



OPEN

Spatio-seasonal variation of water quality influenced by land use and land cover in Lake Muhazi

Edovia Dufatanye Umwali^{1,2,4,5,6}, Alishir Kurban^{1,4,6}✉, Alain Isabwe³, Richard Mind'je^{1,4,5,6}, Hossein Azadi^{1,7,8}, Zengkun Guo^{1,2,4}, Madeleine Udahogora^{1,4}, Anathalie Nyirarwasa^{1,4,6}, Jeanine Umuhoza^{1,2,4,5,6}, Vincent Nzabarinda^{1,2,4}, Aboubakar Gasirabo^{1,2,4,5,6} & Gulnur Sabirhazi^{1,4}

Understanding the influence of land use/land cover (LULC) on water quality is pertinent to sustainable water management. This study aimed at assessing the spatio-seasonal variation of water quality in relation to land use types in Lake Muhazi, Rwanda. The National Sanitation Foundation Water Quality Index (NSF-WQI) was used to evaluate the anthropogenically-induced water quality changes. In addition to Principal Components Analysis (PCA), a Cluster Analysis (CA) was applied on 12-clustered sampling sites and the obtained NSF-WQI. Lastly, the Partial Least Squares Path Modelling (PLS-PM) was used to estimate the nexus between LULC, water quality parameters, and the obtained NSF-WQI. The results revealed a poor water quality status at the Mugarore and Butimba sites in the rainy season, then at Mugarore and Bwimiyange sites in the dry season. Furthermore, PCA displayed a sample dispersion based on seasonality while NSF-WQI's CA hierarchy grouped the samples corresponding to LULC types. Finally, the PLS-PM returned a strong positive correlation (+0.831) between LULCs and water quality parameters in the rainy season but a negative correlation coefficient (−0.542) in the dry season, with great influences of cropland on the water quality parameters. Overall, this study concludes that the lake is seasonally influenced by anthropogenic activities, suggesting sustainable land-use management decisions, such as the establishment and safeguarding protection belts in the lake vicinity.

Water resources and their quality are crucial to human life¹. Water quality monitoring is important for assessing the value of water for a healthy ecosystem, hygienic environment, domestic and other uses such as recreation, agriculture, mining, and consumption². Therefore, effective and consistent information on water quality is relevant and pertinent for the improvement of water management³.

Worldwide, the deterioration of surface water quality due to the increasing effects of human activities and climate change have caused great concern^{4,5}. The water quality is endangered by both point and non-point source pollution. Compared to point-sources that have currently been effectively managed, the non-point source pollutants from various land-use types, mostly agriculture and fluvial waters are the foremost source of water quality change^{6,7}.

On the other hand, climate change can produce heavy storms by changing rainfall patterns, which affect the elements found in the runoff^{8,9}. In aquatic ecosystems, different chemical, biological, and physical factors determine the quality of water¹⁰. Thus, a specific issue in water quality monitoring lies in the difficulty related to the analysis and measurement of numerous variables, as well as high variability owing to both human and nature-related influences¹¹.

¹State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, 818 South Beijing Road, Urumqi 830011, China. ²Key Laboratory of GIS & RS Application Xinjiang Uyghur Autonomous Region, Urumqi 830011, China. ³Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, Fujian, China. ⁴University of Chinese Academy of Sciences, Beijing 100039, China. ⁵University of Lay Adventists of Kigali (UNILAK), Faculty of Environmental Sciences, P.O. Box 6392, Kigali, Rwanda. ⁶Joint Research Center for Natural Resources and Environment in East Africa, P.O. Box 6392, Kigali, Rwanda. ⁷Department of Geography, Ghent University, Ghent, Belgium. ⁸Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Prague, Czech Republic. ✉email: alishir@ms.xjb.ac.cn

With the increase in population, the assessment of water quality induced by land use/land cover (LULC) is an emerging concern due to the trends in water quality in different areas. Consequently, there have been various studies investigating the relationship between LULC and water quality concluding that a significant relationship exists between water quality parameters and land use types at different scales^{12–15}. Some studies reported that LULC at the catchment scale has the most important relationship with the quality of water^{16,17} whereas others explained that LULC at the local scales might provide improved clarifications on the variations in water quality parameters^{18,19}. Accordingly, it is necessary to consider the spatial scales from local to catchment scale for an effective examination of LULC impacts on water quality.

In particular, most tropical lakes in Africa face pollution problems due to rapidly growing populations associated with greater demand for agricultural land, rising economies, and industrialization^{20,21}. These latter have resulted in increased water consumption and wastewater discharge, which causes heavy pollution²². Similarly, in Rwanda, aquatic resources are deteriorating owing to domestic, industrial, and agricultural wastes, irrigation return flows; fertilizers; surface run-off; urban development; deforestation, and mining²³. Of these resources, Lake Muhazi is experiencing a dramatic change in its water quality while it is a major source of water for the neighboring community²⁰. Therefore, it is vital to analyze the presence of different elements to maintain the sustainability of the lake.

Traditionally, water quality has been evaluated via measuring physico-chemical, and biological parameters such as dissolved oxygen (DO), pH, turbidity, total solids (TS), temperature, total phosphate (TP), nitrate, biochemical oxygen demand (BOD), and fecal coliform (FC)²⁴. The measurement of these parameters employs various indices through mathematical and statistical methods that several organizations have suggested and adopted²⁵. Among the various water quality indices, the National Sanitation Foundation related index (NSF-WQI) has become the most effective approach used to collect and process data on water resources to improve management activities²⁶. The strength of using this index, as opposed to the evaluation of individual water quality variables, is mainly the ability to reduce the bulk of information into a single value to convey the data in a simplified and understandable manner²⁷.

Research on the impacts of LULC on water quality have been previously conducted from different scales^{28,29}. However, these studies scarcely have given full consideration to the seasonal differences of water quality parameters while no single study has considered this difference in the study area (Lake Muhazi). Nevertheless, Mupenzi et al.³⁰ showed that the lake is polluted by water flow from mountainsides, the use of agrochemicals in the sugarcane plantations, and other human activities around the lake. Additionally, other studies tracked heavy metals contamination including Cd, Cr, Cu, Fe, Pb, Mn, Zn, along with pH, and temperature in tributary rivers of the same lake^{20,31} and found that the lake's high metal content levels are related to the catchment's land-use practices. By adding chlorophyll-a and transparency to the above-mentioned parameters, the same authors analyzed the nutrient inflows and levels and found that the levels of the nutrient are higher. From the above studies, it is clear that there is a weakness in the provision of comprehensive information on different important parameters including the DO profiles which is very important in the investigation of water quality to better understand nutrient transformations in the lake.

Moreover, in Rwanda where Lake Muhazi is located, we noticed that no single peer-reviewed study has endeavored to inspect the correlation between LULC and water quality indices to provide a quick and simple methodology in the quantification of water quality. The above concerns and drawbacks are discussed in this study to estimate the water quality of Lake Muhazi. Hence, the current study investigates the spatial and seasonal variation of water quality influenced by LULC in Lake Muhazi based on water quality index approach to bridge the gap identified in the literature for effective future water quality monitoring.

This study will hopefully serve as a reference line for further studies concerning water quality monitoring to overcome substantial impacts caused by the poor quality of water. We hypothesized that compared with the rainy season, the dry season would exhibit improved water quality status due to less runoff as a result of less agricultural activities in the watershed while larger cropland areas could result in water deterioration in the nearby sites. The specific objectives of this study are (i) determining the seasonal impacts on water quality in the lake; (ii) mapping the concentration of water quality parameters in the lake; (iii) analyzing the relationship between LULC, water quality parameters, and water quality index based on seasons; and (iv) determining the impact of LULC on seasonal variation of water quality.

Methodology

Study area. Lake Muhazi (Fig. 1) is a shallow lake with a long, thin, and snake-like structure and is one of Rwanda's natural aquatic resources. It is situated about 20 km from the eastern side of Kigali, the capital city³¹ and has a shoreline in three of the five provinces of the country. The western side of the lake forms the border between Kigali (Gasabo district) to the south, and Northern Province (Gicumbi district)³². In the Eastern Province, two-thirds of the eastern lake forms the boundary between the district of Rwamagana to the south and Gatsibo and Kayonza districts to the north.

It is marked by both urban and rural emissions, especially from rural agricultural activities that are poorly controlled. The lake is located in a mountainous region where the primary activities are farming, ecotourism, and mining. The catchment area is about 830 km² while the lake itself occupies an area of approximately 34.1 km² and a volume of about 3.30 × 10⁸ m³^{20,31}. The average depth is 10 m (33 ft) with the deepest point of 14 m (46 ft). The climate in the area is warm and humid with annual rainfall ranging between 858 and 1154 mm and the average temperature ranging between 18 and 21 °C.

Field sampling and sample analysis. Sampling was conducted using the upstream to downstream approach to have a good and clear evaluation of the water quality. Twelve sampling sites (Table 1) were selected

