A multivariate analysis of water quality in Lake Naivasha, Kenya

Article *in* Marine and Freshwater Research · January 2015

DOI: 10.1071/MF14031

Jane Ndungu
Kenya Marine and Fisheries Research Institute (KMFRI)

Suzanne Hulscher
University of Twente

Bernerd M. Fulanda
Pwani University

Nzula Kitaka
Egerton University, Njoro, Kenya

CSIRO PUBLISHING

Marine and Freshwater Research
http://dx.doi.org/10.1071/MF14031

A multivariate analysis of water quality in Lake Naivasha, Kenya

Jane Ndungu[A,B,C,E], *Denie C. M. Augustijn*[A], *Suzanne J. M. H. Hulscher*[A], *Bernard Fulanda*[D], *Nzula Kitaka*[B] *and Jude M. Mathooko*[B]

[A]University of Twente, PO BOX 217, 7500 AE Enschede, The Netherlands.
[B]Egerton University, PO Box 536, Njoro, Kenya.
[C]Kenya Marine and Fisheries Research Institute, PO Box 81651-80100, Mombasa, Kenya.
[D]Pwani University, PO BOX 195-80108, Kilifi, Kenya.
[E]Corresponding author. Email: jandungu@gmail.com

Abstract. Water quality information in aquatic ecosystems is crucial in setting up guidelines for resource management. This study explores the water quality status and pollution sources in Lake Naivasha, Kenya. Analysis of water quality parameters at seven sampling sites was carried out from water samples collected weekly from January to June and biweekly from July to November in 2011. Principal component analysis (PCA) and cluster analysis (CA) were used to analyse the dataset. Principal component analysis showed that four principal components (PCA-1 to PCA-4) explained 94.2% of the water quality variability. PCA-1 and PCA-2 bi-plot suggested that turbidity in the lake correlated directly to nutrients and iron with close association with the sampling site close to the mouth of Malewa River. Three distinct clusters were discerned from the CA analysis: Crescent Lake, a more or less isolated crater lake, the northern region of the lake, and the main lake. The pollution threat in Lake Naivasha includes agricultural and domestic sources. This study provides a valuable dataset on the current water quality status of Lake Naivasha, which is useful for formulating effective management strategies to safeguard ecosystem services and secure the livelihoods of the riparian communities around Lake Naivasha, Kenya.

Additional keywords: cluster analysis, physico-chemical parameters, pollution, principal component analysis.

Received 6 November 2013, accepted 9 May 2014, published online 31 October 2014

Introduction

Lakes and reservoirs are important sources of surface water and livelihood for many rural and urban communities. However, declining water quality in freshwater lakes and reservoirs is an increasing problem that threatens the ecosystem services to the riparian communities, especially in developing countries. One of the major causes of the decline in the quality of water is nutrient enrichment; mainly phosphorus and nitrogen. As a result, massive algal blooms occur, causing a shift from clear to a turbid state in shallow lakes and reservoirs (Lung'Ayia *et al.* 2000; Kitaka *et al.* 2002; Mugidde *et al.* 2005). Consequently, significant changes in the biological structure of the lakes and reservoirs occur which are a major threat to the sources of livelihood of the riparian fisher folks (Harper 1992).

Lake Naivasha is an important inland freshwater lake, especially within the Rift Valley because of the salty nature of the majority of the other water resources in the area. The lake harbours unique faunal and floral biodiversity, which led to it being declared a wetland of international importance in 1994 under the Ramsar convention (Lake Naivasha Riparian Association (LNRA) 1999). The lake is a source of livelihood and supports many socioeconomic activities such as a multimillion horticultural industry, tourism, fishing and domestic water sources (Becht and Harper 2002; Kundu *et al.* 2010). Though still artisanal, the fishing industry of the lake employs over 1000 fishermen and provides a source of protein for people living in the nearby towns (Kundu *et al.* 2010). However, myriad environmental perturbations in Lake Naivasha's ecosystem have transformed the lake from clear to a muddy eutrophic turbid state, which has resulted in a decline in ecological quality, impacting heavily on fish populations and tourism (Hubble and Harper 2001; Mergeay 2004). Sustainable lake management calls for reliable data and information on water quality. However, the quality varies both temporally and spatially. The main causes of the variation include anthropogenic activities, seasonal fluctuations in inflow of nutrients and other substances, and natural variations attributed to biogeochemical processes. Therefore, the need for continuous assessment of water quality is inevitable and calls for continuous monitoring of the lake. This notwithstanding, monitoring programs often result in huge and complex data matrices consisting of many physico-chemical parameters, thus calling for multivariate approaches to the analysis and interpretation of the data.

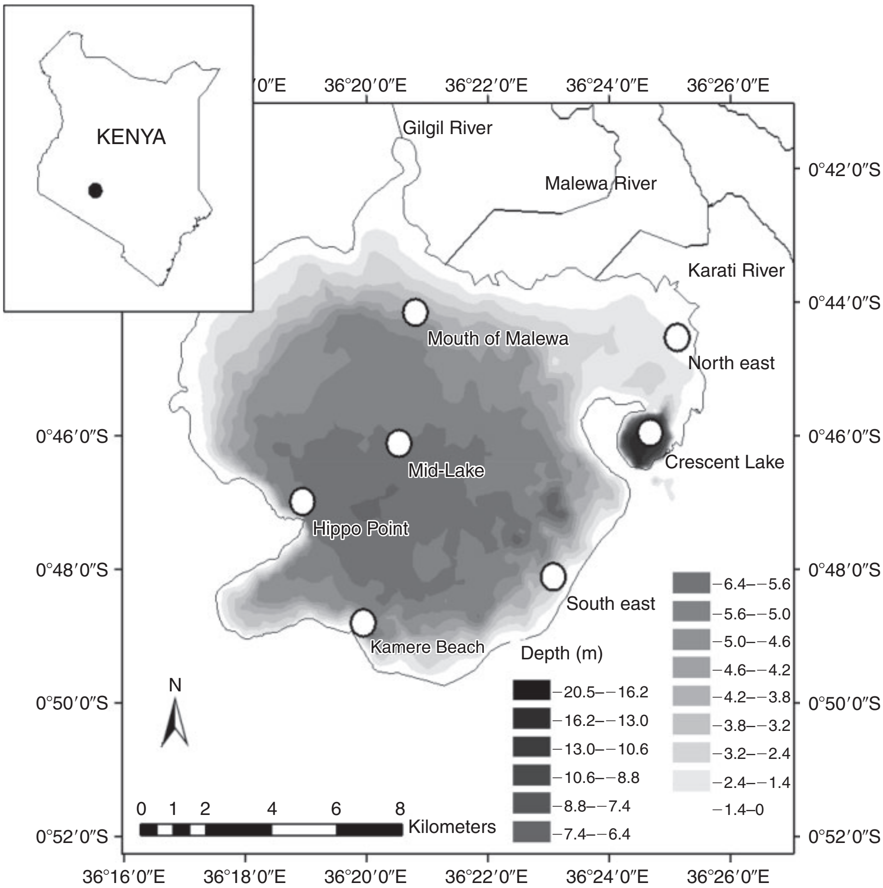

Fig. 1. Map showing the location of Lake Naivasha, input rivers, bathymetry, and the sampling sites used in the present study.

