed a trend of increasing BOF. In Tanzania (Figure 5d), the bloom-affected area grew by 9719.01 km², with 2.02 times growth and a 40.84% rise in BOF. Spatially, except for the southeastern bays where BOF decreased, all other areas exhibited an increasing trend.

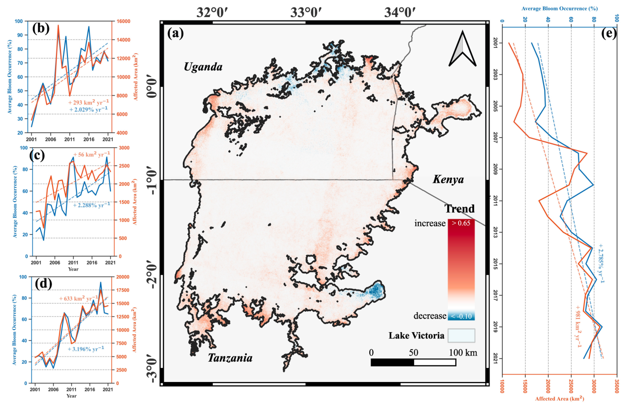

Figure 5. Trends of algal blooms between 2001 and 2021. (a) Spatial patterns of the trends in bloom frequency, and interannual variability and trends in average bloom occurrence and bloom-affected area within (b) Uganda (c) Kenya (d) Tanzania and (c) whole Lake Victoria.

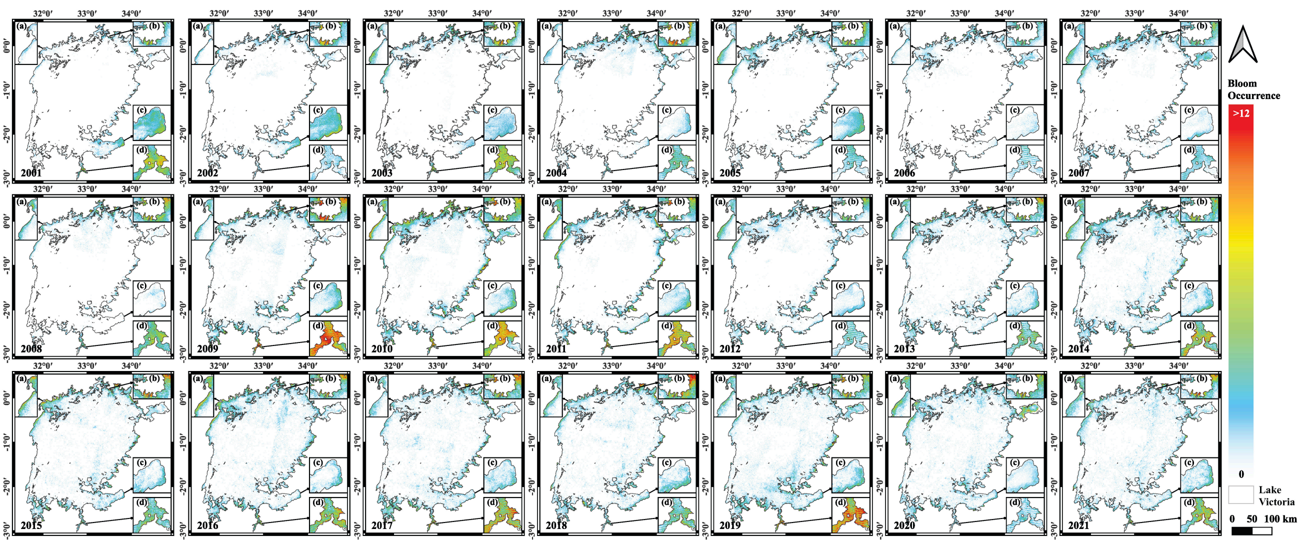

Figure 4. Interannual evolution of bloom occurrence in Lake Victoria with four areas zoomed in. (a) Bukakata. (b) Buluube. (c) Speke Gulf. (d) Mwanza Gulf.