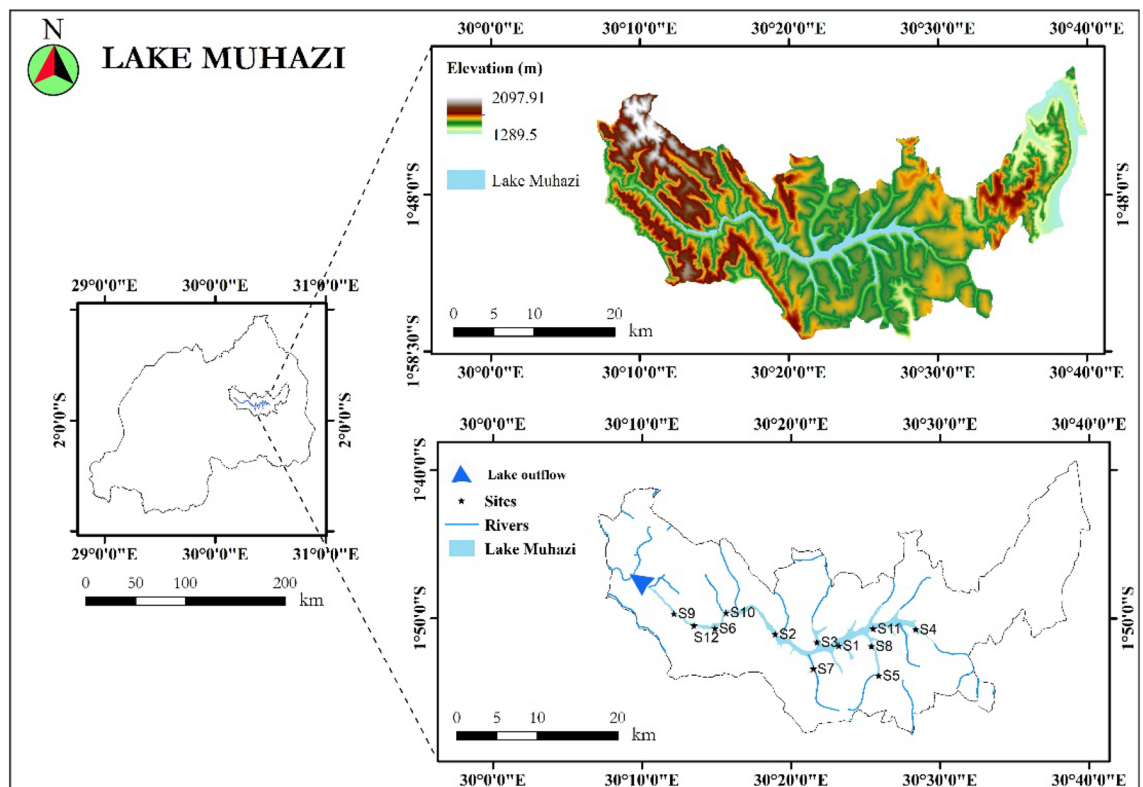


Figure 1. Geographical location of Lake Muhazi and sampling sites along the lake. Source: Authors' self-implementation with the ArcGIS software version 10.8 (<http://www.esri.com>). (Map No: GS (2016) 1663).

Site ID	Site names	Coordinates	
		X	Y
S01	Gasave	30.257123	1.513906
S02	Buburankwi	30.185781	1.500955
S03	Umwiga	30.232620	1.509941
S04	Butimba	30.302810	1.495510
S05	Ubwiza	30.291679	1.547498
S06	Gasharu	30.148464	1.494556
S07	Babasha	30.228533	1.539946
S08	Busharu	30.283410	1.514394
S09	Mugorore	30.102641	1.477750
S10	Karambo	30.160822	1.476896
S11	Kibara	30.285485	1.494822
S12	Bwimiyange	30.125038	1.491283

Table 1. The sampling sites and their geographical coordinate locations.

for the representation of the entire lake (Fig. 1). The study area has four climatic seasons which are referred as long rainy, short rainy, long dry and short dry. The long rainy season lasts from March to May and the short rainy season from September to November, with an average rainfall of 110–200 mm per month. The long dry season lasts from June and runs to early September and the short dry season lasts from December to the end of February. The average temperature vary between 19 and 27 °C⁴. Therefore, water samples were collected during two field campaigns: one in the dry season (from 20 July to 28 Sep 2018), another in the rainy season (from 12 Oct to 22 Dec 2018). The physico-chemical parameters including dissolved oxygen (DO, mg/L), fecal coliform (FC, CFU/100 mL), biochemical oxygen demand (BOD, mg/L), pH, temperature (°C), total phosphate (TP, mg/L), nitrate (mg/L), turbidity (NTU), and total solids (TS, mg/L) were carried out in situ using a different calibrated instrument with standard solutions and were analyzed using a portable water testing kit (Wagtec, Potalab No 2). These instruments are such as a hand-held conductivity tester EC-1382A (Kelilong Electron) a combined pen-type pH and thermometer combination (ATC Pometer), and YSI Professional Plus hand-held

multi-parameter. Water samples were analyzed considering the detection limit and standard technique for water given to us by the laboratory as also specified by the World Health Organization (WHO) and the Rwandan Standards Bureau (RBS). The analysis of major chemical constituents was conducted following the standard laboratory procedures³³.

Details on these detection limit can be found at https://www.who.int/water_sanitation_health/dwq/fulltext.pdf and http://www.rsb.gov.rw/fileadmin/user_upload/files/RS_435_revised_2011.docx.

Land use and land cover (LULC). To obtain LULC types, a Landsat 8 OLI imagery (data acquisition date: 3 July 2018) with respective scene path and row (172 and 61), at 30 m spatial resolution was acquired from the United States Geological Survey (USGS) EROS data center. Before the classification procedure, the image was pre-processed by performing atmospheric and radiometric correction using the Fast Line-of-Sight Atmospheric Analysis of Hypercube (FLAASH) tool in ENVI software version 5.1 to reduce atmospheric effects and radiometric errors; and hence increase the interpretability and quality of the image. Moreover, during the pre-processing phase, a cloud mask was created to mask out the very few clouds cover in the image³⁴. The dataset was then employed to develop an up-to-date LULC map in ArcGIS 10.8 software (<http://www.esri.com>) using the supervised maximum likelihood classification (MLC) approach. However, satellite imageries cannot be fully corrected as they may contain possible errors leading to uncertainties. For this, the accuracy of the classified LULC map was assessed through the overall accuracy and the Kappa coefficient (Eqs. (1) and (2) respectively) using a total of 216 random points in all land use types sampled from a multi-temporal Google Earth aerial imagery and found the total ground true value data of 201 points. The latter were then overlaid on the classified image for validation. This map was classified into six classes namely, forest, grassland, cropland, built-up, wetland, and water bodies, and the percentage contribution of each land-use types were obtained.

$$O.A = \frac{X}{X'} \times 100, \quad (1)$$

$$K = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_i + *x_{+i})}{N^2 - \sum_{i=1}^r (x_i + *x_{+i})}, \quad (2)$$

where O.A represents the overall accuracy, X is the total number of correct samples, X' is the total number of samples, K is the Kappa index, r is the number of rows in the matrix, x_{ii} is the number of observations in row and column i while $x_i +$ and x_{+i} are the marginal totals of row i and column i, respectively, and N is the total number of observations.

Inverse distance weighted (IDW) interpolation. As an extension found in numerous spatial analysis tools, IDW allows the interpolation of water quality parameters at an unknown location from known values to create a continuous surface, and understand the scenarios of water quality parameters in the study area³⁵. Though there are several spatial modeling techniques including kriging available concerning the application in GIS, the IDW approach has been selected as it enables to process of the spatial distribution of water quality for a comprehensive analysis of the results. IDW was found as the best method to estimate EC and pH, respectively, and an effective tool for determining surface water quality³⁶. It is an even reliable and precise method of spatial interpolation to predict the surface water quality in a more accurate format compared to kriging^{37,38}. By using ArcGIS 10.8 software, each physiochemical parameter was interpolated with IDW in both seasons to the entire lake (Fig. 2).

However, to control the significance of surrounding points on the interpolated value, the exponent of distance was set to default as “2” which is commonly used considering that the most reasonable results in the IDW technique are obtained using exponents values ranging from 0.5 to 3. The approach was then used to construct spatial distribution maps of the parameters to the entire lake.

Calculation of water quality index. Among water quality indices, the National Sanitation Foundation Water Quality Index (NSF-WQI) is the most commonly used^{39–41}. The significance of numerous parameters using this index varies based on the use of water. Therefore, the used values include the concentration of each parameter in the 12 selected sites in different seasons (dry and rainy). The parameters' weights (Table 2) have been determined by Brown et al.⁴² in support of the National Sanitation Foundation by requesting the experts to assign values (from 0 to 100) to different concentrations of each of the selected parameters. Hence, Numerical ranges of WQI status were classified into five classes namely Excellent (91–100), Good (71–90), Medium (51–70), Bad (26–50) and Very bad (0–25)^{41,42}.

The NSF-WQI is mathematically expressed as:

$$NSFWQI = \sum_{i=1}^n W_i Q_i, \quad (3)$$

where n is the number of water quality parameters considered for the calculation of WQI, W_i represents the weightage assigned to each parameter associated with Q_i which is the sub-index for each water quality parameter. The NSF sub-index values (NSF-Qis) were calculated using the rating curves and calculated by dividing its known concentration in each water sample after analysis and multiplied by 100.

Statistical analyses. To determine the relationship between all parameters, water quality index, and land-use types, Pearson's correlation analysis was conducted. Multivariate data processing was carried out using PCA

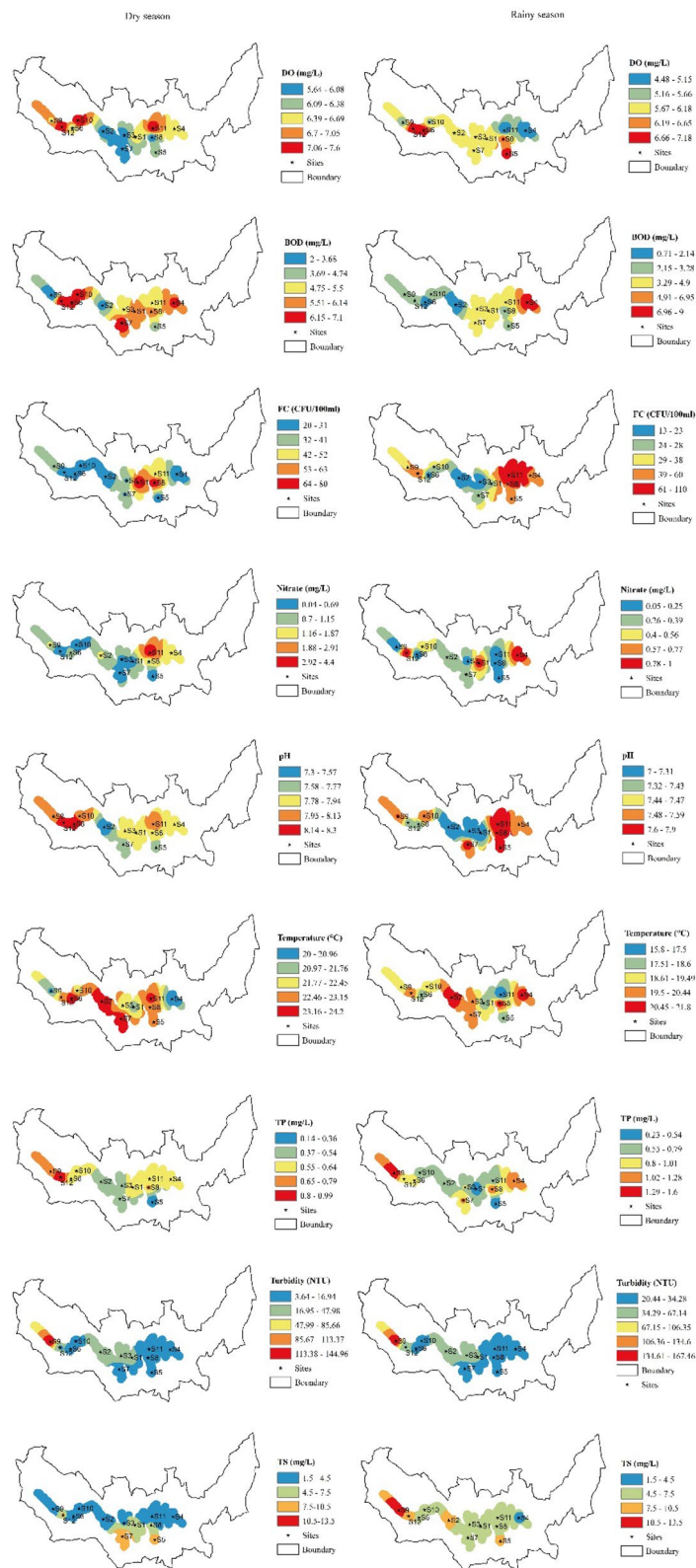


Figure 2. Spatial distribution of the parameters of water quality on Lake Muhazi. Authors' self-implementation with the ArcGIS software version 10.8 (<http://www.esri.com>).

