Winter 2021

Using Principal Component Analysis and Hierarchical Clustering to Explore Trends in the Water Quality, Metal Concentrations, and Algal Taxa Richness of 68 Lakes in Northwest Washington

Jeffrey Pratt
Western Washington University, jeffmpratt@live.com

Using Principal Component Analysis and Hierarchical Clustering to Explore Trends in the Water Quality, Metal Concentrations, and Algal Taxa Richness of 68 Lakes in Northwest Washington

By

Jeffrey Pratt

Accepted in Partial Completion
of the Requirements for the Degree
Master of Science

ADVISORY COMMITTEE

Dr. Ruth Sofield, Chair

Dr. Robin Matthews

Dr. Leo Bodensteiner

GRADUATE SCHOOL

Dr. David L. Patrick, Dean

Master’s Thesis

Jeffrey Pratt

12/1/2021

Abstract

This study analyzed water quality, metals concentrations, and algal taxa richness data from 68 lakes in Northwest Washington that have been sampled by Western Washington University's Institute of Watershed Studies. The primary goals of this analysis were to survey unmonitored lakes to gain a better understanding of the current conditions and to compare how lakes are characterized using combinations of the three data sets. Higher elevation lakes in the North Cascades were expected to have lower concentrations of metals and nutrients and more sensitive algae taxa than low elevation lakes in the Puget Sound Lowlands. When compared against Washington State benchmarks, some lakes exceeded aluminum, iron, copper, temperature, dissolved oxygen, and/or pH values. The most common exceedance was for a modified temperature criteria with 49 of the 68 lakes exceeding the least protective designated use criteria and 66 of the 68 lakes exceeding the most protective designated use criteria. Water quality and metal variables each produced clusters that supported traditional lake trophic state classifications. Algae taxa richness informed clusters showed significant differences between groups and the number of desmid and euglenoid taxa were used to characterize the clustering results. Water quality, metals, and algae richness all produced non-random clusters. Water quality clustering produced clustering largely along productivity variables. Metal clustering produced three clusters representing lakes with high, low, and variable metal concentrations. Cobalt, iron, chromium, and magnesium were significantly different for each cluster and may be useful indicators of total metal loading in lakes. Algae clustering produced less informative clusters than the other data sets due to the use of 419 variables for this analysis and may indicate that a different classification system or more sampling locations could produce a more parsimonious clustering result. Lake groupings between clusters were non-random; overlaying the metal clusters with the water quality clusters produced informative lake descriptions. Use of all three data sets provides a more complete representation of each lake and the groups of lakes as a whole than using only water quality parameters or trophic state. This study found no significant differences between lakes in the North Cascades and Puget Lowlands.

Acknowledgements

I appreciate the time and dedication that many of the faculty and staff at Huxley College have provided me at Western. Michael Hilles, Joan Pickens, and the students at the Institute for Watershed Studies provided much of the data and also ongoing support for my project. Thank you to Dr. Leo Bodensteiner for technical assistance and his thoughtful review of this thesis. Dr. Robin Matthews not only lead the Institute for Watershed Studies and provided taxonomic data, but also offered her endless enthusiasm and breadth of statistical knowledge as guidance. Many thanks to my advisor, Dr. Ruth Sofield, for her support and fortitude throughout the entire process.

Table of Contents

List of Figures

List of Tables

Introduction

Northwest Washington Lakes

While many studies have focused on the relationship between algal communities and water chemistry, few studies have encompassed the Puget Sound lowlands in Washington State and the nearby mountains in the North Cascades (Figures 1). The Puget Lowland ecoregion contains the area of Washington State bounded by the Olympic Mountains and Cascade Range and runs along the entire length of the Interstate 5 corridor from Canada to the Oregon Border (Sleeter et al. 2012). The ecoregion exhibits an average elevation of 150 meters and is dominated by the urban centers along its length. This ecoregion had the highest rate of land-cover change between 1986 to 1992 and 1992 to 2000 compared to all other ecoregions in the western United States. Within this, the largest net change in the Puget Lowland land cover between 1973 and 2000 was the estimated loss of 17.2 percent (1767 km^2) of the forest class. This loss of forested land was mirrored by the 53.8 percent (1186 km^2) increase in developed lands over the same period. These changes in land use towards deforestation and increased urbanization likely contributed to major changes in the quantity, quality, and seasonality of the water that flows through the Puget Lowland ecoregion's streams, lakes, and groundwater. Natural vegetation is the most common and effective shade to reduce heat load in surface waters (Shumar and de Varona, 2009) and the impervious surfaces indicative of urbanized lands transfer heat to stormwater and subsequent receiving waters (Van Buren et al. 2000). Additionally, urbanization contributes to altered flows and impaired water quality (LeBlanc et al. 1997). Some of these impairments include untreated wastewater and runoff that can carry mixtures of pollutants, including metals, as well as nutrients (Tong and Chen 2002, McGrane 2016). None of the lakes in this study had examples of untreated wastewater impairments. These alterations to ground and surface water chemistry can have meaningful impacts on aquatic species and the ability for surface waters to continue under their designated uses for Washington State, which are specified uses in the State's water quality criteria for water bodies or water body segments (WAC 173-201A-200).

The North Cascades ecoregion is adjacent to the Puget Lowland ecoregion and is composed of steep, mountainous terrain including the North Cascades National Park, the Mount Baker-Snoqualmie National Forest, the Okanogan—Wenatchee National Forest, as well as the Pasayten, Glacier Peak, Alpine Lakes, and Henry M. Jackson Wildernesses (Sleeter et al. 2012). All the lakes in my study within the North Cascades ecoregion are in the western or central portions of the ecoregion and are subjected to similar climates. The North Cascades, partially due to the multitude of legally protected areas, experienced roughly one-third of the land-cover change as a percent of the total land area between 1973 and 2000 that the Puget Lowlands did. In 2000, 70.3 percent of the North Cascades ecoregion was estimated to be covered by forest while only 0.6 percent was developed. More than 97 percent of changes in land use were due to timber harvesting, subsequent succession to grassland, and return to forested land. The North Cascades ecoregion is undergoing a very different land-use change to that of the Puget Lowlands and may offer a refuge for more sensitive algae taxa.

Research has been conducted on phytoplankton diversity, influence of watershed and soil parameters on water quality, and cyanobacteria blooms for lakes within the Puget Lowland ecoregion (Llewellyn 2010, Gravon 2013, Horton 2014). Likewise, there has been an effort to study some mountain lakes of the North Cascade ecoregion for the relationship between water quality and algal communities (Wong 2013, Pfannenstein 2016, Lawlor 2019). Mountain lakes are particularly sensitive to changes in climate and water chemistry (Skjelkvale and Wright 1998) whereas small, urban lakes such as those within the Puget Lowlands are often already impacted by eutrophication and other runoff related issues (Jacobsen and Nyholm 1986)

The water quality of individual mountain lakes is greatly influenced by the surrounding watershed. Some watershed characteristics such as terrestrial vegetation, geology, and watershed slope determine levels of nutrients and organic matter in mountain lakes (Wetzel 2001, Fureder et al. 2006, Tolotti et al. 2006). The geology of the North Cascades greatly varies with some lakes being impacted by a mix of glacial deposits, igneous, sedimentary, or metamorphic rocks/formations? from an array of periods. Meanwhile, the geology of the Puget Lowland lakes is dominated by Quaternary glacial deposits (Washington State Department of Natural Resources, 2015). Watershed drainage area, steepness, and aspect all impact the quality and quantity of water flowing into a lake.

Water Chemistry

Metals in surface water occur naturally due to geologic weathering and leaching of metal-laden soils. Alpine systems can also be influenced by atmospheric deposition via precipitation, wildfire smoke, volcanic activity, and industrial activity (Bradl 2005). Some alpine lakes are exposed to acid mine drainage from historical and modern mining activity. Many watersheds within Washington State contain geologic formations high in deposits of gold, silver, copper, lead, zinc, arsenic, and other metals (Broughton 1942). Metal speciation and bioavailability are complex processes and are influenced by a variety of factors including hydrodynamic energy within a lake, water temperature, stratification, biological activity, acidity, and organic carbon availability (Elder 1988). In addition to geologically introduced metals and those distributed by atmospheric processes, many lakes in developed areas are impacted by urban runoff. This runoff can contain elevated levels of metals such as copper, zinc, cadmium, tungsten, nickel, antimony, iron, and lead from vehicles, industry, construction, and leaching from developed surfaces (Brown and Peak 2006, McKenzie et al. 2009, Markiewicz et al. 2017, Muller et al. 2020).

In freshwater systems the two most studied nutrients are phosphorous and nitrogen as they are essential for growth and are often in short supply relative to other nutrients for aquatic photosynthesizing organisms (Schindler 1977). Nitrogen and phosphorous are found in particulate and dissolved phases and both are commonly in various chemical forms. Particulate matter can come from living and dead organic matter or inorganic sources such as minerals. Dissolved forms of nitrogen and phosphorous are often from the same sources as particulate matter but can also include dissolved nitrogen gas. This nitrogen gas and dissolved inorganic forms of nutrients are what plants and algae process. Eutrophication can occur when concentrated sources of nutrients such as fertilizers, septic waste, or soaps enter a water body and is a common issue in urban and peri-urban lake ecosystems (Jacobsen and Nyholm 1986, Putt et al. 2019). This eutrophication can lead to algal blooms that have negative effects on other aquatic species and can potentially be harmful to humans (Schindler 1978, Anderson et al. 2002).

Chlorophyll is a pigment that acts as a photoreceptor and allows photosynthesizing organism such as aquatic macrophytes and algae to convert simple molecules into more complex organic compounds (May 2006). There are six types of chlorophyll (A, B, C, D, E, and F) with A being the most common. Chlorophyll A, the primary photoreceptor for photosynthesis, has peak light absorptions in the red and blue wavelengths. Other forms of chlorophyll and other accessory pigments absorb additional wavelengths of sunlight and assist photosynthesis by transferring energy to chlorophyll A (Croce et al. 2003). Chlorophyll concentrations can be determined by measuring the fluorescence of a water sample and comparing that to a calibration curve or by measuring the absorption of light. This fluorescence measurement is light being emitted as radiation as chlorophyll molecules return to a non-excited state after absorbing an appropriate wavelength of light. Because chlorophyll A is the primary photoreceptive pigment in all freshwater algae, it can be used to estimate phytoplankton population biomass (APHA 2005). Lakes with higher levels of nutrients and productivity would be expected to have increased chlorophyll A levels.

Turbidity is a measure of water clarity based on the amount of light scattered by particles within the water column (EPA 2012; Perlman 2014). Turbidity measurements can be used as an indicator of water quality as they are related to total suspended solids. However, turbidity is not a direct measure of any one particle but instead measures the aggregate light scattering of suspended solids, dissolved organic matter, and other particles or dyes in the water column. As a measure of water clarity, turbidity can provide information on the ability of sunlight to penetrate the water column and the depth of the photic zone (Hakanson 2005). Additionally, turbidity is an important factor for aquatic systems as more suspended particles result in increased water temperature and decreased dissolved oxygen (EPA 2012). Urban runoff contributes to the turbidity of surface waters from direct contributions such as construction dust or particulates. Runoff can also have indirect contributions to turbidity as nutrient inputs can increase phytoplankton populations and detritus.

Similar to turbidity, alkalinity is not the measure of a single chemical. Alkalinity is the acid neutralizing capacity of a body of water (Mattson 2009). In freshwater systems the major buffering components are carbonate, bicarbonate, and hydroxides. Alkalinity of a water body is often due to the weathering of bedrock and soils near that body of water. Minerals rich in carbonates and strong base cations like magnesium and calcium increase the alkalinity of nearby water bodies but the weathering of minerals high in iron and sulfur can reduce alkalinity as ferric hydroxides and sulfuric acid are produced (Mattson 2009). Alkalinity is commonly measured by titrating a sample against a strong acid until the entire acid neutralizing capacity is depleted (USGS 2018).

Wetzel (2001) refers to dissolved oxygen (DO) as an essential factor second only to water itself for the field of limnology. Dissolved oxygen is the amount of oxygen present in water and is necessary for many forms of aquatic life. In areas of eutrophication, high rates of respiration deplete DO and can lead to death. Cold water has the capacity to hold more DO than warmer water. Dissolved oxygen and other factors such as temperature and chlorophyll are often not homogeneous throughout a lake. They are often stratified with each layer having a different composition of respiration, photosynthesis, and temperature. These factors lead to different concentrations of DO (Boehrer and Schultze 2008). Urban lakes which may be experiencing increasing temperatures and nutrient levels would be expected to have lower DO levels than cooler, less productive lakes.

Planktonic Algae as Bioindicators

Algae have an extensive history of use as indicators of water quality conditions as each taxon has specific requirements for, and responses to, physical and chemical conditions (Wetzel 2001; Jarvinen et al. 2012). While macroinvertebrates are generally used as freshwater indicators of environmental quality in the United States, algae can be used in a similar manner based on the ability of freshwater systems to support sensitive taxa (EPA 2017).

The nutrients present in a system can modify the algal community composition. Phosphorous and nitrogen act as limiting nutrients for algal growth (Morris and Lewis 1998; Morales-Baquero et al. 2006; Llewellyn 2010) and are commonly sourced from geologic processes, decomposition of plant and animal material, and human activities (USGS 2017). Cyanobacteria stand out in that they possess the ability to fix dissolved nitrogen gas while other taxa require inorganic salts of nitrogen or organic nitrogen. Taxa unable to fix nitrogen gas can uptake nitrogen and other essential growth nutrients through filter—feeding or engulfing bacteria and particulates (Caron et al. 1993; Herrero et al. 2001).

As lakes become nutrient-rich, especially with phosphorous, cyanobacteria blooms become more common and potentially toxin-forming taxa may be present. On the other side of the trophic state spectrum, some algae taxa prefer oligotrophic conditions. Many desmids are characteristic of low nutrient lakes that may be experiencing slightly acidic and dystrophic conditions (Bellinger and Sigee 2015). Some specific species of desmids such as *Closterium aciculare* are more prevalent in high-nutrient environments while other *Closterium* as well as *Comsarium* and *Staurastrum* desmids are common in minerotrophic fens with acidic, mineral and calcium-rich conditions.

In addition to nutrients, metals can modify algal community composition by acting as nutrients or toxicants. Plancho et al. (2015) found that many algal species were tolerant of metals, such as thallium, cadmium, lead, and zinc, that other organism are highly susceptible to, and that species in Euglenophyta were able to thrive in thallium concentrations up to even 240 ppb which is 120 times the US EPA surface-water standard. While some species may be tolerant of specific metals, a review on metal effects on algal communities by Prasad (2001) reports that both algal abundance and diversity were commonly reduced by highly metal-polluted sites.

Objective

This data was produced through Western Washington University's Institute for Watershed Studies Northwest Lake Monitoring Program. This program has been an ongoing public service since 2006 to provide baseline water quality data for local lakes that did not have ongoing monitoring programs. My first objective for this study was to fulfill the Northwest Lake Monitoring goal and analyze these available data sets for exceedances and trends in the water quality of the lakes. My second goal was to characterize the lakes using combinations of the available data to inform the ongoing monitoring program.

Methods

Sample Collection

Of the 68 lakes represented in this analysis, 62 were sampled by Western Washington University's Institute for Watershed Studies (IWS) as a part of its annual Northwest Lake Monitoring program (Figure 2). The other six lakes were more remote alpine lakes which were sampled by Michael Lawlor in collaboration with IWS for his thesis published in 2019. All lakes were sampled between 6/21/2017 and 8/16/2017. As Lawlor (2019) includes a time series for six lakes, for my work, the single sampling date closest to those of the IWS sampling was selected for each of the six lakes. Of the 68 lakes sampled, 45 fall within the Puget Lowlands ecoregion and 28 are within the North Cascades ecoregion.

The Northwest Lake Monitoring program is a public service established by IWS in 2006 to provide water quality data for lakes that do not have regular water quality monitoring programs. Lakes are sampled from public access points. For the 62 lakes sampled by IWS, a YSI EXO1 multiparameter field meter was used to measure temperature, pH, dissolved oxygen, and conductivity in the field (Table 1). Due to the distance of many sites from the IWS laboratory and the number of sites being collected per day, water samples were stored on ice and in the dark until they were returned to the laboratory. Phytoplankton samples were collected with a 20-µm net and returned to the laboratory unpreserved. In the IWS laboratory, subsamples were created and analyzed for metals, total nitrogen, total phosphorus, turbidity, alkalinity, and chlorophyll. Prior to analysis, subsamples for dissolved metals and dissolved organic carbon (DOC) were filtered through a 0.45—µm filter. Once returned to the IWS laboratory, phytoplankton samples were placed in an environmental chamber at 23 ºC and a 16L/8D light cycle until taxonomic identification could be performed.

Samples collected by Lawlor (2019) had different preservation techniques because they were not able to be quickly returned to IWS for storage and analysis. These samples were all separated into subsamples which were filtered, acid-fixed and chilled when appropriate in the field. These samples were analyzed following the same method, on the same equipment, and with the same quality control procedures once they were returned to IWS.

Laboratory Analysis

Algae

Algal identification was performed by Dr. Robin Matthews, Institute of Watershed Studies, Western Washington University, to the lowest practical taxonomic unit as described in Lawlor (2019). Algal data were formatted in a presence-absence list for each sample. Diatoms were excluded due to time constraints but an analysis including diatoms is available in Pfannenstein (2016).

Metals

Metals analysis was performed by Eric Lawrence (WWU). Samples were acidified with trace metal grade nitric acid and analyzed according to EPA standards (EPA 1994). The analysis was performed for all lakes over three separate days of analysis on Western Washington University's Advanced Materials Science & Engineering Center (AMSEC) Agilent 7500ce Inductively Coupled Plasma-Mass Spectrometer (ICP-MS) following the AMSEC ICP-MS Standard Operating Procedure for liquids based off of EPA (1994). Hardness is expressed in mg/L as calcium carbonate and is determined from the total calcium and magnesium concentrations found at each sampling site using the following equation:

$$\textit{Hardness (mg equivalent CaCO /L)} = 2.497\ \textit{(Ca, mg/L])}\ x\ 4.118\ \textit{(Mg, mg/L)}$$

Water Quality

Field measures by IWS or Lawlor with a YSI probe were made using YSI in-situ methodology and laboratory analyses were performed following American Public Health Association (APHA) methodology as described in Table 1.

Statistical Methodology

Statistical analyses were performed using the statistical software R (R Core Team, 2021). Any metal analyte with more than 40% of samples above 15% relative standard deviation (RSD) or more than 40% below detection limit (BDL) was removed from further analysis (Table 2). Metal analytes with less than 40% BDL had those BDL measurements randomly drawn from a uniform distribution between 0 and the daily limit of detection (LOD) by a two-point calibration (Croghan 2003). Elements with multiple analyzed isotopes that were also highly correlated were pared down to the most common isotope (Table 2). Water quality parameters with over 40% BDL or with 40% of samples above 15% RSD were removed from further analysis (Table 2).

Descriptive statistics and boxplots were used to visualize distinctions among lakes for water quality parameters, metals concentrations, and algal species richness. Bivariate correlations were performed between each water quality variable, metal analyte, and algal species richness. Kendall's tau rank-based correlation analysis was used due to the non-parametric nature of many of the variables (Kendall 1976). A significant correlation was defined as having a p-value equal to or below the α value of 0.05.

Principal component analysis (PCA) was applied to a centered and scaled data matrix derived from all water quality data to evaluate correlations and other relationships between variables (Praus 2005). The method was repeated for metals data and also for the algal species richness data.

Principal component coefficients from each of the three PCAs were applied using squared Euclidean distance with Ward's method to establish stable clusters (Ben-Hur & Guyon 2003). Cluster stability was determined using an iterative approach where the number of principal components used was reduced until clusters became unstable. The most parsimonious clustering results were used as they represented the most informed model while maintaining the ability to interpret the relative input of the original variables. Clusters were defined using the cutree function in R to create a designated number of clusters. Boxplots, univariate descriptive statistics, Kruskal-Wallis tests ($\alpha = 0.05$), and the nonparametric Wilcoxon rank-sum tests ($\alpha = 0.05$) were used to compare original variable data between clusters. A significant result was defined as having a p-value below the α value of 0.05.

Results and Discussion

Univariate

Metals

Boxplots showing range, median, and the first and third quartile for each metal analyte are shown in Figures 3-6. Summary statistics are shown in Table 3. Figure 7 shows all metal analyte boxplots on a single plot for comparative purposes. Calcium had the largest range of values from 390 to 28,072 µg/L. Magnesium and potassium followed with ranges from 59 to 8466 µg/L and 55 to 3293 µg/L respectively. Iron (2.66-1945 µg/L), manganese (0.58-410 µg/L), aluminum (13.1-219 µg/L), and arsenic (0.03-118 µg/L) were the only other elements with a comparatively large range.

The range of calculated hardnesses for the lakes (1.2-103.0 equivalent mg/L of $CaCO_3$) all fell within soft (0-60 mg/L) and moderately hard water ranges (60-120 mg/L) as seen in Figure 8. The relevant benchmark for the sampled metals are the Criteria Maximum Concentration (CMCs) and Criterion Continuous Concentrations (CCCs) from Washington State's aquatic life criteria (WAC 173-201A-200). The metal criteria that require site-specific hardness corrections were nickel, copper, zinc, and lead. In this analysis, samples were compared against the benchmarks based on the hardness value found at each sampling site. For example, the Criteria Maximum Concentration (CMC) for lead at the lowest and least protective hardness is 1.43 µg/L and the CMC at the highest and most protective hardness is 0.79 µg/L (Table 3). All other metal benchmarks in this study were not hardness corrected. Where there were no relevant Washington State criteria, acute and chronic benchmarks were obtained from British Columbia's aquatic life criteria (cobalt and antimony) or Indiana's Department of Environmental Management (molybdenum). None of these alternative criteria were exceeded (Table 3). Due to having more than half of the water samples outside of the model range for the current aluminum DOC-pH-hardness corrected criteria from 2018, the previous 1988 US EPA values were used, which do not correct for site specific water quality. All available benchmarks are included on Figures 3-6 for a visual comparison of recorded total metal values against CMCs and CCCs.

Table 3 gives values for each metal with applicable benchmarks and the number of samples when those criteria were exceeded. There were 20 relevant benchmarks: nine acute benchmarks and 11 chronic benchmarks. There were five instances of lakes exceeding acute benchmarks and 21 instances of lakes exceeding chronic benchmarks (Table 4). Of the 26 exceedances for CCCs and CMCs for any metal, 15 came from lakes within the North Cascades ecoregion and 11 from lakes within the Puget Lowlands ecoregion. With almost double the sampled lakes falling within the Puget Lowland ecoregion, this result is unexpected but not statistically significant ($n = 68$; χ-squared $= 2.79$, p-value > 0.05). Nine lakes exceeded the 1988 US EPA CCC for aluminum (87 µg/L), four exceeded Washington State's current

CCC for iron (1000 µg/L), and eight lakes exceeded Washington State's current copper CCC (hardness corrected) with five of those also exceeding the acute CMC (hardness corrected).

The nine lakes that exceeded the 1988 US EPA CCC for aluminum of 87 µg/L are Bear, Boardman , Evan, Geneva, Gold Mill Pond, Ketchum, Loma, Squalicum, and Vogler (Table 4). Bear Lake, Boardman Lake, Lake Evan, and Gold Mill Pond are all situated near each other on the Mountain Loop Highway within the North Cascades ecoregion and are accessed using forest roads. Lake Geneva and Squalicum Lake are located east of Bellingham, WA within the Puget Lowland ecoregion but immediately adjacent to the North Cascades ecoregion. Lake Geneva is located within a nature reserve while Squalicum Lake (Puget Lowland ecoregion) is surrounded by private homes with public fishing access. Ketchum Lake and Lake Loma are both lakes lined by private homes in northwest Snohomish County, WA within the Puget Lowland ecoregion and experience algae blooms. Vogler Lake is a warm, shallow lake located near Concrete, WA within the North Cascades ecoregion but has public access for non-motorized boats. Four of the five lakes within the North Cascades ecoregion all likely experience similar geologic impacts due to their proximity and Vogler Lake may be within a watershed experiencing similar impacts. Elevated aluminum concentrations within a watershed may be due to the underlying geological formation and the introduction of that formation into surface waters. Aluminum is the most abundant metal in the Earth's crust and nearly ubiquitous within surface waters due to leaching and atmospheric deposition. Another possible influence could be the addition of aluminum in the form of aluminum sulfate, or alum, as a treatment because it can be used to inactivate phosphorous and reduce the likelihood of an algal bloom. Alum treatments can be purchased by the public and are much more likely to have been used in lakes more at risk for cyanobacteria blooms such as those within the Puget Lowlands ecoregion.

Four lakes exceeded the CCC for iron of 1000 µg/L. These lakes were Beaver, Monte Cristo, Wiser, and Geneva (Table 4). Beaver and Wiser Lakes are suburban lakes in the Puget Lowlands which are often used for recreation and both experience algal blooms. Lake Geneva is in a nature reserve adjacent to a low traffic, two-lane road within the Puget Lowlands ecoregion. Beaver Pond is in the same nature reserve near Lake Geneva and it had an elevated iron concentration (715 µg/L), but did not exceed the CCC for iron. Monte Cristo Lake is in the North Cascade ecoregion and is influenced by historical mining activity in the area (Washington Department of Ecology 2011). Myrtle Lake has nearly one tenth the iron level (124 µg/L) of Monte Cristo Lake. Myrtle Lake is across the Mountain Loop Highway from Monte Cristo Lake and is thought to be much less influenced by the mining in the area. Similar to aluminum, iron is the second most prevalent abundant metal in the Earth's crust and is a common metal in freshwater lakes due to the natural weathering of rocks and soils. This natural occurrence can be elevated by mining activity which may have been a source of the high iron level in Monte Cristo Lake.

Of the eight lakes that exceeded their hardness-corrected CCC for copper, five also exceeded their acute hardness-corrected CMC (Table 4). The three lakes that only exceed their CCC are Evan, Loma, and Monte Cristo. Lake Loma is lined by homes north of the Tulalip Reservation in Snohomish County, WA, Lake Evan is located adjacent to a trailhead off the Mountain Loop Highway, and Monte Cristo Lake is a mining-impacted lake on the eastern end of the Mountain Loop Highway. Monte Cristo Lake had the highest measured copper concentration (4.39 µg/L), but the high hardness of the water also resulted in the highest hardness-corrected CCC (4.19 µg/L). The five lakes that exceeded both the chronic and acute criteria are Bear, Highwood, Picture, Sunset, and Terrell (Table 4).

Bear, Highwood, and Picture lakes are all located within the North Cascades ecoregion with Highwood and Picture being across Mt. Baker Highway from each other within the Mt Baker Ski Area. Road salts and brake particles on this highly trafficked road are likely contributing to their elevated levels. Interestingly, Highwood Lake was nearly triple the concentration of copper that Picture Lake does despite their nearly identical hardness-corrected CCCs and CMCs (Table 4). This difference could have been related to the parking lot located near Highwood Lake or the specific runoff dynamics from the encompassing roads into the surface waters. Sunset Pond is an urban lake in Whatcom County adjacent to a mall and is nestled in the joining of Interstate 5 and State Route 542. It was created by excavating material for the construction of Interstate 5. This proximity to impervious surfaces and a constant stream of automobile traffic may be increasing the amount of copper found in the lake. Lake Terrell is also a man-made lake in Whatcom County which provides habitat for stocked fish and waterfowl within a wildlife area. However, it is much larger at 500 acres. This peri-urban lake is surrounded by the wildlife area where the direct impacts from copper building materials and road runoff are not apparent. The size of this shallow lake in combination with atmospheric deposition may have contributed to the excess of copper (Davis et al. 2001). Vehicle emissions are more likely contributors to any local atmospheric copper amounts than the nearby oil refineries (Bosco et al. 2005, Ragothaman and Anderson 2017).

Water Quality

Boxplots for water quality parameters are shown in Figures 9-10. Each plot shows the data range, median, and first and third quartiles. When available, detection limits are included as dashed horizontal blue lines. Summary statistics are shown in Table 5.

The relevant criteria for these selected water quality parameters come from Washington State's aquatic life criteria (WAC 173-201A-200). For temperature and DO, there are different levels of criteria for different categories of freshwater designated uses. Included on Figure 9 for temperature and in the summary table (Table 5) are the most protective 7-day average of the daily maximum temperature (7-

DADMax) for waters designated for Char Spawning and Rearing (12 °C) and the least protective 7-DADMax for waters designated for Indigenous Warm Water Species (20 °C). Of the 68 lakes tested, only two single-point-in-time measurements were below the most protective 7-DADMax of 12°C. Upper Bagley Lake (8.1°C) and Lower Bagley Lake (8.1°C) are alpine lakes in the Heather Meadows Recreation Area near Mt. Baker. Terminal Lake (16 °C) is located between the Bagley Lakes but is a much smaller lake. When compared to the least protective 7-DADMax, 49 of the 68 measurements were above the 20 °C criteria. The highest recorded temperature measurement was from Vogler Lake (25.6 °C) which is a shallow lake near Concrete, WA. It should be noted that the measurements in this study are single data points that were taken at a single-point-in-time that was earlier in the day than the expected daily maximum (Woolway et al. 2016). These measurements are not the intended comparison for the 7-day average maximum required by Washington State's aquatic life criteria for temperature and the designated use of each lake is not considered. However, this failure of single measurements to meet the least protective 7-DADMax aligns with a trend of Washington's waters being impaired due to high temperatures. Of the 3,813 total causes of impairment for Washington's 303(d) Listed Waters in 2008, 988 of those were caused by temperature (US EPA 2021).

As with temperature, dissolved oxygen (DO) has different criteria for different designated uses in Washington State's aquatic life criteria. The relationship was the same; the most protective level was for waters designated for Char Spawning and Rearing (9.5 mg/L) and the least protective level was for Indigenous Warm Water Species (6.5 mg/L) (WAC 173-201A-200). Unlike the temperature criteria which were reviewed for a 7-DADMax, DO is compared against a Lowest 1-Day Minimum. Forty-six lakes were outside the most-protective criteria of 9.5 mg/L. Of those 46, seven were also below the least protective criteria. Beaver Pond had the lowest measured DO (2.21 mg/L) and was joined by Lake Geneva (6.3 mg/L), Gold Mill Pond (6.2 mg/L), Hoag Pond (4.94 mg/L), Monte Cristo Lake(6.13

mg/L), Summer Lake (6.3 mg/L), and Sunday Lake (4.13 mg/L). These are all shallow lakes and ponds.

The aquatic life criteria for pH in Washington State consistently ranges from 6.5-8.5 pH units. There are different allowances in human-caused variation for different designated uses, but those were not considered in this analysis. Only nine lakes fell outside of that range and all were measured at less than 6.5 pH units. The lowest measurement was from Heather Lake (5.43 pH units). Interestingly, of those nine lakes, Beaver Lake (6.1 pH units), Lake Geneva (6.3 pH units), Gold Mill Pond (6.2 pH units), and Summer Lake (6.3 pH units) were also below the least protective DO criteria.

Turbidity has a criterion based on exceedances over background levels that also differ for designated uses in Washington State. As background turbidity data were not available for all sampling sites, measurements were not compared against a criterion. Lower Bagley, Upper Bagley, Lower Twin and Upper Twin lakes all had turbidity levels at 0.2 NTU. All are alpine lakes accessible by short hikes within the North Cascades ecoregion. The site with the highest turbidity was Lake Erie (33.6 NTU) which is located on Fidalgo Island.

Algae

Of the 68 lakes that were sampled, 56 of them had algae samples retrieved (Appendix). From these 56 samples, Dr. Robin Matthews identified 419 algal taxa, not including diatoms, which were excluded due to time constraints. Honeymoon Lake had only 6 taxa while Hoag Pond had 78 taxa observed (Figure 11). Honeymoon Lake is a small residential reservoir located on Whidbey Island while Hoag Pond is a small residential pond with dense vegetation and wildlife populations. The data used for PCA were presence/absence for all 419 taxa at 56 sites. This gives an indication of species richness but not algal biomass or diversity indices that include both species richness and numerical abundance.

Bivariate

A non-parametric correlation matrix (Wei and Simko 2021) for all water quality, metals, and algae richness is shown in Figure 12. The larger circles and deeper colors denote a stronger correlation. There were 378 total correlations with 141 being significant ($\alpha = 0.05$). Non-significant correlations are shown with an 'x'.

The largest cluster of positively correlated variables contained both ionic strength related variables (alkalinity, calcium, magnesium, conductivity and other minor cations) and algal production related variables (total nitrogen, dissolved organic carbon, turbidity, total phosphorus, and chlorophyll). The metals manganese and iron were also included in the positively correlated cluster with algal productivity variables. These metals are known to be limiting factors in algal growth in fresh water (Liu et al. 2018). Algal richness did not have a significant positive correlation with any other variable. This can be accounted for as richness does not represent the productivity of the lake in the way that chlorophyll, total nitrogen, and total phosphorus do. Richness instead represents the number of taxa found at a specific lake.