4.1.3 Seasonal Trend Analysis

Monthly bloom occurrences in Lake Victoria were extracted for a total of 21 years from 2001-2021, and heat maps and monthly and annual statistics of bloom occurrences were plotted (Figure 6). Large-scale algal bloom outbreak occurred between 2008 and 2010. From the monthly variation, it was found that May, June, and October are more prone to algal bloom occurrence, whereas July and August exhibited a counterintuitive opposite trend. This may be due to extremely high temperatures inhibiting algal growth. In terms of annual variation, the evolution of the spatial distribution of bloom occurrence showed an increasing trend, which suggests the water quality of Lake Victoria is gradually deteriorating.

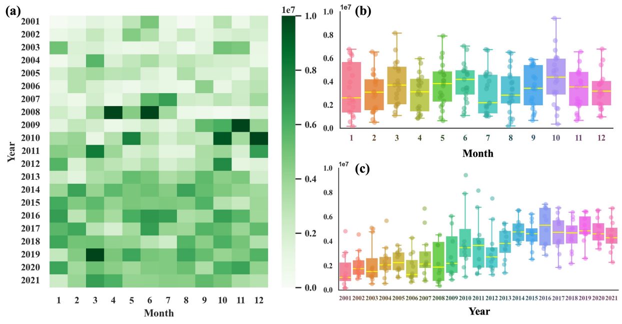

Figure 6. Seasonal analysis of 21-year algal bloom occurrence. (a) Heatmap. (b) Monthly variation. (c) Annually variation.

4.2 Meteorological Impact Analysis

Previous studies have shown that moderate precipitation can promote algal growth, leading to an increase in BOF, while extreme precipitation may negatively affect algal aggregation, thereby impacting bloom formation (Lin, 2017). We first analyzed the relationship between precipitation and BOF. Figure. 7(a) presents the relationship between precipitation and BOF from 2001 to 2019. To more clearly illustrate the impact of precipitation on bloom frequency, the annual bloom frequency was detrended by removing the long-term growth trend (see Figure. 5e). It can be observed that in several years with higher precipitation, there was a corresponding decrease in bloom frequency.

Further analysis was conducted to explore the range of precipitation that influences BOF increase. First, we identified the precipitation values during years when BOF variation was positive. Then, outliers were excluded using 1.5 times the interquartile range (IQR) method. The resulting maximum and minimum precipitation values, approximately 98-118 km^3/yr, are highlighted in light purple in Figure 7(b). Additionally, the time-lag effect of precipitation on BOF was analyzed. To accurately measure the relationship between BOF and its drivers using the Pearson correlation coefficient (PCC), a linear relationship was assumed (Kibena et al., 2014). The PCC between the precipitation and BOF with 1-year lag equals 0.57 ($p<0.05$).

The relationship between water volume and bloom occurrence was also examined. The relative water volume was found to have a positive impact on the bloom occurrence with the PCC equals 0.58 ($p<0.01$) (Figure 7c). The average surface temperature of Lake Victoria has remained between 24°C and 27°C over the past 20 years, showing a gradual decline (Figure 7d). It appears counterintuitive to the observed continuous increase in bloom occurrence (PCC equals -0.26), as many researches have indicated that temperature is positively correlated with algal growth (Ho & Michalak, 2020). Considering Lake Victoria's geographical location, even with the gradual decrease in water surface temperature, it still remains within the optimal temperature range for algal growth, which may explain the absence of a significant decline in bloom occurrence.

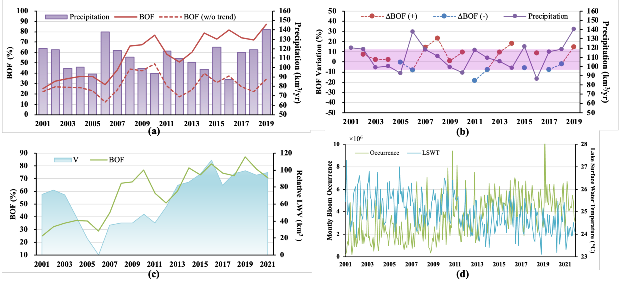

Figure 7. Meteorological factors analysis.
(a) Precipitation and BOF. (b) Precipitation and BOF variation. (c) Relative lake water volume and BOF. (d) Lake water surface temperature and BOF.