Application of multivariate statistical techniques for analysis of environmental data has increased in the past decades (Vega *et al.* 1998; Alberto *et al.* 2001; Reghunath *et al.* 2002; Tariq *et al.* 2005; Lee *et al.* 2006; Pearce *et al.* 2013; Machado *et al.* 2014; Wang *et al.* 2014; Zeng and Arnold 2014). The techniques include: cluster analysis (CA), principal component analysis (PCA), factor analysis (FA) and discriminant analysis (DA). Several studies have demonstrated the usefulness of multivariate approaches in aiding the interpretation of large complex water quality datasets (Singh *et al.* 2004; Shrestha and Kazama 2007; Kazi *et al.* 2009). Despite the numerous management challenges, the multivariate techniques have not been used in the assessment of water quality in many lakes in developing countries including Lake Naivasha. The main aim of this study was therefore to provide information for a basin-wide ecosystem management of Lake Naivasha. Specifically, the objectives of the study were to: (i) assess the status of water quality in relation to physico-chemical parameters, nutrients and major ions in Lake Naivasha; (ii) assess correlations between the different water quality parameters using multivariate analysis; (iii) evaluate any similarities and/or dissimilarities between the different regions of the lake, and (iv) decipher the pollution sources based on physico-chemical parameter associations. This paper shows the mean concentrations of physico-chemical parameters, nutrients and ions in Lake Naivasha. It also gives the correlations between the water quality parameters as analysed using PCA and explains the differences between the different regions as indicated by CA. Probable pollution sources are also discussed.

Methods

Description of the study area

Descriptions of Lake Naivasha have been provided in several publications (Stoof-Leichsenring *et al.* 2011; Ndungu *et al.* 2013*a*, 2013*b*). Therefore, only a brief description will be provided here. Lake Naivasha is a shallow endorheic freshwater lake lying in the Kenyan Rift Valley at 1890 m above sea level in a complex geological arrangement of volcanic rocks and sedimentary deposits. Straddling latitude 00°46′S and longitude 36°22′E, the lake is fed by ephemeral streams and two major perennial rivers, namely, the Malewa and Gilgil rivers; and other semi-permanent rivers such as the Karati River (Fig. 1). To the north-east lies Crescent Lake, a crater lake with a depth of up to 20 m, which occasionally separates from the main lake during low water levels (Childress *et al.* 2002). Lake Naivasha lacks a visible outlet but the lake's water is fresh, likely due to a underground outflow (Åse 1987). During the rainy season, the main lake occupies ~150 km^2 but shrinks to ~100 km^2 during the dry season (LNRA 1999).

The weather in the area is typically tropical with mean temperatures of 25°C. Precipitation is bi-modal in March, April and May and in October and November, at an average of 650 mm year^{-1}. However, the eastern part of the catchment covering the Nyandarua Range receives a higher precipitation, often reaching 2400 mm year^{-1} (Stoof-Leichsenring *et al.* 2011).

Sampling design

Water samples were collected weekly from January to June and bi-weekly from July to November 2011. Sampling was conducted at seven pre-defined stations (Fig. 1). Two sites were located on the northern side of the lake, one in the plume area of the Malewa River (mouth of Malewa site) and the other in an area close to the municipal effluent discharge point (north-east site). Another site was at the central part of the lake (midlake) while three other sites – Hippo Point, Kamere Beach and south-east – were located on the southern side of the lake. Hippo Point was at the deepest point of the main lake and Kamere Beach was close to an informal settlement, whilst south-east was set close to a sector where a large fish kill was experienced in February 2010. The seventh sampling site was at Crescent Lake.

Water temperature, pH, conductivity and turbidity were measured *in situ* during each sampling occasion using pHTestr 30 (pH), ECTestr 11+ (conductivity) and T-100 (turbidity) Oakton waterproof meters (Cole-Parmer, Vernon Hills, IL, USA). Triplicate water samples were collected at ~10 cm below the surface and chilled in ice on site and transported to the laboratory for analysis. The parameters analysed were physico-chemical parameters (total dissolved solids, TDS; total suspended solids, TSS; total hardness, TH; and total alkalinity, TA), nutrients (ammonium nitrogen, NH_4-N; nitrate nitrogen, NO_3-N; nitrite nitrogen, NO_2-N; orthophosphates, PO_4^{3-}-P; and total phosphorus, TP) and the main ions (calcium, Ca^{2+}; magnesium, Mg^{2+}; iron, Fe^{2+}; manganese, Mn^{2+}; chloride, Cl^-; and sulphate, SO_4^{2-}).

Analysis of water samples

Physico-chemical parameters

The TDS was measured directly using an ECTestr 11+ meter, while TSS was determined through the EPA gravimetric method, in which 100 mL of water was filtered onto pre-weighed 0.4 GF/C micron filters and dried in an oven to constant weight. The TSS was calculated as the difference between the weight of the filter and the final dry weight. Total hardness was determined using the ethylene-diamine tetra-acetic acid (EDTA) method while total alkalinity was determined using the titration method, which utilises the phenolphthalein indicator and N/50 sulfuric acid (APHA 2005).

Nutrients

Nitrogen as NH_4-N, NO_3-N and NO_2-N was determined through colorimetric methods as described in APHA (2005). The salicylate method was used to determine NH_4-N, while NO_3-N and NO_2-N were determined using the cadmium reduction method. Total phosphorus was determined using the molybdenum blue-ascorbic acid method, in which duplicate volumes of 50 mL samples were digested with persulfate in an autoclave for 30 min. The digested sample was then topped up with distilled water to 50 mL; the absorbance was read after 30–60 min at 880 nm wavelength using an ultraviolet UVmini-1240 spectrophotometer in 1 cm cells. The TP concentration was then determined using standard calibration curves. PO_4^{3-}-P was also determined by the molybdenum blue-ascorbic acid method by adding phenolphthalein indicator followed by drop-wise addition of 5 N sulphuric acid to discharge the red colour if it develops when the phenolphthalein indicator is added. As described in APHA (2005), 8.0 mL of a solution was made from a combination of 5 N sulfuric acid, ammonium molybdate and ascorbic acid, was added and then mixed thoroughly. After 10 min and not more than 30 min the absorption of each sample was measured at 880 nm wavelength using a reagent blank as a reference solution. The PO_4^{3-}-P concentrations were then determined using calibration curves.

Main ions in water

The Ca^{2+} and Mn^{2+} were determined using the EDTA method and pan-method, respectively (APHA 2005). Fe^{2+} and Cl^- were determined through the HACH portable spectrophotometer procedures, namely, FerroVer and silver nitrate methods, respectively. The SO_4^{2-} concentrations were determined using SulfaVer 4 turbidimetric method as described in the HACH DR2800 series manual (HACH 2005).