Algae richness only had three significant correlations: vanadium, DO, and pH ($\tau = 0.204$; $\tau = 0.273$; $\tau = 0.264$). Vanadium has been shown to have a hormetic effect on algal growth in specific species, but there has not been research on the element's effects on taxa richness (Meisch et al. 1977). Due to the complex nature of water chemistry dynamics, it cannot be concluded that decreasing DO or pH leads to an increase in algae richness. One study of acidic bogs in Slovenia found that changes in pH and DO over time did lead to an increase in tolerant species, but that more parsimonious explanations for differences in algae richness were found in the shade, geology, and altitude of the study sites (Klemenčič et al. 2010). The use of single point measurements in my study differs from Lawlor (2019). Lawlor's temporal study of six lakes provided a much more specific study area where more specific significant correlations were found. The relatively small samples size of six lakes likely had less variability in unmeasured watershed characteristics than the 68 lakes in my study. Klemenčič et al. (2010) found that variables such as the shade, geology, and altitude explain most of the differences in species richness in their data.

Temperature had few significant correlations. Most interesting was the lack of a significant negative correlation between DO and temperature. This well-researched relationship normally is an example of the inverse relationship between temperature and the solubility of gaseous oxygen in water. The relationship may not be evident due to the sampling often being near to shore, at varying times of day, the large differences in ionic strength at sites, the effects of altitude on gas solubility, or the biological influences of respiration and photosynthesis.

Multivariate

PCA and Hierarchical Clustering

The principal components analysis (PCA) of the metal data included all 17 metals in Table 3 for all 68 lakes that were sampled. The first ten principal components (PCs) accounted for 95.0 % of the total variance (Figure 13). The ordination of the first PC accounted for 37.7 % of the total variance. The second PC accounted for 19.6 % of the total variance. The variable loadings for each PC are included in Figure 14 and Table 6.

Hierarchical clustering found that using eight PCs produced the most stable clusters with the fewest PCs (Figure 15). The first eight PCs accounted for 91.2% of the total variance of the metals data. Clustering on the first eight PCs produced three distinct non-random clusters (n= 68; χ-squared =10.5, p-value <0.01). Figures 16-32 show boxplots for all metals for each of the three clusters. Cluster 1 and Cluster 3 were well separated by PC1 (Table 7). Visually, Cluster 3 had the highest concentrations of metals, Cluster 1 had the lowest concentrations, and Cluster 2 was highly variable which limited the ability to distinguish it from the other two clusters. Cluster 1 and Cluster 3 were always significantly different from each other as determined by nonparametric Wilcoxon rank-sum tests. For manganese, zinc, and lead, Cluster 2 was not significantly different from Cluster 1 but was significantly different from Cluster 3. Cobalt was the best separator of all three clusters as each cluster was significantly different from the other two. Note that most of the zinc concentrations detected for Cluster 1 and Cluster 2 were below the limit of quantitation and would be expected to not be significantly different from each other.

All ten water quality variables (Table 5) from all 68 lakes were used for the water quality PCA. The analysis resulted in ten PCs containing 100% of the total variance as there were only 10 water quality parameters used in the PCA. The first PC accounted for 56.2 % of the total variation and was most impacted by total nitrogen, chlorophyll-a, turbidity, and conductivity (Figure 33; Table 8). The second PC accounted for an additional 15.1 % of the total variance and was influenced by pH, DO, and total phosphorus. This separation of influential variables in the first two PCs can also be seen visually by the separation of variables in Figure 34.

Hierarchical clustering based on the water quality PCA resulted in four clusters based off five PCs being the most parsimonious and stable clustering option (Figure 35). The first five PCs from the water quality PCA contained 94.2% of the total variance in the water quality data. These four clusters were found to be non-random (n= 68; χ-squared =10.5, p-value <0.05). Figures 36-45 show boxplots for all water quality parameters for each of the four clusters. Cluster 1 contains only four lakes, but those four lakes had significantly higher levels of every parameter except temperature (Table 9). Temperature was the least informative parameter for PC 1 (Figure 34) and had only eight significant correlations with all other variables in the total analysis (Figure 12). For total nitrogen, chlorophyll, turbidity, and DOC, Cluster 1 was elevated, Clusters 2 and 3 were not significantly different from each other, and Cluster 4 had lower levels (Figures 36-38 and 41;Table 9). The variables that separated Clusters 2 and 3 were conductivity, alkalinity, total phosphorus, pH, DO, and temperature (Figures 39-10, 42, 43-45; Table 9). Cluster 4 had significantly higher pH and DO values than Cluster 3. These are also the only two variables where all four clusters were significantly different from each other although there are visual overlaps between the ranges of each cluster for both pH and DO (Figures 43-44).

Algae samples were analyzed for 56 of the 68 lakes. For those 56 lakes, all 419 algal taxa identified during the sampling season were recorded as present or absent. The first ten PCs shown in Figure 53 represent 50.4% of the total variance in the data set. The first PC comprised 8.3% of the total variance and there is much less of a visual distinction between PCs in the algae PCA (Figure 46) than the metal (Figure 13) and water quality (Figure 33) PCAs. Figure 47 and Table 10 show that PC 1 was largely influenced by desmid taxa (“DE”) while PC2 was mainly impacted by euglenoids (“EU”).

Hierarchical clustering from the algae PCA produced 6 stable clusters from the first eight PCs (Figure 48). These eight PCs accounted for only 44.2 % of the total variance of the algae data. These six clusters were non-random (n = 56; χ-squared =18.4; p-value <0.01). Figure 49 shows boxplots for total taxa identified by cluster. With 56 lakes used in the algal analysis and six clusters formed, there are fewer lakes per cluster than the metals and water quality data clusters which limits the ability to compare clusters against each other. There are significant differences between the number of lakes purely based on the number of algae taxa identified. Clusters 1 and 3 had the highest number of taxa followed by Cluster 2 and Cluster 4 with Clusters 5 and 6 having the fewest algae taxa identified per lake (Table 11). There were many significant groups of clusters formed which are best understood visually in Figure 49.

As desmid and euglenoid taxa had greater impacts than chlorophytes and others on clustering, they were included in the algae cluster comparison. Numbers of desmid taxa observed had significant differences among algae clusters with Clusters 1 and 2 having significantly more desmids observed than Clusters 4, 5, and 6 (Figure 50; Table 11). Fewer significant differences were observed in numbers of euglenoid taxa between clusters with only Cluster 3 being higher than Clusters 2, 5, and 6 (Figure 51; Table 11).

An attempt to group sites known to produce harmful algal blooms by Dr. Robin Matthews was unsuccessful as there is not enough toxin data available from Washington State Ecology or Western Washington University's Institute for Watershed Studies. This supports previous work that has been able to predict algae blooms based on water quality parameters but not cyanobacteria blooms or harmful algal blooms (Llewellyn 2010). The addition of metals concentrations and algae taxa identification did not show any significant trends and the lakes known to have produced harmful algal blooms based on the information from the IWS did not have a higher proportion of potentially toxin-forming species to cyanobacteria algae than other lakes included in the analysis (Institute for Watershed Studies 2021).

Cluster Comparison

When comparing the clusters produced from each data set against each other, a mosaic plot can be used to show where there are positive, non-random associations and where there are negative, non-random associations. For example, when comparing the three clusters from the metals data against the four clusters from the water quality data, the association is non-random (χ-squared =56.2; p-value <<0.001). The overlap between clusters can be visualized as the size of the squares in Figure 52. The positive residuals (solid outlined) indicate a positive association and the negative residuals (dashed outline) indicate a negative association. There are large positive associations between metal Cluster 1 and water quality Cluster 3; metal Cluster 2 and water quality Cluster 1; and metal Cluster 3 and water quality Cluster 2. Water quality clusters were used for the following lake grouping discussion as they are the more prevalent form of understanding lake systems and provide the most coherent explanations for cluster overlaps.

A comparison between water quality clusters and algae clusters shows a non-random pattern (χ-squared =38.6; p-value <<0.001). This comparison is shown visually in Figure 53. There are positive associations between water quality Cluster 2 and algae Cluster 1 as well as between water quality Cluster 3 and algae Cluster 2. Algae Clusters 1 and 2 only have 6 and 4 lakes in them, respectively, so the trends that are observed are likely not applicable across the entire set of water quality clusters or the entire data set. These two algae clusters have high to average numbers of total algae taxa identified in them (Figure 53; Table 11), and they do have more desmids than other algae clusters (Figure 50).

Performing an association analysis on the metal clusters and algae clusters shows a non-random association (χ-squared =26.1; p-value <0.01) with a positive association between metal Cluster 3 and algae Cluster 4 (Figure 54). Algae Cluster 4 shows middling total taxa, desmid taxa, and euglenoid taxa richness (Table 11). Metal Cluster 3 shows elevated concentrations of almost all metals that have the greatest impact on clustering and is often significantly higher than Metal Clusters 1 and 2 (Table 7).

Lake Grouping Descriptions

Water Quality Cluster 1

These four low elevation lakes had the highest levels of almost all water quality parameters including chlorophyll, pH, DOC, and total phosphorus which indicate high productivity lakes. All four lakes are urban or peri-urban lakes located within the Puget Lowlands ecoregion. They are classified as hypereutrophic or eutrophic by Carlson's chlorophyll and phosphorus index (Carlson 1976). The lakes are classified in metals Clusters 2 and 3 and have elevated levels of hardness-related cations and iron. High counts of cyanobacteria taxa and the nutrients present indicate strong possibilities of cyanobacteria blooms that may contain toxin—forming taxa (EPA 2014, Bennett 2017; Appendix).

Water Quality Cluster 2

Of the 25 lakes in this cluster, 24 were also included in metals Cluster 2. Sunset Lake was included in metals Cluster 1. They are mostly classified as oligotrophic or mesotrophic by Carlson's chlorophyll parameter with two lakes classified as eutrophic. Using the phosphorus parameter provides similar classifications with the same two lakes classified as eutrophic but three other lakes join them for a total of five lakes classified as eutrophic by the phosphorus parameter. These lakes had elevated conductivity and metal concentrations similar to water quality Cluster 1. These intermediate elevation lakes are also prone to cyanobacteria blooms that may be toxic even if they are classified as oligotrophic. The lakes classified as oligotrophic are Gold Mill Pond, Monte Cristo, and Myrtle which are all within the North Cascades ecoregion. The only other lake within this water quality cluster is Evan Lake which was classified as mesotrophic by both parameters.

Water Quality Cluster 3

These 16 lakes did not significantly overlap with any metals groups or classifications using Carlson's index. They are distributed to all three metal clusters and have chlorophyll-based classifications from oligotrophic to eutrophic and phosphorus-based classifications from oligotrophic to hypereutrophic. However, this group of lakes is similar in their low DO, low pH, elevated metals such as arsenic and iron, and increased presence of euglenoid taxa. These lakes could be described as acidic bogs or having bog-like watersheds in the case of Monte Cristo. The presence of Monte Cristo, the lake with by far the highest concentration of mining-related metals, serves as an exaggerated example of the defining characteristics.

Water Quality Cluster 4

Most of the 23 lakes in this cluster were also in metals Cluster 1 with only five being classified into metals Cluster 2. Almost all the lakes were classified as oligotrophic by both of Carlson's parameters with six being classified as mesotrophic by the chlorophyll parameter and 3 being classified as mesotrophic by the phosphorus parameter. There are generally clear, unproductive lakes at high elevation. They have low turbidity, low pH, and higher presence of desmid taxa. This cluster contains almost all the alpine lakes within the North Cascades and includes some of the less impacted intermediate elevation peri-urban lakes such as Lake Samish and Reed Lake.

Conclusions

The premise of my study was to define similarities and differences in individual and groups of lakes across Northwest Washington. Sixty-eight lakes were analyzed for water quality parameters, metal concentrations, and 56 lakes for alga taxa richness. Univariate analysis showed the variability among lakes for each group of analytes with most analytes having a relatively small interquartile range with a high total range. By comparing the single analytes against Washington State benchmarks, a survey of impacted waters offers agencies an opportunity for succinct monitoring. A comparison of lakes exceeding metals benchmarks by ecoregion shows that lakes within the more urbanized Puget Lowlands ecoregion are not more likely to have water quality or metals parameters exceeding benchmarks than lakes within the North Cascades ecoregion. The underlying geology of Washington State may be contributing to elevated concentrations of metals in lakes within the North Cascades ecoregion.

A bivariate analysis reinforced expectations of certain water chemistry parameters. Nutrients, dissolved organic carbon, ionic strength, and chlorophyll were all correlated and clustered together by hierarchical clustering. Metals were mostly positively correlated with each other.

Multivariate analysis and clustering emphasized specific variables that are most useful for differentiating lakes. Principal component analysis and hierarchical clustering of water quality parameters, metal concentrations, and algal taxa richness each produced distinct clusters but all three data sets had significant overlaps with each other. While the algal taxa richness was informative, much of the information that may have been gleaned was lost when reducing the data to rank order or total richness. The unique knowledge of and experience with algae in Northwest Washington that Dr. Robin Matthews provided this study may limit the widespread application of this approach beyond fundamental measures of richness and diversity. The metals data set provided an interesting contrast as some analytes are biological building blocks while others, or even the same analytes in high enough concentrations, are measured in excess of state benchmarks. Due to the high correlation among metals concentrations, the abundance of essential minerals is often tied to a proportional presence of regulated and concerning elements.

Water quality parameters are the most common approach to understanding biological conditions in a lake and were the most informative approach to lake grouping. However, the additional information offered by the metals and algal taxa richness produced a much more complete analysis and elucidated similarities that may have not been identified otherwise. Use of all three data sets provided a more complete representation of each lake and the groups of lakes as a whole than using only water quality parameters or an index of trophic state.

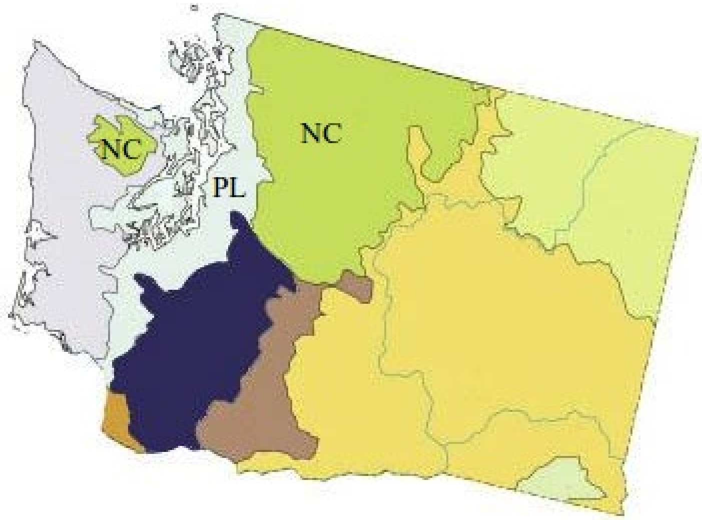


Figure 1. Washington State ecoregions with Puget Lowlands labeled as PL and North Cascades labeled as NC. Adapted from Thorson et al. 2003.

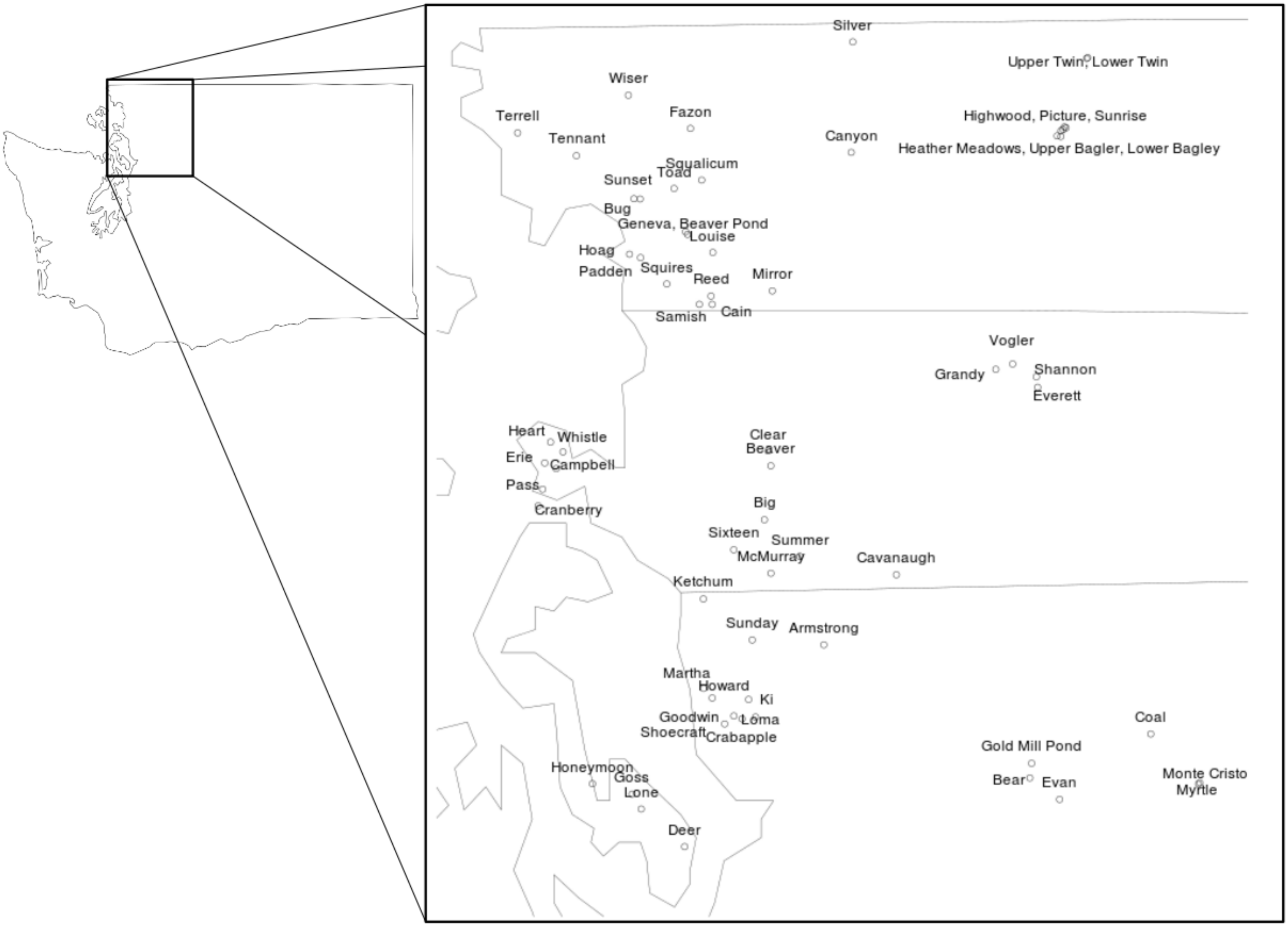


Figure 2. Map of sampling locations for the Northwest Lake Monitoring project Western Washington University's Institute for Watershed Studies. The delineation between lakes classified as Puget Lowlands and North Cascades is shown in red. An interactive version with site descriptions and water quality data is available at https://www.wwu.edu/iws/.

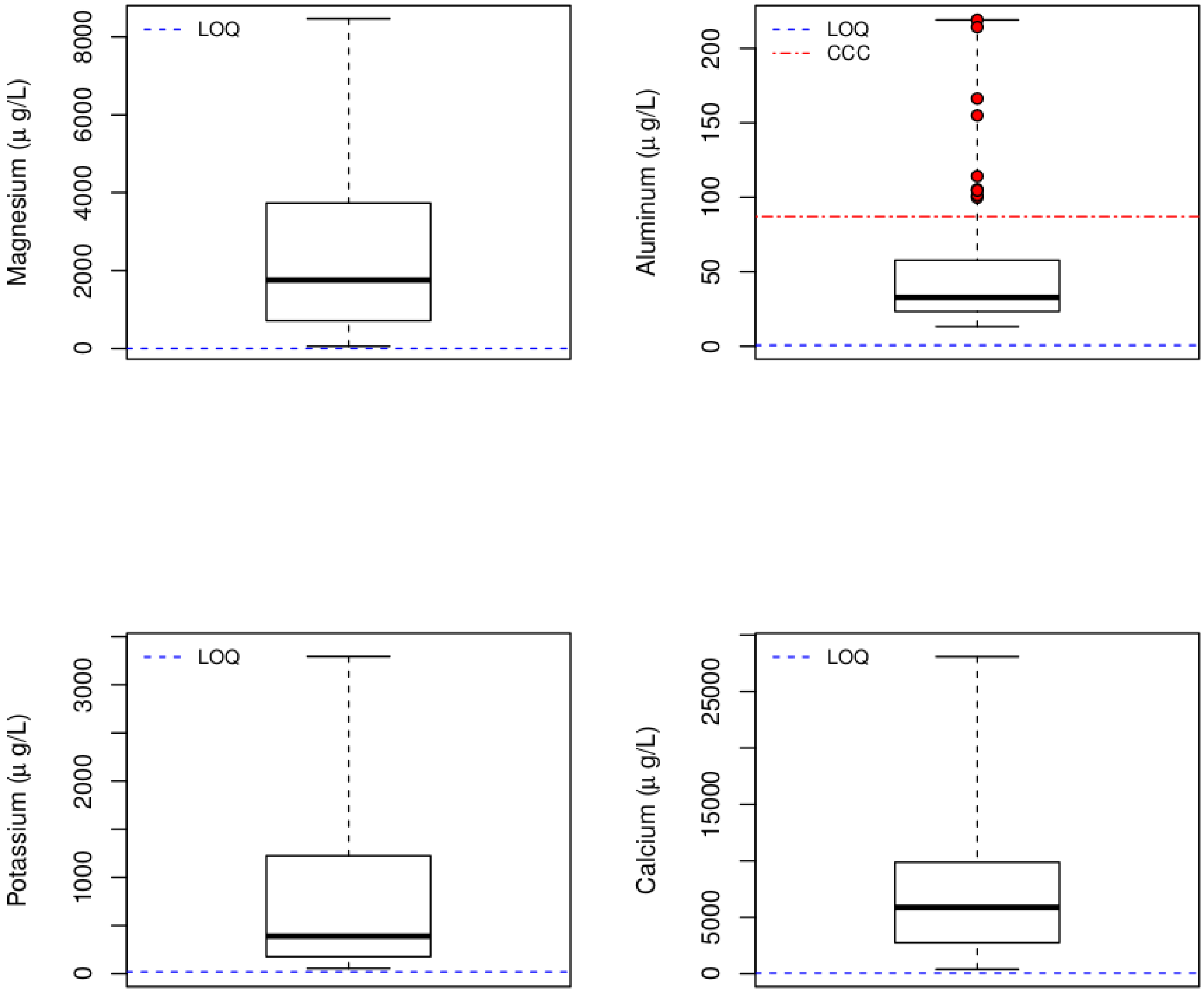
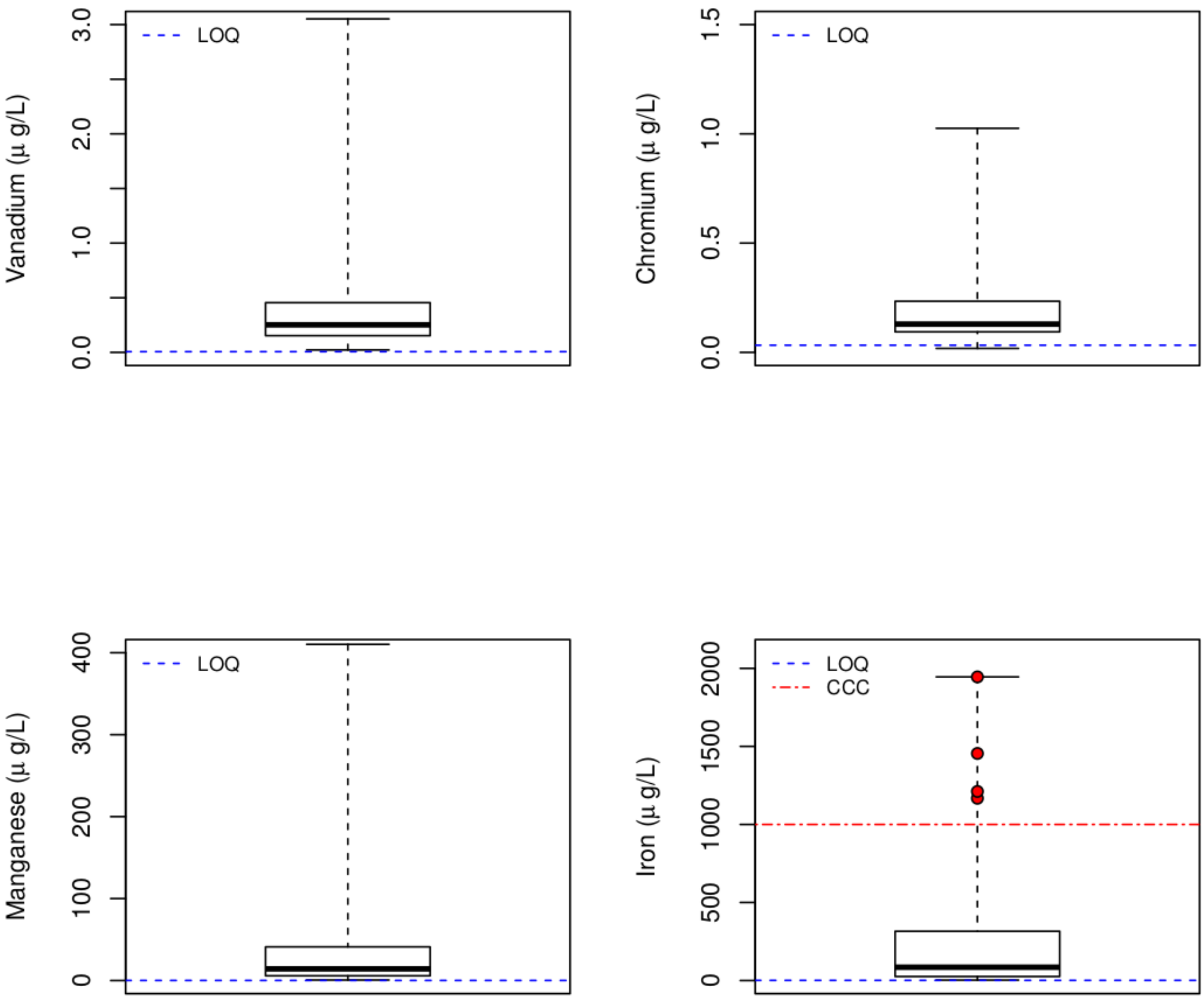
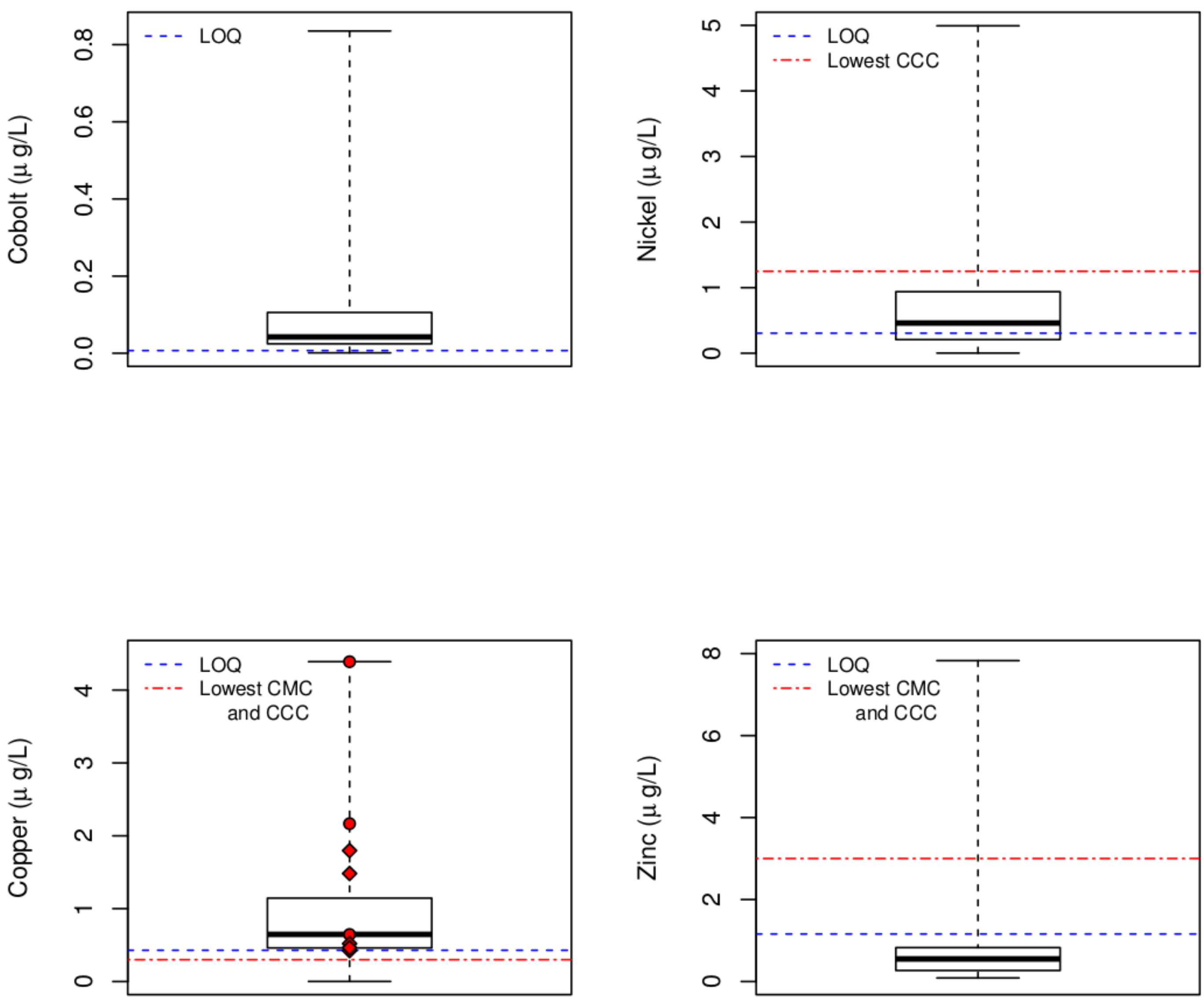


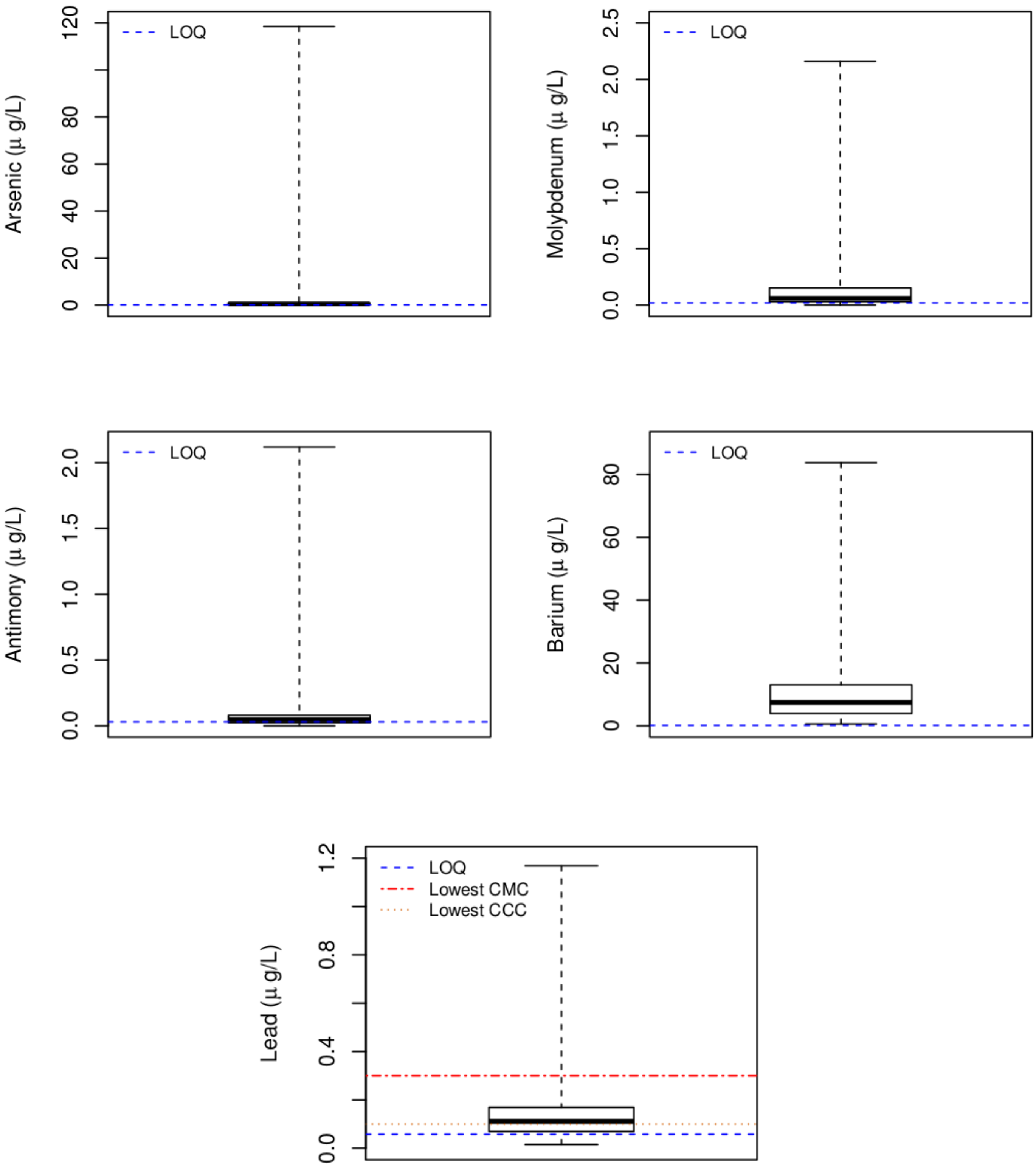
Figure 3. Boxplots of magnesium, aluminum, potassium, and calcium concentrations at all 68 lake sampling sites. Shown on the boxplots are minimum, lower quartile, median, upper quartile, and maximum. The average limit of quantitation is shown as a dashed horizontal line. Benchmarks are shown as a dot-dash or dotted horizontal line. Lakes exceeding their chronic benchmark are shown as points and those that exceed their acute benchmarks are shown as diamonds.