Parameters	W _i	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10	S11	S12
Dry season													
FC	0.17	47	63	59	63	61	59	58	48	55	58	54	63
pH	0.16	90	93	87	90	92	73	91	87	84	80	80	73
BOD	0.11	50	66	56	49	59	46	47	51	80	48	57	46
DO	0.10	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	6.0	6.0	6.0
Temp	0.10	20	17	19	22	18	17	16	18	20	19	17	18
TP	0.10	53	66	60	54	94	58	59	52	51	57	58	40
Nitrates	0.10	97	96	96	96	97	96	97	96	96	97	68	97
Turbidity	0.08	61	52	51	75	73	83	89	86	5.0	82	80	70
TS	0.07	80	80	80	80	81	80	81	80	80	80	80	81
Rainy season													
FC	0.17	61	65	68	50	52	63	61	50	56	59	43	51
pH	0.16	88	88	88	93	90	93	91	92	92	92	87	93
BOD	0.11	64	97	57	38	66	92	60	68	69	68	64	66
DO	0.10	5.0	5.0	5.0	4.0	6.0	6.0	5.0	5.0	5.0	5.0	4.0	6.0
Temp	0.10	26	19	22	20	26	26	21	20	24	24	30	21
TP	0.10	85	54	47	35	89	46	39	35	30	52	57	45
Nitrates	0.10	96	97	97	96	97	97	97	97	97	97	97	96
Turbidity	0.08	49	39	40	58	58	58	61	55	5.0	56	57	51
TS	0.07	81	81	81	80	81	81	81	81	83	81	81	82

Table 2. Weights and sub-index per each parameter.

and CA techniques to analyze the spatial and seasonal changes. PCA was used to summarize the overall variation in sample grouping by reducing the dimensionality for all water quality parameters (log-transformed, except pH). Nonetheless, as one of the statistical approaches used in water quality analysis, CA is used to categorize entities including sampling sites into discrete different classes based on a given objective. For this, the among-group similarity is minimized while within-group similarity is maximized⁴³. In this study, it was used to analyze whether the classification of sampling sites using physico-chemical parameters may be compatible with the NSF-WQI variance, separately in the dry and rainy seasons. Furthermore, PLS-PM is reported to be a soft modeling method applied to model fundamental paths among blocks of parameters⁴⁴. It is preferred by different researchers as it is simple with no strong assumptions regarding the sample size, distributions, and the estimation scale requirement⁴⁵. Its application in this study eased the estimation of the complex cause-effect correlation among LULC variables, water quality parameters in both seasons, and the NSF-WQI values.

Such mathematical studies have been carried out in R (version 4.0.0) (<https://www.r-project.org/>) supplemented by the packages including *Hmisc*, *ggplot2*, *pls*, *plspm*, *superheat*, and *factoextra*^{46–50}.

Results

The spatial and seasonal distribution of water quality parameters in the lake. The spatial distribution and concentration of measured parameters in the lake water samples at different sites for the dry and rainy seasons were assessed (Fig. 2). The highest amount of DO was 7.6 mg/L in the dry season, while the lowest was 4.48 mg/L at S11 (Kibara) in the rainy season. The FC number was the largest (108 CFU/100 mL) at the S11 (Kibara), and the lowest was 13 CFU/100 mL at the S03 (Umwiga), both in the rainy season. The pH value in the lake was the highest (pH = 8.3) at the S06 (Gasharu) in the dry season and the lowest (pH = 7) at S01 (Gasave) in the rainy season. The highest amount of BOD (9 mg/L) was found at S04 (Butimba), and the lowest (0.71 mg/L) was found at S02 (Buburankwi). The temperature of the lake water was the highest (24.2 °C) at the S07 (Babasha) in the dry season and the lowest (15.8 °C) at S11 (Kibara) during the rainy season. The highest amount of TP (1.6 mg/L) was found at S09 (Mugorore) during the dry season, and the lowest was 0.14 mg/L at S05 (Ubwiza). Additionally, the highest amount of nitrate was 4.4 mg/L at the S11 (Kibara), and the lowest amount was 0.04 mg/L recorded at S10 (Karambo), both during the dry season. The results also showed that the amount of turbidity (167.5 NTU) was high at the S09 (Mugorore) in the rainy season and was low (3.64 NTU) at S04 (Butimba) in the dry season. TS was high at S09 (Mugorore) with 13 mg/L in the rainy season and low (1.8 mg/L) at the S06 (Gasharu) in the dry season.

Besides the maximum and minimum, Table 3 displays the quantified results of the water quality parameters where the computed statistical variables such as the mean and standard deviations in both seasons are also shown. However, average FC, TP, Turbidity, and TS were higher in the rainy season compared to the dry season.

Correlation between LULC, water quality parameters, and water quality index. *Land use and land cover map.* By percentages, cropland (59.10%), forest (13.76%), grassland (21.49%), wetland (0.46%), built-up (1.01%), and waterbodies (4.18%) were identified around the Lake Muhazi watershed (Fig. 3).

	Dry season			Rainy season		
	Min	Max	Mean ± Std.dev	Min	Max	Mean ± Std.dev
DO	5.64	7.60	6.52 ± 0.71	4.48	7.18	5.99 ± 0.89
FC	20.00	80.00	36.75 ± 20.33	13.00	108.00	41.33 ± 27.00
pH	7.30	8.30	7.90 ± 0.29	7.00	7.90	7.46 ± 0.31
BOD	2.00	7.10	5.42 ± 1.60	0.71	9.00	3.47 ± 2.07
Temp	20.00	24.20	22.50 ± 1.36	15.80	21.8	19.29 ± 1.79
TP	0.14	0.99	0.57 ± 0.19	0.23	1.60	0.83 ± 0.41
Nitrate	0.04	4.40	1.00 ± 1.17	0.05	1.00	0.39 ± 0.39
Turbidity	3.64	145.00	24.52 ± 39.17	20.44	167.50	42.14 ± 40.68
TS	1.80	6.80	4.08 ± 1.53	4.00	13.00	7.30 ± 2.39

Table 3. Statistical description of Lake Muhazi water quality parameters during the dry and rainy seasons.

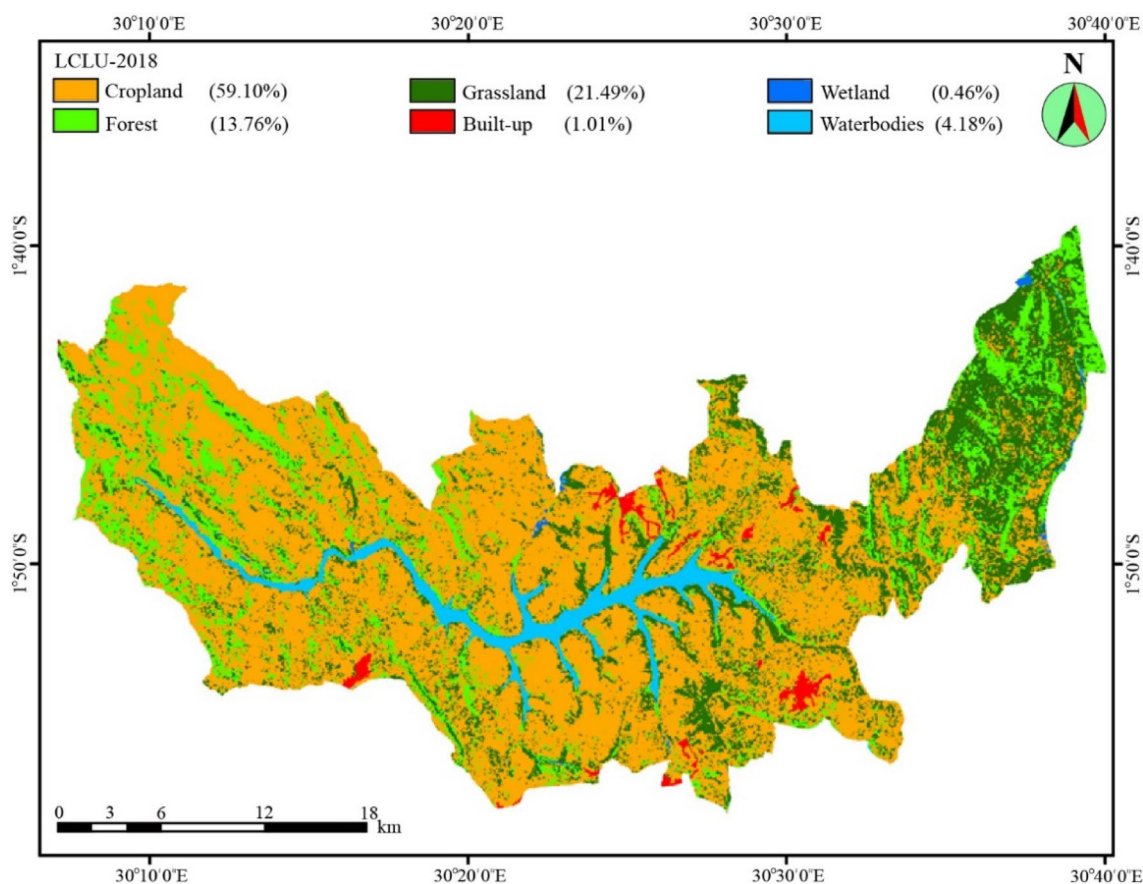


Figure 3. LULC classification map of the Lake Muhazi watershed (2018). Source: authors’ self-implementation with the ArcGIS software version 10.8 (<http://www.esri.com>).