Data analysis

The multivariate analysis of the data using PCA and CA enabled the identification of the sources of constituents and the distinguishing of the natural and anthropogenic contributions of pollutants into the lake system based on the level of association of the variables. The PCA and CA applied correlation ($\alpha = 0.05$) matrices to the variables in order to establish possible associations and input sources among polluting elements, as described by Delgado *et al.* (2010). In PCA, the eigenvalues of the principal components are a measure of their associated variances (Mellinger 1987; Meglen 1992; Wenning and Erickson 1994). Correlation of principal components and original variables is given by loadings. This treatment provides a small number of factors that usually account for approximately the same amount of information as the original set of observations. Cluster analysis uncovers intrinsic structure or underlying behaviour of a dataset without making *a priori* assumptions about the data. It further classifies objects of the system into categories (clusters) based on their similarity. In hierarchical CA, the distance between samples is used as a measure of similarity. In the present analysis hierarchical agglomerative CA was performed on the normalised data by means of the complete linkage (furthest neighbour), average linkage (between and within groups) and Ward's (1963) Euclidean distance method. The outputs were displayed as bi-plots in which the plotted points for sites were related to water quality parameters presented as rays. Both PCA and CA were done using XLSTAT 2013.2.04 package for Microsoft Excel.

Results

Physico-chemical parameters

Several physico-chemical parameters were considered in this study: temperature, pH, conductivity, turbidity, total dissolved solids (TDS), total suspended solids (TSS), total hardness (TH), and total alkalinity (TA) (Table 1). The temperatures at the study sites ranged from 18.1 to 29.6°C over the study duration. Crescent Lake recorded the lowest mean temperature (22.4°C) followed by the north-east site (22.5°C). During the study period pH ranged from 7.2 to 9.5, with the mouth of Malewa showing the highest variations (standard deviation, s.d. = 0.57). Mouth of Malewa showed the lowest mean conductivity (251 µS cm^{-1}) while Crescent Lake and north-east recorded the highest mean values of 421 and 358 µS cm^{-1}, respectively. The turbidity was lowest at Crescent Lake (mean = 10.05 nephelometric turbidity units (NTU)) and highest at mouth of Malewa (67.17 NTU) and the north-east (43.94 NTU). Total suspended solids covered a wide range of 1.0–432.0 mg L^{-1}, with highest and lowest values recorded at Kamere Beach and Crescent Lake, respectively. Total suspended solids ranged from 37.0 to 415.0 mg L^{-1}, with the low limit at mouth of Malewa and the high limit at Crescent Lake. Total hardness ranged from 0 to 120.0 mg $CaCO_3$ L^{-1}. The mouth of Malewa, Hippo Point and Kamere Beach stations recorded the lowest values of TH, while the highest values were recorded at the Crescent Lake and midlake stations. Total alkalinity ranged from 20 to 220 mg $CaCO_3$ L^{-1}, with lowest values at Kamere Beach, midlake and mouth of Malewa. Crescent Lake and north-east recorded the highest values for TA.

Nutrients

Nutrients analysed in the present study were nitrogen (as NH_4-N, NO_3-N and NO_2-N) and phosphorus (as PO_4^{3-}-P and TP) (Table 2). Concentrations of NH_4-N ranged from 0 to 0.51 mg L^{-1} and was highest at mouth of Malewa and north-east, while the south-east and midlake sites recorded the lowest values. The mean NO_3-N concentration ranged from 0.17 to 0.25 mg L^{-1}. The highest mean NO_3-N was recorded at mouth of Malewa followed by north-east, while Kamere Beach recorded the lowest values. Concentrations of NO_2-N ranged from 0 to 0.09 mg L^{-1}, with the lowest recorded at Hippo Point and the highest at south-east. Concentrations of PO_4^{3-}-P ranged from 0 to 0.08 mg L^{-1} and were similar at all the sampling sites. However, the mean TP, which ranged from 0.06 to 0.082 mg L^{-1}, was highest at the mouth of Malewa, south-east and north-east.

Ion concentrations

The main ions analysed were Ca^{2+}, Mg^{2+}, Fe^{2+}, Mn^{2+}, Cl^- and SO_4^{3-}, and their details are summarised in Table 3. The concentration of Ca^{2+} ranged from 0 to 43.2 mg L^{-1} and was highest at Crescent Lake and north-east, and lowest at Hippo Point. The Mg^{2+}concentration ranged from 0.0 to 24.0 mg L^{-1}, with lower values at Kamere Beach, midlake and mouth of Malewa. The south-east site recorded the highest Mg^{2+} concentration (24.0 mg L^{-1}). The Fe^{2+} concentration was lowest at Crescent Lake (0.01 mg L^{-1}), while the mouth of Malewa site recorded the highest value (1.98 mg L^{-1}). The Cl^- concentration ranged from 4 to 64 mg L^{-1}, with lower values at south-east, midlake, mouth of Malewa and north-east, and higher values at Crescent Lake and Hippo Point. Generally, SO_4^{3-} concentrations were highest at Hippo Point and Kamere Beach (16 mg L^{-1}), while Crescent Lake and mouth of Malewa recorded lower values (4 mg L^{-1}).

Multivariate analysis

Principal component analysis

Principal component analysis performed on the correlation matrix of means of the analysed water quality parameters by site

Table 1. Mean and range of physico-chemical parameters from the sampling sites of Lake Naivasha, Kenya between January and November, 2011

	Crescent Lake	Hippo Point	Kamere Beach	South-east	Midlake	Mouth of Malewa	North-east
Temperature (°C)							
Mean	22.4	23.3	24.0	24.4	24.4	22.9	22.5
Range	18.9–26.6	18.1–29.6	19.0–28.1	18.7–29.3	19.6–29.0	18.1–28.4	18.2–26.7
pH							
Mean	8.51	8.98	8.93	8.84	8.95	8.13	8.01
Range	7.97–8.95	7.98–9.28	7.85–9.27	7.97–9.5	8.16–9.3	7.20–9.28	7.26–9.11
Conductivity (µS cm^{-1})							
Mean	421	276	271	268	271	251	358
Range	384–526	226–322	216–310	159–313	225–307	74–305	289–392
Turbidity (NTU)							
Mean	10.05	22.83	23.06	25.50	23.47	67.17	43.94
Range	2.17–16.40	6.76–51.50	7.04–57.50	7.30–60.80	6.85–47.10	4.94–282.00	5.97–124.00
Total dissolved solids (mg L^{-1})							
Mean	205	144	138	140	138	124	175
Range	120–415	110–274	84–200	112–177	110–263	37–160	132–240
Total suspended solids (mg L^{-1})							
Mean	18	48	82	32	29	56	34
Range	1–101	4–152	1–432	4–124	2–112	1–211	3–93
Total hardness (mg $CaCO_3$ L^{-1})							
Mean	48	26	27	32	30	27	41
Range	0–118	0–62	0–62	0–74	0–108	0–62	0–120
Total alkalinity (mg $CaCO_3$ L^{-1})							
Mean	153	107	99	101	99	96	138
Range	68–220	72–180	20–144	38–204	50–140	32–136	68–192

Table 2. Mean and range of nutrients concentrations measured in the sampling sites of Lake Naivasha, Kenya during January through November, 2011