4.3 Anthropogenic Impact Analysis

4.3.1 LUCC Impact on Bloom Occurrence

Within the buffer zone, croplands are widespread distributed throughout the buffer zone surrounding Lake Victoria (Figure 8). The forest areas cluster in the southwestern part of the lake, and also present in the northern and northeastern regions. Built-up areas around Lake Victoria are predominantly concentrated in the Kampala region to the north of the lake, and have significantly expanded over the past 20 years. The eastern part of the lake has seen a reduction in forest, much of which has been converted into cropland, alongside an increase in built-up land patches. The 20-year LUCC pattern suggests that the growing demands for urban living and agricultural irrigation are likely to alter the nutrient status of Lake Victoria, thereby affecting the bloom occurrence. Additionally, other ecological land types along the western and eastern shores of the lake, such as grasslands and wetlands, have undergone slight spatial changes, though they continue to shrink in area.

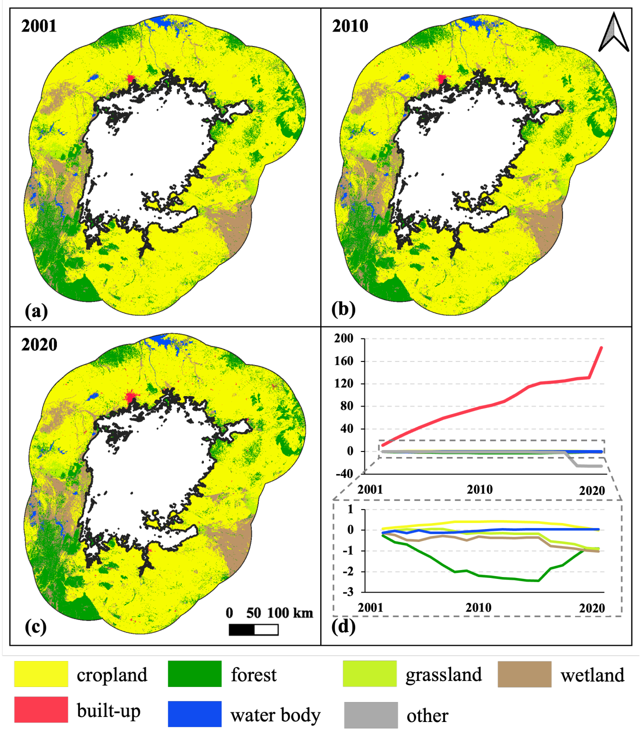

Figure 8. LULC types and spatial distribution within the buffer zone of Lake Victoria in (a)2001, (b)2010 and (c)2020 , and (d) changing rate of 7 LULC types between 2001 and 2020 with 5 types zoomed in.

Figure 8(d) shows the changing rate of area for each land use/cover type over 20 years. Built-up areas have experienced the most dramatic changes, with an increase of 184% from 2001 to 2020. Except for built-up areas and other types, the area change rates for the remaining land use/cover types are all within 3%. The Mann-Kendall trend analysis indicates that forest, grassland, wetland, and other land types have significantly decreased in area (Table. 4), while built-up areas and water bodies have shown a growth trend, although the rate of water body change is low and spatially insignificant. As urbanization progresses and agricultural production expands, the per capita demand for land continues to increase, leading to the conversion of more land into productive use. This includes the reclamation of some wetlands, which has contributed to increased soil erosion.

	s	p-value	trend
cropland	0.1158	0.4957	—
forest	-0.3895*	0.0179	↓
grassland	-0.7421***	0.0000	↓
wetland	-0.6211***	0.0001	↓
built-up	1.0000***	0.0000	↑
water body	0.6053***	0.0002	↑
other	-0.7632***	0.0000	↓

(*: p-value < 0.05, **: p-value < 0.01, ***: p-value < 0.001)

Table 4. MK-test for 7 LULC types.