Figures 4. Boxplots of vanadium, chromium, manganese, and iron concentrations at all 68 lake sampling sites. Shown on the boxplots are minimum, lower quartile, median, upper quartile, and maximum. The average limit of quantitation is shown as a dashed horizontal line. Benchmarks are shown as a dot-dash or dotted horizontal line. Lakes exceeding their chronic benchmark are shown as points and those that exceed their acute benchmarks are shown as diamonds.



Figures 5. Boxplots of cobolt, nickel, copper, and zinc concentrations at all 68 lake sampling sites. Shown on the boxplots are minimum, lower quartile, median, upper quartile, and maximum. The average limit of quantitation is shown as a dashed horizontal line. Benchmarks are shown as a dot-dash or dotted horizontal line. Lakes exceeding their chronic benchmark are shown as points and those that exceed their acute benchmarks are shown as diamonds.



Figures 6. Boxplots of arsenic, molybdenum, antimony, barium, and lead concentrations at all 68 lake sampling sites. Shown on the boxplots are minimum, lower quartile, median, upper quartile, and maximum. The average limit of quantitation is shown as a dashed horizontal line. Benchmarks are shown as a dot-dash or dotted horizontal line. Lakes exceeding their chronic benchmark are shown as points and those that exceed their acute benchmarks are shown as diamonds.

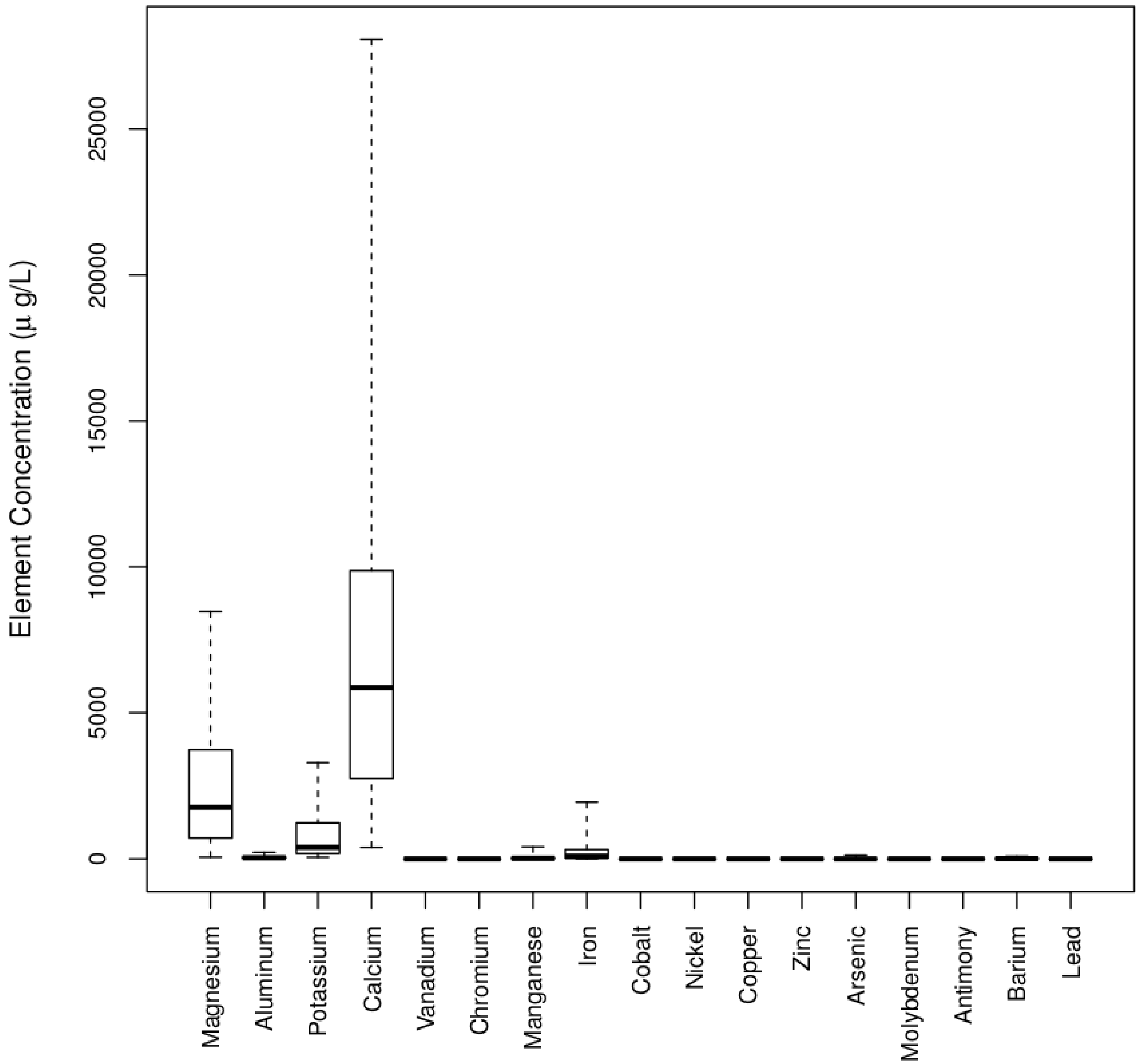


Figure 7. Boxplots of all metal analytes. Shown on the boxplots are minimum, lower quartile, median, upper quartile, and maximum.

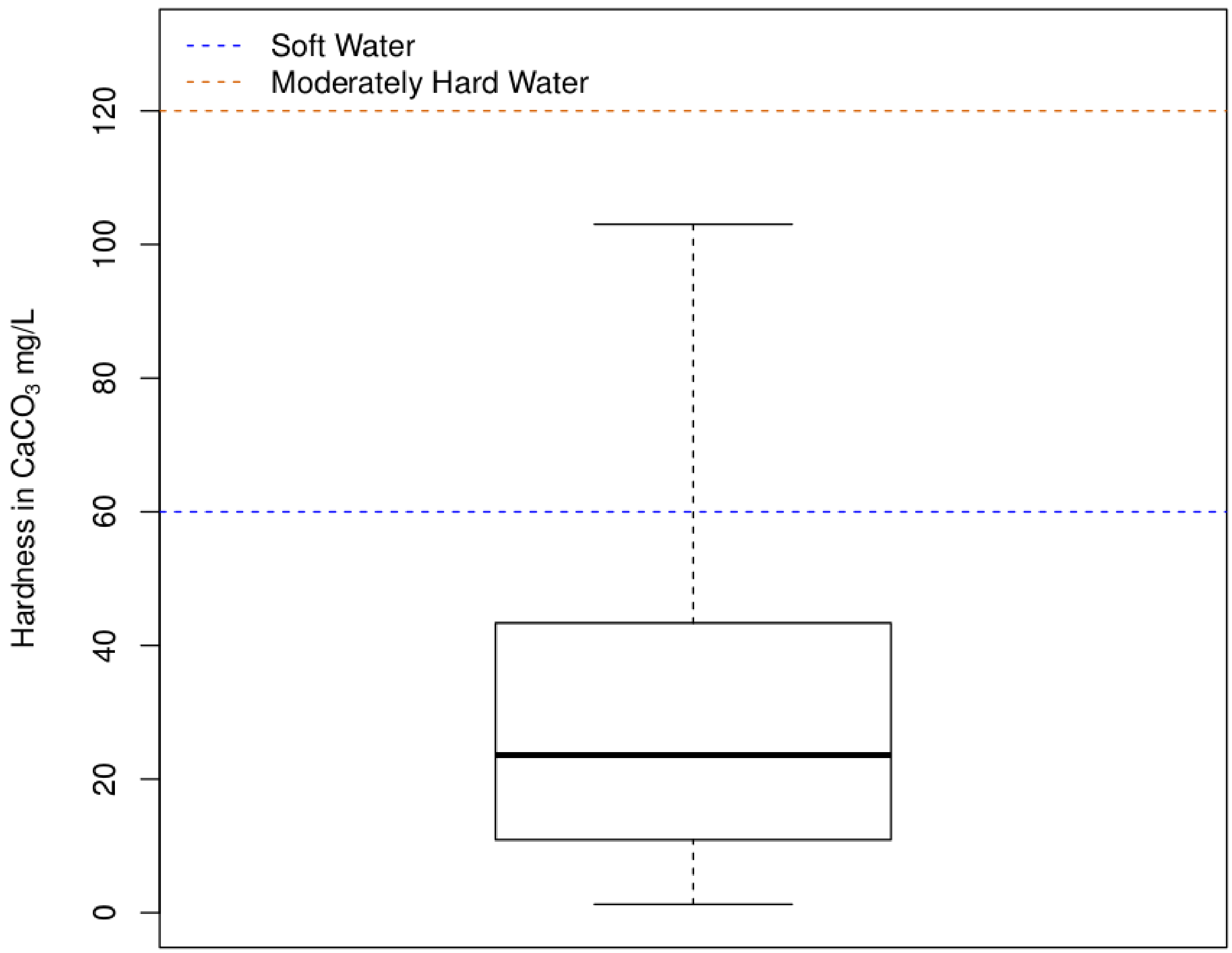
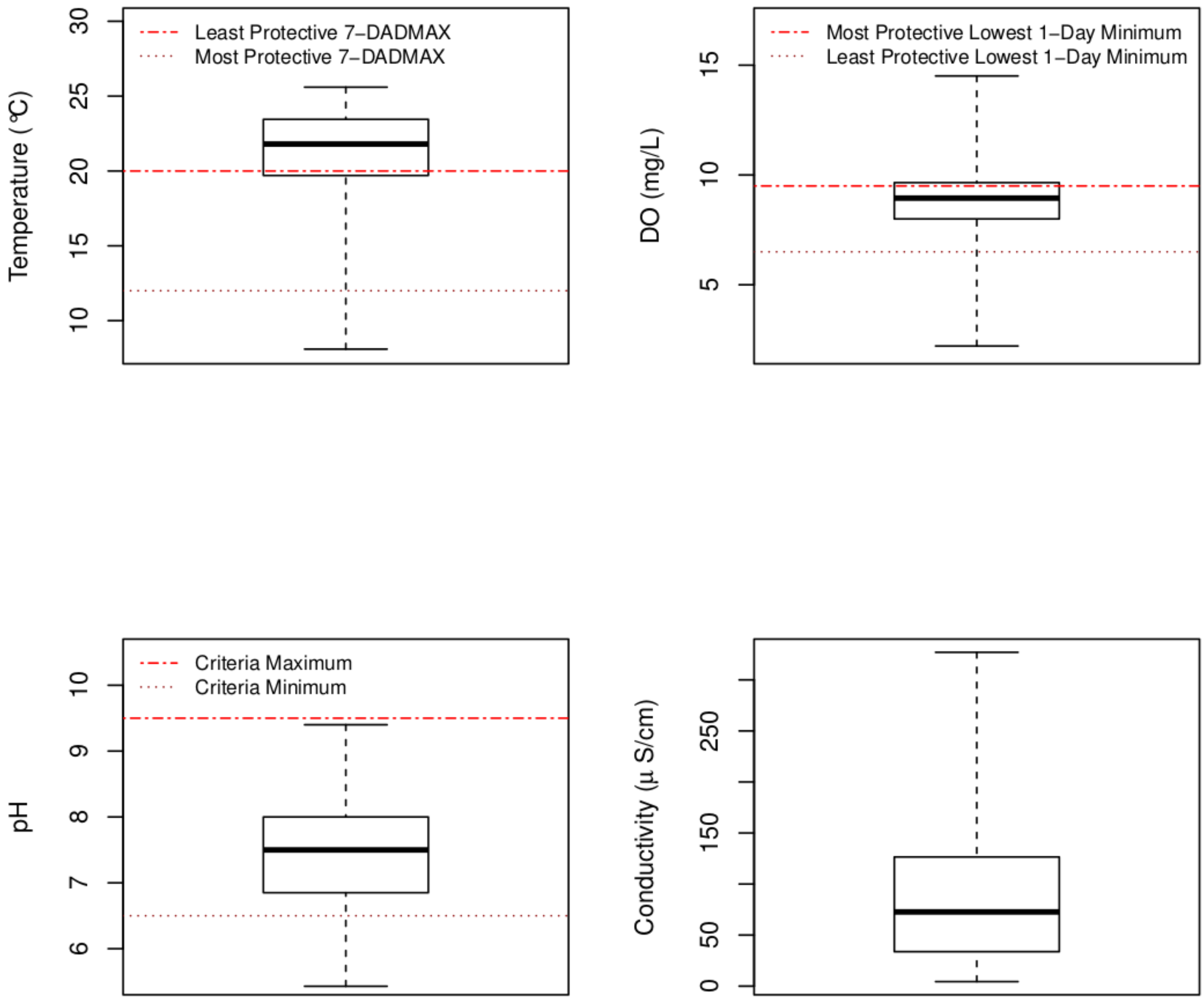
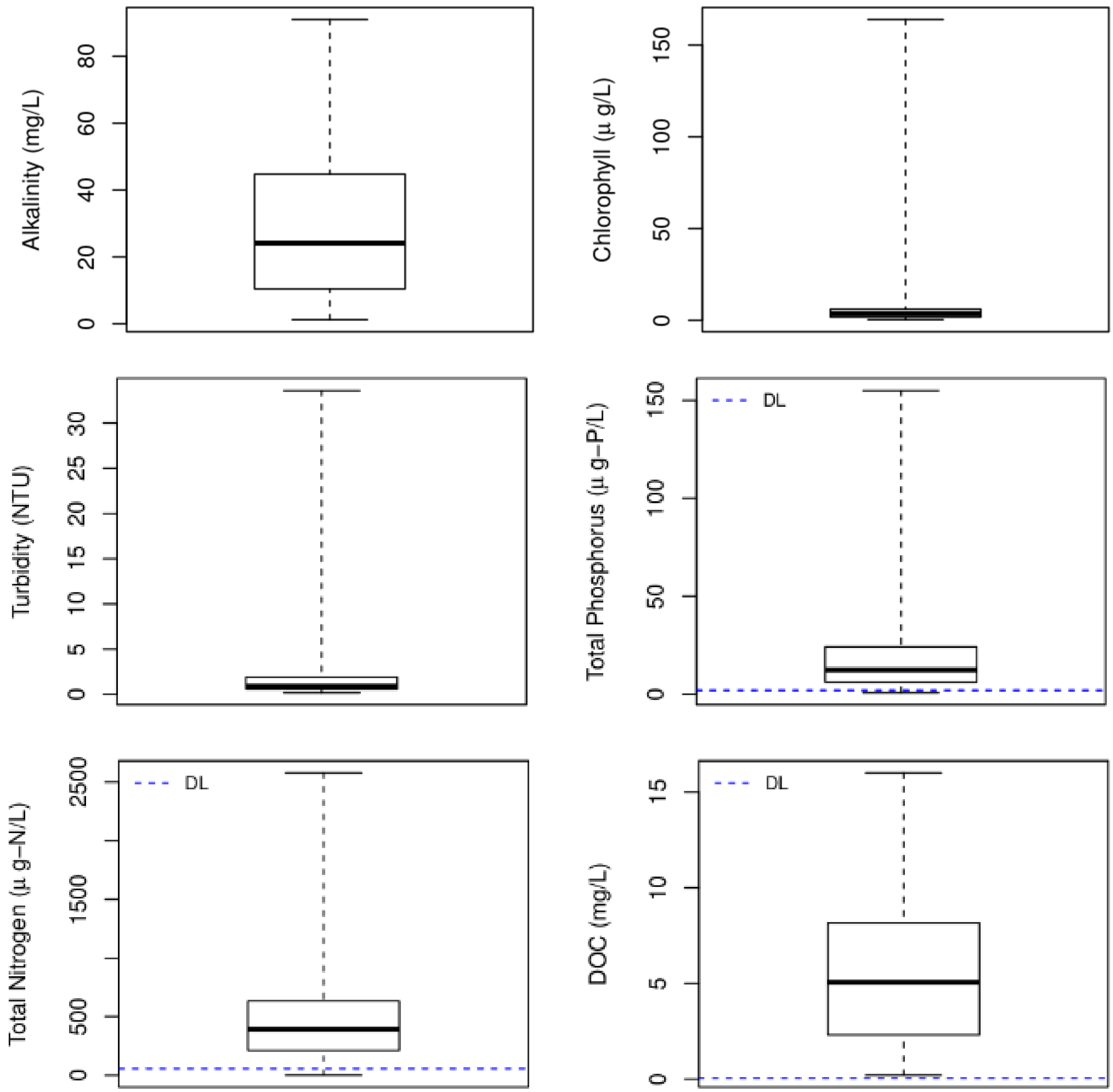


Figure 8. Boxplot of hardness at lake sampling sites determined from calcium and magnesium concentrations. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Nine lakes were characterized as moderately hard water (60-120 mg/L) and fifty-nine were characterized as soft water (0-60 mg/L).



Figures 9. Boxplot of water quality measurements at all 68 lake sampling sites. Shown on the boxplots are minimum, lower quartile, median, upper quartile, and maximum. Relevant Washington State criteria are shown as dotted and dot-dash horizontal lines.



Figures 10. Boxplots of water quality measurements at all 68 lake sampling sites. Shown on the boxplots are minimum, lower quartile, median, upper quartile, and maximum. Detection limits are shown as dashed horizonatal lines.

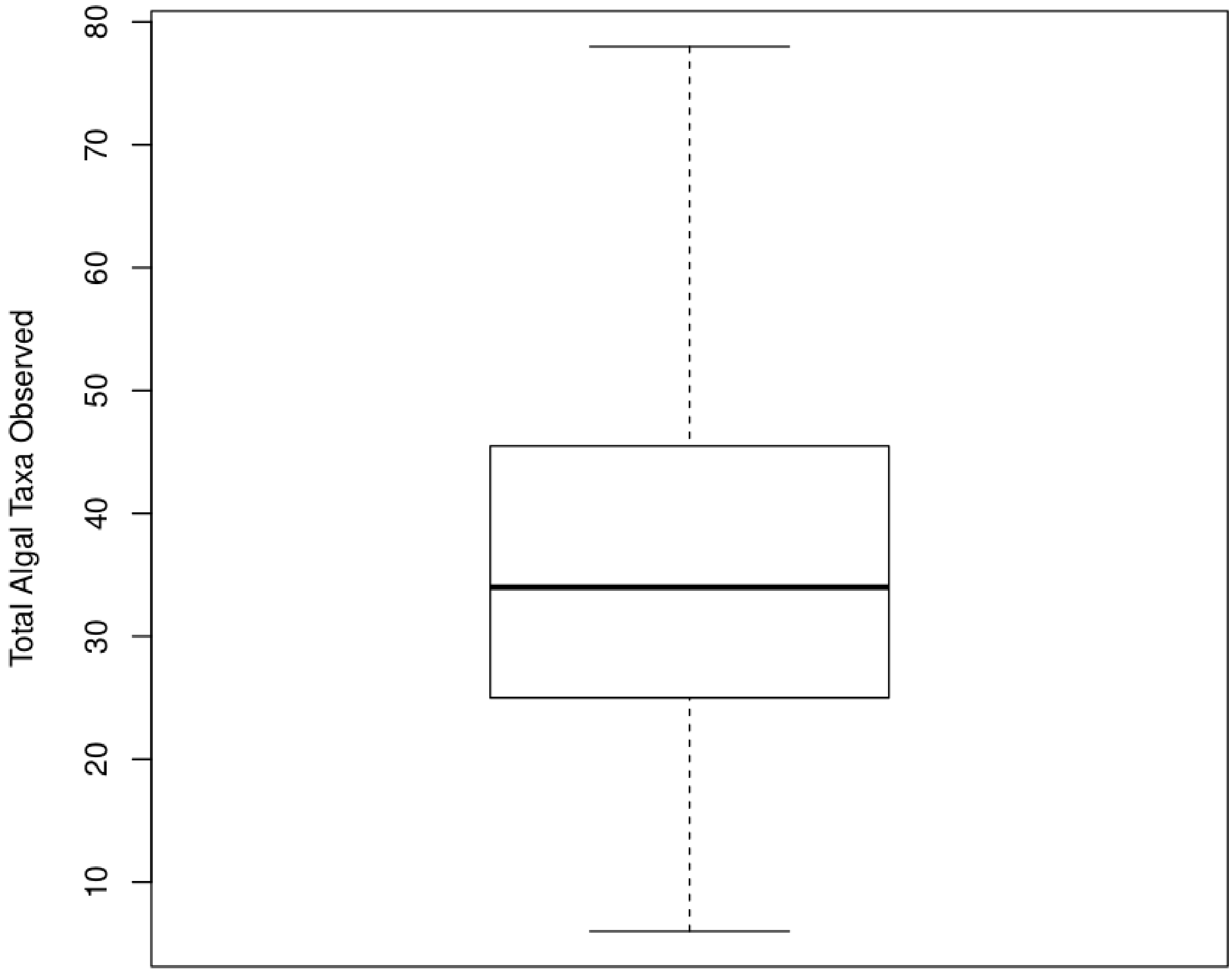


Figure 11. Boxplot of total algal taxa observed in each lake. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Algae were identified to the lowest practical taxinomic unit.

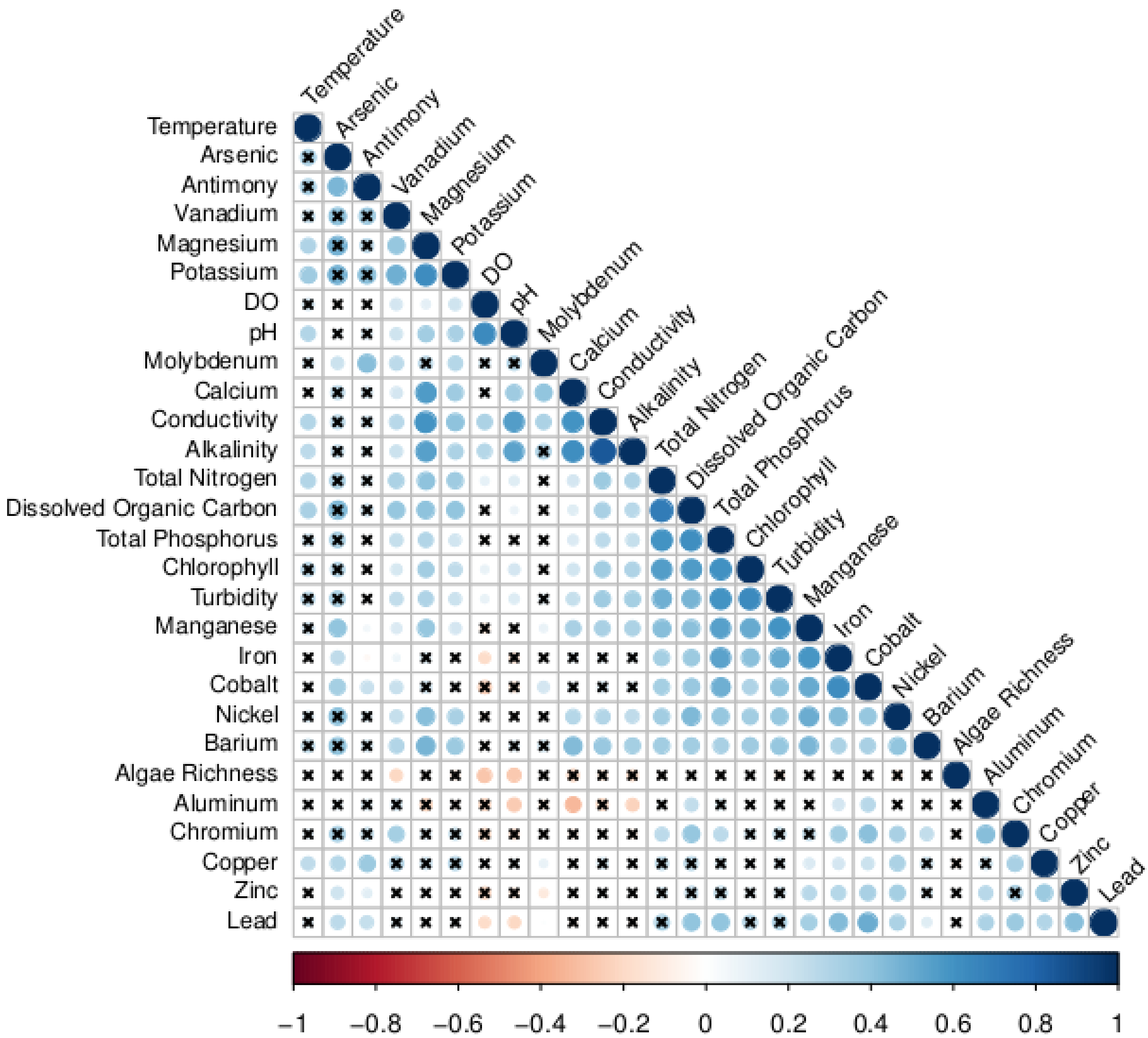


Figure 12. Correlation plot for all metal and water quality data. Larger circle and deeper color means more positively/negatively correlated. 'x' indicates not significant p-value (a = 0.05). There are 378 total correlations of which 141 are significant. Hierarchical clustering was applied to group variables with those with the most similar correlations. Correlation variables and significance statistics are available in Appendix.

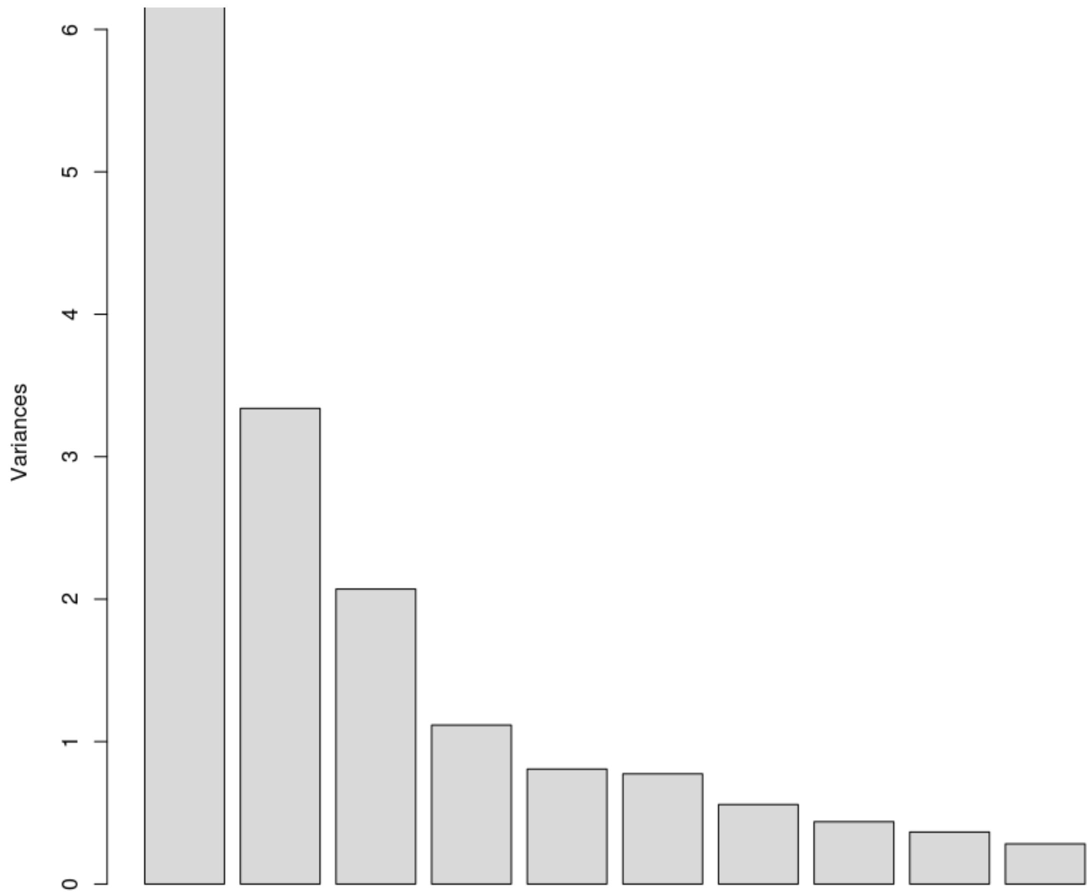


Figure 13. Variance plot of the first ten principal components (PCs) for the PCA based on the total metals data. These ten PCs account for 95.0% of the total variance.

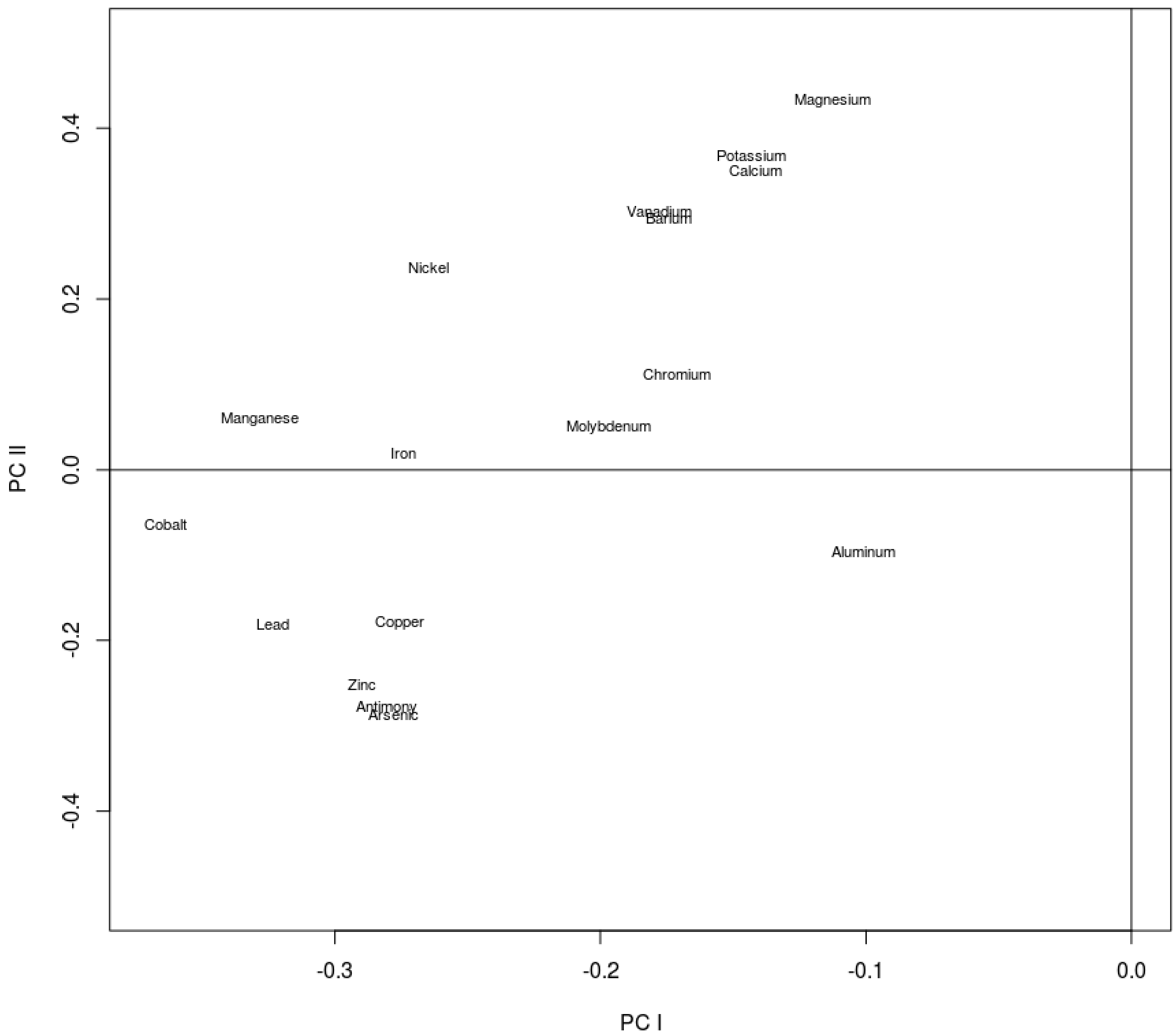


Figure 14. Ordination along PC I and PC II from the PCA of the total metals data.