	Crescent Lake	Hippo Point	Kamere Beach	South-east	Midlake	Mouth of Malewa	North-east
NH_4-N ($mg\,L^{-1}$)							
Mean	0.057	0.068	0.055	0.045	0.048	0.085	0.083
Range	0.010–0.250	0.000–0.510	0.010–0.240	0.010–0.160	0.000–0.220	0.010–0.300	0.010–0.400
NO_3-N ($mg\,L^{-1}$)							
Mean	0.186	0.181	0.167	0.189	0.202	0.247	0.230
Range	0.100–0.800	0.050–0.700	0.100–0.600	0.060–0.800	0.020–0.600	0.100–1.100	0.100–0.700
NO_2-N ($mg\,L^{-1}$)							
Mean	0.012	0.010	0.012	0.014	0.013	0.013	0.013
Range	0.002–0.061	0.001–0.015	0.001–0.065	0.001–0.085	0.001–0.071	0.002–0.062	0.001–0.059
PO_4^3-P ($mg\,L^{-1}$)							
Mean	0.021	0.025	0.023	0.023	0.022	0.022	0.022
Range	0.001–0.079	0.001–0.079	0.001–0.077	0.003–079	0.000–0.079	0.001–0.079	0.004–0.079
Total phosphorus ($mg\,L^{-1}$)							
Mean	0.061	0.066	0.065	0.075	0.064	0.082	0.074
Range	0.031–0.174	0.030–0.176	0.032–0192	0.027–0.410	0.030–0.179	0.031–0.342	0.042–0.192

Table 3. Mean and range of main ions measured in the sampling sites of Lake Naivasha, Kenya, during January through November, 2011

	Crescent Lake	Hippo Point	Kamere Beach	South-east	Midlake	Mouth of Malewa	North-east
Ca^{2+} ($mg\,L^{-1}$)							
Mean	25.5	14.6	15.2	16.3	15.5	15.8	22.6
Range	16.0–43.2	4.8–24.0	8.0–20.6	6.4–28.8	8.0–22.4	0.0–32.0	8.0–33.6
Mg^{2+} ($mg\,L^{-1}$)							
Mean	1.5	0.7	1.1	2.3	1.2	1.3	1.7
Range	0.0–7.6	0.0–3.8	0.0–4.3	0.0–24.0	0.0–8.2	0.0–7.7	0.0–8.6
Fe^{2+} ($mg\,L^{-1}$)							
Mean	0.32	0.48	0.42	0.47	0.53	1.05	0.59
Range	0.06–1.03	0.17–1.19	0.18–0.90	0.04–1.46	0.15–1.22	0.31–1.98	0.01–1.03
Mn^{2+} ($mg\,L^{-1}$)							
Mean	0.126	0.206	0.229	0.236	0.312	0.292	0.236
Range	0.044–0.361	0.091–0.387	0.110–0.371	0.054–0.718	0.099–2.880	0.150–0.690	0.059–0.490
Cl^- ($mg\,L^{-1}$)							
Mean	21	22	19	18	18	18	18
Range	4–32	14–64	12–43	14–22	12–48	8–45	14–22
SO_4^{3-} ($mg\,L^{-1}$)							
Mean	1	2	2	1	1	3	1
Range	0–4	0–16	0–16	0–10	0–4	0–13	0–7

showed that four principal components represented ~94.2% of the total variation in the entire dataset. The actual eigenvalue and the percentage cumulative variability are shown in Fig. 2, and Table 4 summarises the corresponding eigenvectors. The first principal component accounted for 41.3% of the variation between sites and comprised the parameters TDS, conductivity, TA, TH, Ca^{2+}, Cl^- and Mg^{2+}. The second principal component accounted for 29.8% of the variation with temperature, pH, TSS, orthophosphates and Cl^- as the associated parameters. The third principal component explained 18% of the total variation between sites comprising ions (Ca^{2+}, Mn^{2+} and Mg^{2+}), nutrients (NO_2-N and TP) and physico-chemical parameters (pH and temperature). A further 5% of the total variation was explained by the fourth principal component, and 5.8% of the site variation was explained by the fifth and sixth principal components.

The bi-plot of the first and second principal components showed that turbidity in Lake Naivasha was closely associated with the nutrients (NH_4-N, NO_3-N, NO_2-N and TP) and Fe^{2+} and showed inverse relation to Cl^- (Fig. 3*a*). Most of these parameters mainly characterised the mouth of Malewa (Fig. 3*b*). Crescent Lake's distinctiveness was attributed to TDS, conductivity, TA, TH and Ca^{2+}. The parameter influencing the distinction in the north-east site was mainly Mg^{2+}, while Hippo Point, Kamere Beach, midlake and south-east sites were influenced by pH, temperature, orthophosphate and TSS, respectively.

Fig. 4, showing the results of the CA, indicates that the sampling sites varied and clustered into three distinct regions as follows: (i) northern region (mouth of Malewa and north-east), (ii) Crescent Lake and (iii) main lake (Kamere Beach, midlake, Hippo Point and the south-east).

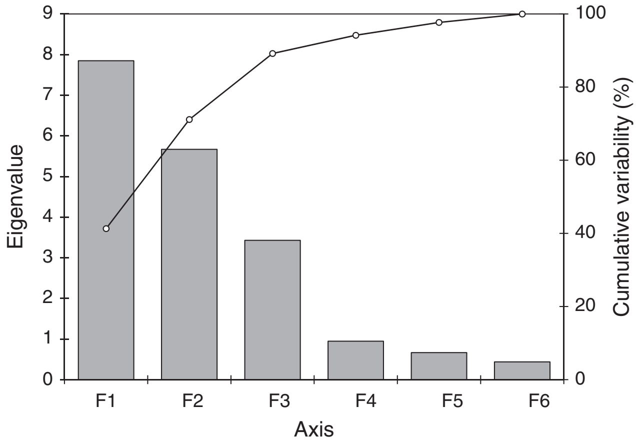

Fig. 2. Results of the principal component analysis showing eigenvalues (histogram) and cumulative variability (line with markers).

Table 4. Eigenvectors of the principal components

	F1	F2	F3	F4	F5	F6
Turbidity	−0.233	−0.309	−0.092	−0.033	0.027	−0.021
pH	−0.025	0.411	0.080	0.00011	−0.116	0.137
Temperature	−0.177	0.266	0.320	−0.00005	−0.008	−0.002
Conductivity	0.339	−0.125	−0.041	−0.023	0.072	−0.032
NH_4-N	−0.071	−0.313	−0.323	−0.045	−0.043	−0.324
NO_3-N	−0.117	−0.378	−0.012	−0.198	−0.260	0.013
NO_2-N	−0.071	−0.216	0.445	0.079	0.099	0.057
PO_4^{3-}-P	−0.125	0.172	−0.276	0.506	−0.404	−0.472
TP	−0.224	−0.291	0.022	0.354	−0.066	0.024
TSS	−0.215	0.107	−0.209	0.024	0.770	−0.246
Fe^{2+}	−0.251	−0.269	−0.098	−0.131	−0.079	0.306
Mn^{2+}	−0.293	−0.057	0.178	−0.410	−0.179	−0.197
Mg^{2+}	0.042	−0.158	0.406	0.537	−0.004	0.094
TA	0.322	−0.171	0.073	0.025	0.047	0.016
Ca^{2+}	0.308	−0.208	0.010	0.011	0.111	0.000
TA	0.327	−0.152	−0.069	0.026	0.006	−0.145
Cl^{-}	0.180	0.160	−0.376	−0.048	−0.232	0.414
SO_4^{3-}	−0.229	−0.080	−0.312	0.302	0.182	0.501
TDS	0.348	−0.089	−0.036	0.025	0.033	−0.049