Table 5 presents the correlation analysis between changes in LULC types and the bloom occurrence. Grassland and wetland exhibit negative correlations (p<0.5 both) with bloom occurrence, while built-up area (p<0.001) and water body (p<0.5) show positive correlations, which suggests that the expansion of built-up areas and the reduction of vegetation will significantly increase the frequency of bloom occurrence. This may increase the extent of impervious surfaces, which affects water quality and surface processes, and thus in turn, reduces the lake's purification capacity (Lin et al., 2024), leading to an increase in bloom occurrence within the lake. Also, the decrease of grassland and wetland areas weakens nitrogen/phosphorus fixation by vegetation and enhanced nutrient transport, thus reducing the self-purification capacity of lake (Dar et al., 2021). Although the increase of surrounding water body has positive correlation with bloom occurrence, the changing rate of water body is under 0.5%, where we thought it should be less causal connection.

	r	p-value (2-side)	N
Cropland	0.3066	0.3591	20
Forest	-0.5470	0.0816	20
Grassland	-0.7232*	0.0119	20
Wetland	-0.6705*	0.0240	20
Built-up	0.8653***	0.0006	20
Water body	0.6055*	0.0484	20
Other	-0.4880	0.1278	20

(*: p-value < 0.05, **: p-value < 0.01, ***: p-value < 0.001)

Table 5. Pearson correlation between LUCC and bloom occurrence.

4.3.2 Socio-economic Impact on Bloom Occurrence

Figure. 10(a) and 10(b) illustrate the spatial distribution of population (POP) and gross domestic product (GDP) respectively within the buffer zone in 2020, with pixel-level GDP spatialization calculating by Eq. (10). The distributions of POP and GDP are commonly consistent, with high values concentrated in areas with dense populations and built-up regions (see Figure 8). Over the past 20 years, the population around Lake Victoria has nearly doubled, rising from 25 million to approximately 46 million, making it the most populous basin among the world's five largest lakes (Dobiesz et al., 2010). This growth rate significantly exceeds that of other regions in Africa (UNEP, 2006). Meanwhile, the GDP of the region has increased from 14 billion USD to 54 billion USD. As the population grows, economic conditions are expected to develop accordingly, so thus neither individual metric is suitable for analyzing their impact on bloom occurrence. In this study, GDP per capita is calculated for the impact analysis. Figure 10(d) presents the correlation plot between per capita GDP and bloom occurrence frequency (BOF). Over the past 20 years, per capita GDP within the buffer zone has shown a steadily increasing trend, with a linear fit result of $37.42/yr. Pearson correlation analysis shows a strong positive correlation between per capita GDP and BOF, with a correlation coefficient of 0.86 (p<0.001). The growth in socioeconomic indicators reflects the ongoing urban development and expansion of economic activities around Lake Victoria. Economic activities in this region are predominantly focused on agriculture and fisheries, both of which are major contributors to water quality deterioration (Defersha and Melesse, 2012; Nyamweya et al., 2022).

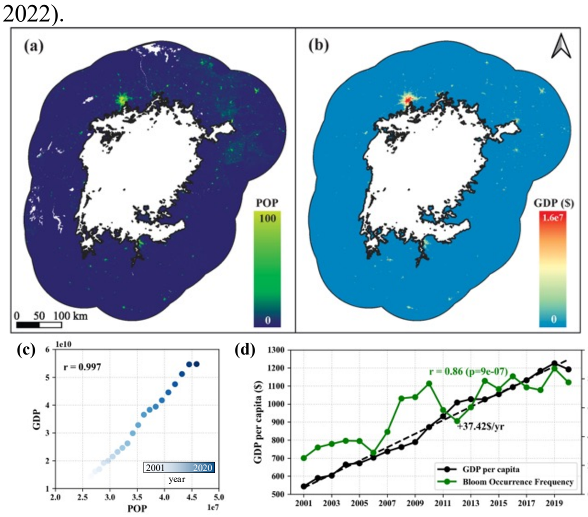

Figure 10. Socio-economic indicators. (a) POP in 2020. (b) Spatialized GDP in 2020. (c) Scatter plot of GDP and POP from 2001 to 2020 with darker marker representing later year. (d) GDP per capita and BOF from 2001 to 2020.