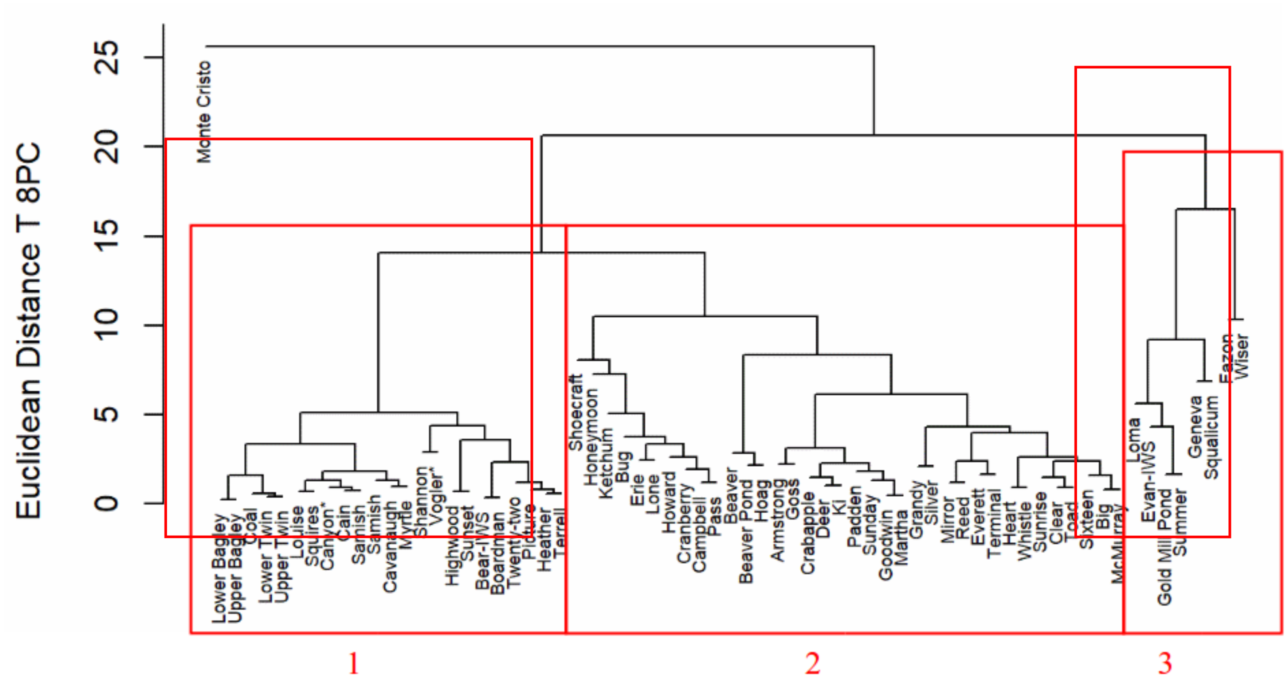


Figure 15. Hierarchical clustering of sampled lakes using total metals concentrations. Using the first eight principal components produced the most parsimonious clustering. These components account for 91.2 % of the total variance. Monte Cristo was included in Cluster 3. Clusters are non-random. n = 68; χ-squared =10.5; p-value <0.001.

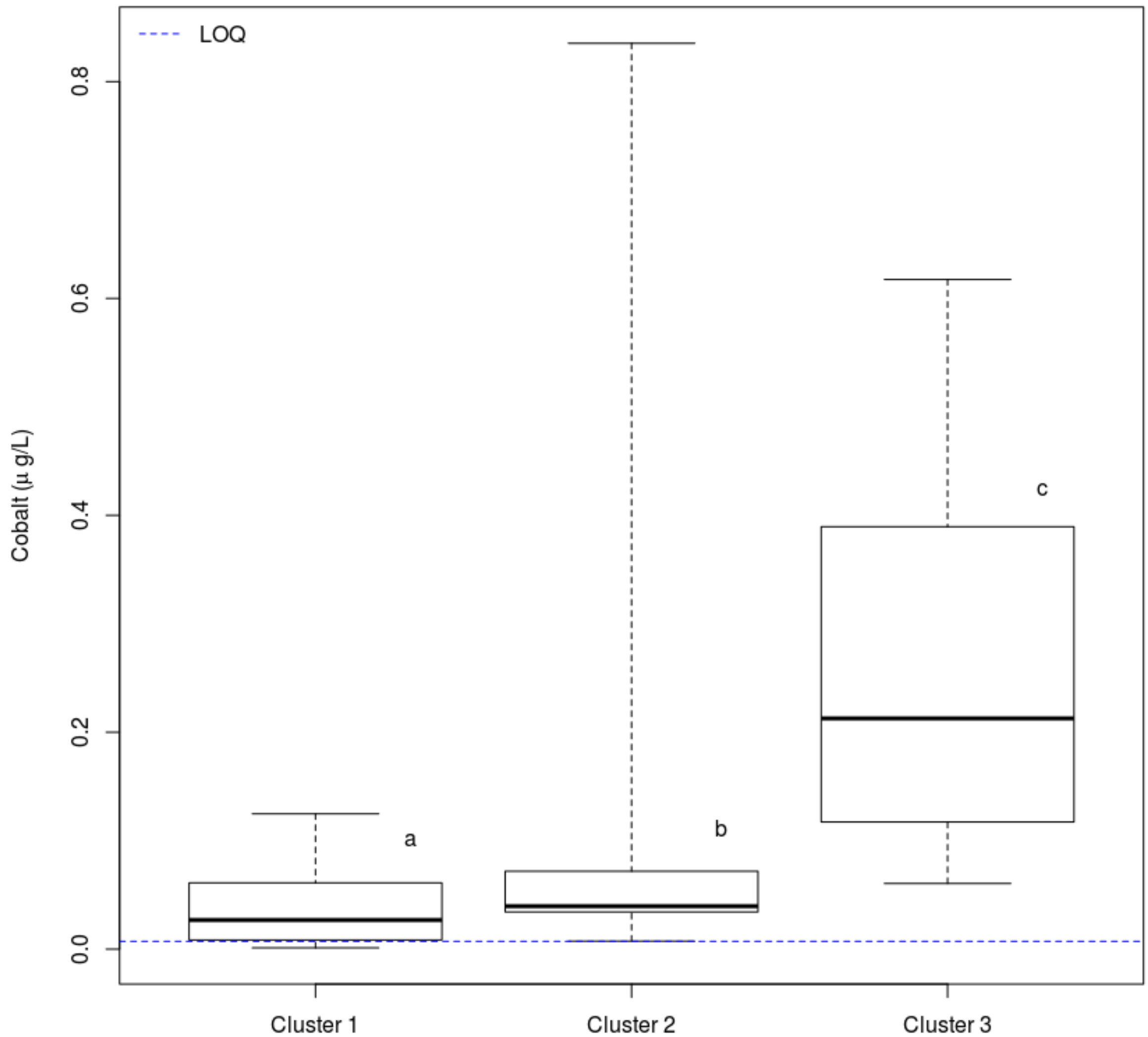


Figure 16. Boxplot of cobalt concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

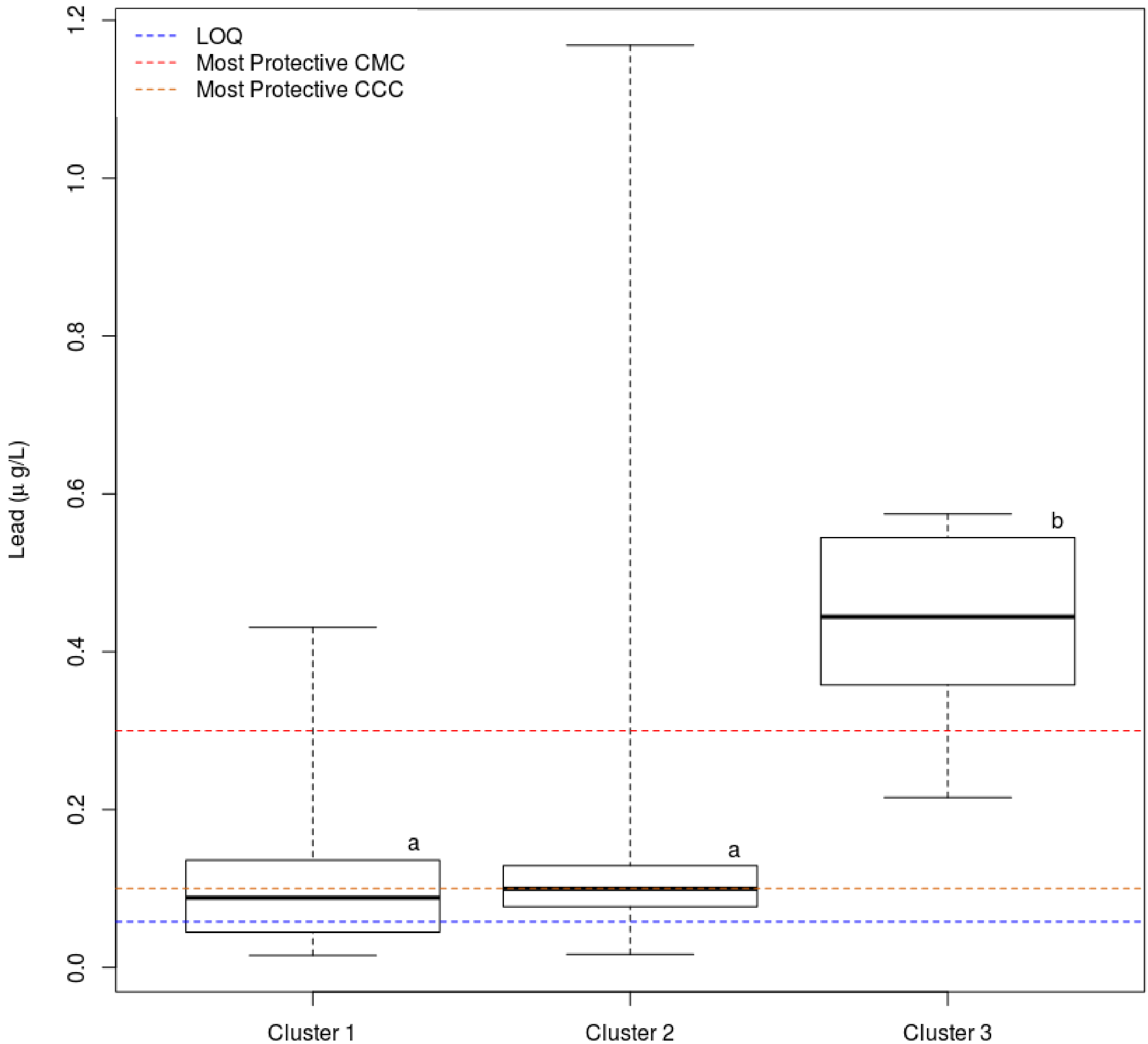


Figure 17. Boxplot of lead concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

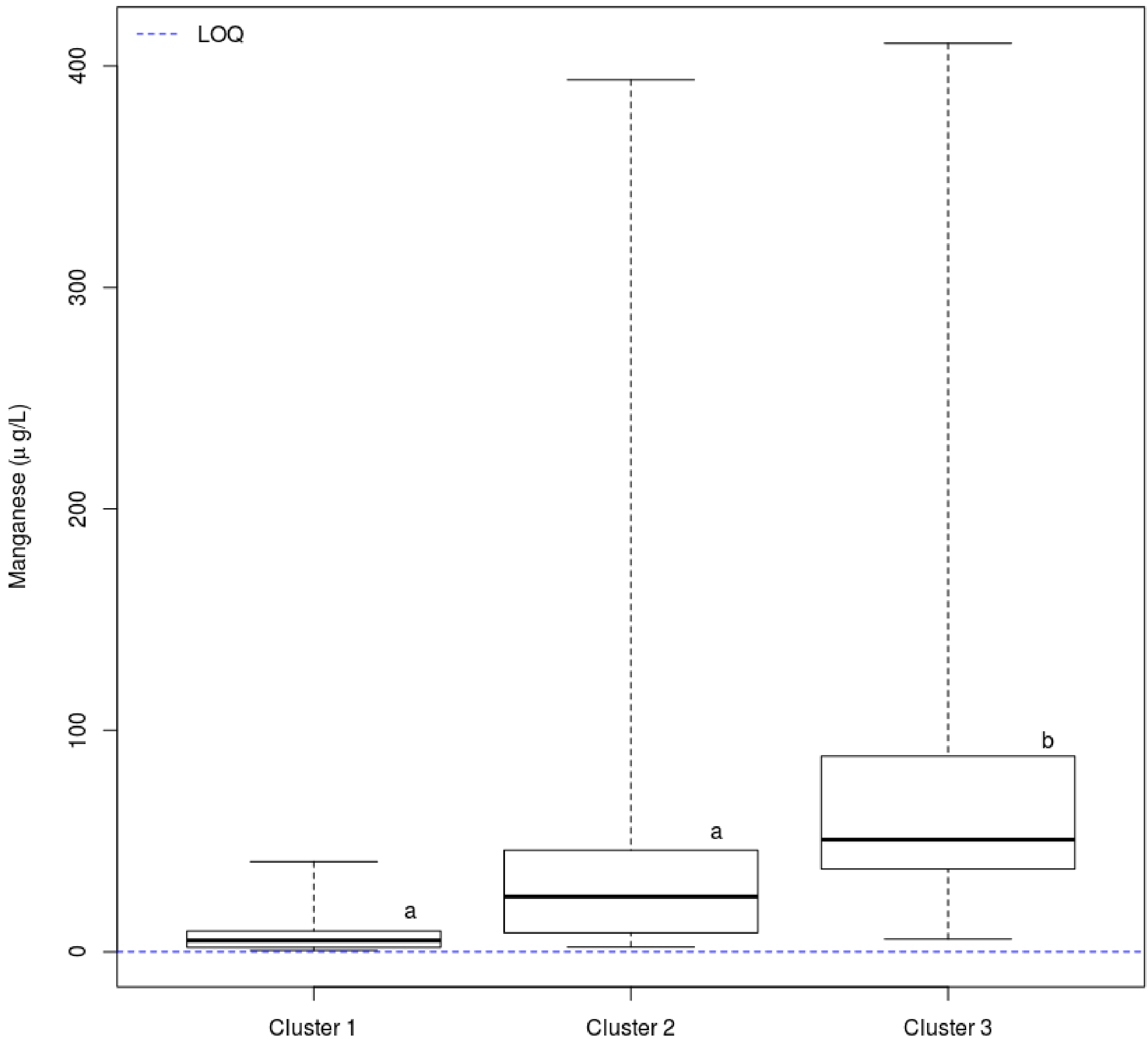


Figure 18. Boxplot of manganese concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

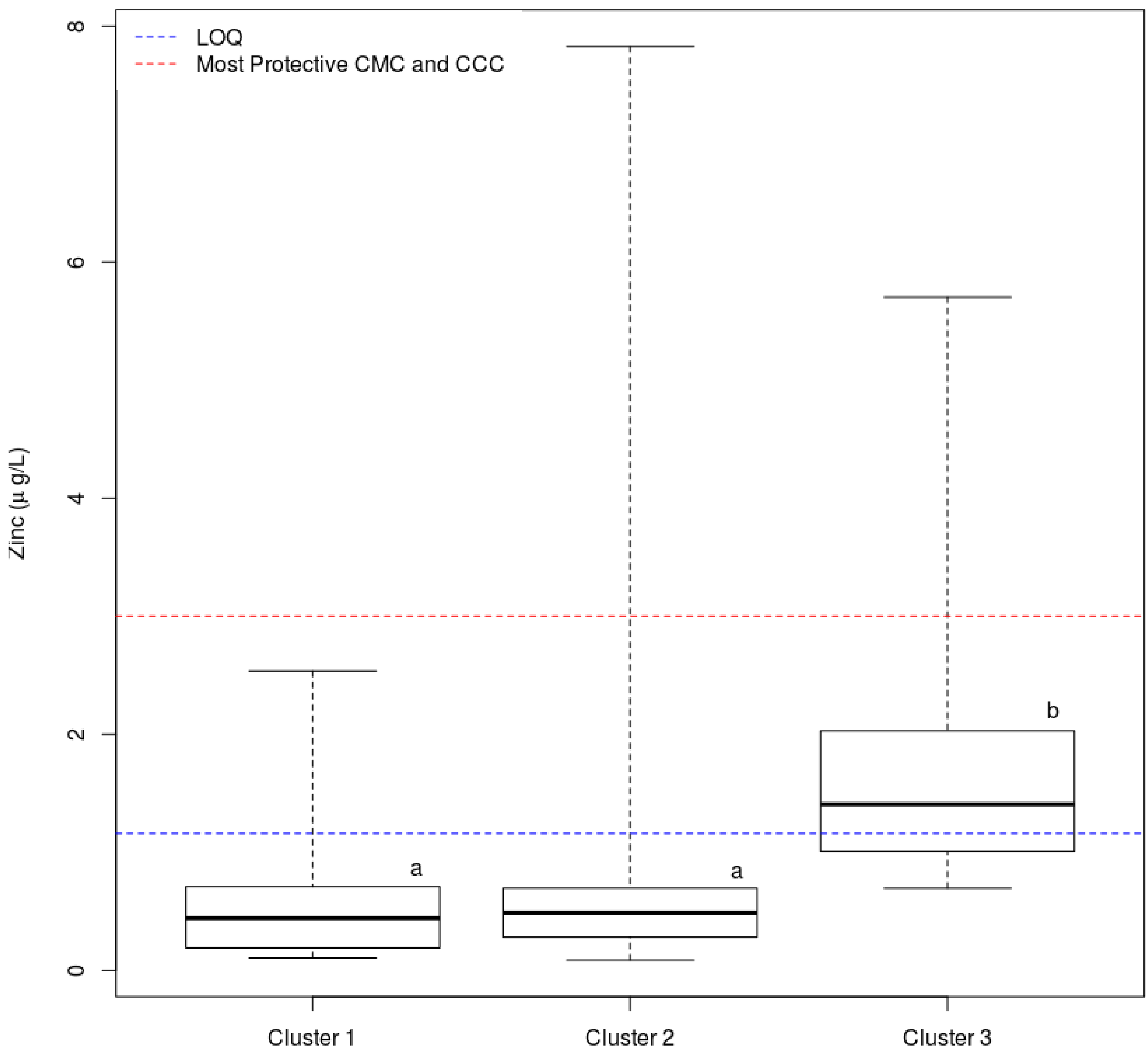


Figure 19. Boxplot of zinc concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

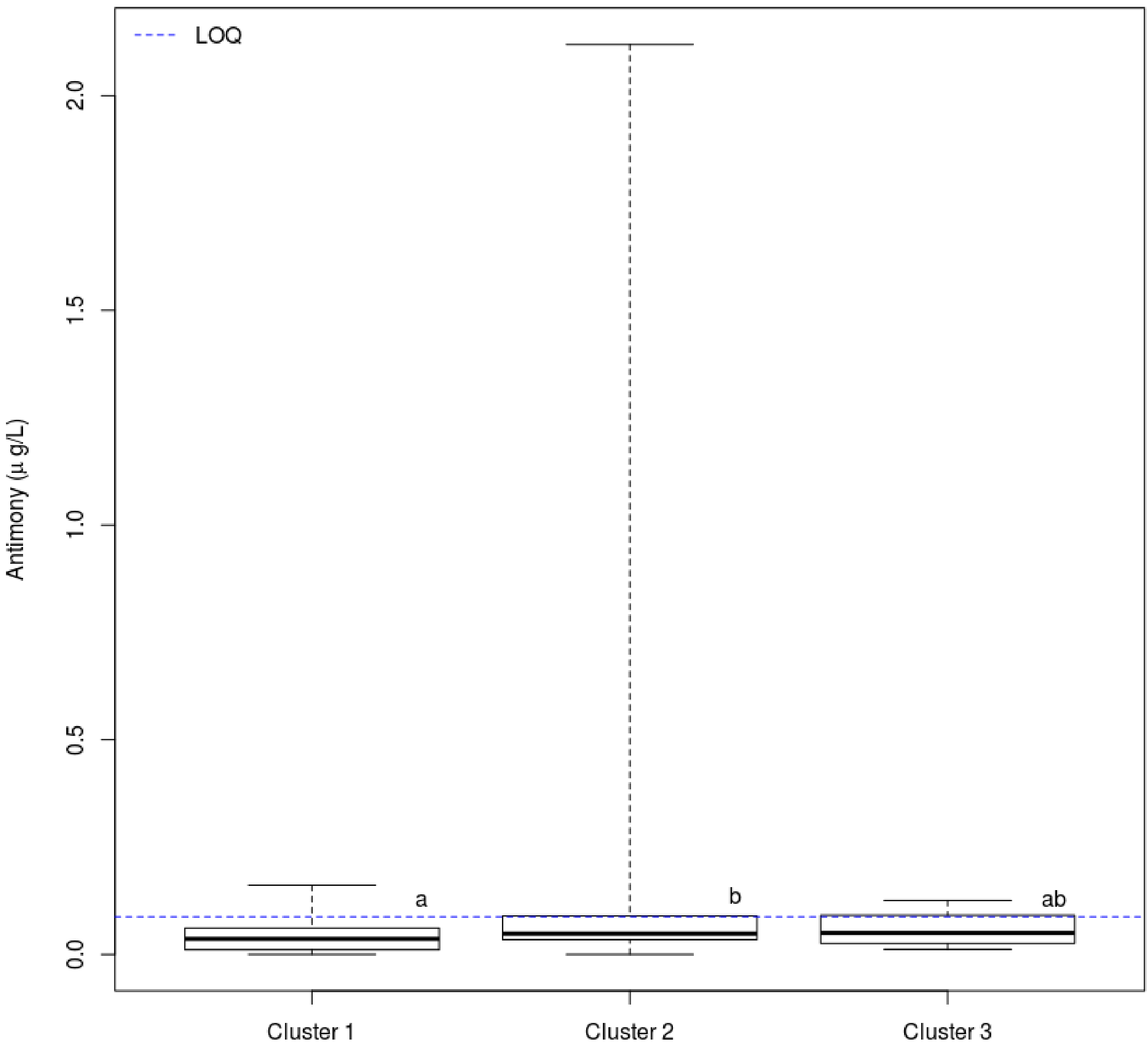


Figure 20. Boxplot of antimony concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

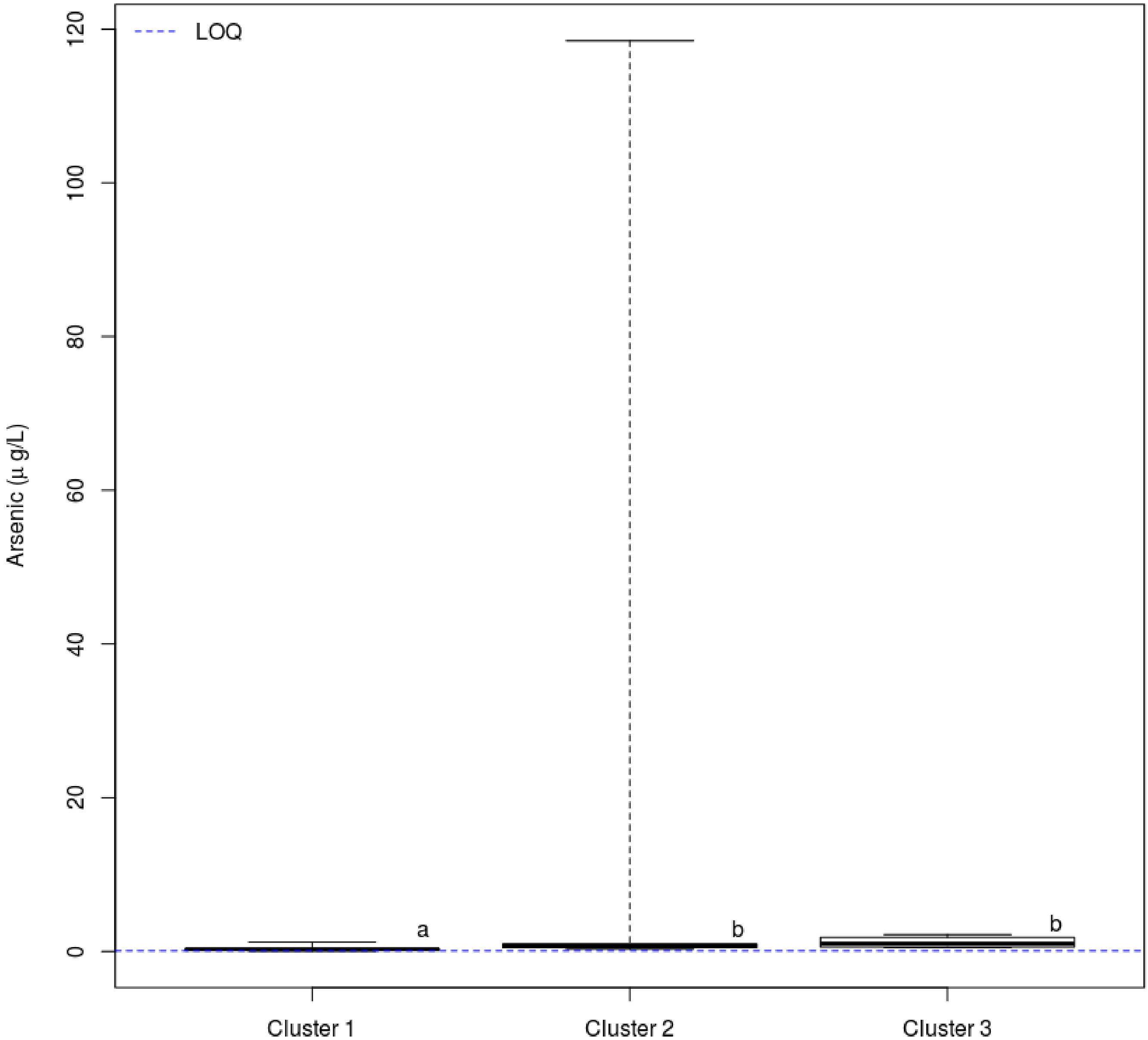


Figure 21. Boxplot of arsenic concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

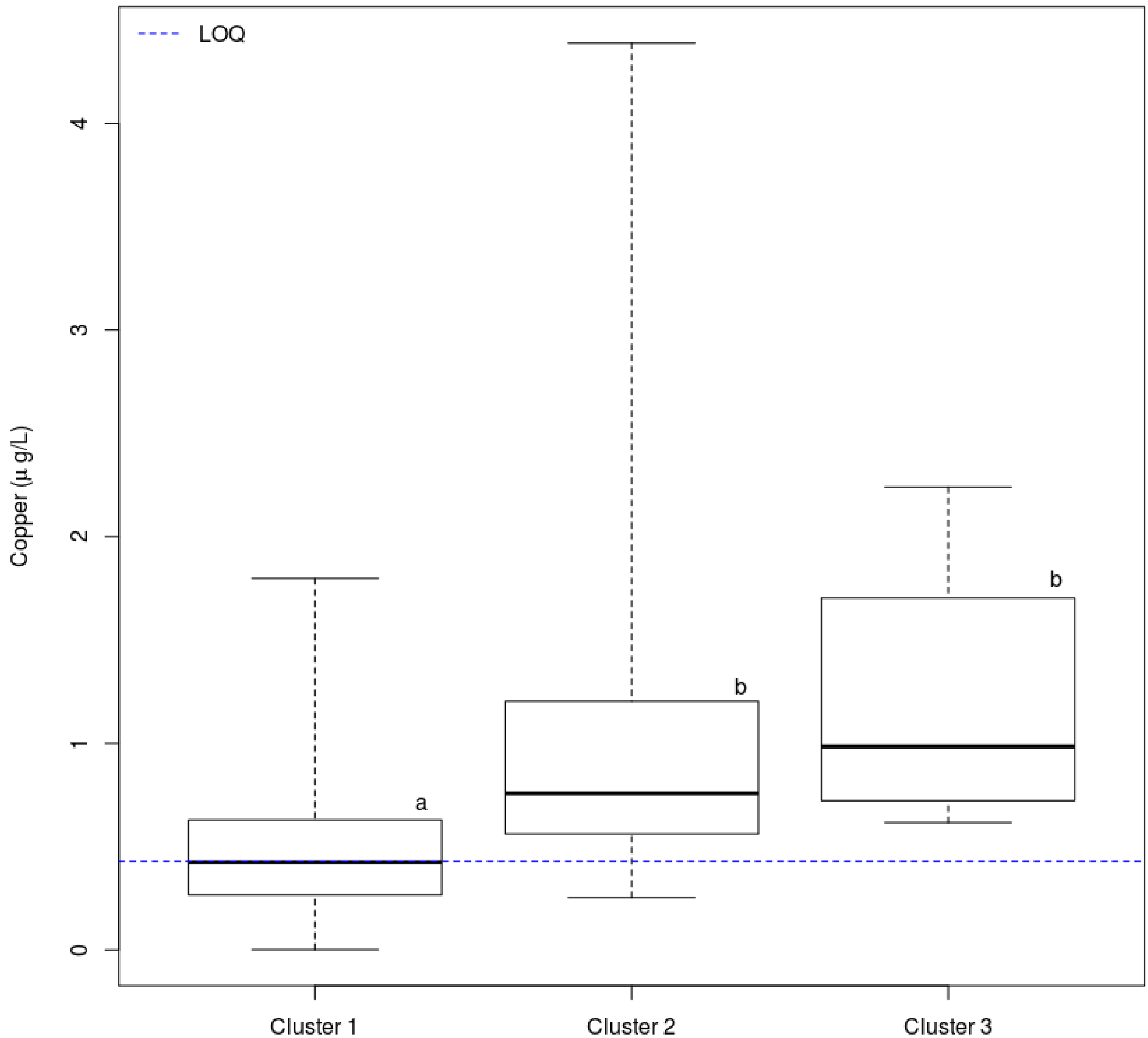


Figure 22. Boxplot of copper concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