LULC types	Cropland	Forest	Grassland	Wetland	Built-up	Water	User accuracy (%)
Cropland	40	4	1	0	1	0	86.96
Forest	1	32	1	0	0	0	94.12
Grassland	4	2	56	0	0	0	90.32
Wetland	0	0	0	18	0	0	100
Built-up	0	1	0	0	37	0	97.37
Water	0	0	0	0	0	18	100
Producer accuracy (%)	88.89	82.05	96.55	100	97.37	100	

Table 4. Confusion matrices of LULC classification.

	DO	FC	pH	BOD	Temp	TP	Nitrates	Turbidity	TS
Dry season									
Cropland	0.532*	0.257	0.303	0.409	0.445	0.590*	0.349	0.384	0.370
Forest	0.278	0.306	-0.062	-0.299	-0.431	-0.178	-0.105	0.579*	-0.372
Grassland	0.120	-0.109	0.272	0.158	-0.436	0.173	0.324	0.460	0.261
Wetland	-0.082	0.482	-0.227	0.240	-0.580*	-0.097	-0.266	-0.178	0.070
Built-up	0.124	0.581*	0.138	0.176	0.013	-0.140	0.276	-0.227	-0.124
Waterbodies	-0.461	0.637*	-0.358	-0.023	0.318	-0.487	-0.089	-0.256	0.416
Rainy season									
Cropland	0.215	0.312	-0.258	0.632*	0.048	-0.174	-0.353	0.029	0.262
Forest	-0.409	-0.174	0.006	0.198	0.291	0.570	0.169	0.579*	0.270
Grassland	-0.410	0.467	0.316	0.780**	-0.050	0.063	0.332	-0.302	-0.430
Wetland	-0.241	-0.187	-0.337	0.488	-0.142	-0.332	0.524*	-0.220	-0.375
Built-up	-0.029	0.144	-0.082	-0.161	-0.711**	-0.576*	0.094	-0.245	-0.117
Waterbodies	0.385	-0.081	-0.044	-0.398	-0.25	-0.507*	-0.296	-0.258	-0.040

Table 5. Pearson correlation coefficients between LULC types and water quality parameters in the dry and rainy season. *Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

Site ID	Site name	Dry season		Rainy season	
		WQI	Status	WQI	Status
S01	Gasave	51.25	Medium	57.62	Medium
S02	Buburankwi	56.08	Medium	57.39	Medium
S03	Umwiga	53.20	Medium	53.15	Medium
S04	Butimba	55.02	Medium	48.43	Bad
S05	Ubwiza	59.63	Medium	58.01	Medium
S06	Gasharu	52.72	Medium	58.66	Medium
S07	Babasha	55.30	Medium	53.47	Medium
S08	Busharu	52.79	Medium	51.72	Medium
S09	Mugorore	50.39	Bad	48.83	Bad
S10	Karambo	53.84	Medium	55.34	Medium
S11	Kibara	51.03	Medium	52.80	Medium
S12	Bwimiyange	50.96	Bad	52.69	Medium

Table 6. The water quality status based on the NSFQI during the dry and rainy seasons in Lake Muhazi.

The calculation of the assessed accuracy through the random sampling process for the image revealed an overall accuracy and Kappa coefficient of 93.05% and 91.37%, respectively (Table 4). These values were judged satisfactory for this study. Of all the LULC classes, we noted that cropland covers the highest proportion (59.10%), which explains the availability of different pollutants in the lake coming from agricultural sources. Hence, the classified LULC was worth being used in comparison with water quality parameters per sub-catchments for each site to facilitate the analysis of different correlations.

Correlation between LULC types and water quality parameters. Table 5 presents the correlation analysis between water quality parameters and LULC classes based on the seasons. In compliance with^{51,52} the interpretation of Pearson correlation level was followed, where the ranges between 0.5 and 1.0 indicate a strong positive correlation, a medium positive correlation (0.5–0.3, a small positive correlation (0.3–0.1 and a null (0.0–0.09). While values ranging between -1.0 to -0.5 (a strong negative correlation, a medium negative correlation (-0.5 to -0.3, a small negative correlation (-0.3 to -0.1 and a null correlation (0.0 to -0.09). In the dry season, a strong positive correlation was revealed between DO, TP and cropland; turbidity and forest; FC and built-up and between FC and waterbodies. Moreover A medium correlation between BOD, temperature, nitrate, turbidity, pH and cropland; FC and forest; nitrate, turbidity and grassland and also between FC and wetland while temperature and wetland were negatively correlated. On the other hand, a strong positive correlation was found between BOD and cropland; turbidity and forest; BOD and grassland while TP and built-up showed a strong negative correlation, temperature and built-up. Finally, nitrates and wetland were strongly positively correlated while TP and built-up unveiled a negative correlation. Besides, positive but at a small level and negative correlation between different water quality parameters and land use types were noticed (Table 5).

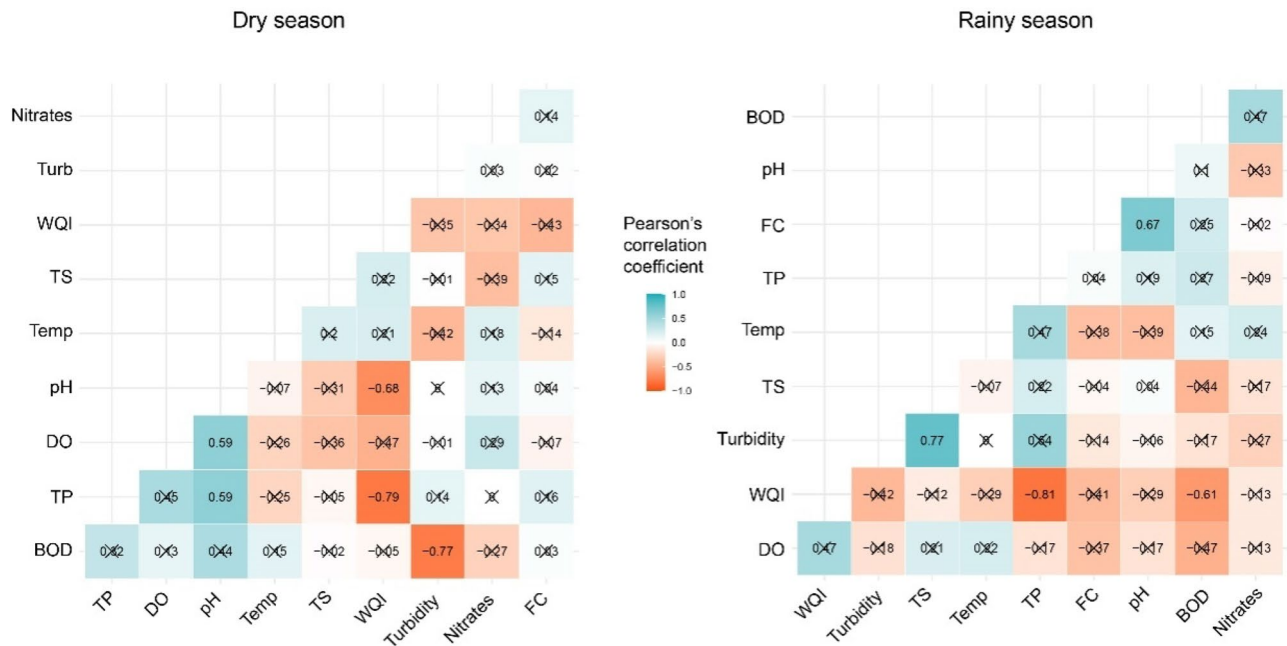


Figure 4. The correlation plot between water quality parameters and the NSFQWI in the seasons. Source: authors' self-implementation with R software version 4.0.0 (<https://www.r-project.org/>).

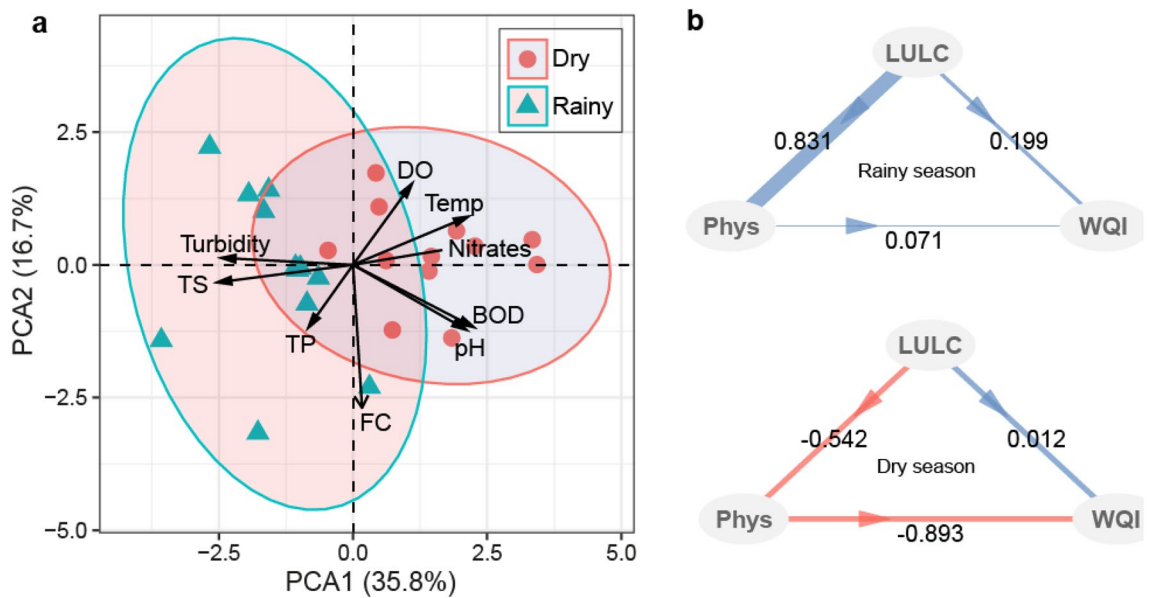


Figure 5. Principal component analysis of all sampled water quality parameters (a) and the correlation between land-use types, water quality parameters, and water quality index in seasons (b). Source: authors' self-implementation with R software version 4.0.0 (<https://www.r-project.org/>).