Discussion

Physico-chemical parameters

Being a shallow tropical lake situated very close to the equator, the mean water temperature in Lake Naivasha did not vary much between the sites. However, Crescent Lake showed lower temporal variability in comparison with the rest of the sites because of the deep water at this site, which translates to relatively larger water mass which takes longer to warm up and cool down. The pH is relatively high which is common for lakes of volcanic origin (Chernet *et al.* 2001; Costantini *et al.* 2007). The pH at mouth of Malewa and north-east was somewhat lower in comparison with the other sites, which can be attributed to the high influx of fresh water from the Malewa, Gilgil and Karati rivers (Gaudet 1979; Stoof-Leichsenring *et al.* 2011). The range of conductivity in lake Naivasha seems to have widened from 282–374 µS cm^{-1}, as measured by Ballot *et al.* (2009) in 2001–05, to 74–526 µS cm^{-1} in the present study. Since the 1980s, emergence of multi-million dollar horticultural farms in the lake Naivasha catchment, as well as around the lake, led to urbanisation problems such as rapid informal settlements for the growing population and large water abstractions for irrigation; thus, exerting agricultural and domestic pollution pressures on Lake Naivasha (Becht and Harper 2002). Therefore, turbidity and TSS were generally high at mouth of Malewa due to the effect of surface runoff from the agriculturally rich catchment area. The Kamere Beach site recorded high turbidity which was attributed to the discharges from the Kamere informal settlements.

Being a crater lake, Crescent Lake recorded the highest TDS, probably due to its volcanic origin, which is often associated with high concentrations of dissolved minerals (Ayenew 2005). Furthermore, Crescent Lake station also recorded the highest levels of TH and TA, suggesting the presence of high concentrations of ions associated with the volcanic geology.

Nutrients

The mean concentrations and seasonal variations of NH_4-N, NO_3-N and TP were highest at the mouth of Malewa and north-east sites compared with the other sites. Presence of NH_4-N is an indication of domestic waste pollution while the other nutrients are closely associated with agricultural effluents from surface runoff (Kazi *et al.* 2009).This suggests that Lake Naivasha is experiencing high influxes of phosphorus and nitrogen from exogenous sources. Nitrites and orthophosphates concentrations showed little variation between the sampled sites. However, the turnover rate of orthophosphates in phosphorus-limited aquatic environments is extremely rapid, making TP the most informative measurement of phosphorus in surface waters (Wetzel 2001). In Lake Naivasha the mean phosphorus loading was estimated to be 0.6 g m^{-2} yr^{-1} in 1997–98 (Kitaka *et al.* 2002). During this study, total phosphorus loading data were not collected. However, the trophic state of the lake was found to have deteriorated based on comparative assessment of the total phosphorus trophic state index (TSI-TP) of 1998–97 and 2011 (Ndungu *et al.* 2013*a*).

Main ions

The concentration of Mg^{2+} was about half the concentration of Ca^{2+}; a phenomenon observed in other parts of the world (Grochowska and Tandyrak 2009). Both cations were generally higher in the Crescent Lake and north-east sites, explaining the high TH, TA and conductivity at these two sites. Studies in other parts of the world have expressed similar association between Ca^{2+} and Mg^{2+}, and TH, TA and conductivity (Prepas *et al.* 2001). The volcanic origin of Crescent Lake and the proximity of the north-east site to this satellite lake could plausibly explain the high Ca^{2+} and Mg^{2+} concentrations in the two sites. Concentrations of Fe^{2+} were generally high at the mouth of Malewa site, which is an indication that the high levels of Fe^{2+} in the lake were mainly emanating from surface runoff from the iron-rich catchment soils. The higher temporal variations (s.d. = 0.44) of Fe^{2+} in the mouth of Malewa sampling site, compared with other studied sites, may

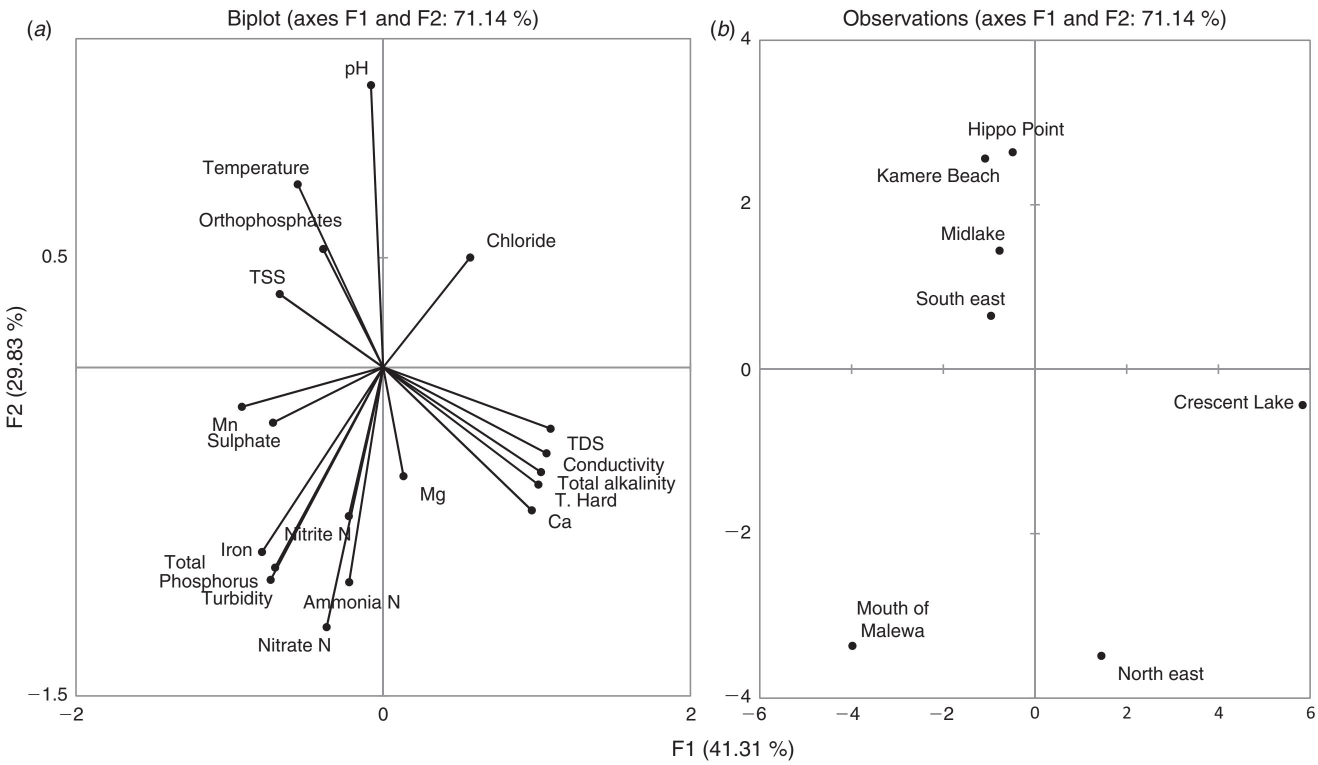

Fig. 3. Results of the principal component analysis for various water quality parameters measured in Lake Naivasha from January to November 2011: (*a*) bi-plot of the correlation between the water quality parameters in this study; (*b*) correlation between the studied sites in respect to the water quality parameters.