4.4 Grey Relational Analysis of Meteorological and Anthropogenic Impact

The Grey Relational Analysis (GRA) results indicate the degree of influence of the selected 11 factors on algal blooms in the Lake Victoria region, as shown in Table.6 arranged in descending order. Among these factors, lake water volume, GDP per capita, and built-up area exert the greatest influence, highlighting the dominant factors in climate change and human activities. In terms

Factor	Grey Relation Coefficient
LWV	0.6904
GDP per capita	0.6778
Built-up	0.6634
Precipitation	0.6608
Cropland	0.6506
Water body	0.6353
Wetland	0.5981
Grassland	0.5753
Other	0.5657
LSWT	0.5569
Forest	0.5358

Table 6. GRA results of 11 factors.

of hydrological and meteorological conditions, lake water volume and precipitation have a significant impact on bloom occurrence frequency. Larger lake volumes likely provide a more favorable environment for algal growth. Additionally, variations in precipitation can affect water quality, with increased rainfall potentially flushing nutrients into the lake, thereby promoting algal growth. Anthropogenic influence is primarily driven by socioeconomic development and the expansion of built-up and agricultural land. Furthermore, changes in water bodies, wetlands, and grasslands, may facilitate the transfer of nutrients from land to the lake, increasing the degree of water body eutrophication.

5. Conclusion

A continuous increase is revealed in the frequency and affected area of algal blooms in Lake Victoria from 2001 to 2021. The BOF and affected areas are growing at a trend of 2.875% per year and 981km^2 per year, respectively. Additionally, the influence of climate change and human activities on BOF are quantitatively analyzed. The key findings are as follows: (i) Based on GRA, the dominant meteorological factors affecting BOF are lake water volume and precipitation, while socioeconomic development and agricultural expansion play a significant role among human factors. (ii) Among meteorological factors, lake water volume is positively correlated with BOF, and proper precipitation is linked to an increase in BOF between 98 and 118 km^3/yr approximately. The precipitation has a positive correlation with BOF variation, with an annual time lag of about one year. The surface water temperature of Lake Victoria remained between 24°C and 27°C, showing a weaker influence on BOF. (iii) Regarding human activities, economic development, the expansion of built-up areas and croplands, and the reduction of certain ecological land types (e.g., wetlands and grasslands) contribute to the growth in BOF.

Our study provides a comprehensive framework for analyzing bloom occurrence trends and influencing factors, offering valuable insights for lake water resource management decisions. However, due to the limitations of satellite remote sensing in capturing in-lake hydrological conditions, this study lacks hydrological impact analysis. Additionally, the observational cycle of satellite imagery may introduce systematic errors in bloom frequency calculations. In future work, integrating multi-source satellite data with higher temporal resolution could improve the accuracy of the analysis.

Acknowledgements

We sincerely thank the anonymous reviewers for reviewing this paper and providing valuable suggestions.

References

Akurut, M., Willems, P., & Niwagaba, C. (2014). Potential Impacts of Climate Change on Precipitation over Lake Victoria, East Africa, in the 21st Century. *Water*, 6(9), 2634–2659. https://doi.org/10.3390/w6092634

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome*, *300*(9), D05109.

Amanullah, Khalid, S., Imran, Khan, H. A., Arif, M., Altawaha, A. R., Adnan, M., Fahad, S., Shah, A., & Parmar, B. (2020). Effects of climate change on irrigation water quality. *Environment, Climate, Plant and Vegetation Growth*, 123–132.

Amorim, C. A., & do Nascimento Moura, A. (2021). Ecological impacts of freshwater algal blooms on water quality, plankton biodiversity, structure, and ecosystem functioning. *Science of the Total Environment*, *758*, 143605.

Baltodano, A., Agramont, A., Reusen, I., & van Griensven, A. (2022). Land Cover Change and Water Quality: How Remote Sensing Can Help Understand Driver–Impact Relations in the Lake Titicaca Basin. *Water*, 14(7), 1021. https://doi.org/10.3390/w14071021

Carrea, L.; Crétaux, J.-F.; Liu, X.; Wu, Y.; Bergé-Nguyen, M.; Calmettes, B.; Duguay, C.; Jiang, D.; Merchant, C.J.; Mueller, D.; Selmes, N.; Simis, S.; Spyrakos, E.; Stelzer, K.; Warren, M.; Yesou, H.; Zhang, D. (2024): ESA Lakes Climate Change Initiative (Lakes_cci): Lake products, Version 2.1. NERC EDS Centre for Environmental Data Analysis, *04 April 2024*. doi:10.5285/7fc9df8070d34cacab8092e45ef276f1.

Dar, S. A., Rashid, I., & Bhat, S. U. (2021). Linking land system changes (1980–2017) with the trophic status of an urban wetland: Implications for wetland management. *Environmental Monitoring and Assessment*, *193*(11), 710.