The relationship between water quality parameters and water quality index. During the rainy season, all sites indicated medium status except S04 and S09 while S09 and S12 indicated bad status during the dry season (Table 6). Nonetheless, the water quality of Lake Muhazi was classified as medium regardless of the seasons. In the rainy season, both the inflow (Butimba) and outflow (Mugore) showed worse quality than the dry season. The correlation coefficients between parameters for the dry and rainy season and WQI were presented in Fig. 4.

Effects of LULC on the seasonal variation of water quality. The PCA of all water quality parameters showed a sample separation based on seasonality. The first PCA axis correlated with turbidity, nitrate, temperature, and DO. The second PCA axis generally separated both the dry and rainy season samples (Fig. 5a). Similar variations among parameters associated with rainfall such as turbidity and TS increasing toward the

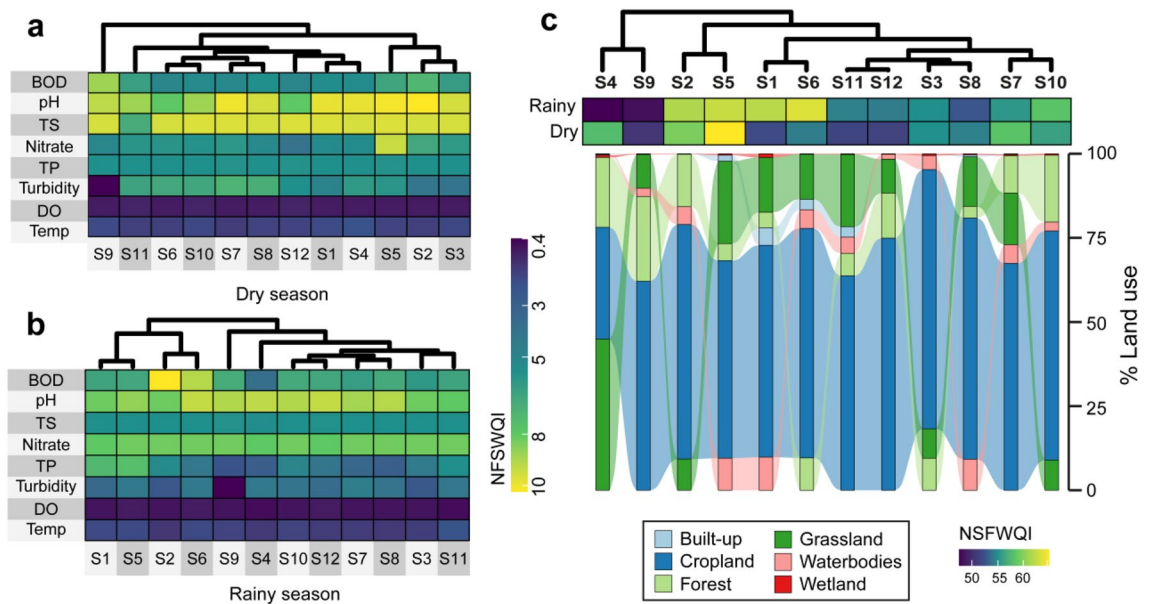


Figure 6. Heatmaps and hierarchical cluster analyses of the water quality parameters in seasons (a,b) and sample clustering based on NSF-WQI and the percentage of land use types at each sampling site (c). Source: authors' self-implementation with R software version 4.0.0 (<https://www.r-project.org/>).

rainy season samples were observed. In contrast, water temperature generally increased in dry season samples along with nitrates, BOD, and pH. However, FC varied evenly among the two-season samples. The PLS-PM demonstrated that the associations between the variables of land use types, water quality, and water quality index were all positive in the rainy season. Particularly, a strong positive connection between all land-use types and the water quality parameters was observed (Fig. 5b). This correlation was negative in the dry season (Fig. 5b); water quality parameters had a strong negative association with the water quality index. The hierarchical cluster analysis of water quality parameters (except FC) indicated a strong association between S06 and S10, S07 and S08, S01 and S4, S02 and S03, and the unique clustering of samples from the S09 (Fig. 6a). This isolation of S09 was also observed in the rainy season, a season during which S01 and S05, S10 and S12, and S07 and S08 were noted (Fig. 6b). Cluster analysis of NSF-WQI in both seasons returned sample groups corresponding with the dominance of cropland in the lake (Fig. 6c). The goodness of fit was 0.31 and 0.22 for rainy and dry seasons, respectively.

Discussion

Water quality evaluation of the lake. The core concept of the WQI is to turn variables into a single numerical value²⁶. The results show that the rainfall in the region has a direct influence on water quality and this can be confirmed by the bad quality status of water in the inflow and outflow of the study area. In previous studies, Gradilla-Hernández et al.³⁹, and Khodzher et al.⁵³ claimed that during the rainy season, high quantities of pollutants enter the lake through the inflow. Water quality drops as a result of the area being located near population centers and due to the entry of microbial contaminants. Human activities generally affect the distribution, quantity, and chemical quality of water resources throughout the sedimentation process^{54,55}. Therefore, the quality of water inflows to Lake Muhazi has possibly affected the differences in the quality of water on a certain level. This study classified the water quality status of the lake as medium. The level of water quality deterioration and the extent of the impacts varies based on the area's characteristics with a wide range of possible changes in LULC. These changes result in the removal of forests, increased cropland (which is positively correlated with different parameters in the lake), substituting grasslands, and urban expansion on a large-scale.

Seasonal variation of water quality based on parameters. Water quality is crucial to the health security of any ecosystem, which is a determination of the biochemical and physical features of the water^{41,56}. For instance, DO was used in this study as an important water quality indicator of seasonal variations. Although oxygen is difficult to dissolve, it is needed by all kinds of life forms in water bodies. The results of this research contradict an earlier study whereby the highest value of DO was observed in the dry season and the lowest value in the rainy season⁵⁷. The possible reason for the result obtained from our study can be attributed to the dilution effect as also found by Umer et al.⁵⁸. The physical characteristic of the study area is related much more where greater photosynthetic events and the reduction in turbidity (Table 3) occur in the dry season and result into the generation of high oxygen amount in the lake.

The greatest biological water parameter is FC (*E. coli*) colonies⁵⁹. They can enter groundwater by a direct release of feces from warm-blooded mammals and birds. Although this type of bacteria does not induce a particular disease directly, its presence in water indicates a low level of sanitation. Using the WHO standards, the quality

of drinking water in open wells from zones with low precipitation is in the intermediate risk (100 CFU/100 mL) category during both seasons. The primary sources of FC are animal wastes and runoff in the rainy season. Similarly, a high FC amount was found in the rainy season probably as a result of the livestock effluents process. Confirming the presence of FC bacteria in the lake normally implicates recent fecal contamination, which may create an instant health risk when consumed⁶⁰.

Nevertheless, the best pH range for sustainable aquatic life ranges from 6.5 to 8.5⁶¹. Sewage into water changes the hydrogen ion concentration in water and becomes more acidic or more alkaline depending on the type of waste and chemical substances contained in it⁶². Thus, the quality of the lake water changes in different sites based on the seasons as the changing concentration of pH is possibly due to the influence of lake water penetration, heavy metal pollution, and high biological activities⁶³ while the amount of pH during the dry season is scientifically justified by high water volume and greater water retention. In addition, some researcher has reported that the concentration of pH into the water is influenced by meteorological conditions such as ambient atmospheric warming leading to the variation in water quality⁶⁴. For instance, an increase of pH concentration was noticed in the Rhine and Meuse rivers in the dry season following the warming of water by about 2 °C after the severe drought of 2003^{65,66}. In the same study, the lower Mekong River, negative significant correlations were generally found between precipitation and pH. Hence, besides anthropogenically-driven factors, this seems to be verified in our study since the atmospheric thermal condition in the dry season might have influenced the pH concentration.

BOD is the quantity of DO required to cut down organic materials present in a given sample of water at a certain temperature in a given time by aerobic bio-organisms in a water body⁶⁷. According to Lokhande et al.⁶⁸, the decrease in BOD indicates a good water quality while its increase indicates polluted water bodies. In agreement with previous studies, the minimum and maximum amounts of BOD were both recorded in the rainy season^{41,69}. It was revealed that an increase in BOD within Lake Muhazi comes from commercial wastewater effluents, agricultural wastes, and household sewage from heavy discharge. As BOD affects the quantity of DO directly in lakes and streams, the higher amount of BOD in the lake justifies the critical reduction in aquatic lives⁷⁰.

Temperature exerts a major influence on biochemical activities^{71,72}. However, it also affects the level of DO in water, which directs the types of organisms living in lakes and rivers. As temperatures go far above or below the range, the individuals of the species reduce until none remains present. The link between temperature and DO is a natural process as warmer water becomes more easily saturated with oxygen and it can hold less DO. Therefore, based on the previously described results, the temperature might cause serious effects on the lake in the dry season compared to the rainy season which goes always with a low amount of water temperature^{73,74}. For any highly significant seasonal temperature change to be observed, data must be collected from the lake over a longer period.

The excess amounts of phosphate and Nitrate into water cause eutrophication leading to excessive algal blooms^{75,76}. As essential elements for plant life, phosphorous and Nitrate concentrations in Lake Muhazi (Fig. 2) are higher in dry season which may be justified by relatively more farming activities executed around the area moving into the lake as runoff. Soil erosion that occurs during floods takes away these parameters from adjacent land and the river banks into the water bodies^{77,78}. Cropland influence is explained by inappropriate use of inputs, mostly chemical fertilizers specifically urea, and NPK that is hugely used in the study area. According to Najjar and Khan⁷⁹ and Solanki⁸⁰, sometimes, the low amount of nitrate during the dry season happens due to the increase of adaptation by biota owing to their peak growth and the use of plankton or aquatic plants for metabolic activities.

Regarding the NTU parameter, its measurement is frequently applied as water quality indicators depending on the clarity and measured total suspended solids (clay or silt), organic matter, inorganic materials, and decaying material⁸¹. The increased turbidity in the lake, especially in the rainy season, is generally attributed to high amounts of decaying vegetation; which creates a discoloration of water. Consequently, more unclear water leads to less photic zone and less potential for photosynthetic production. Furthermore, the amounts of NTU found in different sites are caused by the loosened soils from cropland around the area taken away by rain and wind to the lake. The effect of this high turbidity in the lake may cause siltation and sedimentation of the habitat of the fisheries and other aquatic life.