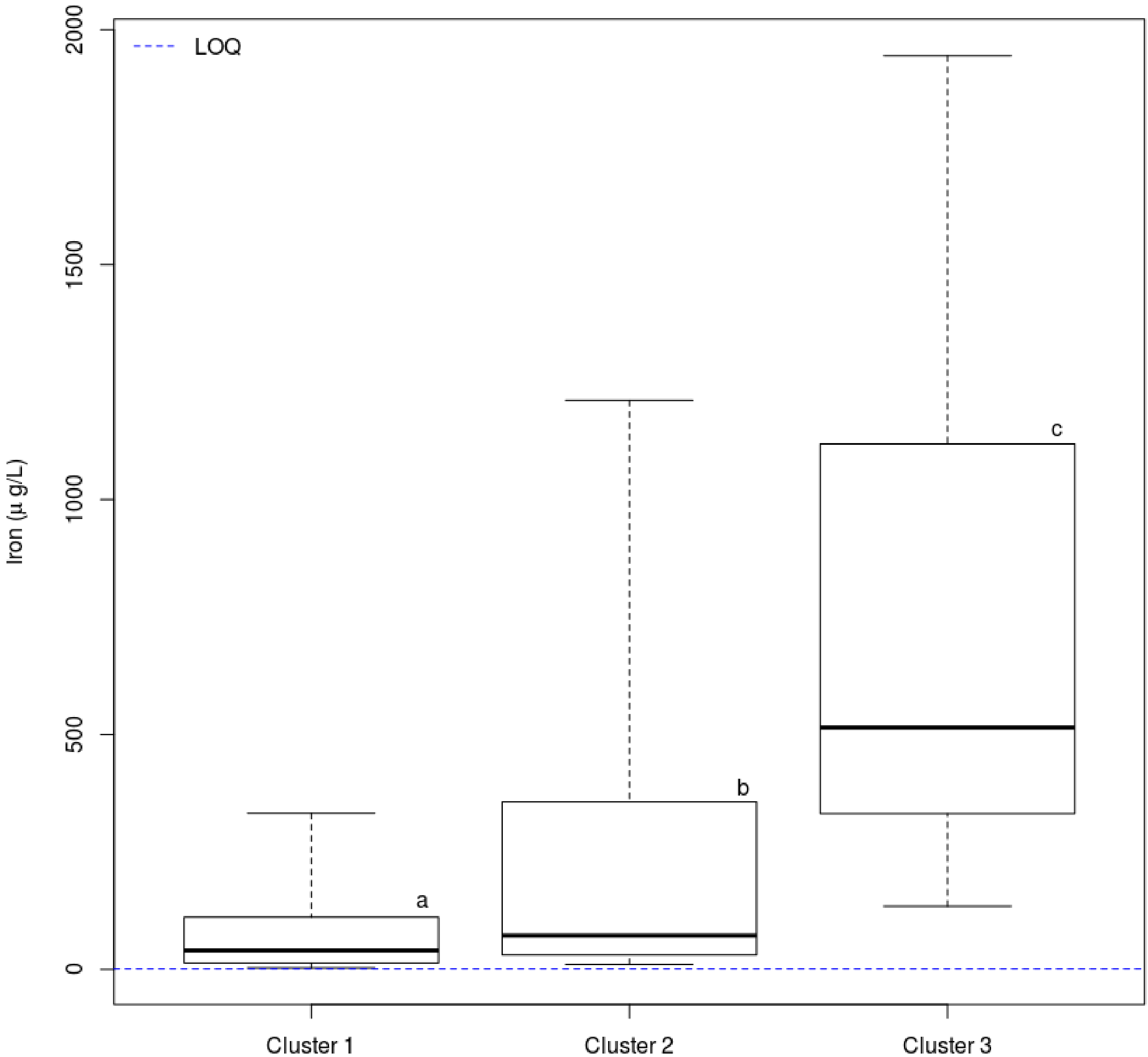


Figure 23. Boxplot of iron concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

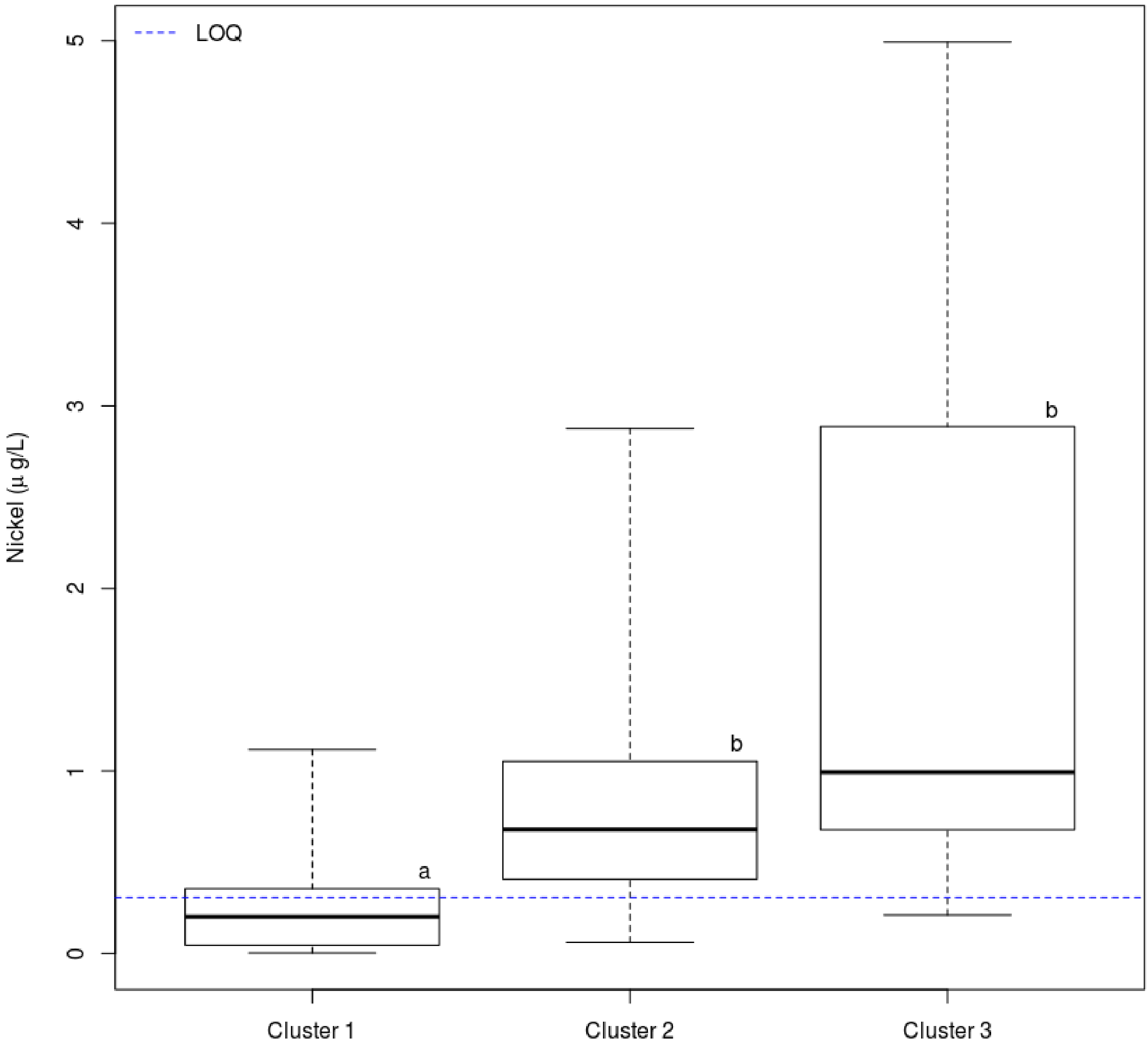


Figure 24. Boxplot of nickel concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

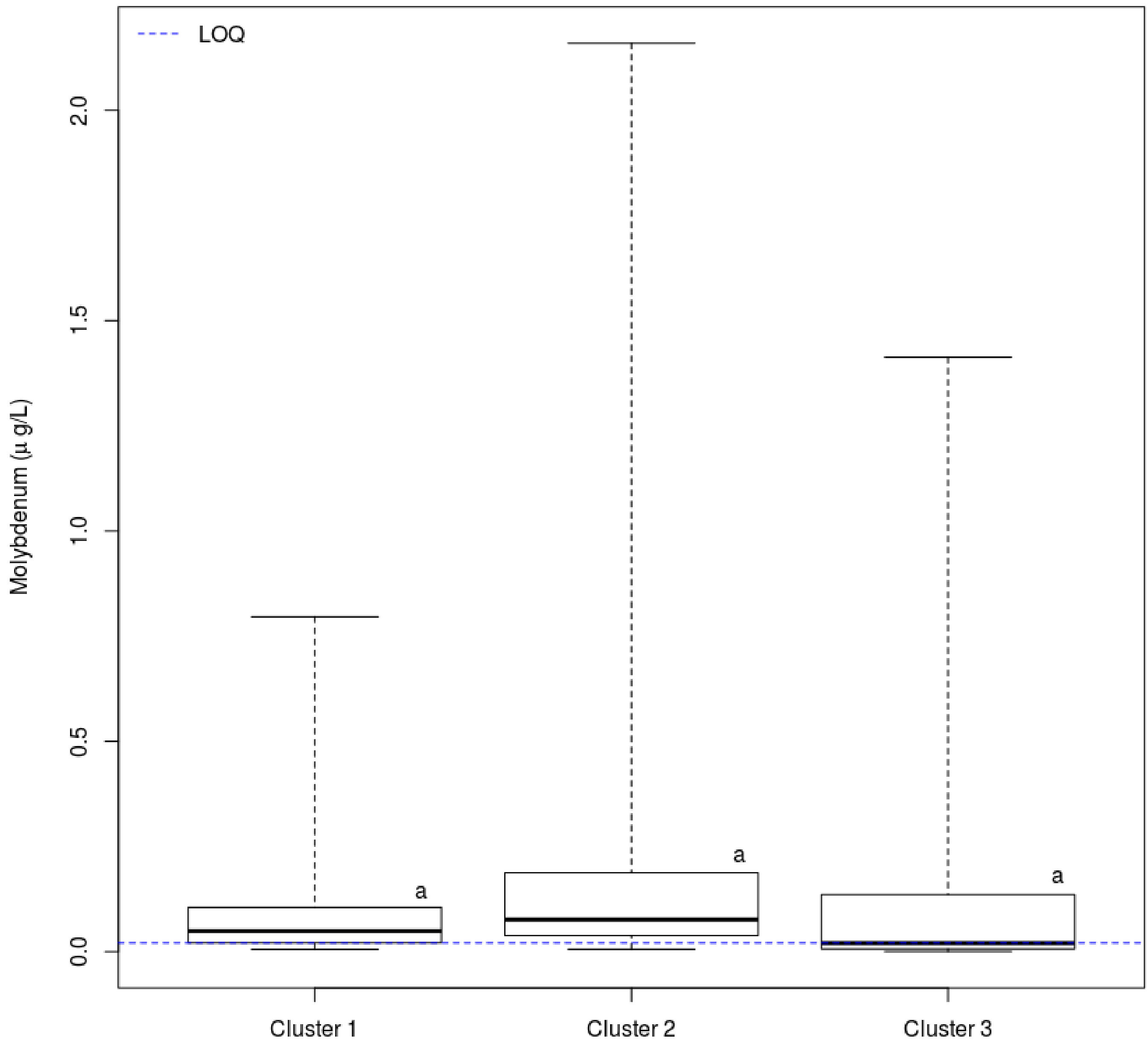


Figure 25. Boxplot of molybdenum concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

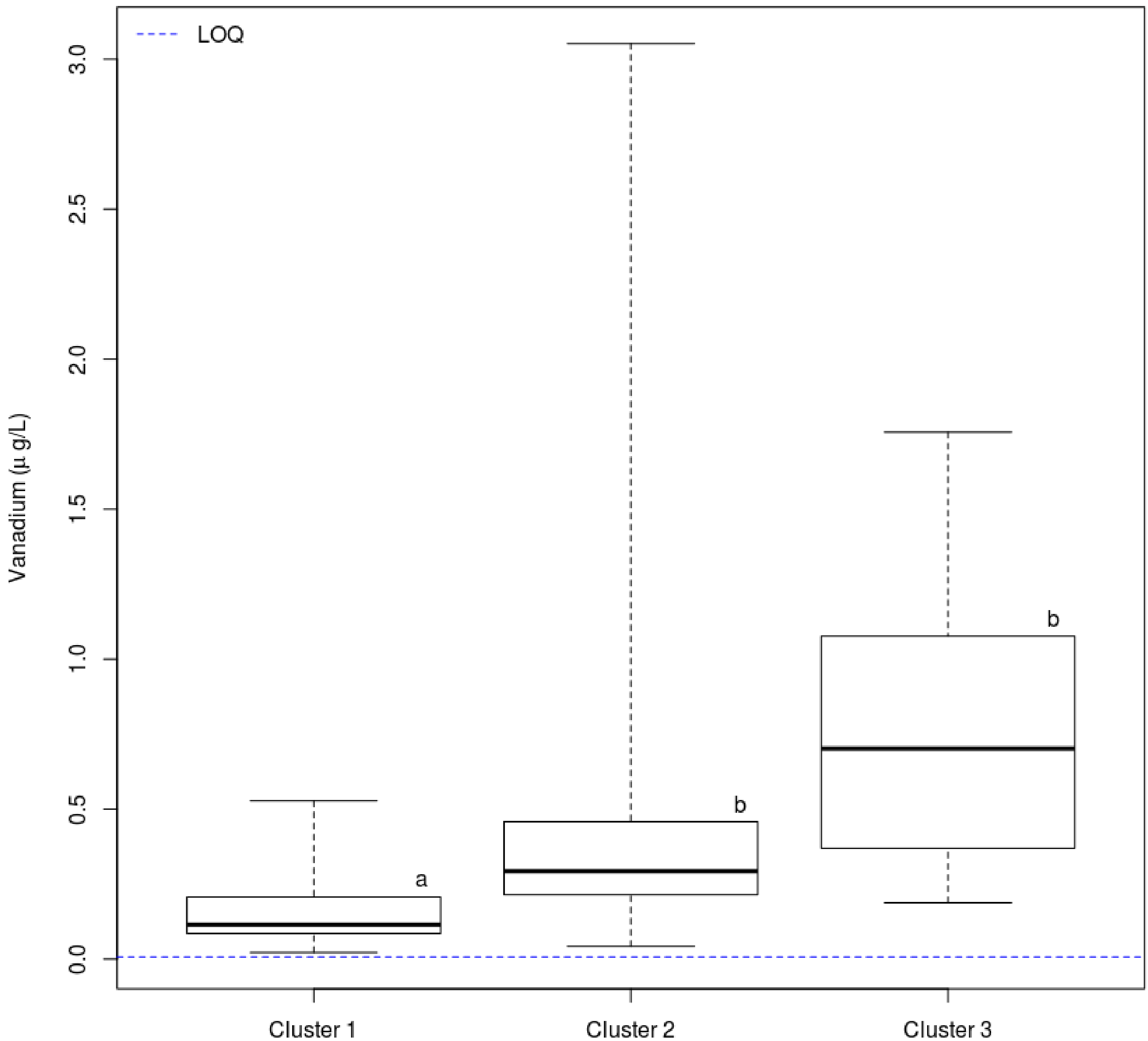


Figure 26. Boxplot of vanadium concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an α = 0.05.

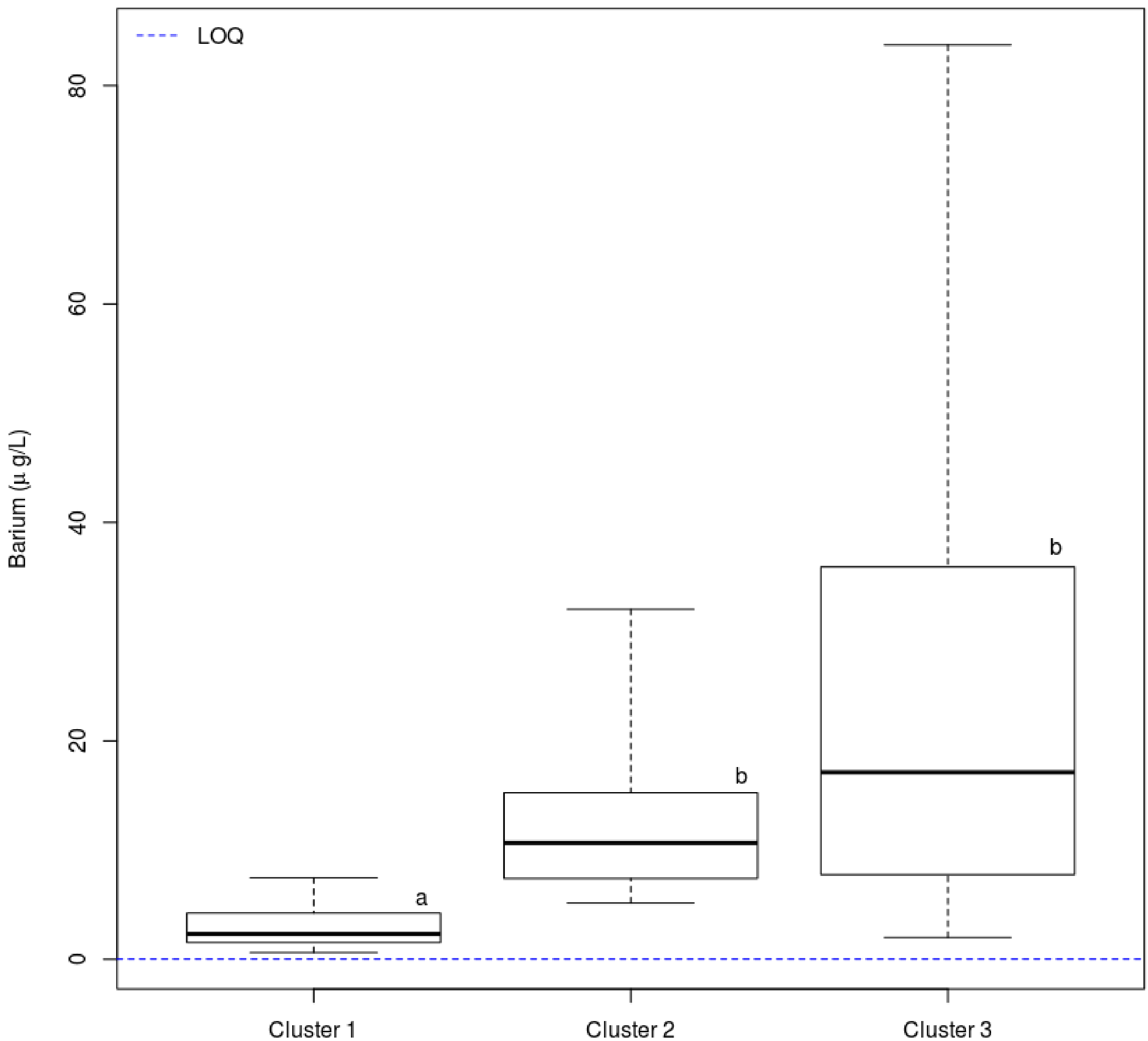


Figure 27. Boxplot of barium concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

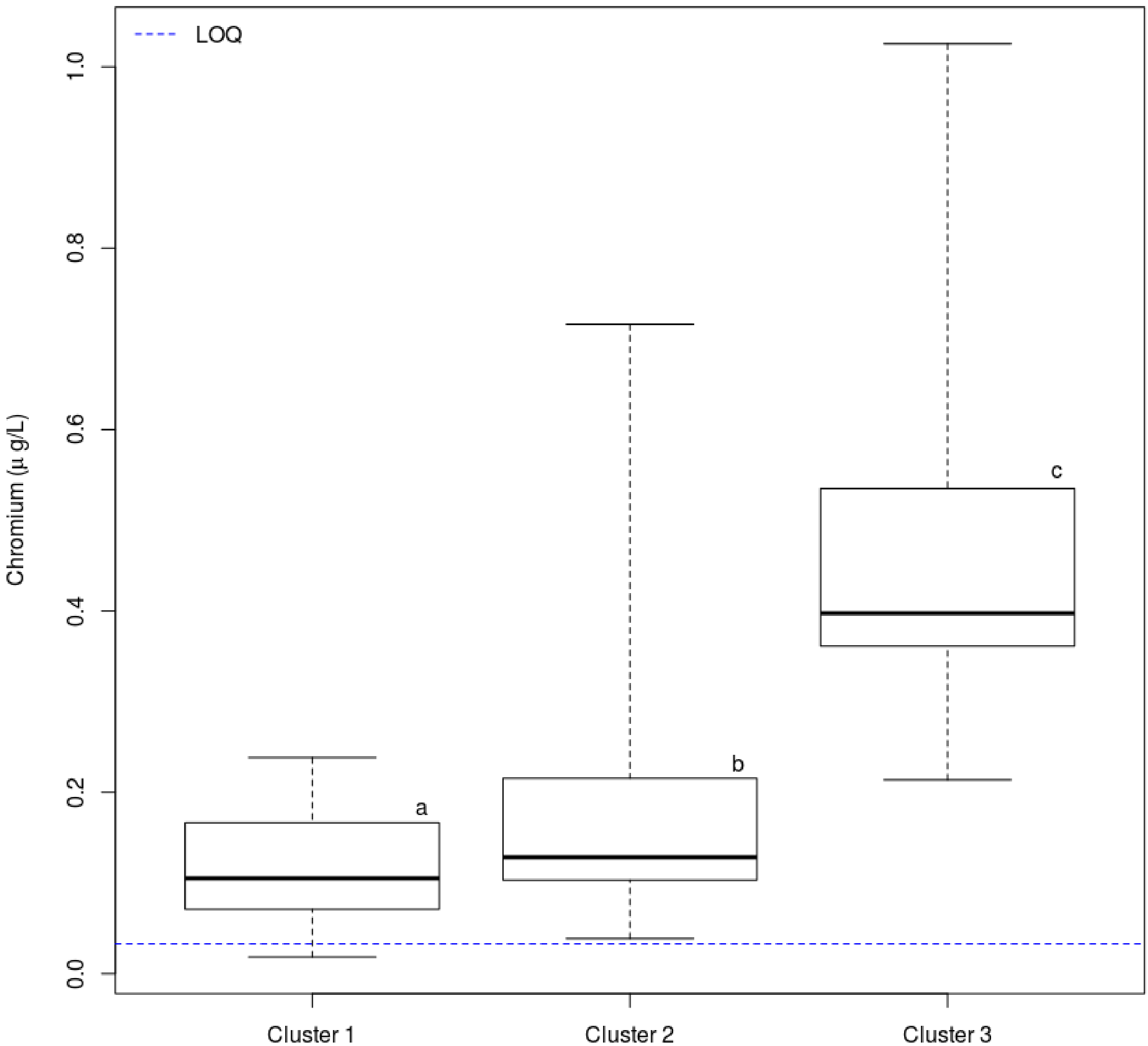


Figure 28. Boxplot of chromium concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

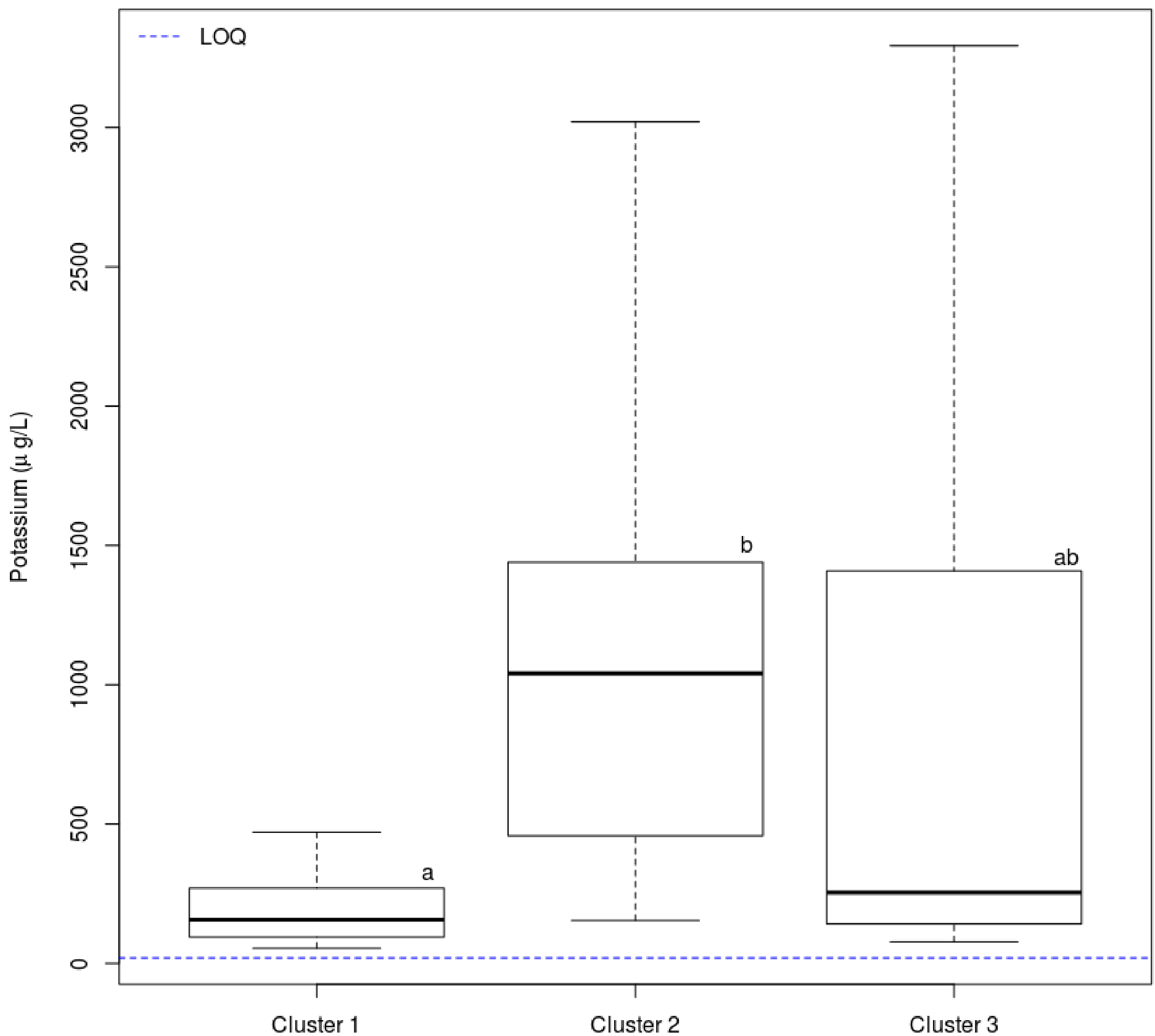


Figure 29. Boxplot of potassium concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

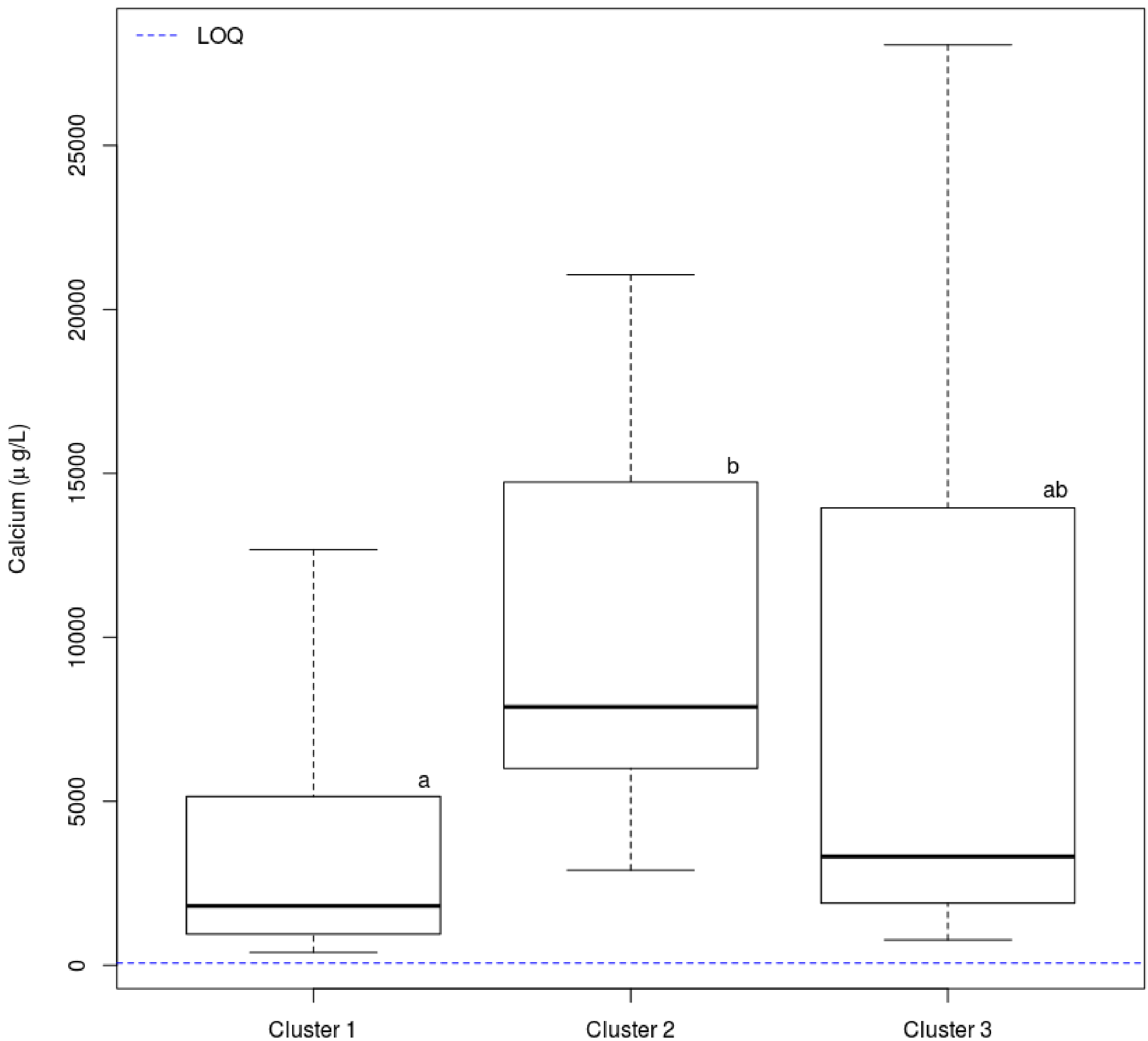


Figure 30. Boxplot of calcium concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

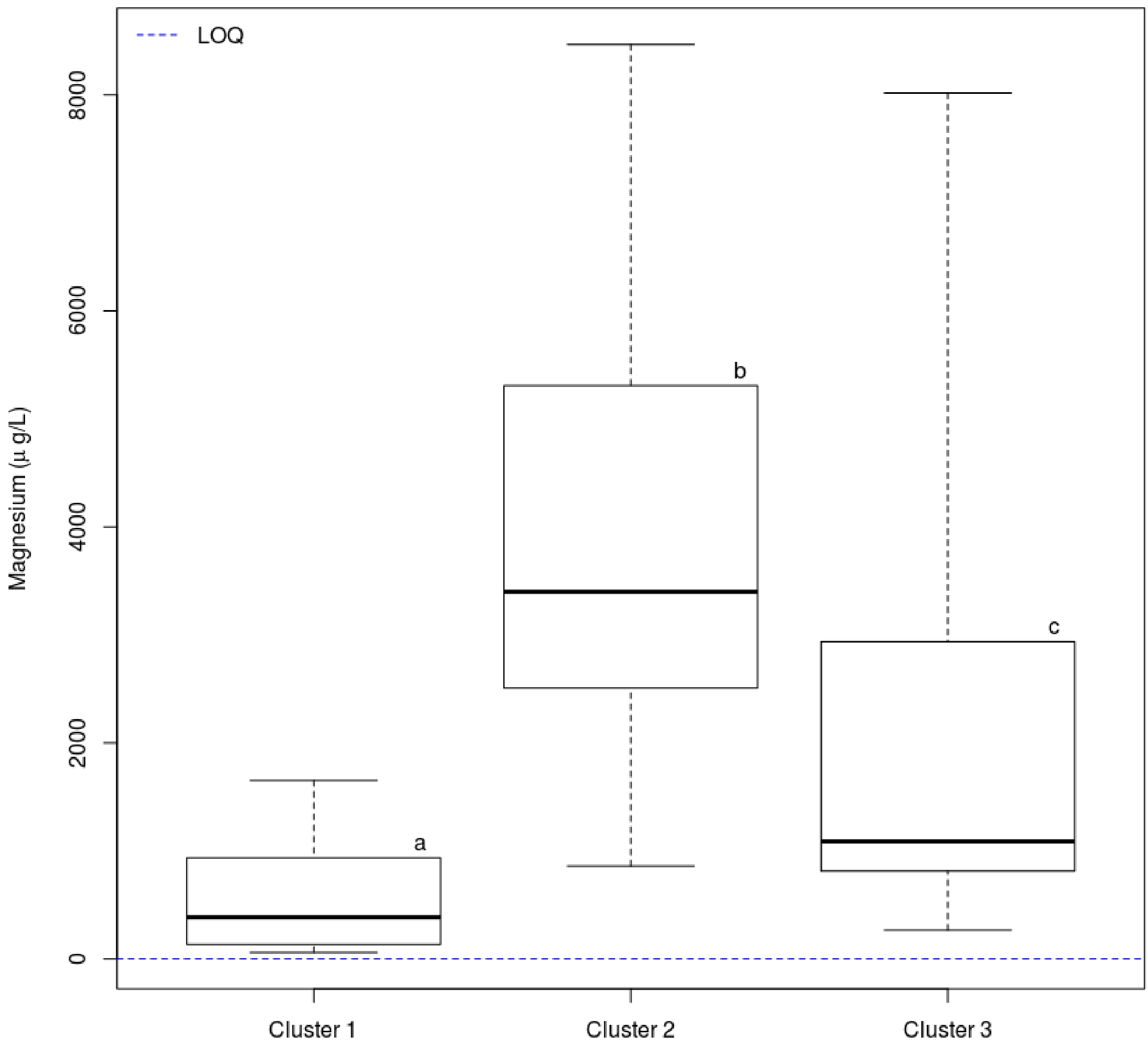


Figure 31. Boxplot of magnesium concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