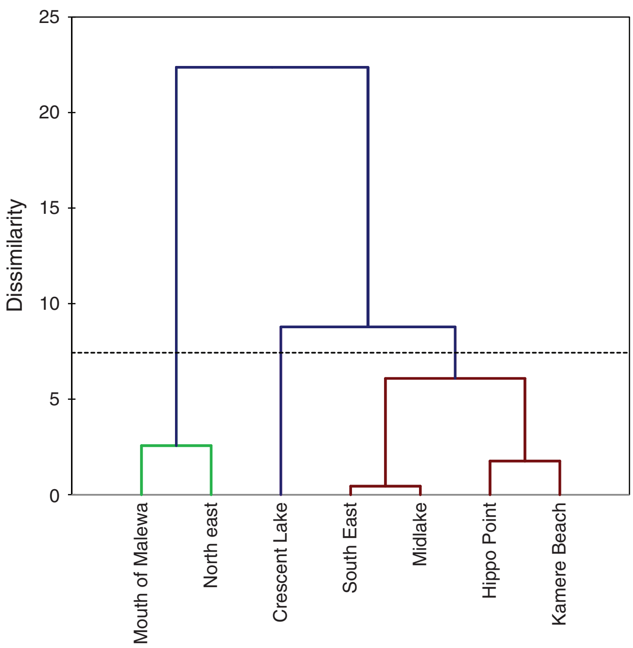

Fig. 4. Dendrogram of the dissimilarity between the three distinct areas of Lake Naivasha based on water quality parameters (dotted line denotes the truncation line that represents the stations that are somewhat homogeneous).

also be explained by the inflow variations between the wet and dry seasons. Lower Fe^{2+} concentrations were recorded in the Crescent Lake, which may be attributed to limited exchange of the river water into the area; and the occasional disconnection between the Crescent Lake and the main lake. The range of Fe^{2+} concentrations observed in this study (40–198.0 $\mu g L^{-1}$) falls within the range found in Ethiopian Rift Valley lakes (3.2–4699 $\mu g L^{-1}$) (Zinabu and Pearce 2003). However, Mn^{2+} concentrations observed in this study were found to be higher than observed in Kenyan and Ethiopian Rift Valley lakes (Zinabu and Pearce 2003; Ochieng *et al.* 2007). Ochieng *et al.* (2007) measured Mn^{2+} in Lake Naivasha sediment but could not detect dissolved Mn^{2+} in the water samples collected.

Multivariate analysis

Principal components analysis associated water turbidity with nutrients (NH_4-N, NO_3-N, NO_2-N and TP), SO_4^{2-}, Mn^{2+} and Fe^{2+}, which were the key parameters characterising the mouth of Malewa sampling site (Fig. 3), which suggests an influence of agricultural activities in the catchment. Higher turbidity was associated with mouth of Malewa in comparison with other sites, while TSS was higher in the main lake, implying presence of more dissolved particles around the river input region (mouth of Malewa). This is because dissolved particles are detected using the turbidity meter but passes through filter paper and are therefore not reflected in the TSS. The north-east region of the lake was associated with NH_4-N and Mg^{2+}. The association with NH_4-N can be explained by the close proximity to the Naivasha municipal treatment plant and Kihoto informal settlements which are sources of fresh organic material with high ammonium content. The high influence of Mg^{2+} in the characterisation may be indicative for interaction with Crescent Lake, whose natural mineral composition is associated with its volcanic origin. Conductivity, TDS, TA, TH and Ca^{2+} were more associated with Crescent Lake than with the other sites. Kilham (1990) also found close association between chemical composition of African lakes and volcanic rocks. Gaudet and Melack (1981) also associated African waters with the chemical composition of the underlying rocks. There appears to be no close association of Cl^{-} between the sites, and so may not be a strong discriminating parameter.

Cluster analysis resulted in three major regions: (i) northern region (mouth of Malewa, and the north-east), (ii) Crescent Lake and (iii) main lake (Kamere Beach, midlake, Hippo Point and the south-east). This grouping agrees with the results of discriminant analysis done on the trophic state variables (Ndungu *et al.* 2013*a*). The northern region is more influenced by external discharge, which consists of runoff from the agricultural activities in the catchment, urban waste water from the Naivasha municipal treatment plant and Kihoto informal settlement. Crescent Lake is a volcanic crater lake and therefore expected to have the chemical composition of the leached underlying volcanic rocks (Gaudet and Melack 1981). The main lake receives the external input from the northern region but it is characterised by wind-driven mixing (confirmed by a hydrodynamic study, unpub. data) making it quite homogeneous.

Other water quality studies that applied PCA and CA analysis found the techniques helpful in the interpretation of large datasets. Kazi *et al.* (2009) used PCA and CA in the analysis of water quality in Manchar Lake in Pakistan and found the techniques useful in apportionment of pollution sources based on parameter association. Though Kazi *et al.* (2009) used 36 parameters, their findings agree with the present study, particularly in the association between nutrients and catchment runoff, and NH_4 with domestic wastes. Magyar *et al.* (2013) used PCA and CA as well but the study had 33 sampling sites and 13 physico-chemical and biological water quality parameters, which helped in identifying the underlying processes responsible for the heterogeneity in different parts of Lake Neusiedler See in Hungary. The study also found the river input region to be significantly different. Sheela et al. (2012) applied PCA and CA to identify the factors influencing the water quality in the different seasons in Akkulam–Veli Lake in India. The study revealed that organic pollution was more significant during the dry season compared with the rainy season and decreased during the rainy season because of precipitation and tidal influence causing dilution. Principal component analysis and CA have also been applied in water quality studies in other lakes (Simeonov *et al.* 2010; Wenchuan *et al.* 2001).

Conclusions

The present study aimed to assess the status of the water quality in Lake Naivasha. Concentrations of different physico-chemical parameters, nutrients and the main ions were determined through field measurements. Turbidity was high in mouth of Malewa in comparison with other sites, implying presence of more suspended particles around the river input region. Nutrient concentrations were also high in the region around the river inputs and effluent of the waste water treatment (mouth of Malewa, north-east). Main ions were found to be high in Crescent Lake, which is a more or less isolated volcanic crater lake, due to the influence of underlying volcanic rocks.