Concerning the TS parameter, the high amount results from forest removal around water bodies⁶². Findings disclosed the highest value of TS in the rainy season and the lowest in the dry season. It is important to note that high TS amount can harm the reproduction level and survival of organisms in water^{82,83}. However, very low levels can also cause swelling of cells organism, affecting cell density balance of aquatic life⁸⁴. The mixture of settlement and cropland causes the increase of TS especially in Mugarore site with a district road polluting the lake with runoff.

However, water intended for human consumption and other activities must be free from organisms and concentrations of chemical substances that may be a hazard to health⁸⁵. Unfortunately, Lake Muhazi is still used for different purposes including drinking, domestic uses, or recreation. Thus, based on the concentration of parameters found in the lake compared to the international WHO Drinking Water Requirements. As an aid to enhance the care and perception of water quality, the activities should currently be avoided as the lake needs to be well considered and currently be subjected to the conservation and restoration process by improving techniques for minimizing the concentrations of the parameters.

The link between of land use types and water quality parameters. One of the most serious threats to water quality is the transition of land use in riparian areas⁸⁶. Anthropogenic activities are directly reflected in land use characteristics⁸⁷. Understanding the correlations between water quality and land-use types assists in identifying the primary threats to the quality of water. These correlations are essential for appropriate management of water quality since they can be used to tackle critical areas of land use and therefore, establish appropriate activities to reduce the load of pollutants¹². The study found positive correlations between different param-

ters and land use types (Table 5). Ngoye and Machiwa⁸⁸ associated the lack of vegetation cover, with the presence of different parameters in adjacent waterbodies.

The above relationship was confirmed in this study as the results exhibited a shortage of forest (13.76%) and grassland (21.49%) around the area. Generally, vegetation covers such as forest and grassland around the area affect the concentration of some parameters as both acts as nutrient detention mechanisms and eventually contributor to the betterment of water quality⁸⁹. Yet, a negative correlation was found between forest and FC probably attributed to livestock wastage which is not detained during runoff processes as a result of less vegetation cover. Consequently, it can be said that the less amount of forest and grassland in the area play a role in the concentration of pollutants in the lake.

Researches have shown that the urban environment offers a realistic influence on the deterioration of water quality in all parameters^{90,91}. The main factor determining the negative impact of development on water quality is the level of impervious surfaces directly connected to a stream. The above can justify the concentration of different parameters found in the lake (Fig. 2) and implicates that built-up land probably plays an important role in the pollution of the lake. The reason for this relates to the fact that built-up areas and their associated increase in impervious surface facilitate the transport of more pollutants by overland flow particularly in rainy season⁹².

However, cropland was the dominant type of land use with 59.10% of the total area. Consequently, the inappropriate use of input fertilizers is the major cause of water quality deterioration of Lake Muhazi. In crop management and production, nutrient changes in water quality are due to the use of fertilizers (organic or inorganic) at a higher rate than those developed by soil particles. As the primary cause of non-point source pollution, fertilization activities dominant in farming activities is one of the main crop management practices that have been reported to be the core nutrients 'source especially nitrate and phosphorous amount into surface water. The rainfall following the fertilization period in the study area (when crops are at their peak growing phase in the long and short rainy season) results in amplified overflow toward the lake either in solution or linked with eroded soil particles. Further, the transfer of the aforementioned nutrients from fertilization to the lake takes place during subsurface drainage or base flow. The management of fertilization techniques performed in the area to lessen fertilizer release into the lake leads to the major impacts that they are used in the specific site with an increased ability to deteriorate water quality. In this case, the application of fertilizers or manure without integration can trigger the persistent nutrient dispersal and cause intrusion of the lake water into aquifer⁹³. The above discussion justifies the correlations between water quality and cropland type which is indicated by the major influence in the increase of water pollution because of the increased influence of agricultural runoff⁹⁴. In conjunction with earlier studies, the concentration of parameters in a water body as a result of LULC depends on the land use planning of an area. Considering the effects of distance from the lake can improve the water quality change^{95,96}. Huang et al.⁹⁷ suggested that the different land use types exclusively croplands and built-up should be put to quite some distance away from the water bodies. For this, a set of equidistant buffer zones of some distances to sampling sites can support the understanding of the impacts of land-use types on the concentration of different pollutants because the association between these types and the water quality parameters may change as the distance to water body increases. Therefore, the creation of riparian buffer zones particularly with vegetated areas located adjacent to the investigated lake is one way to effectively lessen the concentration of the measured nutrient in the lake. Thus, increasing the vegetation within large buffer zones should be considered to reduce the concentrations of different parameters in the lake.

Uncertainties. Uncertainties in water quality-related researches are unavoidable due to difficulties in the identification of the best approach as well as the quality of input datasets under different required conditions⁹⁸. Regrettably, the collection and analysis of water quality data are expensive and require careful handling and analysis in laboratories. Thus, data on these types of research are often scarce to support the model and this contributes to possible uncertainties⁹⁹. For instance, the lack of streamflow data as in this study could be significant in water quality assessment as the concentration of parameters could have been coupled with the collected data to assess the transport mechanisms of different nutrients.

Moreover, data processing and management can also present a certain level of uncertainty. Results may be limited by missing data, methods adopted to estimate missing values, and errors in data management. This study did not address this source of uncertainty as it did not arise from random statistical variation but rather from individual mistakes or equipment faults. Another important uncertainty of the lake water quality data is associated with the sampling measurement. This uncertainty relates to the selection of representative sampling sites as they can affect the optimization of a water sample. The sampling site choice might have substantial effects on the measured concentration of a given parameter considering that some parameters have high spatio-temporal variability within the cross-section of a given reach. Therefore, high-frequency sampling efforts at multiple sampling stations, are needed to get robust results for the whole lake. Nevertheless, the effect of these above-mentioned possible uncertainties is considered not significant for this study since the significance of this study is not to represent an exact condition of the area's pollution level but to guide the comprehension of the possible future scenarios for policymakers. Hence, reducing uncertainties to improving the understanding of the vulnerability of water quality in light of land use types should be prioritized in future researches.

Conclusion

Using GIS analysis showed the spatial distribution and the condition of water quality of the Lake. It has been shown that Lake Muhazi is influenced by anthropogenic activities driven by LULC of the area. Furthermore, the condition of the lake's water quality varies depending on the seasons. From the overall obtained results, the following conclusions and recommendations were drawn:

1. Based on our analysis, water pollution was caused by the increase in human activities especially agriculture around the area.
2. Of all the sites, none has shown an excellent or good status of water quality, therefore water should be properly treated before usage.
3. To retain the runoff before releasing it into the lake and minimize water pollution, rainwater retention ponds can be constructed near developed zones around the study area.
4. The study area needs effective land use planning and formulates non-point source pollution control policies taking into consideration the increase in forest and grassland.
5. It will be significant to consider buffer-zone scales in assessing the effects of land use on water quality in the areas by controlling the anthropogenic activities within and around the watershed by taking necessary actions against polluters.
6. The study is suggested considering the runoff flow direction or incoming flow from the lake Feeder Rivers as part of the strategic sampling approach in future studies.

We believe that this study contains sufficient information to keep making efforts for strong monitoring of pollutants sources. An epidemiological study factor is required in future studies to evaluate the spatial distribution and association analysis between waterborne diseases and the concentrations of different parameters.