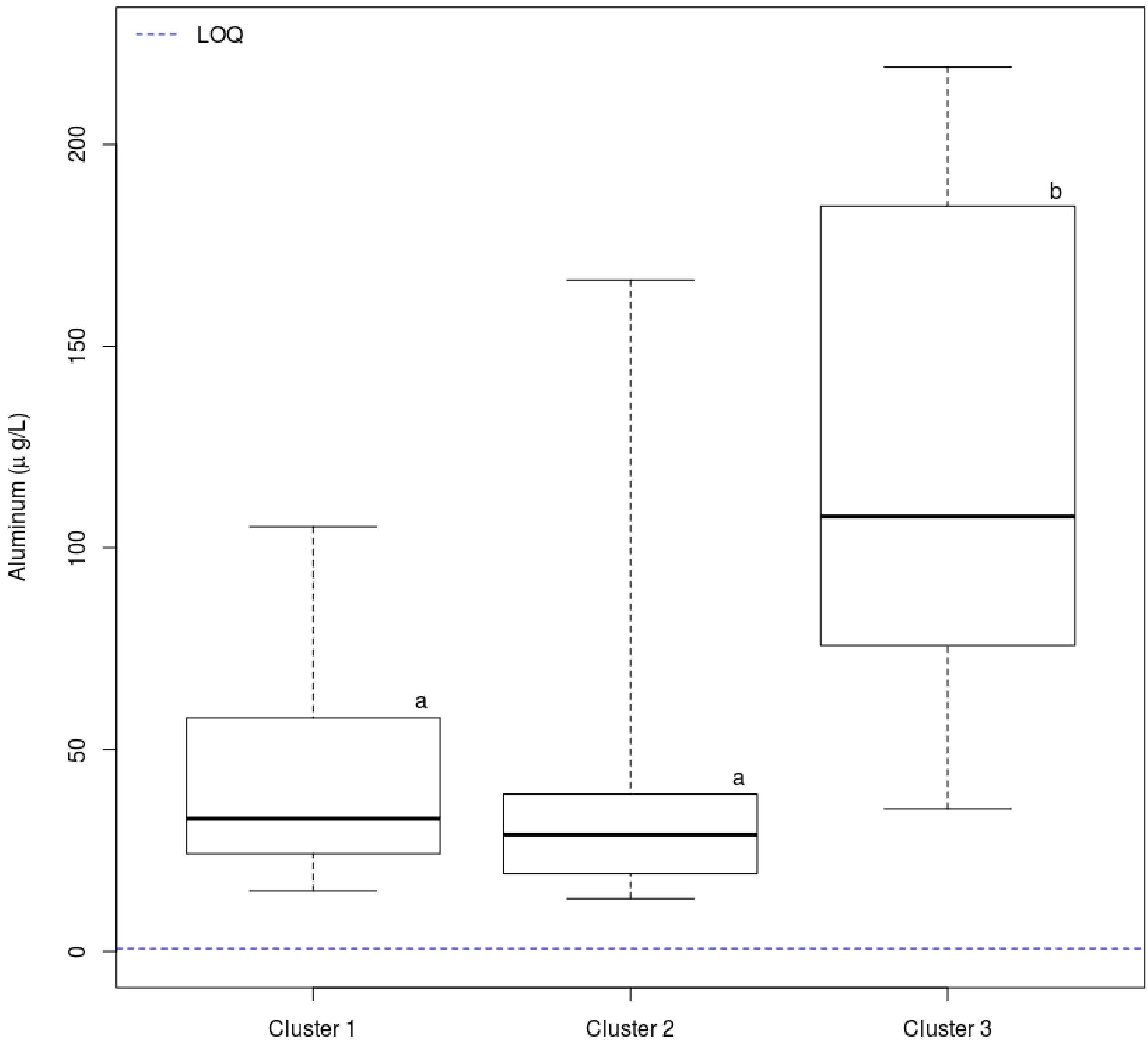


Figure 32. Boxplot of aluminum concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown as a dashed horizontal line. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

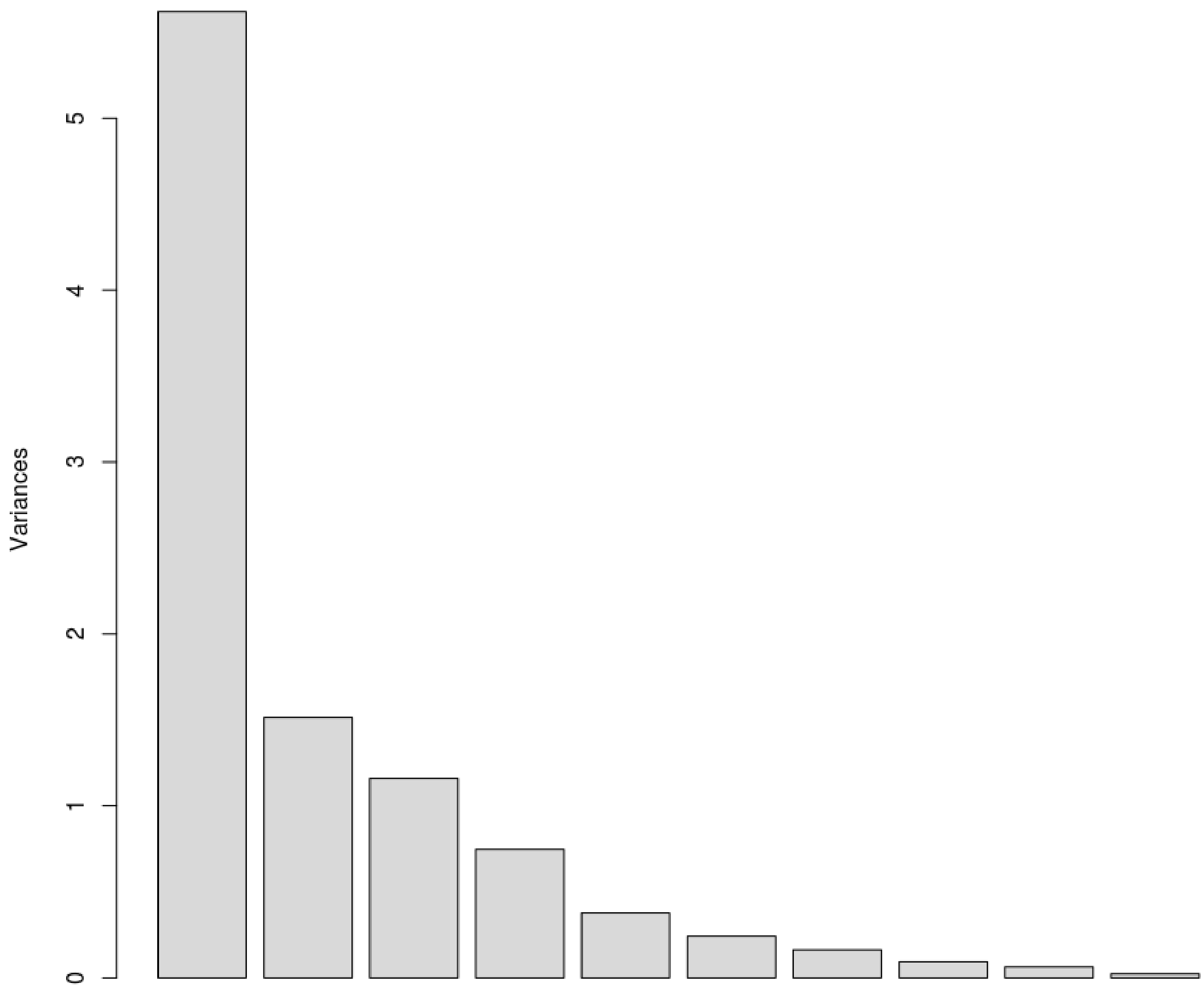


Figure 33. Variance plot of all ten PCs for the PCA based on the water quality data. These ten PCs account for 100 % of the total variance.

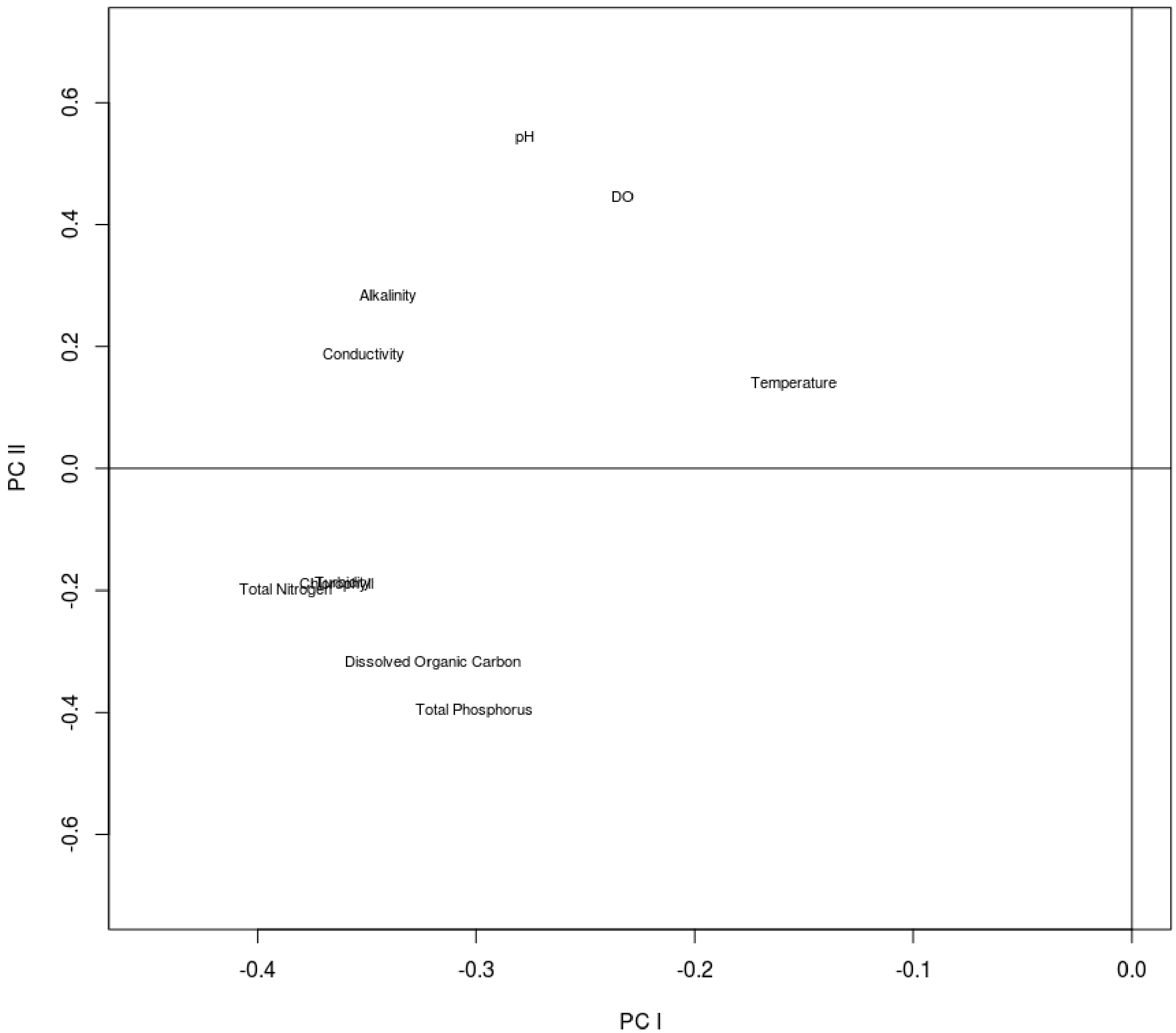


Figure 34. Ordination along PC I and PC II from the PCA of the water quality data. The grouping at approximately (-.4, -.2) contains the variables total nitrogen, turbidity, and chlorophyll.

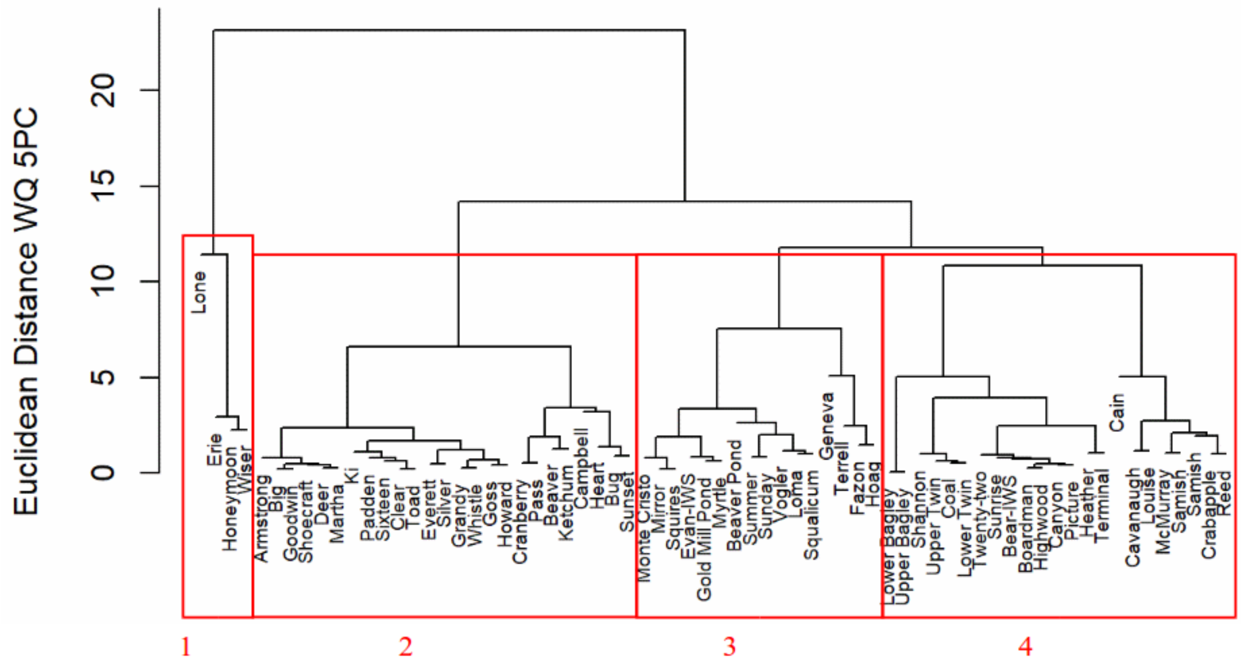


Figure 35. Hierarchical clustering of sampled lakes using water quality parameters. Using the first five principal components produced the most parsimonious clustering. These five variables account for 94.2 % of the total variance. Clusters are non-random. n = 68; χ-squared =10.5; p-value <0.05.

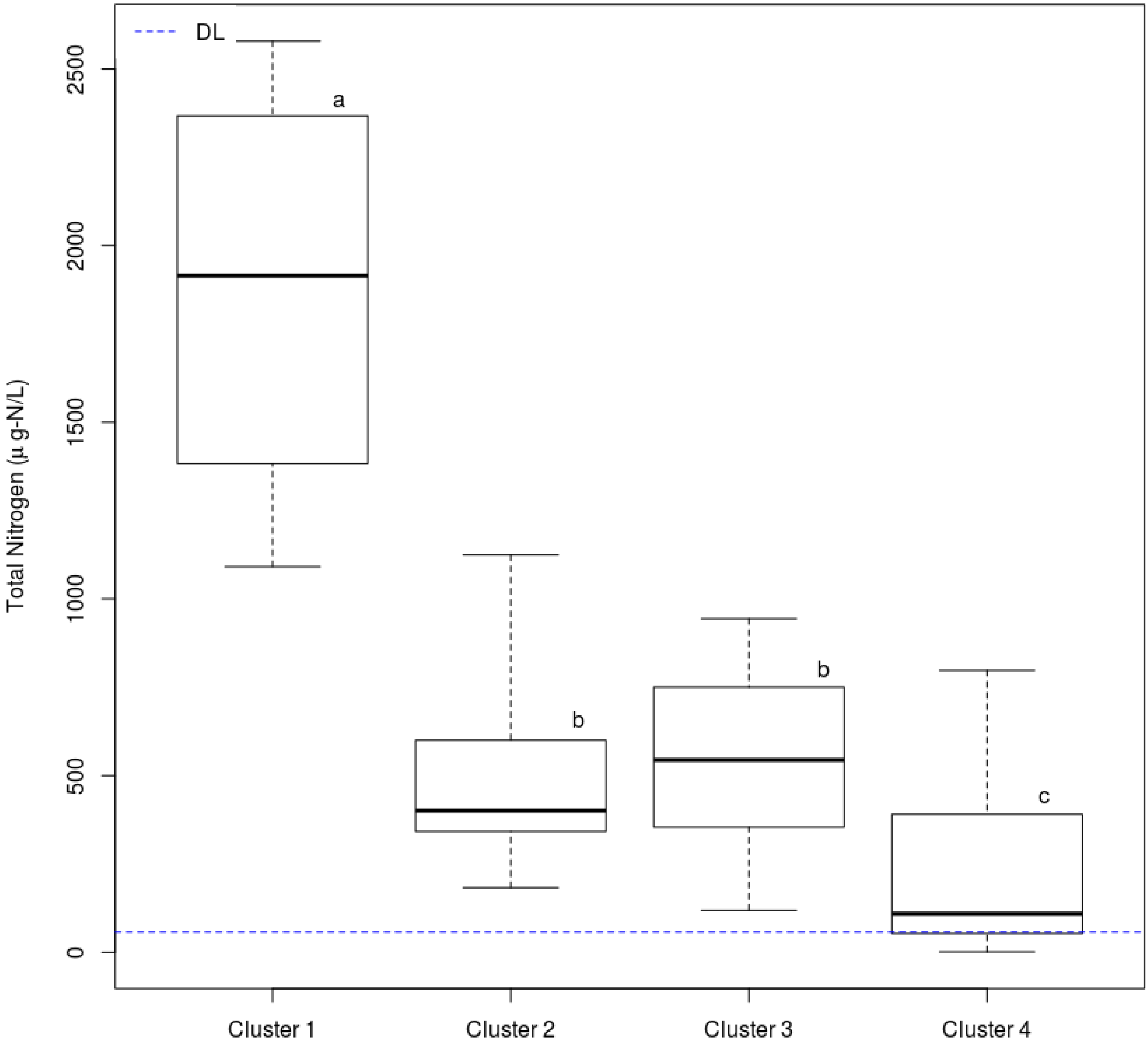


Figure 36. Boxplot of total nitrogen concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. The detection limit is shown at 57.8 µg-N/L. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

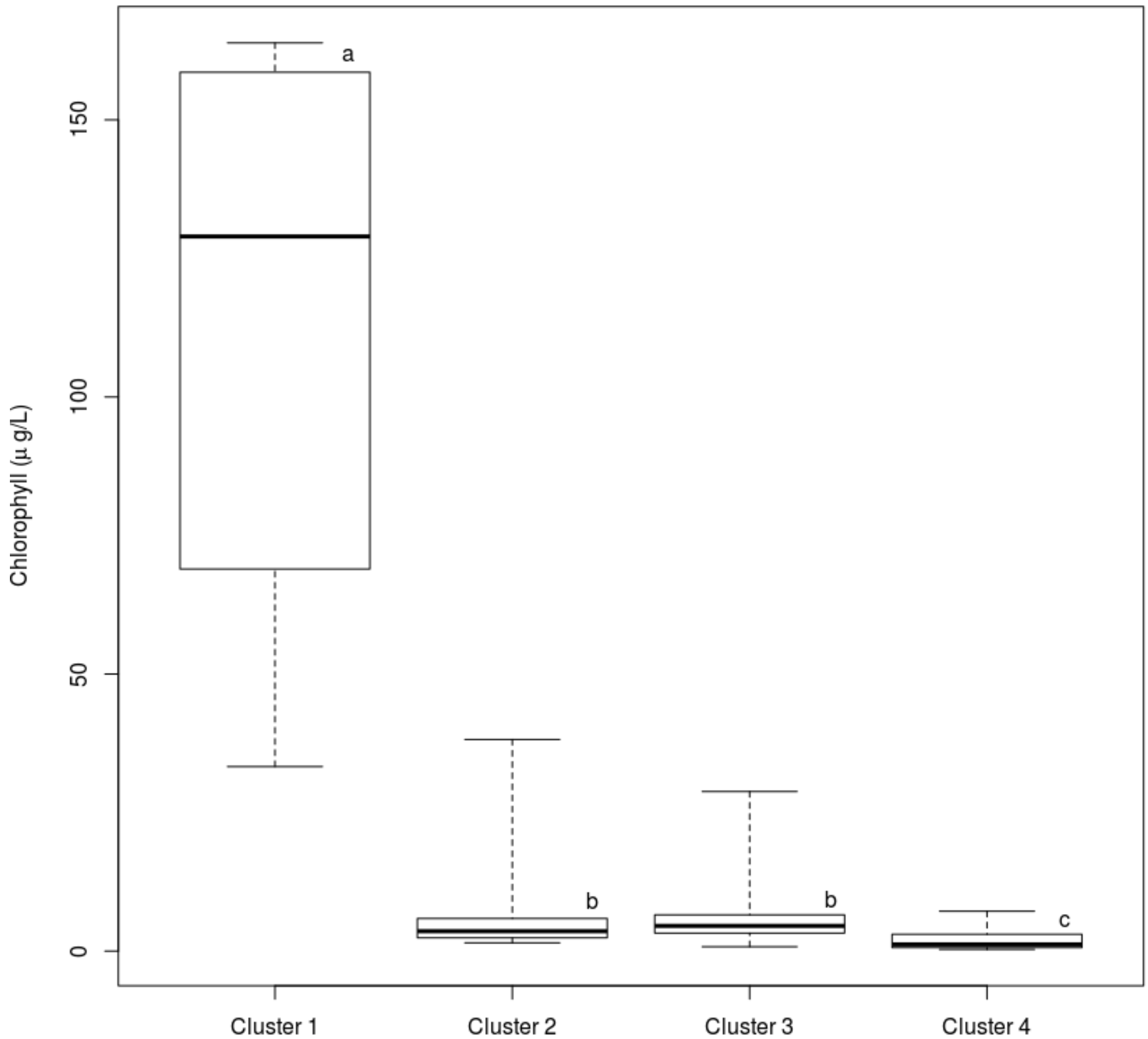


Figure 37. Boxplot of chlorophyll concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

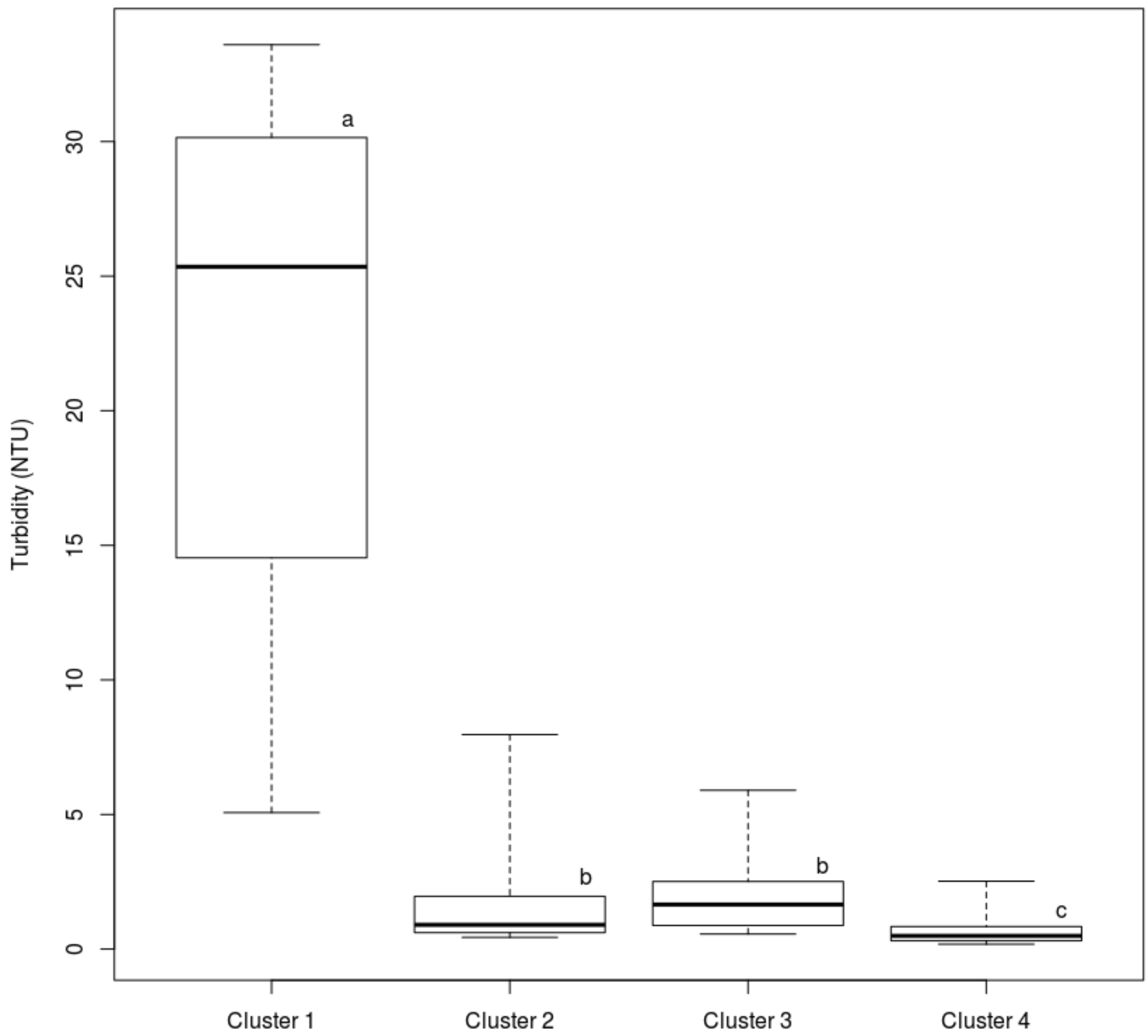


Figure 38. Boxplot of turbidity measurements by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

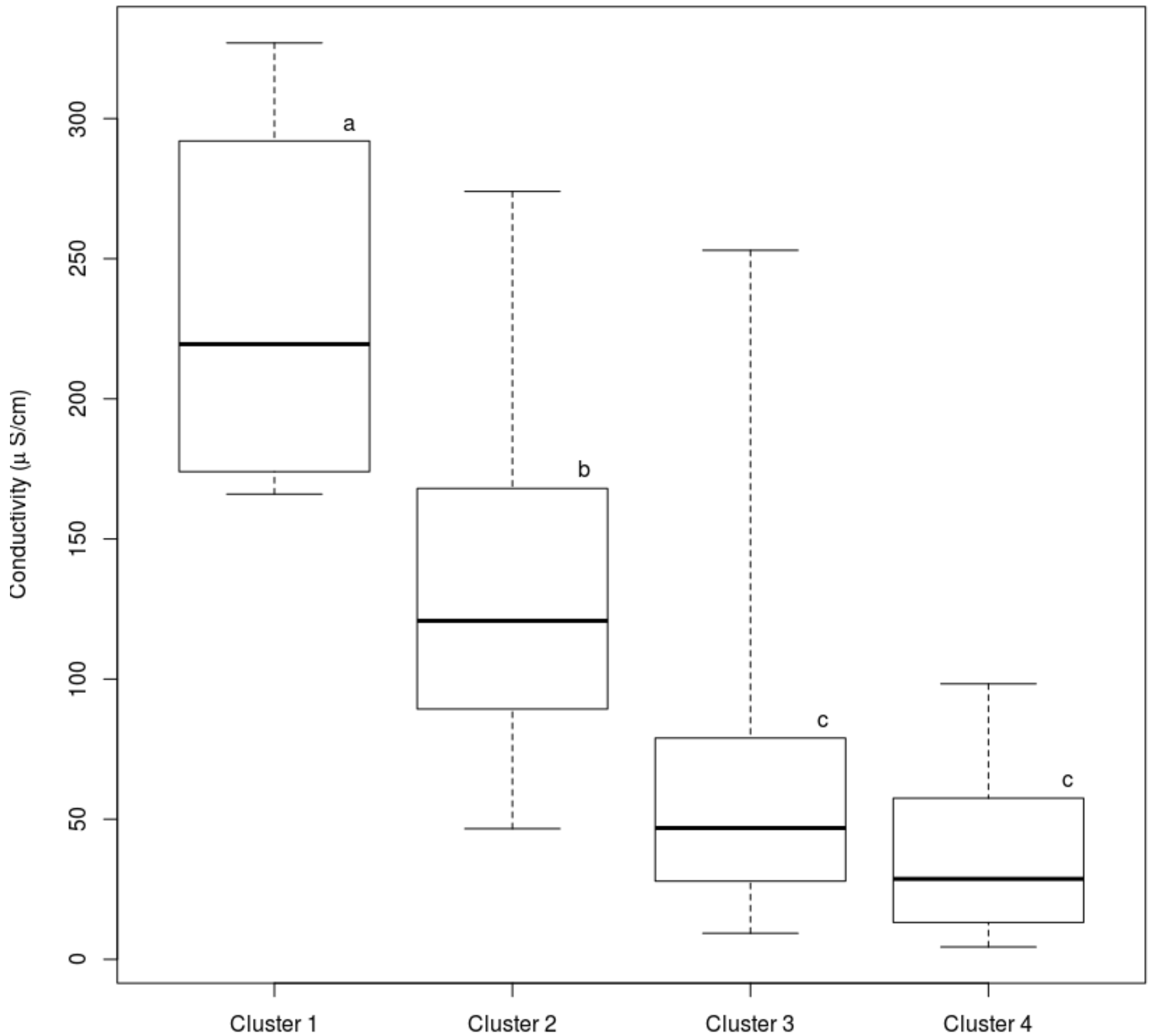


Figure 39. Boxplot of conductivity measurements by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

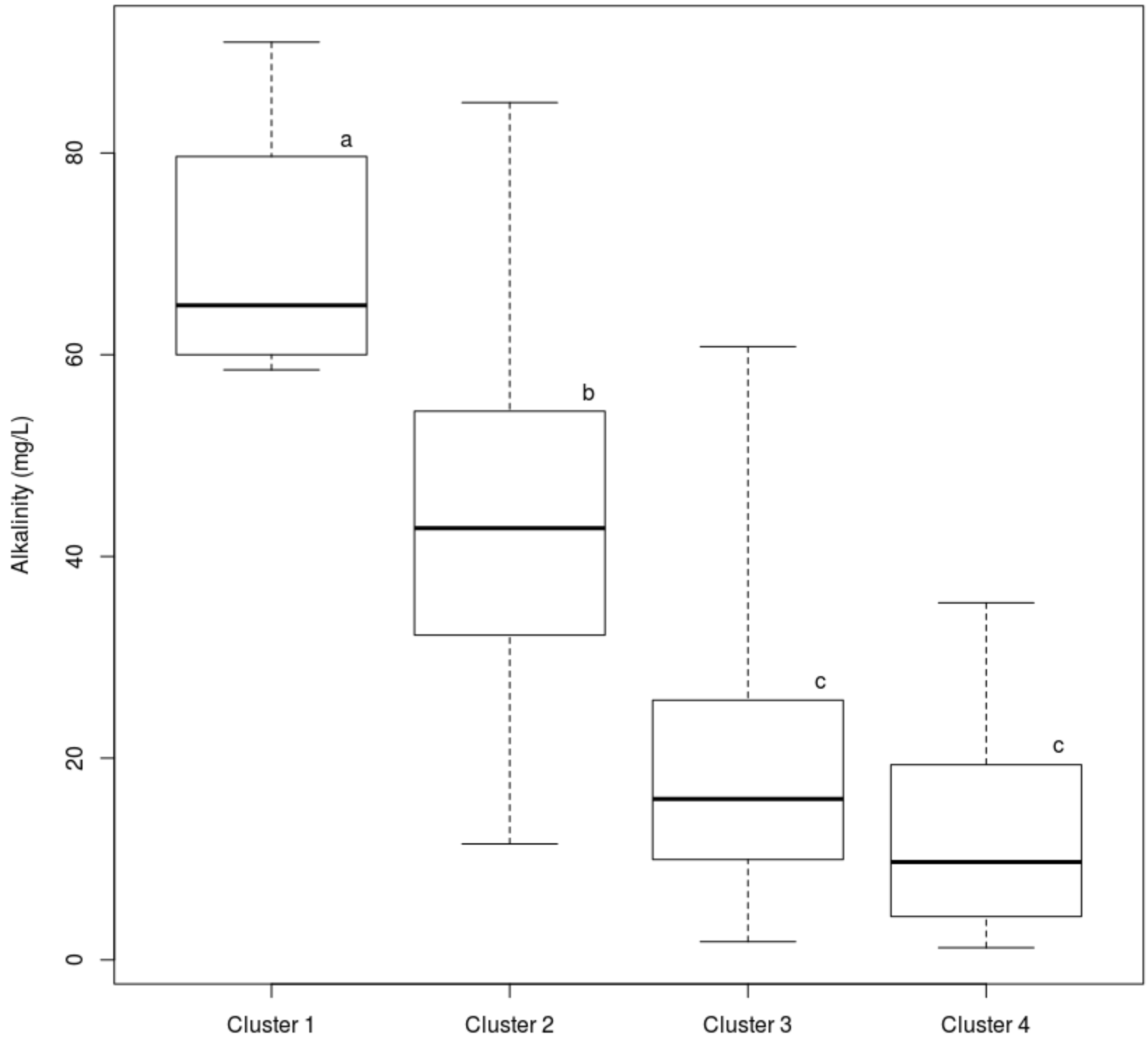


Figure 40. Boxplot of alkalinity concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