The use of PCA and CA to provide an insight on water quality in Lake Naivasha showed the usefulness of such multivariate analysis in establishing the characteristics of different regions in aquatic ecosystems based on numerous water quality parameters. Three distinct regions were observed: (i) northern region (mouth of Malewa and north-east), (ii) Crescent Lake and (iii) main lake (Kamere Beach, midlake, Hippo Point and the south-east). Apart from provision of the distinction between the studied sites, water quality parameter associations provided information on the factors that influence the water quality of Lake Naivasha, which include agricultural activities and domestic effluent. The northern region (mouth of Malewa and north-east) of the Lake was dominated by agricultural activities in the catchment and domestic effluent. In Crescent Lake, natural mineral composition associated with its volcanic origin played a major role. The main lake is the recipient of the northern region effects but is somewhat well mixed. This information is fundamental especially in setting guidelines for effective ecosystem management particularly in the control of eutrophication.

Acknowledgement

We gratefully acknowledge the WOTRO Science for Global Development for funding this research work. We also wish to extend our gratitude to the administrative assistance provided by the Faculty of Geo-information and Earth Observation in the University of Twente, The Netherlands. The work would not have been completed without support and contributions from the Water Resource Management Authority (WRMA), Kenya; Fisheries Department (FiD), Kenya; the Kenya Marine and Fisheries Research Institute (KMFRI); and Egerton University, Kenya.

References

Alberto, W. D., María del Pilar, D., María Valeria, A., Fabiana, P. S., Cecilia, H. A., and María de los Ángeles, B. (2001). Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquía River Basin (Córdoba–Argentina). *Water Research* **35**(12), 2881–2894. doi:10.1016/S0043-1354(00)00592-3

APHA (2005). 'Standard Methods, 21st Edition.' (American Public Health Association: Washington, DC.)

Åse, L.-E. (1987). A note on the water budget of Lake Naivasha, Kenya. Especially the role of *Salvinia molesta* Mitch and *Cyperus papyrus* L. *Geografiska Annaler. Series A. Physical Geography* **69**, 415–429. doi:10.2307/521355

Ayenew, T. (2005). Major ions composition of the groundwater and surface water systems and their geological and geochemical controls in the Ethiopian volcanic terrain. *SINET: Ethiopian Journal of Science* **28**(2), 171–188.

Ballot, A., Kotut, K., Novelo, E., and Krienitz, L. (2009). Changes of phytoplankton communities in Lakes Naivasha and Oloidien, examples of degradation and salinization of lakes in the Kenyan Rift Valley. *Hydrobiologia* **632**(1), 359–363. doi:10.1007/S10750-009-9847-0

Becht, R., and Harper, D. M. (2002). Towards an understanding of human impact upon the hydrology of Lake Naivasha, Kenya. *Hydrobiologia* **488**, 1–11. doi:10.1023/A:1023318007715

Chernet, T., Travi, Y., and Valles, V. (2001). Mechanism of degradation of the quality of natural water in the lakes region of the Ethiopian rift valley. *Water Research* **35**(12), 2819–2832. doi:10.1016/S0043-1354(01)00002-1

Childress, R. B., Bennun, L. A., and Harper, D. M. (2002). Population changes in sympatric great and long-tailed cormorants (*Phalacrocorax carbo* and *P. africanus*): the effects of niche overlap or environmental change? *Hydrobiologia* **488**(1–3), 163–170. doi:10.1023/A:1023338816801

Costantini, M. L., Rossi, L., Scialanca, F., Nascetti, G., Rossi, D., and Sabetta, L. (2007). Association of riparian features and water chemistry with reed litter breakdown in a volcanic lake (Lake Vico, Italy). *Aquatic Sciences* **69**(4), 503–510. doi:10.1007/S00027-007-0917-Y

Delgado, J., Nieto, J., and Boski, T. (2010). Analysis of the spatial variation of heavy metals in the Guadiana Estuary sediments (SW Iberian Peninsula) based on GIS-mapping techniques. *Estuarine, Coastal and Shelf Science* **88**(1), 71–83. doi:10.1016/J.ECSS.2010.03.011

Gaudet, J. J. (1979). Seasonal changes in nutrients in a tropical swamp: North Swamp, Lake Naivasha, Kenya. *Journal of Ecology* **67**, 953–981. doi:10.2307/2259223

Gaudet, J. J., and Melack, J. M. (1981). Major ion chemistry in a tropical African lake basin. *Freshwater Biology* **11**(4), 309–333. doi:10.1111/J.1365-2427.1981.TB01264.X

Grochowska, J., and Tandyrak, R. (2009). The influence of the use of land on the content of calcium, magnesium, iron and manganese in water, exemplified in three lakes in the Olsztyn vicinity. *Limnological Review* **9**(1), 9–16.

HACH (2005). 'DR 2800 User Manual.' 1st edition. (HACH Chemical Company: Loveland, Colorado.)

Harper, D. (1992). The ecological relationships of aquatic plants at Lake Naivasha, Kenya. *Hydrobiologia* **232**(1), 65–71. doi:10.1007/BF00014613

Hubble, D. S., and Harper, D. M. (2001). Impact of light regimen and self-shading by algal cells on primary productivity in the water column of a shallow tropical lake (Lake Naivasha, Kenya). *Lakes and Reservoirs: Research and Management* **6**(2), 143–150. doi:10.1046/J.1440-1770.2001.00133.X

Kazi, T., Arain, M., Jamali, M., Jalbani, N., Afridi, H., Sarfraz, R., Baig, J., and Shah, A. Q. (2009). Assessment of water quality of polluted lake using multivariate statistical techniques: a case study. *Ecotoxicology and Environmental Safety* **72**(2), 301–309. doi:10.1016/J.ECOENV.2008.02.024

Kilham, P. (1990). Mechanisms controlling the chemical composition of lakes and rivers: data from Africa. *Limnology and Oceanography* **35**(1), 80–83. doi:10.4319/LO.1990.35.1.0080

Kitaka, N., Harper, D. M., and Mavuti, K. M. (2002). Phosphorus inputs to Lake Naivasha, Kenya, from its catchment and the trophic state of the lake. *Hydrobiologia* **488**(1–3), 73–80. doi:10.1023/A:1023362027279

Kundu, R., Aura, C. M., Muchiri, M., Njiru, J. M., and Ojuok, J. E. (2010). Difficulties of fishing at Lake Naivasha, Kenya: is community participation in management the solution? *Lakes and Reservoirs: Research and Management* **15**(1), 15–23. doi:10.1111/J.1440-1770.2010.00419.X

Lee, C. S.-l., Li, X., Shi, W., Cheung, S. C.-n., and Thornton, I. (2006). Metal contamination in urban, suburban, and country park soils of Hong Kong: a study based on GIS and multivariate statistics. *The Science of the Total Environment* **356**(1–3), 45–61. doi:10.1016/J.SCITOTENV.2005.03.024

Lake Naivasha Riparian Association (LNRA) (1999). 'Lake Naivasha Management Plan.' (LNRA: Naivasha, Kenya.)