Defersha, M. B., & Melesse, A. M. (2012). Field-scale investigation of the effect of land use on sediment yield and runoff using runoff plot data and models in the Mara River basin, Kenya. *Catena*, *89*(1), 54-64.

Dobiesz, N. E., Hecky, R. E., Johnson, T. B., Sarvala, J., Dettmers, J. M., Lehtiniemi, M., ... & Witte, F. (2010). Metrics of ecosystem status for large aquatic systems–A global comparison. *Journal of Great Lakes Research*, *36*(1), 123-138.

Downing, A. S., Van Nes, E. H., Balirwa, J. S., Beuving, J., Bwathondi, P. O., Chapman, L. J., ... & Mooij, W. M. (2014). Coupled human and natural system dynamics as key to the sustainability of Lake Victoria's ecosystem services. *Ecology and Society*, *19*(4).

Egessa, R., Pabire, G. W., & Ocaya, H. (2018). Benthic macroinvertebrate community structure in Napoleon Gulf, Lake Victoria: effects of cage aquaculture in eutrophic lake. *Environmental monitoring and assessment*, *190*, 1-10.

Gidudu, A., Letaru, L., & Kulabako, R. N. (2021). Empirical modeling of chlorophyll a from MODIS satellite imagery for trophic status monitoring of Lake Victoria in East Africa. *Journal of Great Lakes Research*, *47*(4), 1209-1218.

Ho, J. C., Michalak, A. M., & Pahlevan, N. (2019). Widespread global increase in intense lake phytoplankton blooms since the 1980s. *Nature*, *574*(7780), 667-670.

Ho, J. C., & Michalak, A. M. (2020). Exploring temperature and precipitation impacts on harmful algal blooms across continental US lakes. *Limnology and Oceanography*, *65*(5), 992-1009.

Hou, X., Feng, L., Dai, Y., Hu, C., Gibson, L., Tang, J., ... & Zheng, C. (2022). Global mapping reveals increase in lacustrine algal blooms over the past decade. *Nature Geoscience*, *15*(2), 130-134.

Khan, M. N., & Mohammad, F. (2014). Eutrophication: challenges and solutions. *Eutrophication: Causes, Consequences and Control: Volume 2*, 1-15.

Kibena, J., Nhapi, I., & Gumindoga, W. (2014). Assessing the relationship between water quality parameters and changes in landuse patterns in the Upper Manyame River, Zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C*, *67*, 153-163.

Kizza, M., Rodhe, A., Xu, C. Y., Ntale, H. K., & Halldin, S. (2009). Temporal rainfall variability in the Lake Victoria Basin in East Africa during the twentieth century. *Theoretical and applied climatology*, *98*, 119-135.

Lin, S. (2017). *Climate change and algal blooms*. Michigan State University.

Lin, Y., Gao, C., Li, X., Zhang, T., Yu, J., Rong, Y., ... & Sneeuw, N. (2023). Attribution Analysis of Climate Change and Human Activities to Water Volume Variation in Large Lakes. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *48*, 443-450.

Lin, Y., Li, X., Zhang, T., Chao, N., Yu, J., Cai, J., and Sneeuw, N., 2020: Water volume variations estimation and analysis using multisource satellite data: a case study of lake victoria. *Remote Sensing*, 12(18), 3052.

Lin, Y., Zhang, T., Liu, X., Yu, J., Li, J., & Gao, K. (2022). Dynamic monitoring and modeling of the growth-poverty-inequality trilemma in the Nile River Basin with consistent night-time data (2000–2020). *International Journal of Applied Earth Observation and Geoinformation*, *112*, 102903.

Lin, Y., Zhong, D., Rong, Y., Yu, J., Li, L., Chen, X., ... & Wang, H. (2024). A comprehensive framework for assessing the spatiotemporal impact of land use and cover change on lake eutrophication. *International Journal of Applied Earth Observation and Geoinformation*, *131*, 103989.

Liu, M., Zhang, Y., Shi, K., Zhang, Y., Zhou, Y., Zhu, M., ... & Liu, M. (2020). Effects of rainfall on thermal stratification and dissolved oxygen in a deep drinking water reservoir. *Hydrological Processes*, *34*(15), 3387-3399.