Received: 8 December 2020; Accepted: 11 August 2021

Published online: 30 August 2021

References

1. Water, U. N. Policy brief: water quality. United Nations. (2011). http://www.unwater.org/downloads/waterquality_policybrief.pdf.
2. Iqbal, K., Ahmad, S. & Dutta, V. Pollution mapping in the urban segment of a tropical river: Is water quality index (WQI) enough for a nutrient-polluted river?. *Appl. Water Sci.* **9**(8), 197. <https://doi.org/10.1007/s13201-019-1083-9> (2019).
3. Tyagi, S., Sharma, B., Singh, P. & Dobhal, R. Water quality assessment in terms of water quality index. *Am. J. Water Resour.* **1**(3), 34–38. <https://doi.org/10.12691/ajwr-1-3-3> (2013).
4. Mukanyandwi, V. *et al.* Seasonal assessment of drinking water sources in Rwanda using GIS, Contamination Degree (Cd), And Metal Index (Mi). *Environ. Monit. Assess.* **191**, 734. <https://doi.org/10.1007/s10661-019-7757-9> (2019).
5. Wu, Z. *et al.* Water quality assessment based on the water quality index method in Lake Poyang: The largest freshwater lake in China. *Sci. Rep.* **7**, 1–10. <https://doi.org/10.1038/s41598-017-18285-y> (2017).
6. Lai, Y., Yang, C., Hsieh, C., Wu, C. & Kao, C. Evaluation of non-point source pollution and river water quality using a multimedia two-model system. *J. Hydrol.* **409**, 583–595. <https://doi.org/10.1016/j.jhydrol.2011.08.040> (2011).
7. Rudra, R. P. *et al.* Currents status, challenges, and future directions in identifying critical source areas for non-point source pollution in Canadian conditions. *Agriculture* **10**(10), 468 (2020).
8. Neogi, B., Tiwari, A. K., Singh, A. K. & Pathak, D. D. Evaluation of metal contamination and risk assessment to human health in a coalmine region of India: A case study of the North Karanpura coalfield. *Hum. Ecol. Risk Assess. Int. J.* **24**(8), 2011–2023. <https://doi.org/10.1080/10807039.2018.1436434> (2018).
9. Tabari, H. Climate change impact on flood and extreme precipitation increases with water availability. *Sci. Rep.* **10**(1), 1–10 (2020).
10. Cuffney, T. F., Kennen, J. G., & Waite, I. R. Aquatic ecosystems as indicators of status and trends in water quality. **1**, 122–126 (2014).
11. Boyacioglu, H. Surface water quality assessment using factor analysis. *Water SA* **32**(3), 389–393. <https://doi.org/10.4314/wsa.v32i3.5264> (2006).
12. Ding, J. *et al.* Impacts of land use on surface water quality in a subtropical River Basin: A case study of the Dongjiang River Basin, Southeastern China. *Water* **7**(8), 4427–4445. <https://doi.org/10.3390/w7084427> (2015).
13. Madriñán, M. J. M., Al-Hamdan, M. Z., Rickman, D. L. & Ye, J. Relationship between watershed land-cover/land-use change and water turbidity status of Tampa Bay major tributaries, Florida, USA. *Water Air Soil Pollut.* **223**, 2093–2109. <https://doi.org/10.1007/s11270-011-1007-2> (2012).
14. Namugize, J. N., Jewitt, G. & Graham, M. Effects of land use and land cover changes on water quality in the uMngeni river catchment, South Africa. *Phys. Chem. Earth Parts A/B/C* **105**, 247–264. <https://doi.org/10.1016/j.pce.2018.03.013> (2018).
15. Teixeira, Z., Teixeira, H. & Marques, J. C. Systematic processes of land use/land cover change to identify relevant driving forces: Implications on water quality. *Sci. Total Environ.* **470**, 1320–1335 (2014).
16. Julian, J. P., Beurs, K. M. D., Owsley, B., Davies-Colley, R. J. & Ausseil, A.-G.E. River water quality changes in New Zealand over 26 years: Response to land use intensity. *Hydrol. Earth Syst. Sci.* **21**, 1149–1171 (2017).
17. Rodrigues, V. *et al.* Effects of land use and seasonality on stream water quality in a small tropical catchment: The headwater of Córrego Água Limpa, São Paulo (Brazil). *Sci. Total Environ.* **622**, 1553–1561 (2018).
18. Shukla, A. K. *et al.* Population growth, land use and land cover transformations, and water quality nexus in the Upper Ganga River basin. *Hydrol. Earth Syst. Sci.* **22**, 4745–4770 (2018).
19. Vrebos, D., Beauchard, O. & Meire, P. The impact of land use and spatial mediated processes on the water quality in a river system. *Sci. Total Environ.* **601**, 365–373 (2017).
20. Nhapi, I. *et al.* Heavy metals inflow into Lake Muhazi, Rwanda. *Eng. J.* **5**(1), 10. <https://doi.org/10.2174/1874829501205010056> (2012).
21. Wang, H. *et al.* Water and wastewater treatment in Africa—Current practices and challenges. *Clean: Soil, Air, Water* **42**(8), 1029–1035. <https://doi.org/10.1002/clean.201300208> (2014).
22. Omos, I. B., Wang, H., Cheng, S., & Li, F. Sustainable tertiary wastewater treatment is required for water resources pollution control in Africa. *Environ. Sci. Technol.* **46**(13), 7065–7066. <https://doi.org/10.1021/es3022254> (2012).
23. Nahayo, L. *et al.* Seasonal drinking water quality monitoring for the community wellbeing in the Eastern Rwanda. *J. Environ. Prot. Sustain. Dev.* **4**(1), 1–6 (2018).
24. Hernández-Romero, A. H., Tovilla-Hernández, C., Malo, E. A. & Bello-Mendoza, R. Water quality and presence of pesticides in a tropical coastal wetland in southern Mexico. *Mar. Pollut. Bull.* **48**(11–12), 1130–1141. <https://doi.org/10.1016/j.marpolbul.2004.01.003> (2004).
25. Halder, D., Halder, S., Das, P. & Halder, G. Assessment of water quality of Damodar River in South Bengal region of India by Canadian Council of Ministers of Environment (CCME) Water Quality Index: A case study. *Desalin. Water Treat.* **57**(8), 3489–3502. <https://doi.org/10.1080/19443994.2014.987168> (2016).
26. Kachroud, M., Trolard, F., Kefi, M., Jebari, S. & Bourrié, G. Water quality indices: Challenges and application limits in the literature. *Water* **11**(2), 361. <https://doi.org/10.3390/w11020361> (2019).

27. Mnisi, L. N. *Assessment of the state of the water quality of the Lusushwana River, Swaziland, using selected water quality indices* (2012).
28. Santy, S., Mujumdar, P. & Bala, G. Potential impacts of climate and land use change on the water quality of Ganga River around the industrialized Kanpur region. *Sci. Rep.* **10**, 1–13 (2020).
29. Zhang, J., Li, S. & Jiang, C. Effects of land use on water quality in a River Basin (Daning) of the Three Gorges Reservoir Area, China: Watershed versus riparian zone. *Ecol. Indic.* **113**, 106226 (2020).
30. Mupenzi, J. P., Ge, J., & Habiyaemye, G. Major elements in lake Muhazi, Rwanda, East Africa. *Acta Geol. Sin. (Eng. Edn.)* **83**, 927–931 (2009).
31. Nhapi, I. *et al.* Distribution of heavy metals in Lake Muhazi, Rwanda. *Open Environ. Eng. J.* <https://doi.org/10.2174/1874829501205010096> (2012).
32. National Institute of Statistics of Rwanda, Ministry of Finance and Economic Planning (MINECOFIN) [Rwanda]. *Rwanda Fourth Population and Housing Census. Thematic Report: Population Size, Structure and Distribution* (2012).
33. Federation, Water Environmental, & APH Association. *Standard methods for the examination of water and wastewater* (American Public Health Association (APHA), Washington, DC, 2005).
34. Stillinger, T., Roberts, D. A., Collar, N. M. & Dozier, J. Cloud masking for Landsat 8 and MODIS Terra over snow-covered terrain: Error analysis and spectral similarity between snow and cloud. *Water Resour. Res.* **55**, 6169–6184 (2019).
35. Ogbozige, F. J., Adie, D. B. & Abubakar, U. A. Water quality assessment and mapping using inverse distance weighted interpolation: A case of River Kaduna, Nigeria. *Niger. J. Technol.* **37**(1), 249–261. <https://doi.org/10.4314/njt.v37i1.33> (2018).
36. Sapna, K., *et al.* Spatial analysis of river water quality using inverse distance weighted interpolation in Noyyal watershed in Coimbatore, Tamilnadu, India. *Res. J. Life Sci. Bioinform. Pharm. Chem. Sci.* **4**, 150 (2018).
37. Kumar, S., & Sangeetha, B. Assessment of ground water quality in Madurai city by using geospatial techniques. *Groundw. Sustain. Dev.* **10**, 100297 (2020).
38. Paul, R., *et al.* Identification of hydrogeochemical processes controlling groundwater quality in Tripura, Northeast India using evaluation indices, GIS, and multivariate statistical methods. *Environ. Earth Sci.* **78**, 1–16 (2019).
39. Gradilla-Hernández, M. S. *et al.* Assessment of the water quality of a subtropical lake using the NSF-WQI and a newly proposed ecosystem specific water quality index. *Environ. Monit. Assess.* **192**, 1–19. <https://doi.org/10.1007/s10661-020-08265-7> (2020).
40. Noori, R., Berndtsson, R., Hosseinzadeh, M., Adamowski, J. F. & Abyaneh, M. R. A critical review on the application of the National Sanitation Foundation Water Quality Index. *Environ. Pollut.* **244**, 575–587. <https://doi.org/10.1016/j.envpol.2018.10.076> (2019).
41. Sharifinia, M., Ramezani, Z., Imanpour, J., Mahmoudifard, A. & Rahmani, T. Water quality assessment of the Zarivar Lake using physico-chemical parameters and NSF-WQI indicator, Kurdistan Province-Iran. *Int. J. Adv. Biol. Biomed. Res.* **1**(3), 302–312 (2013).
42. Brown, R. M., McClelland, N. I., Deininger, R. A. & Tozer, R. G. Water quality index-do we dare?. *Water Sew. Works* **117**(10), 339–343 (1970).
43. Kettenring, J. R. The practice of cluster analysis. *J. Classif.* **23**, 3–30 (2006).
44. Fernandes, L. S., Fernandes, A., Ferreira, A., Cortes, R. & Pacheco, F. A partial least squares—Path modeling analysis for the understanding of biodiversity loss in rural and urban watersheds in Portugal. *Sci. Total Environ.* **626**, 1069–1085 (2018).
45. Patrícia Salgado Terêncio, D., *et al.* Can land cover changes mitigate large floods? A reflection based on partial least squares-path modeling. *Water*, **11**, 684 (2019).
46. Barter, R. L. & Yu, B. Superheat: An R package for creating beautiful and extendable heatmaps for visualizing complex data. *J. Comput. Graph. Stat.* **27**, 910–922 (2018).
47. Harrell, Jr. F. E., Harrell, Jr. M. F. *Package 'rms'*. CRAN2018 (Harrell Miscellaneous, 2021).
48. Kassambara, A. *Practical guide to principal component methods in R: PCA, M (CA), FAMD, MFA, HCPC, factoextra* (STHDA, 2017).
49. Sanchez, G., Trinchera, L., & Russolillo, G. *Introduction to the R package pls* (2014).
50. Wickham, Hadley, W. Chang, and L. Henry. "ggplot2." <http://ggplot2.org> (2012).
51. Mukaka, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med. J.* **24**, 69–71 (2012).
52. Mupenzi, C. *et al.* Spatial pattern assessment of Lake Kivu basin rivers water quality using National Sanitation Foundation water quality and rivers pollution indices. *Desalin. Water Treat.* **95**, 118–127. <https://doi.org/10.5004/dwt.2017.21566> (2017).
53. Khodzher, T. V., Domyshcheva, V. M., Sorokovikova, L. M., Sakirko, M. V. & Tomberg, I. V. Current chemical composition of Lake Baikal water. *Inland Waters* **7**(3), 250–258. <https://doi.org/10.1080/20442041.2017.1329982> (2017).
54. Carr, G. M. & Neary, J. P. *Water Quality for Ecosystem and Human Health* (UNEP/Earthprint, 2008).
55. Hamid, A., Bhat, S. U. & Jehangir, A. Local determinants influencing stream water quality. *Appl. Water Sci.* **10**(1), 24. <https://doi.org/10.1007/s13201-019-1043-4> (2020).
56. Mustapha, M. K. Assessment of the water quality of Oyun Reservoir, Offa, Nigeria, using selected physico-chemical parameters. *Turk. J. Fish. Aquat. Sci.* **8**(2), 309–319 (2008).
57. Saad, A. S., Massoud, M. A., Amer, R. A. & Ghorab, M. A. Assessment of the physicochemical characteristics and water quality analysis of Mariout Lake, Southern of Alexandria, Egypt. *J. Environ. Anal. Toxicol.* **7**(1), 421. <https://doi.org/10.4172/2161-0525.1000421> (2017).
58. Umer, A., Assefa, B. & Fito, J. Spatial and seasonal variation of lake water quality: Beseka in the Rift Valley of Oromia region, Ethiopia. *Int. J. Energy Water Resour.* **4**, 47–54 (2020).
59. Rajendran, A. & Mansiya, C. Physico-chemical analysis of ground water samples of coastal areas of south Chennai in the post-Tsunami scenario. *Ecotoxicol. Environ. Saf.* **121**, 218–222. <https://doi.org/10.1016/j.ecoenv.2015.03.037> (2015).
60. Benouara, N., Laraba, A. & Hachemi Rachedi, L. Assessment of groundwater quality in the Seraidi region (northeast of Algeria) using NSF-WQI. *Water Sci. Technol. Water Supply* **16**(4), 1132–1137. <https://doi.org/10.2166/ws.2016.030> (2016).
61. Edition, F. Guidelines for drinking-water quality. *WHO Chron.* **38**(4), 104–108 (2011).
62. Ichwana, I., Syahrul, S. & Nelly, W. Water quality index by using national sanitation foundation-Water quality index (NSF-WQI) method at krueng tamiang aceh. In *International Conference on Technology, Innovation, Universitas Syiah Kuala, Indonesia* (2016).
63. Deepa, P., Raveen, R., Venkatesan, P., Arivoli, S. & Samuel, T. Seasonal variations of physicochemical parameters of Korattur lake, Chennai, Tamil Nadu, India. *Int. J. Chem. Stud.* **4**(3), 116–123 (2016).
64. Delpla, I., Jung, A.-V., Baures, E., Clement, M. & Thomas, O. Impacts of climate change on surface water quality in relation to drinking water production. *Environ. Int.* **35**, 1225–1233 (2009).
65. Van Vliet, M. & Zwolsman, J. Impact of summer droughts on the water quality of the Meuse river. *J. Hydrol.* **353**, 1–17. <https://doi.org/10.1016/j.jhydrol.2008.01.001> (2008).
66. Zwolsman, J. & Van Bokhoven, A. Impact of summer droughts on water quality of the Rhine River—A preview of climate change?. *Water Sci. Technol.* **56**, 45–55. <https://doi.org/10.2166/wst.2007.535> (2007).
67. Li, D. & Liu, S. *Water Quality Monitoring and Management: Basis* (Academic Press, 2018).
68. Lokhande, R. S., Singare, P. U. & Pimple, D. S. Study on physico-chemical parameters of wastewater effluents from Talaja industrial area of Mumbai, India. *Int. J. Ecosyst.* **1**(1), 1–9. <https://doi.org/10.5923/j.ije.20110101.01> (2011).
69. Sudarshan, P., Mahesh, M. K. & Ramachandra, T. V. Assessment of seasonal variation in water quality and water quality index (WQI) of Hebbal Lake, Bangalore, India. *Environ. Ecol.* **37**, 309–317 (2019).