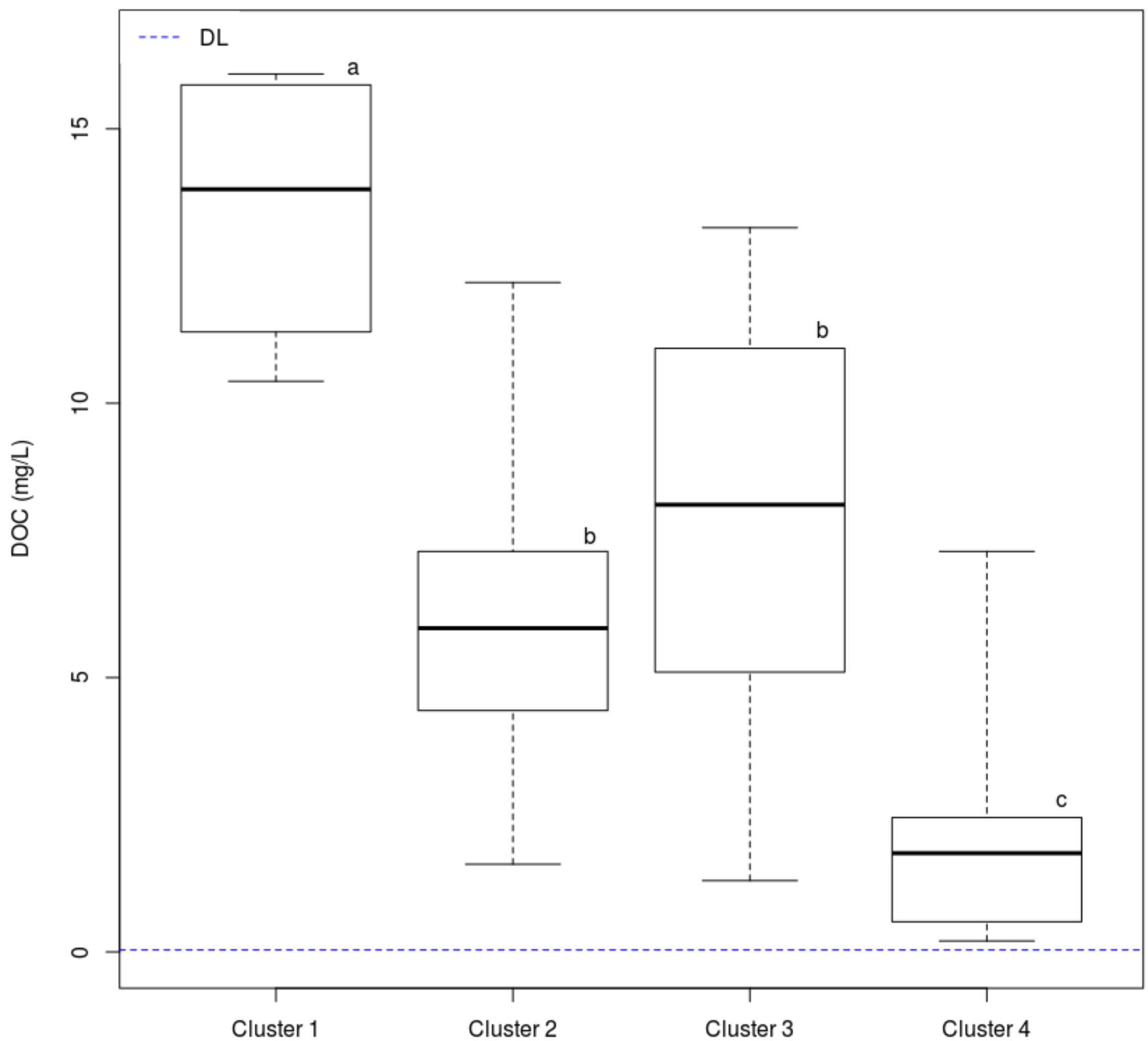


Figure 41. Boxplot of DOC concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

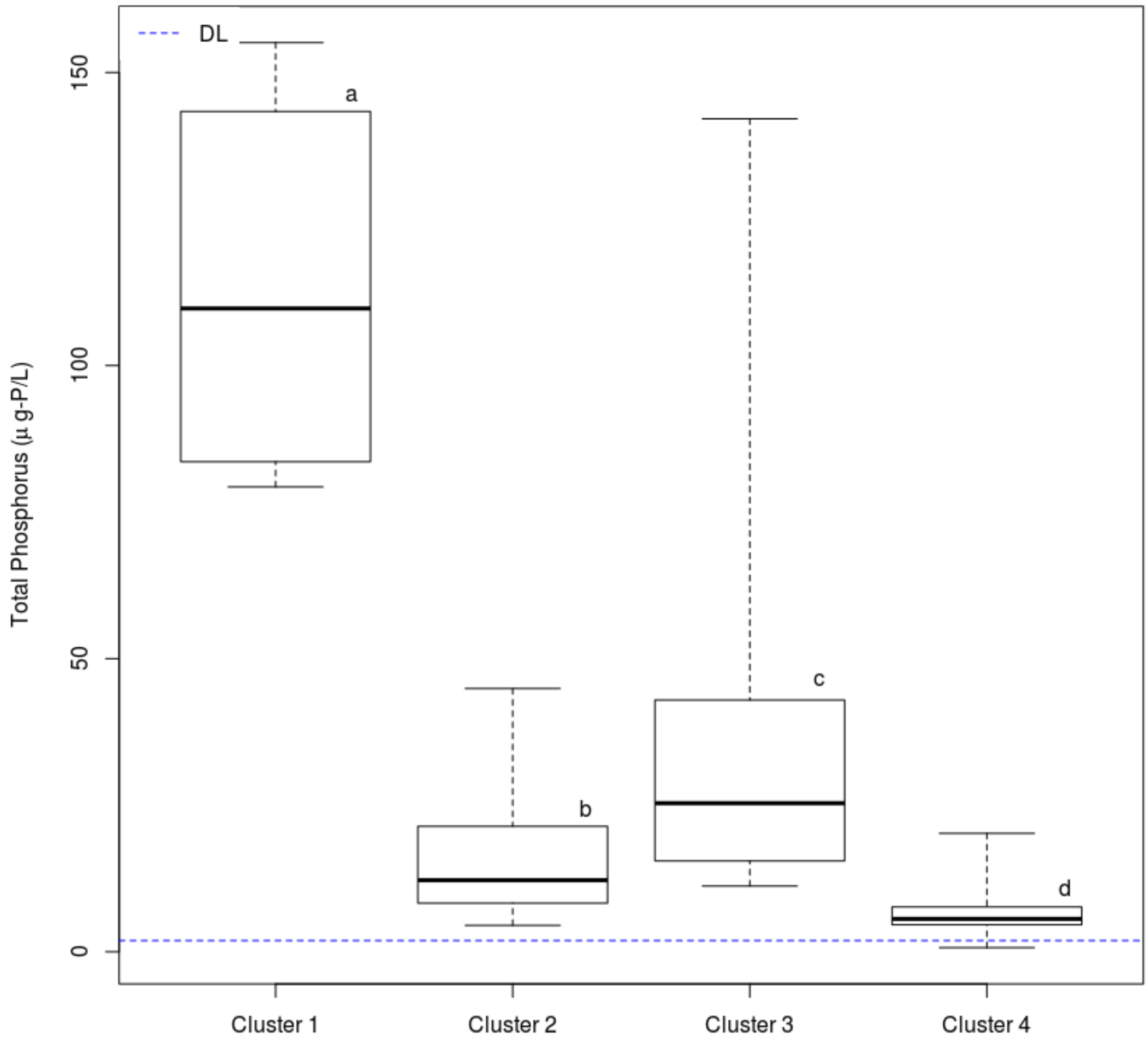


Figure 42. Boxplot of total phosphorus concentrations by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

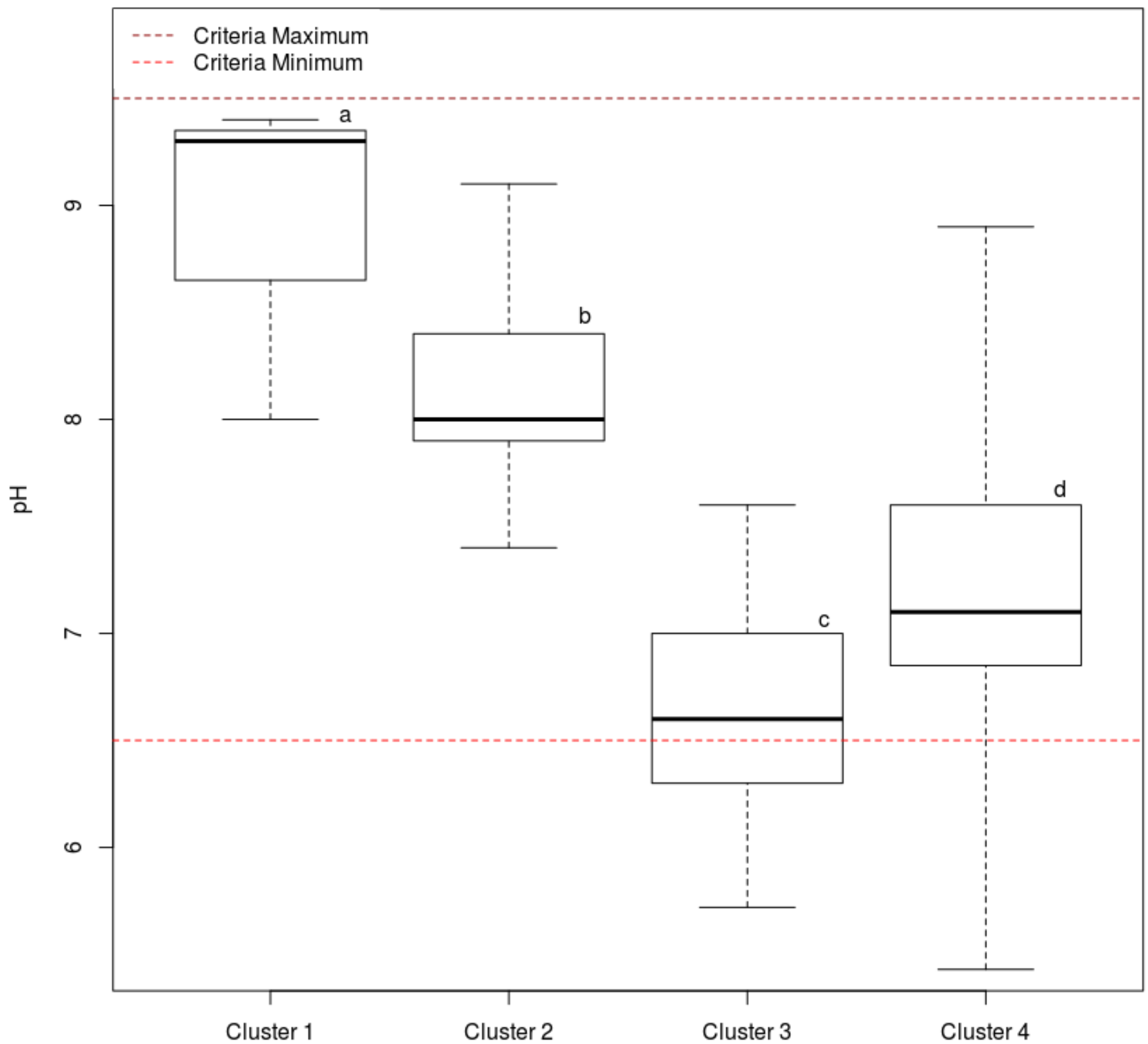


Figure 43. Boxplot of pH measurements by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an α = 0.05.

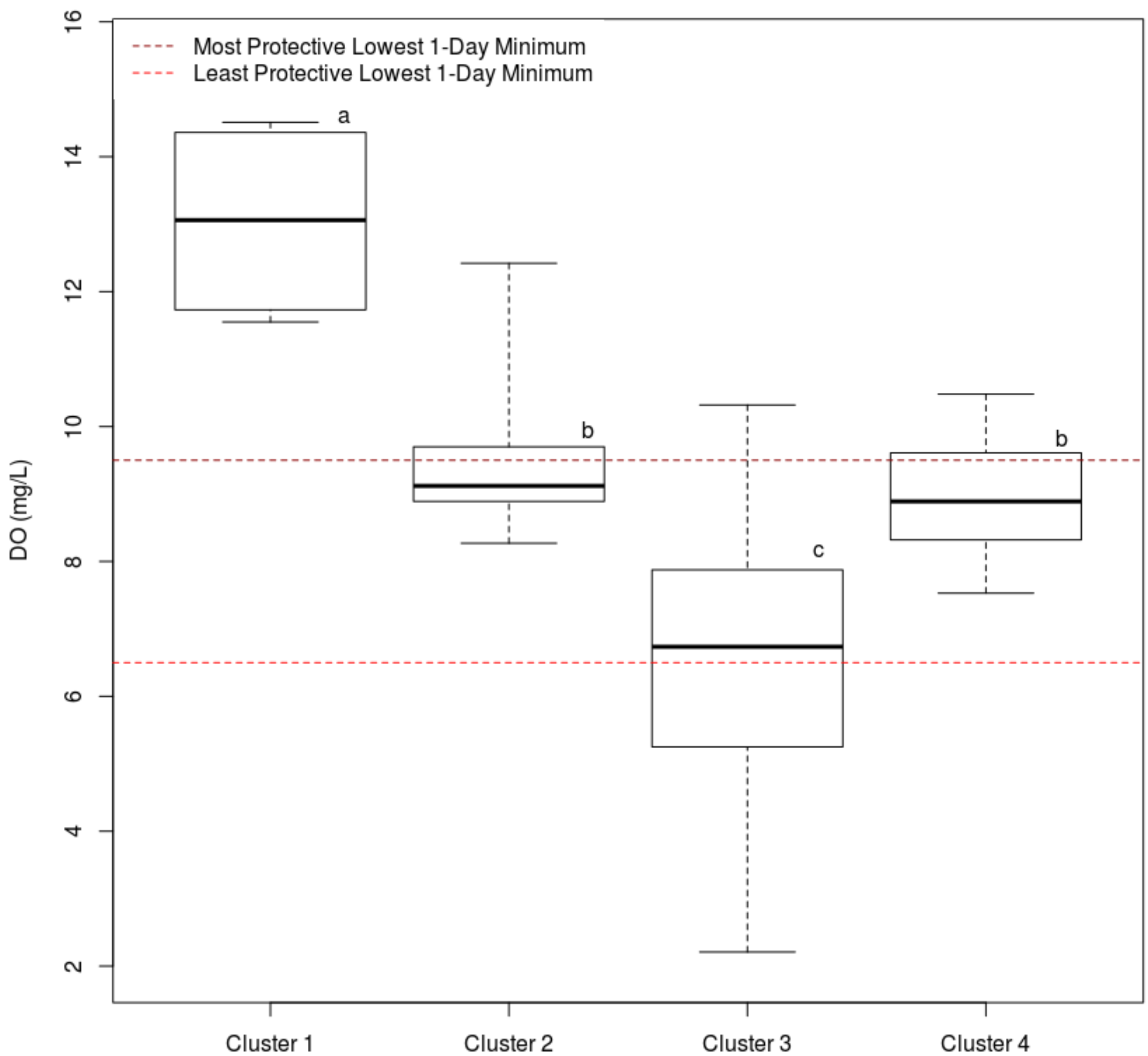


Figure 44. Boxplot of DO measurements by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an α = 0.05.

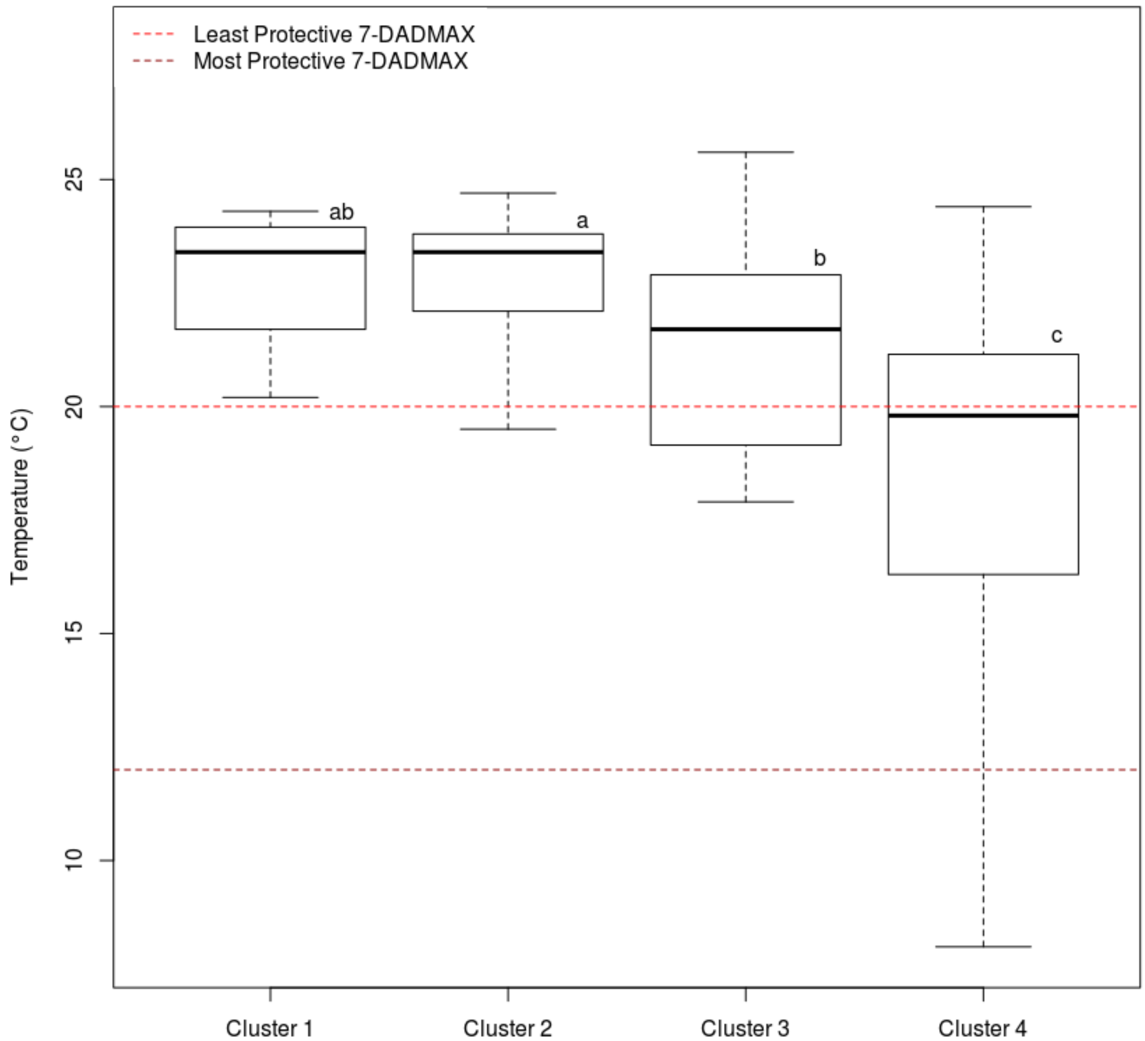


Figure 45. Boxplot of temperature measurements by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

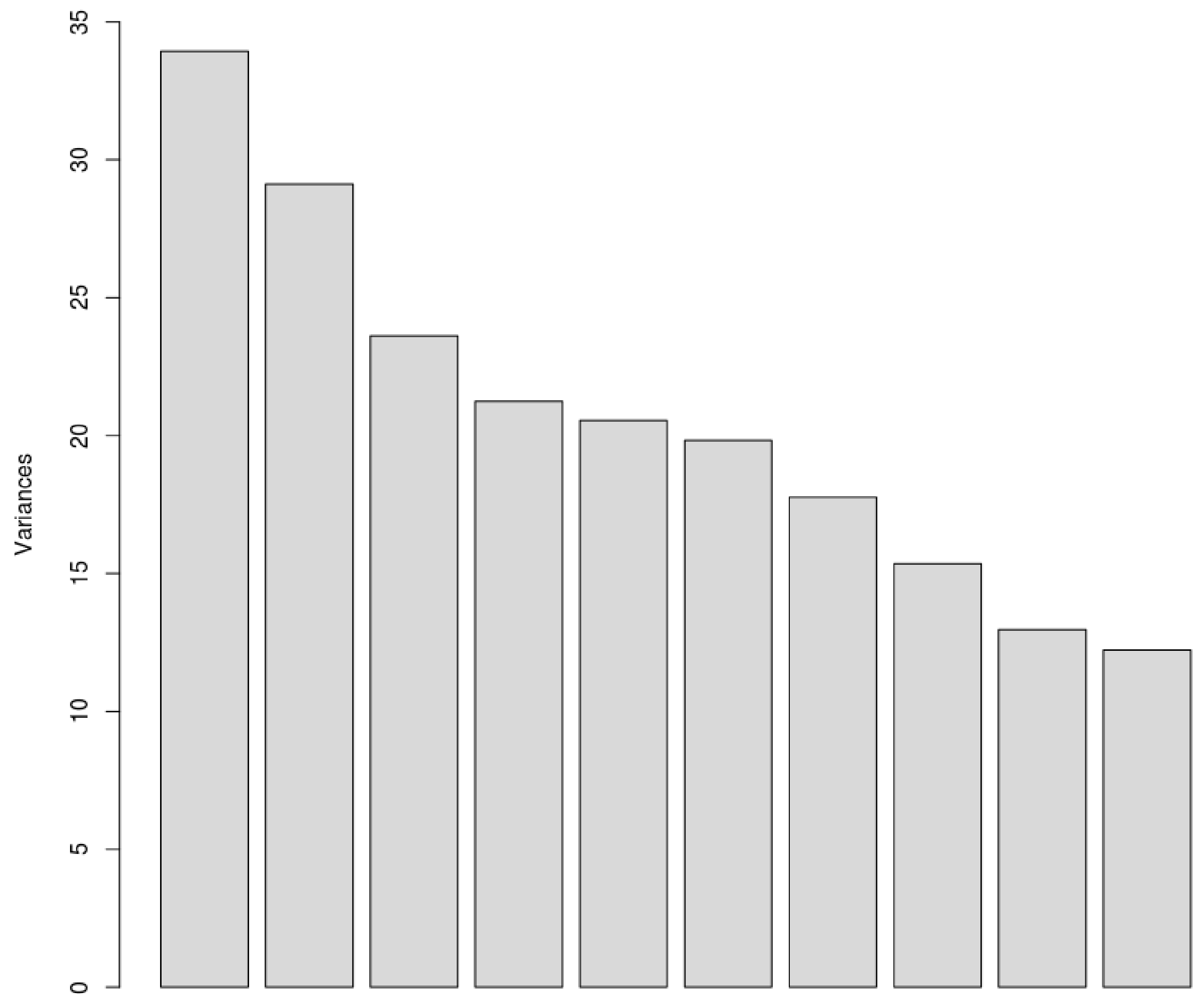


Figure 46. Variance plot of the first ten PCs for the PCA based on the algal richness data. These ten PCs account for 50.4% of the total variance.

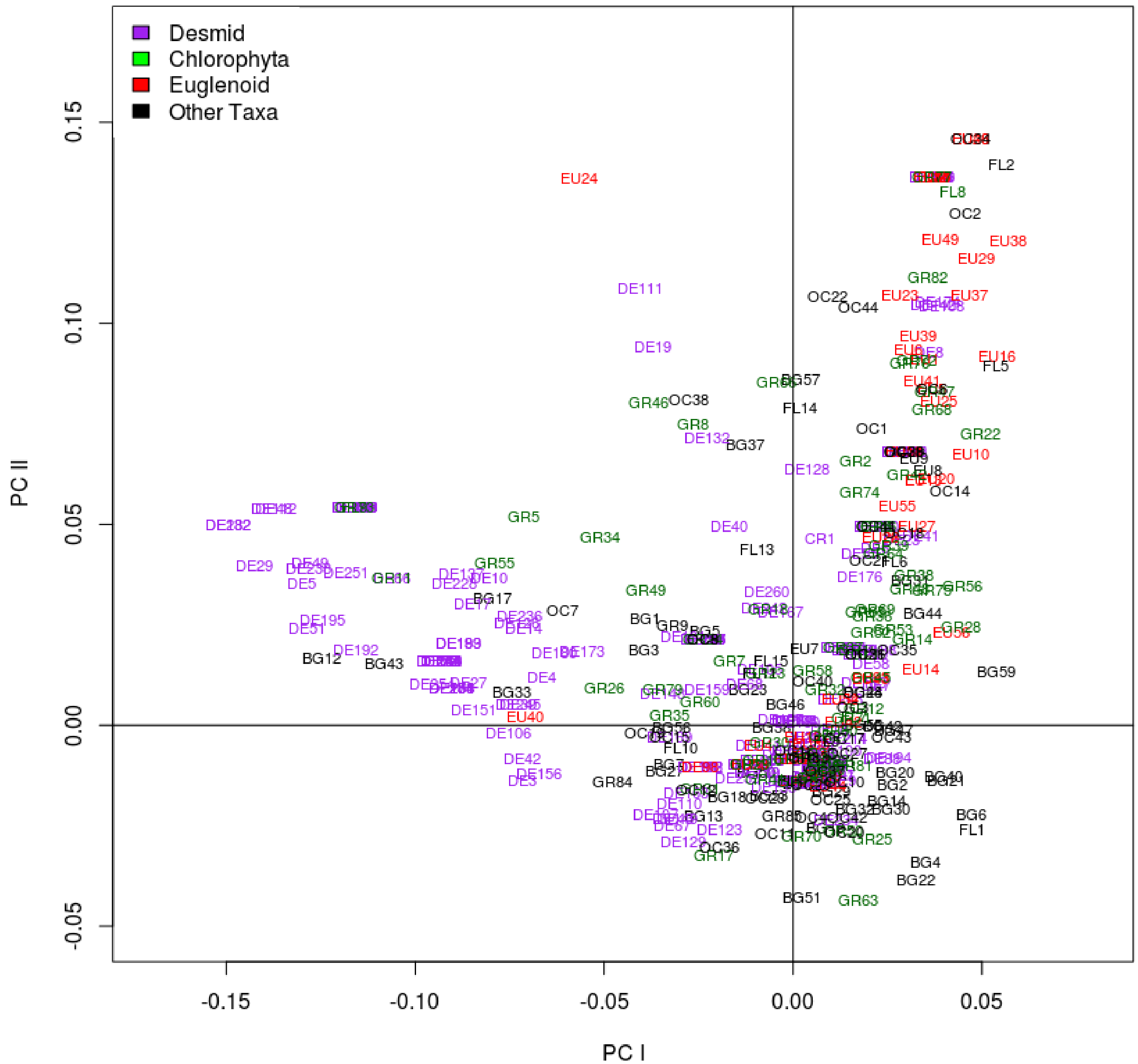


Figure 47. Ordination along PC I and PC II from the PCA of the algal richness data. Key taxa have been colored for visual comparison.

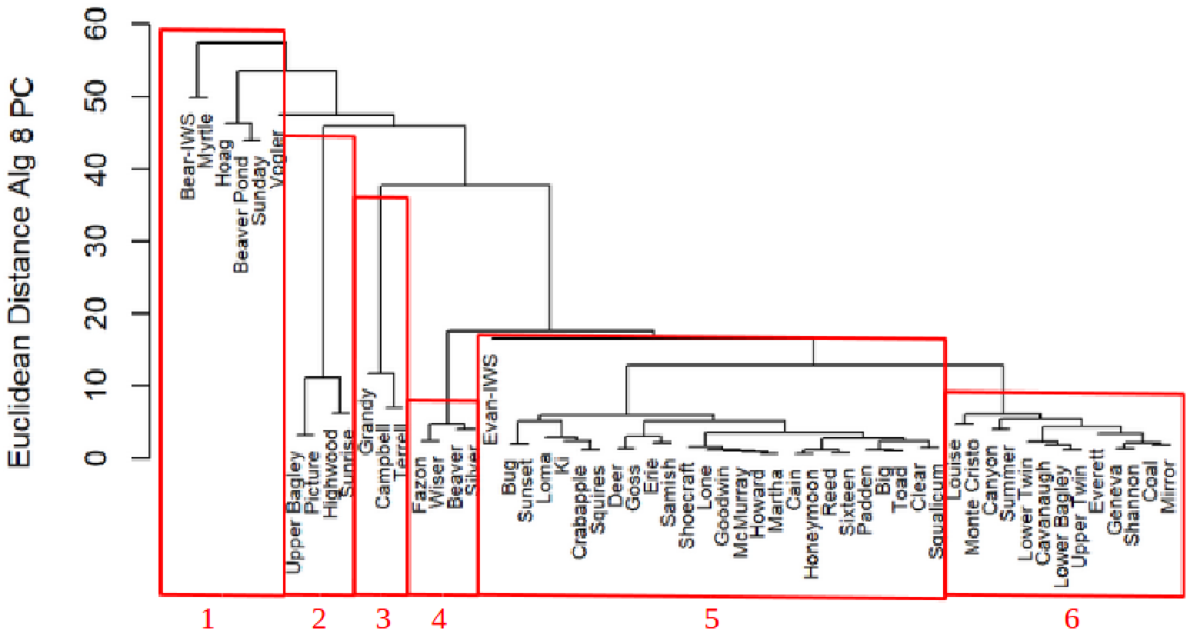


Figure 48. Hierarchical clustering of sampled lakes using algal presence data. Using the first eight principal components produced the most parsimonious clustering. These components represent 44.2 % of the total variance. Clusters are non-random. n = 56; χ-squared =18.4; p-value <0.01.