Lung'Ayia, H., M'Harzi, A., Tackx, M., Gichuki, J., and Symoens, J. (2000). Phytoplankton community structure and environment in the Kenyan waters of Lake Victoria. *Freshwater Biology* **43**(4), 529–543. doi:10.1046/J.1365-2427.2000.T01-1-00525.X

Machado, L., Magnusson, M., Paul, N. A., de Nys, R., and Tomkins, N. (2014). Effects of marine and freshwater macroalgae on in vitro total gas and methane production. *PLoS ONE* **9**(1), e85289. doi:10.1371/JOURNAL.PONE.0085289

Magyar, N., Hatvani, I. G., Székely, I. K., Herzig, A., Dinka, M., and Kovács, J. (2013). Application of multivariate statistical methods in determining spatial changes in water quality in the Austrian part of Neusiedler See. *Ecological Engineering* **55**(0), 82–92. doi:10.1016/J.ECOLENG.2013.02.005

Meglen, R. R. (1992). Examining large databases: a chemometric approach using principal component analysis. *Marine Chemistry* **39**(1–3), 217–237. doi:10.1016/0304-4203(92)90103-H

Mellinger, M. (1987). Multivariate data analysis: its methods. *Chemometrics and Intelligent Laboratory Systems* **2**(1–3), 29–36. doi:10.1016/0169-7439(87)80083-7

Mergeay, J. (2004). Two hundred years of a diverse *Daphnia* community in Lake Naivasha (Kenya): effects of natural and human-induced environmental changes. *Freshwater Biology* **49**(8), 998–1013. doi:10.1111/J.1365-2427.2004.01244.X

Mugidde, R., Gichuki, J., Rutagemwa, D., Ndawula, L., and Matovu, X. (2005). Status of water quality an its implication on the fishery production. In 'The State of the Fisheries Resources of Lake Victoria and Their Management. Proceedings of the Regional Stakeholders' Conference'. (Ed. L.V.F.O. Secretariat.) pp. 106–112. (Jinja, Uganda.)

Ndungu, J., Augustijn, D. C. M., Hulscher, S. J. M. H., Kitaka, N., and Mathooko, J. (2013*a*). Spatio-temporal variations in the trophic status of Lake Naivasha, Kenya. *Lakes and Reservoirs: Research and Management* **18**(4), 317–328. doi:10.1111/LRE.12043

Ndungu, J., Monger, B. C., Augustijn, D. C. M., Hulscher, S. J. M. H., Kitaka, N., and Mathooko, J. M. (2013*b*). Evaluation of spatio-temporal variations in chlorophyll-a in Lake Naivasha, Kenya: remote-sensing approach. *International Journal of Remote Sensing* **34**(22), 8142–8155. doi:10.1080/01431161.2013.833359

Ochieng, E., Lalah, J., and Wandiga, S. (2007). Analysis of heavy metals in water and surface sediment in five rift valley lakes in Kenya for assessment of recent increase in anthropogenic activities. *Bulletin of Environmental Contamination and Toxicology* **79**(5), 570–576. doi:10.1007/S00128-007-9286-4

Pearce, A. R., Rizzo, D. M., Watzin, M. C., and Druschel, G. K. (2013). Unraveling associations between cyanobacteria blooms and in-lake environmental conditions in Missisquoi Bay, Lake Champlain, USA, using a modified self-organizing map. *Environmental Science & Technology* **47**(24), 14267–14274. doi:10.1021/ES403490G

Prepas, E. E., Planas, D., Gibson, J., Vitt, D., Prowse, T., Dinsmore, W., Halsey, L., McEachern, P., Paquet, S., and Scrimgeour, G. (2001). Landscape variables influencing nutrients and phytoplankton communities in Boreal Plain lakes of northern Alberta: a comparison of wetland- and upland-dominated catchments. *Canadian Journal of Fisheries and Aquatic Sciences* **58**(7), 1286–1299. doi:10.1139/F01-081

Reghunath, R., Murthy, T. R. S., and Raghavan, B. R. (2002). The utility of multivariate statistical techniques in hydrogeochemical studies: an example from Karnataka, India. *Water Research* **36**(10), 2437–2442. doi:10.1016/S0043-1354(01)00490-0

Sheela, A.M., Letha, J., Joseph, S., Chacko, M., Sanal kumar, S. P., and Thomas, J. (2012). Water quality assessment of a tropical coastal lake system using multivariate cluster, principal component and factor analysis. *Lakes and Reservoirs: Research and Management* **17**(2), 143–159. doi:10.1111/J.1440-1770.2012.00506.X

Shrestha, S., and Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. *Environmental Modelling & Software* **22**(4), 464–475. doi:10.1016/J.ENVSOFT.2006.02.001

Simeonov, V., Simeonova, P., Tsakovski, S., and Lovchinov, V. (2010). Lake water monitoring data assessment by multivariate statistics.

Journal of Water Resource and Protection **2**(4), 353–361. doi:10.4236/JWARP.2010.24041

Singh, K. P., Malik, A., Mohan, D., and Sinha, S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India): a case study. *Water Research* **38** (18), 3980–3992. doi:10.1016/J.WATRES.2004.06.011

Stoof-Leichsenring, K., Junginger, A., Olaka, L., Tiedemann, R., and Trauth, M. (2011). Environmental variability in Lake Naivasha, Kenya, over the last two centuries. *Journal of Paleolimnology* **45**(3), 353–367. doi:10.1007/S10933-011-9502-4

Tariq, S. R., Shah, M. H., Shaheen, N., Khalique, A., Manzoor, S., and Jaffar, M. (2005). Multivariate analysis of selected metals in tannery effluents and related soil. *Journal of Hazardous Materials* **122**(1–2), 17–22. doi:10.1016/J.JHAZMAT.2005.03.017

Vega, M., Pardo, R., Barrado, E., and Debán, L. (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research* **32**(12), 3581–3592. doi:10.1016/S0043-1354(98)00138-9

Wang, Y.-B., Liu, C.-W., Liao, P.-Y., and Lee, J.-J. (2014). Spatial pattern assessment of river water quality: implications of reducing the number of monitoring stations and chemical parameters. *Environmental Monitoring and Assessment* **186**(3), 1781–1792. doi:10.1007/S10661-013-3492-9

Ward, J. H. (1963). Hierarchical Grouping to optimize an objective function. *Journal of American Statistical Association* **58**, 236–244.

Wenchuan, Q., Dickman, M., and Sumin, W. (2001). Multivariate analysis of heavy metal and nutrient concentrations in sediments of Taihu Lake, China. *Hydrobiologia* **450**(1–3), 83–89. doi:10.1023/A:1017551701587

Wenning, R. J., and Erickson, G. A. (1994). Interpretation and analysis of complex environmental data using chemometric methods. *TrAC Trends in Analytical Chemistry* **13**(10), 446–457. doi:10.1016/0165-9936(94)85026-7

Wetzel, R.G. (2001). 'Limnology: lake and river ecosystems.' (Access Online via Elsevier.)

Zeng, T., and Arnold, W. A. (2014). Clustering chlorine reactivity of haloacetic acid precursors in inland lakes. *Environmental Science & Technology* **48**(1), 139–148. doi:10.1021/ES403766N

Zinabu, G., and Pearce, N. J. (2003). Concentrations of heavy metals and related trace elements in some Ethiopian rift-valley lakes and their inflows. *Hydrobiologia* **492**(1–3), 171–178. doi:10.1023/A:1024856207478