70. Dhinamala, K., Pushpalatha, M., Samuel, T. & Raveen, R. Spatial and temporal variations in the water quality parameters of Pulicat Lake, Tamil Nadu, India. *Int. J. Fish. Aquat. Stud.* **3**(2), 255–259 (2015).
71. Katakwar, M. Water quality and pollution status of Narmada river's Anjan tributary in Madhya Pradesh, India. *Int. J. Curr. Res. Acad. Rev.* **2**(11), 93–98 (2014).
72. Saksena, D. N., Garg, R. K. & Rao, R. J. Water quality and pollution status of Chambal river in National Chambal sanctuary, Madhya Pradesh. *J. Environ. Biol.* **29**(5), 701–710 (2008).
73. Akpe, A. R., Femi, I. J., Okwu, G. I. & Obiazi, H. Seasonal variations in the physicochemical parameters of Ikpoba river water samples. *Int. J. Life. Sci. Sci. Res.* **2455**(1716), 1716. <https://doi.org/10.21276/ijlssr.2018.4.3.12> (2018).
74. Zhao, Z., Liu, G., Liu, Q., Huang, C. & Li, H. Studies on the spatiotemporal variability of river water quality and its relationships with soil and precipitation: A case study of the Mun River Basin in Thailand. *Int. J. Environ. Res. Public Health* **15**(11), 2466. <https://doi.org/10.3390/ijerph15112466> (2018).
75. Gayathri, S., Latha, N. & Ramachandra Mohan, M. Impact of climate change on water quality of Shoolkere Lake, Bangalore. *J. Acad. Ind. Res.* **2**, 362 (2014).
76. Tasdighi, A., Arabi, M. & Osmond, D. L. The relationship between land use and vulnerability to nitrogen and phosphorus pollution in an urban watershed. *J. Environ. Qual.* **46**(1), 113–122. <https://doi.org/10.2134/jeq2016.06.0239> (2017).
77. Riemersma, S., Little, J., Ontkean, G., & Moskal-Hébert, T. *Phosphorus sources and sinks in watersheds: A review. Alberta soil phosphorus limits project 5* (2006).
78. Ross, D. S. *et al.* Impact of an extreme storm event on river corridor bank erosion and phosphorus mobilization in a mountainous watershed in the northeastern United States. *J. Geophys. Res. Biogeosci.* **124**(1), 18–32. <https://doi.org/10.1029/2018JG004497> (2019).
79. Najjar, I. A. & Khan, A. B. Assessment of seasonal variation in water quality of Dal Lake (Kashmir, India) using multivariate statistical techniques. In Conference on Water Pollution-XI: Modeling, Monitoring and Management. *WIT Trans. Ecol. Environ.* **164**, 123–134 (2012).
80. Solanki, H. A. Physico-chemical parameters of water in Bibi Lake, Ahmedabad, Gujarat, India. *J. Pollut. Effects Control* 1–5 (2015).
81. World Health Organization. *Water Quality and Health-Review of Turbidity: Information for Regulators and Water Suppliers.* <https://apps.who.int/iris/bitstream/handle/10665/254631/WHO-FWC-WSH-17.01-eng.pdf> (2017).
82. Environmental and Fondriest. *Conductivity, salinity and total dissolved solids: fundamentals of environmental measurements* (2014).
83. Phyllis, K. & Lawrence, K. Effects of total dissolved solids on aquatic organisms: A review of literature and recommendation for salmonid species. *Am. J. Environ. Sci.* **3**, 1–6 (2007).
84. Kodom, K., *et al.* Physicochemical parameters assessment and seasonal variations of surface water quality of the Volta lake in Ghana. *Int. J. Res.* **5** (2018).
85. Sharma, S. & Bhattacharya, A. Drinking water contamination and treatment techniques. *Appl. Water Sci.* **7**(3), 1043–1067. <https://doi.org/10.1007/s13201-016-0455-7> (2017).
86. Ding, S., Zhang, Y., Liu, B., Kong, W. & Meng, W. Effects of riparian land use on water quality and fish communities in the headwater stream of the Taizi River in China. *Front. Environ. Sci. Eng.* **7**(5), 699–708. <https://doi.org/10.1007/s11783-013-0528-x> (2013).
87. Kang, J. H. *et al.* Linking land-use type and stream water quality using spatial data of fecal indicator bacteria and heavy metals in the Yeongsan river basin. *Water Res.* **44**(14), 4143–4157. <https://doi.org/10.1016/j.watres.2010.05.009> (2010).
88. Ngoye, E. & Machiwa, J. F. The influence of land-use patterns in the Ruvu river watershed on water quality in the river system. *Phys. Chem. Earth Parts A/B/C* **29**(15–18), 1161–1166. <https://doi.org/10.1016/j.pce.2004.09.002> (2004).
89. Huang, J., Zhan, J., Yan, H., Wu, F. & Deng, X. Evaluation of the impacts of land use on water quality: A case study in the Chaohu Lake basin. *Sci. World J.* <https://doi.org/10.1155/2013/329187> (2013).
90. Permatasari, P. A., Setiawan, Y., Khairiah, R. N. & Effendi, H. The effect of land use change on water quality: A case study in Ciliwung Watershed. In *IOP Conference Series: Earth and Environmental Science*, vol. 5, 012026 (IOP Publishing, 2017).
91. Wang, L., Li, H., Dang, J., Zhao, Y., Zhu, Y. E. & Qiao, P. Effects of urbanization on water quality and the macro benthos community structure in the Fen he River, Shanxi Province, China. *J. Chem.* <https://doi.org/10.1155/2020/8653486> (2020).
92. Seo, Y. & Schmidt, A. R. The effect of rainstorm movement on urban drainage network runoff hydrographs. *Hydrol. Process.* **26**, 3830–3841 (2012).
93. Kleinman, P. J. *et al.* Managing agricultural phosphorus for water quality protection: Principles for progress. *Plant Soil* **349**, 169–182 (2011).
94. Mishra, S., Sharma, M. P. & Kumar, A. Assessment of water quality in Surha Lake based on physicochemical parameters, India. *J. Mater. Environ. Sci.* **6**, 2446–2452 (2015).
95. De Oliveira, L. M., Maillard, P. & De Andrade Pinto, É. J. Modeling the effect of land use/land cover on nitrogen, phosphorus and dissolved oxygen loads in the Velhas River using the concept of exclusive contribution area. *Environ. Monit. Assess.* **188**, 333 (2016).
96. Uriarte, M., Yackulic, C. B., Lim, Y. & Arce-Nazario, J. A. Influence of land use on water quality in a tropical landscape: A multi-scale analysis. *Landsc. Ecol.* **26**, 1151–1164 (2011).
97. Huang, W., Mao, J., Zhu, D. & Lin, C. Impacts of land use and land cover on water quality at multiple buffer-zone scales in a Lakeside City. *Water* **12**, 47 (2020).
98. Moreno-Rodenas, A. M., Tscheikner-Gratl, F., Langeveld, J. G. & Clemens, F. H. Uncertainty analysis in a large-scale water quality integrated catchment modelling study. *Water Res.* **158**, 46–60 (2019).
99. McIntyre, N. R., Wagener, T., Wheeler, H. S. & Yu, Z. S. Uncertainty and risk in water quality modelling and management. *J. Hydroinform.* **5**, 259–274 (2003).

Acknowledgements

The authors express their sincere gratitude for the research funding supported by the key program for the bureau of international cooperation, Chinese Academy of Science (CAS), Grant number 151542KYSB20200018. In addition, we express our appreciation to the Rwanda Environment Management Authority (REMA) for facilitating the water sample analysis.

Author contributions

A.K.: supervision; E.D.U.: conceptualization, data curation, formal analysis, writing original draft, methodology, validation and reviewed the manuscript; A.I.: data curation, formal analysis, software, methodology and reviewed the manuscript; R.M.: formal analysis, validation, editing and reviewed the manuscript; H.A.: visualization and editing; Z.G.: visualization; M.U.: visualization; A.N.: visualization; J.U.: visualization; A.G.: resources; V.N.: visualization; G.S.: Editing.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



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

© The Author(s) 2021