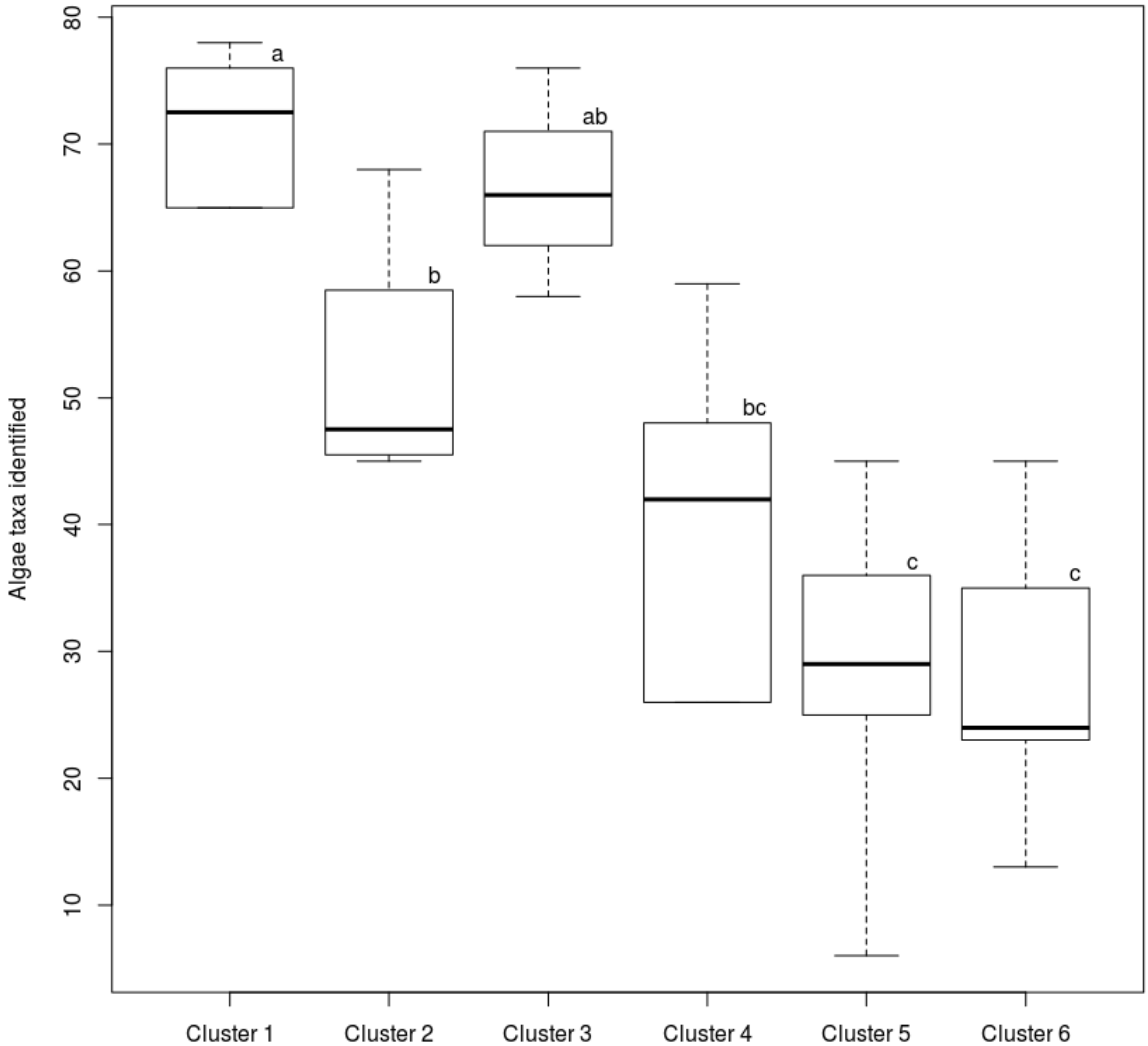


Figure 49. Boxplot of algae taxa identified by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

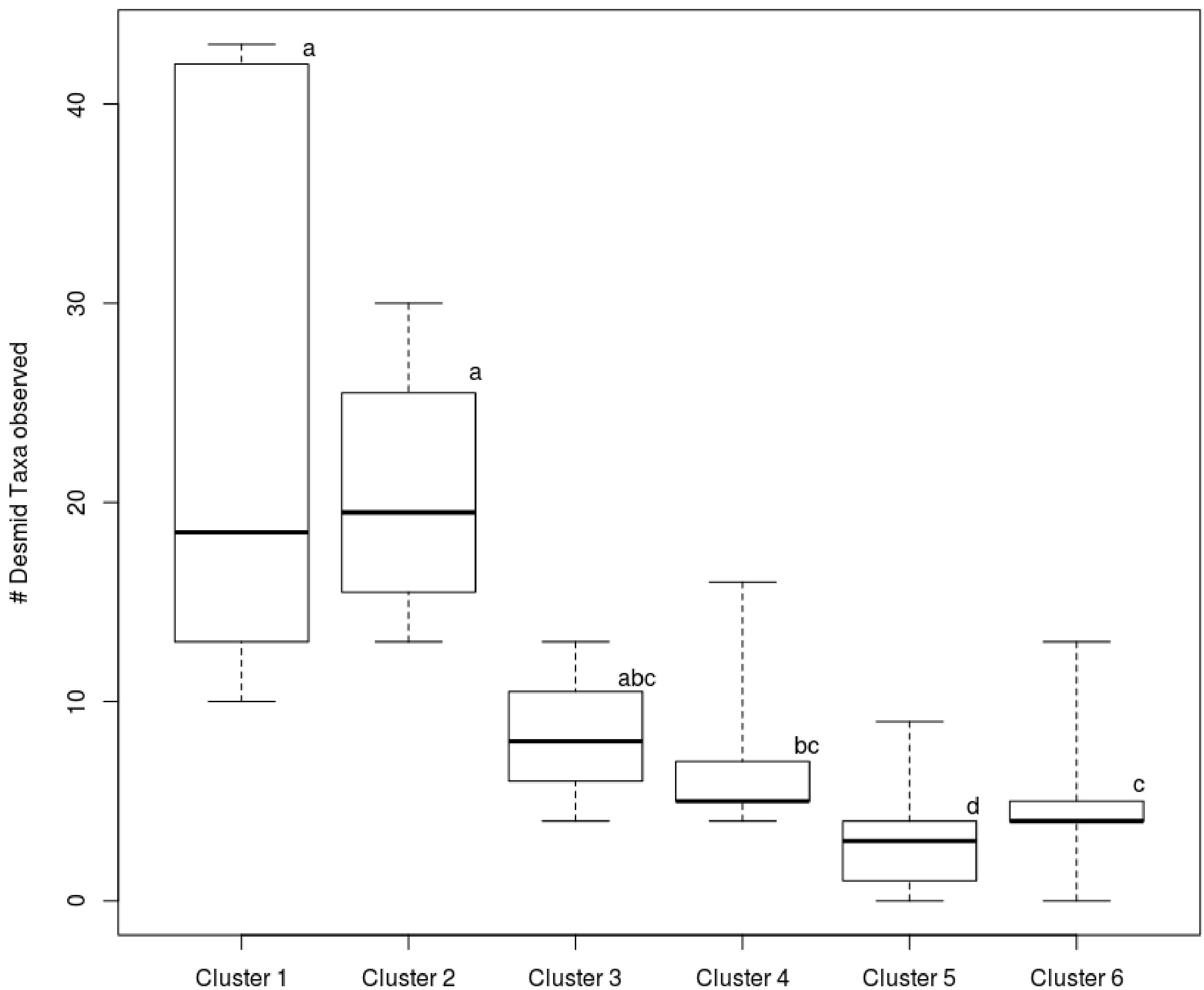


Figure 50. Boxplot of desmid taxa identified by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

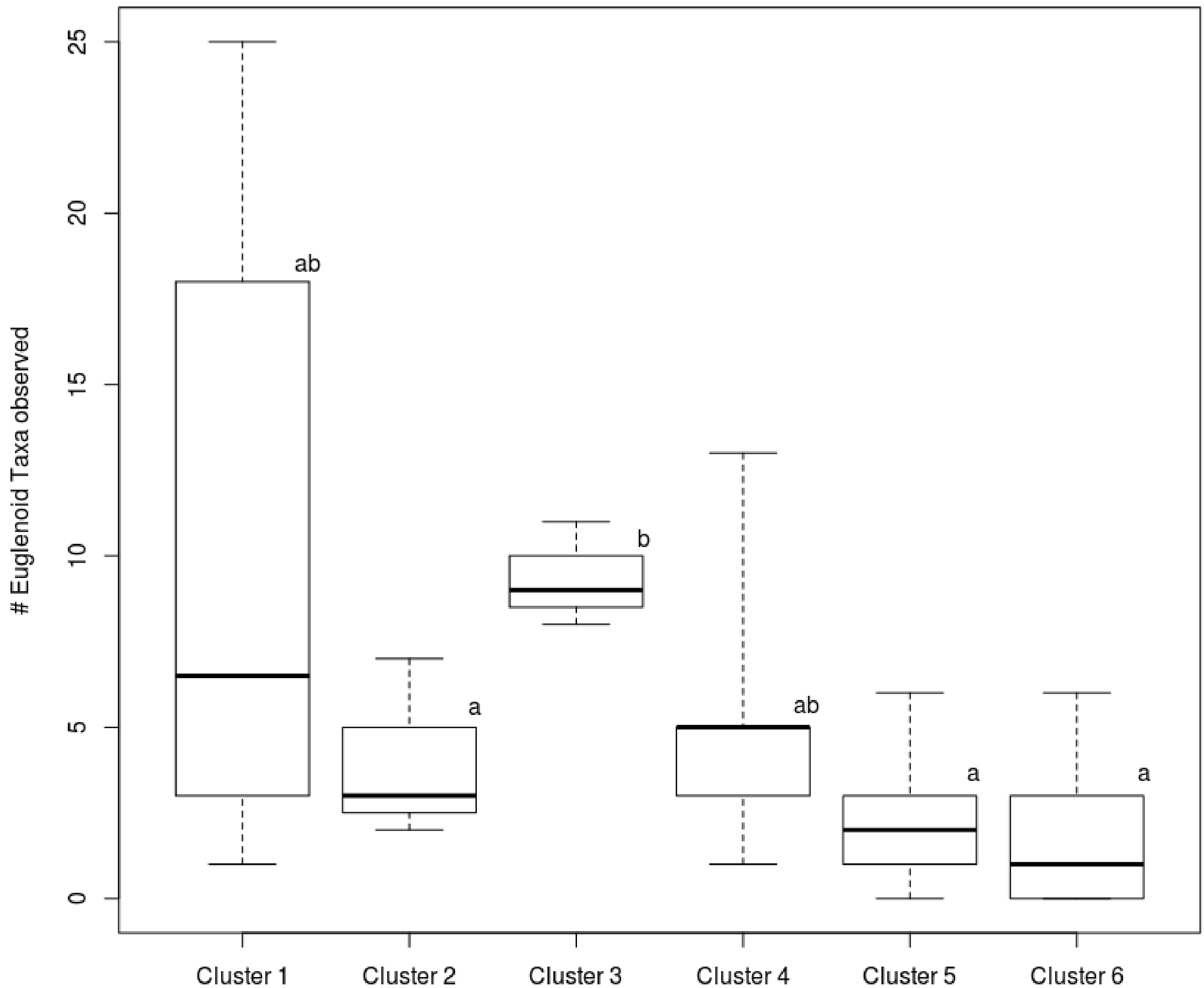


Figure 51. Boxplot of euglenoid taxa identified by hierarchical cluster. Shown on the boxplot are minimum, lower quartile, median, upper quartile, and maximum. Letters are used to denote significant differences between clusters by nonparametric Wilcoxon rank-sum tests with an $\alpha = 0.05$.

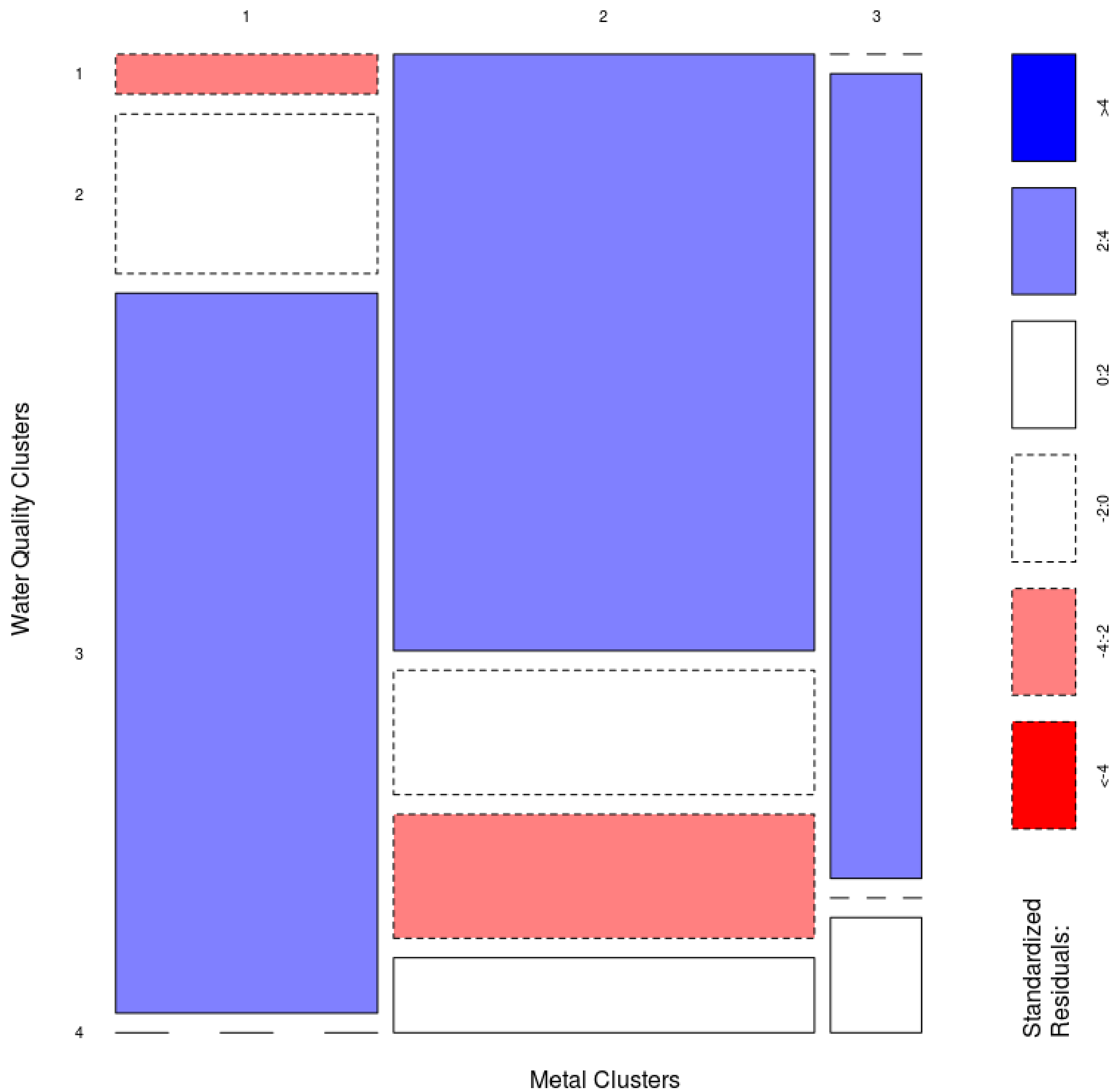


Figure 52. A mosaic plot of the three metal clusters and four water quality clusters. The size of the rectangle indicates the number of lakes that are shared between the two groups. The shade of the rectangles indicates the magnitude of the Pearson residuals and a dashed outline indicates a negative Pearson residual.

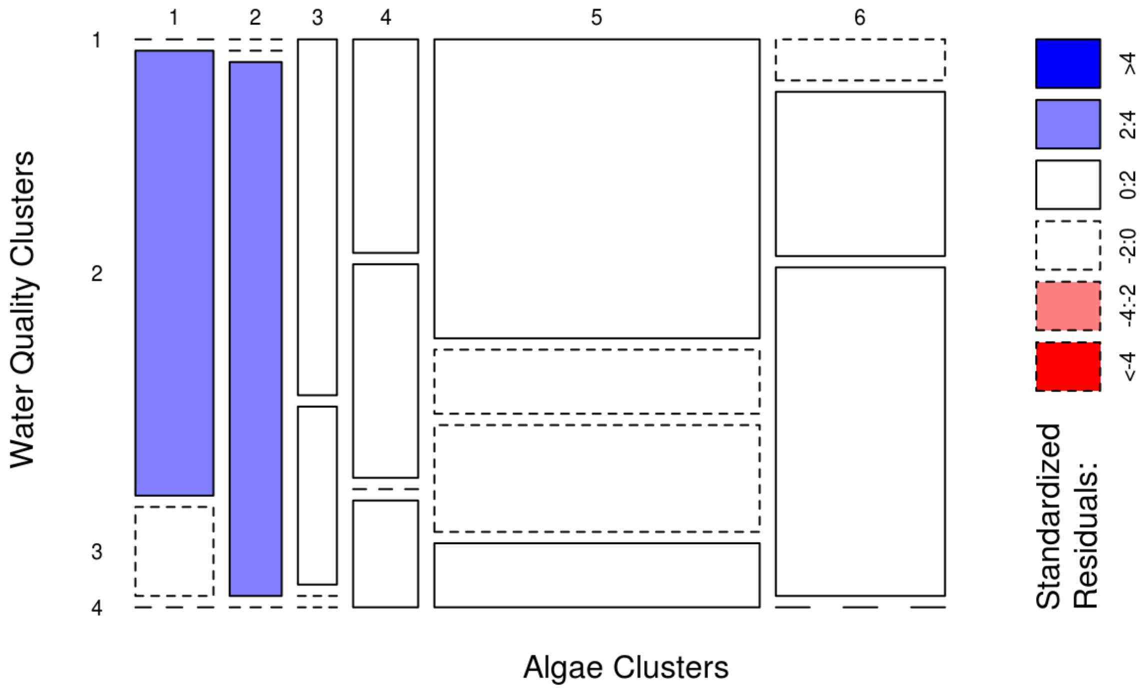


Figure 53. A mosaic plot of the six algae clusters and four water quality clusters. The size of the rectangle indicates the number of lakes that are shared between the two groups. The shade of the rectangles indicates the magnitude of the Pearson residuals and a dashed outline indicates a negative Pearson residual.

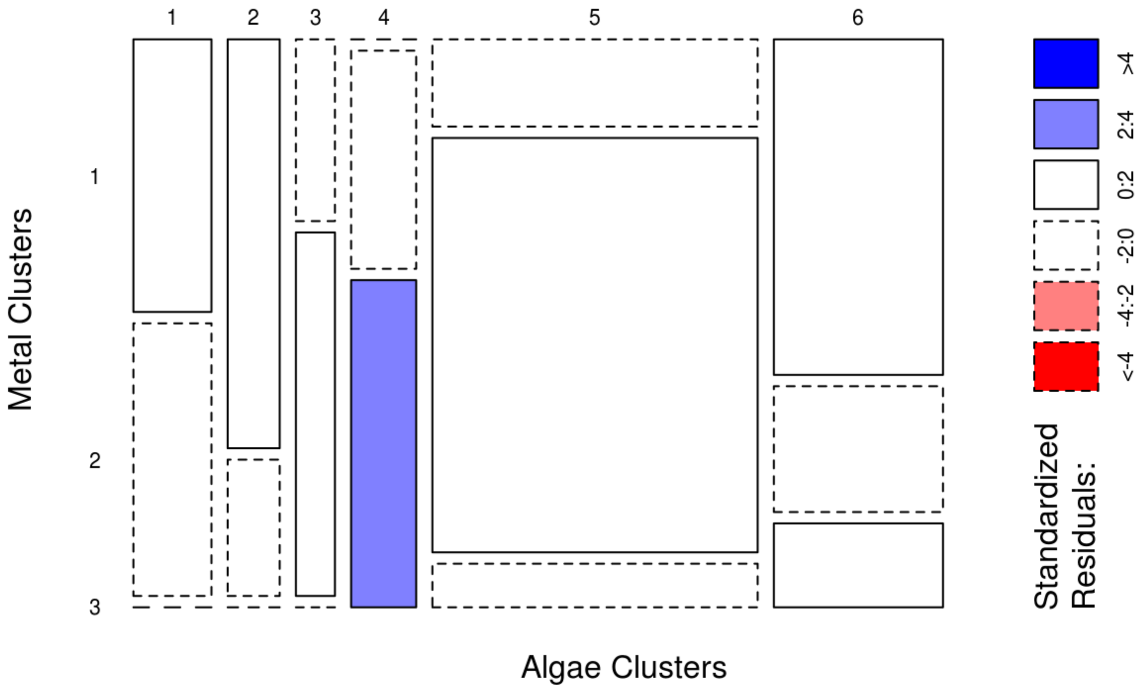


Figure 54. A mosaic plot of the six algae clusters and three metal clusters. The size of the rectangle indicates the number of lakes that are shared between the two groups. The shade of the rectangles indicates the magnitude of the Pearson residuals and a dashed outline indicates a negative Pearson residual.

Table 1. Parameter analysis location, methodology, detection limit, and confidence interval for all water quality parameters and metals. Parameters were either measured in situ, in the Institute for Watershed Studies Laboratory, or the Advanced Materials Science & Engineering Center. Adapted from Lawlor (2019).

Table 2. A list of analyzed elements and water quality parameters. Included are any analytes that were excluded for having too high of a relative standard deviation (RSD) between repeated measurements, having too many measurements below the detection limit (BDL) or isotopes that were removed to avoid redundancy within the principal components analysis.

Variable	Retained	Removed for RSD	Removed for BDL	Removed Isotope
Element	Magnesium, aluminum, potassium, calcium, vanadium, chromium, manganese, iron, cobalt, nickel, copper, zinc, arsenic, molybdenum, antimony, barium, lead	Beryllium	Beryllium, selenium, silver, cadmium, thallium, thorium, uranium	Iron 57, barium 137, lead 207
Water Quality Parameter	Temperature, dissolved oxygen, pH, conductivity, alkalinity, chlorophyll, turbidity, total phosphorus, total nitrogen, dissolved organic carbon.	Nitrate, soluble reactive phosphate	Nitrate	

Table 3. Summary statistics for metal concentrations for all 68 lakes sampled. Also included are the average detection limit (DL) and average limit of quantitation (LOQ) from the three days of metals analysis. Criteria Maximum Concentrations (CMCs) and Criterion Continuous Concentrations (CCCs) from Washington State's aquatic life criteria and the number of lakes that exceed those criteria are included where available. Some metal criteria require hardness corrections and their values are included at the lowest measured hardness and at the (highest measured hardness). The number of lakes exceeding those hardness-corrected criteria are for their specific water conditions. Other criteria are also included when appropriate Washington State values were not available. Table 3 shows measurements for lakes exceeding criteria. Raw data available in Appendix.

[a] Values are from US EPA 1988 National Recommended Aquatic Life Criteria.
[b] Values are acute and chronic criteria from British Columbia's 2004 Water Quality Guidelines.
[c] Values are Tier II acute and chronic values from the Indiana Department of Environmental Management.
[d] Value is from British Columbia's 2021 Working Water Quality Guidelines: Aquatic Life, Wildlife, & Agriculture.

Table 4. Measurements of specific metals in lakes that exceed criteria. Criteria Maximum Concentrations (CMCs) and Criterion Continuous Concentrations (CCCs) from Washington State's aquatic life criteria are included where available. Some metal criteria require hardness corrections and their values are included at that lake's specific criteria. Another criteria is included when appropriate Washington State values were not available.

[a] Value from US EPA 1988 National Recommended Aquatic Life Criteria.

Table 5. Summary statistics for water quality parameters for all 68 lakes sampled. Detection limit (DL) and confidence interval (CI) are included where available. Washington State aquatic life criteria and the number of lakes outside of the criteria are included. Due to the different criteria for specific water use designations, the least protective criteria and the (most protective criteria) have been included for temperature and dissolved oxygen.

Table 6. The relative contribution by each element to the variance accounted for in principal components one (PC1) and two (PC2). Relative contributions have been reported as absolute values.

Table 7. Comparison of cluster medians for the four metals that contribute most to principal component 1. Significantly different cluster values for each metal are denoted by different letters adjacent to each value. Significance was determined by a Mann-Whitney test using R (R Core Team 2021) with an $\alpha = 0.05$.

Table 8. The relative contribution by each water quality parameter to the variance accounted for in principal components one (PC1) and two (PC2).

Table 9. Comparison of cluster medians for all water quality parameters. Parameters are ordered by decreasing impact on principal component 1. Cluster medians not sharing a superscripted letter are significantly different. Significance was determined by a Mann-Whitney test using R (R Core Team 2021) with an $\alpha = 0.05$.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
	n=4	n=25	n=16	n=23
Water Quality Parameter	Median value			
Total Nitrogen (µg-N/L)	1915^{a}	401.3^{b}	544.4^{b}	109.4^{c}
Chlorophyll a (µg/L)	129^{a}	3.6^{b}	4.55^{b}	1.21^{c}
Turbidity (NTU)	25.35^{a}	0.9^{b}	1.655^{b}	0.49^{c}
Conductivity (µS/cm)	219.5^{a}	120.8^{b}	46.85^{c}	28.7^{c}
Alkalinity (mg/L)	64.9^{a}	42.8^{b}	15.95^{c}	9.7^{c}
DOC (mg/L)	13.9^{a}	5.9^{b}	8.15^{b}	1.8^{c}
Total Phosphorous (µg-P/L)	109.8^{a}	12.2^{b}	25.35^{c}	5.6^{d}
pH (pH unit)	9.3^{a}	8^{b}	6.6^{c}	7.1^{d}
DO (mg/L)	13.06^{a}	9.12^{b}	6.74^{c}	8.89^{b}
Temperature (°C)	23.4^{ab}	23.4^{a}	21.7^{b}	19.8^{c}

Table 10. Taxonomic units with the greatest contribution to the variance accounted for in principal components one (PC1) and two (PC2).

Table 11. Comparison of cluster medians for algae taxa identified. Cluster medians not sharing a superscripted letter are significantly different. Significance was determined by a Mann-Whitney test using R (R Core Team 2021) with an $\alpha = 0.05$.

	Cluster 1 n=6	Cluster 2 n=4	Cluster 3 n=3	Cluster 4 n=5	Cluster 5 n=25	Cluster 6 n=13
	Median					
Algae Taxa Identified	73^{a}	48^{b}	66^{ab}	42^{bc}	29^{c}	24^{c}
Desmid Taxa Identified	19^{a}	20^{a}	8^{abc}	5^{bc}	3^{d}	4^{c}
Euglenoid Taxa Identified	7^{ab}	3^{a}	9^{b}	5^{ab}	2^{a}	1^{a}

Literature Cited

Ben-Hur, A., Guyon, I. 2003. Detecting stable clusters using principal component analysis. Methods Mol Biol. 224:159-82. doi: 10.1385/1-59259-364-X:159. PMID: 12710673.

Bennett, L. 2017. Algae, cyanobacteria blooms, and climate change. Climate Institute Report, April 2017.

Bosco, M., Varrica, D., and Dongarrà, G. 2005. Case study: Inorganic pollutants associated with particulate matter from an area near a petrochemical plant. Environmental Research, Volume 99, Issue 1, Pages 18-30. ISSN 0013-9351. https://doi.org/10.1016/j.envres.2004.09.011.

British Columbia Water Protection Section. 2004 Technical report – Water quality guidelines for cobalt. Available online at: https://www2.gov.bc.ca/assets/gov/environment/air-land-water/water/waterquality/water-quality-guidelines/approved-wqgs/cobalt_tech.pdf.

Brown, J., and Peake, B. 2006. **Sources of heavy metals and polycyclic aromatic hydrocarbons in urban stormwater runoff.** Sci. Total Environ., 359 (1–3) (2006), pp. 145-155. doi: 10.1016/j.scitotenv.2005.05.016. Epub 2005 Jul 12. PMID: 16014309.

Caron, D., Sanders, R., Lim, E Marrase, C., Amaral, L., Whitney, S., Aoki, R., and Porters, K. 1993. Light-dependent phagotrophy in the freshwater mixotrophic chrysophyte Dinobryon cylindricum. Microb Ecol 25, 93–11. https://doi.org/10.1007/BF00182132\.

Davis, A., Shokouhian, M., Ni, S. 2001. Loading estimates of lead, copper, cadmium, and zinc in urban runoff from specific sources. Chemosphere,Volume 44, Issue 5, Pages 997-1009. ISSN 0045-6535. https://doi.org/10.1016/S0045-6535(00)00561-0.

Indiana Department of Environmental Management. 2015. Tier II acute and chronic aquatic life values. Available online at: https://www.epa.gov/sites/production/files/2015-06/documents/in_al_463_10011998.pdf.

Institute for Watershed Studies. 2021. Northwest Lake Monitoring. Institute for Watershed Studies, Western, Washington University, Bellingham, Washington. Available online at: https://www.wwu.edu/iws/

Klemenčič, A., Smolar-Žvanut, N., Istenič, D., and Griessler Bulc, T. 2010. Algal community patterns in Slovenian bogs along environmental gradients. Biologia. 65. 422-437. 10.2478/s11756-010-0033-7.

Lawlor, M. 2019. A Multivariate Analysis of the Relationship Between the Water Quality Conditions and Algal Species Composition of Six Mountain Lakes in the North Cascades, WA, USA. M.S. Thesis, Western Washington University, Bellingham, Washington.

LeBlanc, R., Brown, R., and FitzGibbon, J. 1997. Modeling the effects of land use change on the water temperature in unregulated urban streams. Journal of Environmental Management, 49, 445-469.

Liu, J., Tan, K., Linjuan, H., Qiu, Y., Tan, W., Guo, Y., Wang, Z., and Sun, W. 2018. Effect of limitation of iron and manganese on microalgae growth in fresh water. Microbiology Society 164:12.

Llewellyn, C. 2010. Predicting cyanobacteria blooms in 50 lakes of Northwest Washington. M.S. Thesis, Western Washington University, Bellingham, Washington.

Markiewicz, A., Björklund, K., Eriksson, E., Kalmykova, Y., Strömvall, A., and Siopi, A. 2017. Emissions of organic pollutants from traffic and roads: Priority pollutants selection and substance flow analysis. Sci Total Environ. 15;580:1162-1174. doi: 10.1016/j.scitotenv.2016.12.074. Epub 2016 Dec 27. PMID: 28038877.

McGrane, S. 2016. Impacts of urbanisation on hydrological and water quality dynamics, and urban water management: a review. Hydrological Sciences Journal, 61 (13), 2295–2311. doi:10.1080/02626667.2015.1128084.

McKenzie, E., Money, J., Green, P., and Young, T. 2009. Metals associated with stormwater-relevant brake and tire samples. *The Science of the total environment*, *407*(22), 5855–5860. https://doi.org/10.1016/j.scitotenv.2009.07.018.

Meisch, H., Benzschawel, H., and Bielig, H. 1977. The role of vanadium in gree plants. II. Vanadium in green algae--two sites of action. Arch Microbiol 26;114(1):67-70. doi: 10.1007/BF00429632. PMID: 20864.

Müller, A., Österlund, H., Marsalek, J., Viklanderm M. 2020. The pollution conveyed by urban runoff: A review of sources. Science of The Total Environment, Volume 709, 136125, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2019.136125.

Pfannenstein, K. 2016. Water quality and algal diversity of ten lakes along the Mountain Loop Highway, Washington. M.S. Thesis, Western Washington University, Bellingham, Washington.

Praus, P. 2005. Water quality assessment using SVD-based principal component analysis of hydrogeological data. Water SA., 31: 417-422.

Putt, A., MacIsaac, E., Herunter, H., Cooper, A., and Selbie D. 2019. Eutrophication forcings on a peri-urban lake ecosystem: Context for integrated watershed to airshed management. PLoS ONE 14(7): e0219241. https://doi.org/10.1371/journal.pone.0219241.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at https://www.R-project.org/.

Ragothaman, A., and Anderson, W. 2017. Air Quality Impacts of Petroleum Refining and Petrochemical Industries. Environments, 4, 66. https://doi.org/10.3390/environments4030066

Shumar, M. and de Varona, J. 2009. The potential natural vegetation (PNV) temperature total maximum dailyload (TMDL) procedures manual. Surface Water Program, Idaho Department of Environmental Quality.

Sleeter, B.M., Wilson, T.S., and Acevedo, W., eds. 2012. Status and trends of land change in the Western United States—1973 to 2000: U.S. Geological Survey Professional Paper 1794–A, 324 p. Available online at: http://pubs.usgs.gov/pp/1794/a/.

Thorson, T., Bryce, S., Lammers, D., Woods, A., Omernik, J., Kagan, J., Pater, D., and Comstock, J. 2003. Ecoregions of Washington (color poster with map, descriptive text, summary tables, and photographs): Reston, Virginia, U.S. Geological Survey (map scale 1:1,500,000).

Tong, S.. and Chen, W. 2002. Modeling the relationship between land use and surface water quality. Journal of Environmental Management, 66, 377–393. doi:10.1006/jema.2002.0593

US Environmental Protection Agency. 1994. Method 200.8: Determination of Trace Elements in Waters and Wastes by Inductively Coupled Plasma-Mass Spectrometry, Revision 5.4.

US Environmental Protection Agency. 2014. Cyanobacteria and Cyanotoxins: Information for Drinking Water Systems. U. S. Environmental Protection Agency, Office of Water, EPA-810F11001.

US Environmental Protection Agency. 2018 Final 2018 aquatic life ambient water quality criteria for aluminum in freshwaters. U. S. Environmental Protection Agency, Office of Water, EPA-822-F-18-003. Available online at: https://www.epa.gov/sites/production/files/2018-12/documents/aluminum-criteria-final-factsheet.pdf.

US Environmental Protection Agency. 2021. Washington water quality assessment report. Available online at: https://ofmpub.epa.gov/waters10/attains_state.control?p_state=WA#causes_303d.

Van Buren, M., Watt, W., Marsalek, J., and Anderson, B. 2000. Thermalenhancement of stormwater runoff by paved surfaces. Water Resources,34(4), 1359-1371.

Washington State Department of Ecology. 2011. Monte Cristo Mining Area. Toxics Cleanup Program. Available online at: https://apps.ecology.wa.gov/gsp/CleanupSiteDocuments.aspx?csid=4550

Washington State Department of Natural Resources. 2015. Geologic provinces: Puget Lowland. Available online at: https://www.dnr.wa.gov/programs-and-services/geology/explore-popular-geology/geologic-provinces-washington/puget-lowland#geology.

Wei, T., Simko, V. 2021. R package 'corrplot': Visualization of a Correlation Matrix. (Version 0.90), Available online at: https://github.com/taiyun/corrplot.

Wong, S. 2013. Phytoplankton ecology in four high-elevation lakes of the North Cascades, WA. M.S. Thesis, Western Washington University, Bellingham, Washington.

Woolway, R., Jones, I., Maberly, S., French, J., and Livingston, D. 2016. Diel Surface Temperature Range Scales with Lake Size. PLOS ONE 11(3): e0152466.