Mbonde, A. S., Sitoki, L., & Kurmayer, R. (2015). Phytoplankton composition and microcystin concentrations in open and closed bays of Lake Victoria, Tanzania. *Aquatic ecosystem health & management*, *18*(2), 212-220.

Merel, S., Walker, D., Chicana, R., Snyder, S., Baurès, E., & Thomas, O. (2013). State of knowledge and concerns on cyanobacterial blooms and cyanotoxins. *Environment international*, *59*, 303-327.

Michalak, A. M., Anderson, E. J., Beletsky, D., Boland, S., Bosch, N. S., Bridgeman, T. B., ... & Zagorski, M. A. (2013). Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proceedings of the National Academy of Sciences*, *110*(16), 6448-6452.

Muggidde, R., Hecky, R. E., & Ndawula, L. (2005). Eutrophication of Lake Victoria, Uganda. *Water Quality and Quantity Synthesis Final Report, LVEMP*.

Nyamweya, C., Lawrence, T. J., Ajode, M. Z., Smith, S., Achieng, A. O., Barasa, J. E., ... & Nkalubo, W. (2023). Lake Victoria: Overview of research needs and the way forward. *Journal of Great Lakes Research*.

Nyamweya, C., Nyaboke, H. M., Aura, C. M., Momanyi, K. N., Mlaponi, E., Odoli, C. O., & Njiru, J. M. (2022). Lake Victoria’s bounty: A case for riparian countries’ blue economic investment. *Frontiers in Environmental Science*, *10*, 952654.

Olokotum, M., Mitroi, V., Troussellier, M., Semyalo, R., Bernard, C., Montuelle, B., ... & Humbert, J. F. (2020). A review of the socioecological causes and consequences of cyanobacterial blooms in Lake Victoria. *Harmful algae*, *96*, 101829.

Ojok, W., Wasswa, J. and Ntambi, E. (2017) Assessment of Seasonal Variation in Water Quality in River Rwizi Using Multivariate Statistical Techniques, Mbarara Municipality, Uganda. *Journal of Water Resource and Protection*, 9, 83-97.

Osoro, E., Wandiga, S. O., Abongo, D. A., Madadi, V. O., & Macharia, J. W. (2016). Organochlorine pesticides residues in water and sediment from Rusinga Island, Lake Victoria, Kenya. *Journal of Applied Chemistry*, *9*(9), 56-63.

Paerl, H. W., Gardner, W. S., Havens, K. E., Joyner, A. R., McCarthy, M. J., Newell, S. E., ... & Scott, J. T. (2016). Mitigating cyanobacterial harmful algal blooms in aquatic ecosystems impacted by climate change and anthropogenic nutrients. *Harmful Algae*, *54*, 213-222.

Svirčev, Z., Lalić, D., Bojadžija Savić, G., Tokodi, N., Drobac Backović, D., Chen, L., ... & Codd, G. A. (2019). Global geographical and historical overview of cyanotoxin distribution and cyanobacterial poisonings. *Archives of toxicology*, *93*, 2429-2481.

Treuer, G., Kirchhoff, C., Lemos, M. C., & McGrath, F. (2021). Challenges of managing harmful algal blooms in US drinking water systems. *Nature Sustainability*, *4*(11), 958-964.

Tungaraza, C., Eliapenda, E., Osewe, K. O., & Palapala, P. M. (2012). Long-term climate impact on the Lake Victoria region influences water level fluctuation and resource availability. *International Journal of Environmental Sciences*, *2*(3), 1717-1732.

United Nations, 2015. About the Sustainable Development Goals [WWW Document]. Accessed 10.15.24. https://www.un.org/sustainabledevelopment/sustainable-development-goals/.

United Nations Environment Program, 2006. UNEP 2006 annual report. https://www.unenvironment.org/resources/annual-report/unep-2006-annual-report.

Wang, S., Li, J., Zhang, B., Spyrakos, E., Tyler, A. N., Shen, Q., ... & Peng, D. (2018). Trophic state assessment of global inland waters using a MODIS-derived Forel-Ule index. *Remote sensing of environment*, *217*, 444-460.

Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). SegFormer: Simple and efficient design for semantic segmentation with transformers. *Advances in neural information processing systems*, *34*, 12077-12090.

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote sensing of environment*, *118*, 83-94.