Agricultural and urban land use explain substantial variation in lake water quality across national-scale physiographic and climatic gradients

Joe R. Sánchez Schacht*[1], Paul W. MacKeigan[1,4], Zofia E. Taranu[2,4], Yannick Huot[3,4], Irene Gregory-Eaves[1,4]

Author affiliations:

[1] Department of Biology, McGill University, Montreal, Quebec, Canada

[2] Aquatic Contaminant Division, Environment and Climate Change Canada, Montreal, Quebec, Canada

[3] Département de géomatique appliquée, Université de Sherbrooke, Montreal, Quebec, Canada

[4] Group for Interuniversity Research in Limnology and aquatic environment, Montreal, Quebec, Canada

* Corresponding author:

Email: joe.sanchezschacht@mail.mcgill.ca.

Abstract

Despite decades of research and mitigation efforts, declines in freshwater quality resulting from anthropogenic nutrient input remain a persistent issue worldwide. Canada has the greatest number of freshwater lakes in the world, yet we have a limited understanding of the magnitude and scale at which most lakes have been affected by human activities, namely Land Use/Land Cover (LULC) alterations. In response, the NSERC Canadian Lake Pulse Network has compiled the first nationwide systematic database of Canadian lake quality metrics by surveying 664 lakes across 12 ecozones over three years. To assess the influence of human impact on water quality and its spatial variation across this lake set, we built models quantifying the association between watershed LULC and physiochemical water parameters. We found that agricultural and urban land use explained the greatest proportion of variation in water quality among LULC categories (R^2 = 0.20–0.29), although threshold values of water quality metrics (i.e., concentrations of major ions, total phosphorus [TP], total nitrogen [TN], dissolved organic carbon [DOC], and Chlorophyll *a* [Chl-*a*]) differed among ecozones. Generally, when the percentage of catchment agricultural land use exceeded 30–35%, we noted significant increases in TP, TN, DOC, and Chl-*a* concentrations. Similarly, when urban land use surpassed 1–6%, we observed significant increases in ion concentrations and specific conductance. Overall, our study highlights that while drivers of water quality may be similar across regions, management strategies for freshwater ecosystems must consider their geographic context to better predict when and where threshold responses will occur.

Keywords

Agriculture; Canada; eutrophication; lake; urban; water quality.

Introduction

Despite decades of research into the dynamics of lake water quality and evolving land management strategies, cultural eutrophication remains one of the leading threats to freshwater ecosystems across the globe (Carpenter et al. 1998; Beusen et al. 2016; Reid et al. 2019). Influx of limiting nutrients, specifically nitrogen (N) and phosphorus (P), from point sources (i.e., sewage effluent and/or factory discharges) was first recognized as an important contributor to the eutrophication of freshwater ecosystems. Often, the diversion of these point-source nutrients leads to improvements in water quality (Jeppesen et al. 2005; Kronvang et al. 2005). However, mitigation of point sources does not always result in reduced nutrient and algal biomass concentrations in lakes, even in cases where combined (N and P) nutrient loading reduction programs have been applied to lakes and their catchments (Özkundakci et al. 2011; McCrackin et al. 2017). Nonpoint, or diffuse, sources of nutrient inputs remain the main cause of eutrophication worldwide (Reid et al. 2019) partly due to the difficulties in quantifying and regulating them: fluxes can vary temporally (Medalie et al. 2012) and spatially (Lencha et al. 2022) and are often regulated by management strategies with conflicting purposes such as food production and natural resource protection (Wunderlich and Martinez 2018). Furthermore, annual inputs of N and P from agricultural activities often exceed the removal by harvested crops (MacDonald et al. 2011; Spiess 2011), leading to a surplus of highly mobile nutrients in soils (GOC 2016) that can leach into downstream water bodies (Cassman and Dobermann 2022). The

accumulation of soil P and N can result in legacy loading that can take decades to deplete (Grimvall et al. 2000; Van Meter et al. 2016).

LULC in the watershed has previously been linked to variation in lake water quality across the world. More specifically, agriculture and urbanization have been linked to poorer water quality (i.e., eu- or hyper-eutrophic status, increases in Chl-*a*, TP, TN, ions, conductivity, DOC; decrease in Secchi depth) compared to natural landscapes (Taranu and Gregory-Eaves 2008; Nielsen et al. 2012; Read et al. 2015; de Mello et al. 2020), primarily through higher concentrations of limiting nutrients and their symptomatic cyanobacterial blooms (Doubek et al. 2015; Almanaza et al. 2019). Nevertheless, the relationship between watershed LULC and lake water quality varies significantly over time and space (de Mello et al. 2020) and is often modulated by factors such as precipitation and temperature (Collins et al. 2019), flow regime (Mbonde et al. 2015; Abirhire et al. 2016), lake physiography (i.e., area, depth, altitude; Taranu and Gregory-Eaves 2008; Read et al. 2015), and bedrock geology (Dillon and Kirchner 1975; Griffith 2014). Therefore, studies at a large spatial scale must strive to disentangle the regional and local drivers of water quality. Ultimately, quantifying the effect of nonpoint nutrient loading on water quality across different Canadian ecozones will be instrumental in helping us predict which regions will be more susceptible to lake eutrophication in the event of future LULC modifications.

Nonetheless, agricultural production is an important contributor to Canada's economy and a critical part of society. For example, Canada's agricultural sector generated 39.8 billion dollars in 2020 (GOC 2021). Given the economic importance of its domestic agricultural production, the

Canadian federal, provincial, and territorial governments formed the Canadian Agricultural Partnership, investing three billion dollars in federal programs and strategic initiatives over five years (2018–2023) to help strengthen the industry, of which 23% is dedicated to research and innovation (GOC 2019). Other investments by the Canadian government over the next 10 years are targeted for the adoption, testing, and monitoring of best management practices (BMPs; e.g., no-till, cover crops) that directly or indirectly improve inland water quality (Faust et al. 2018). Yet, voluntary participation in unsubsidized BMPs remains low (Liu and Brouwer 2022), with Canadian agricultural producers citing costs as an important factor when deciding to implement BMPs (Hassanzadeh et al. 2019). Failing to adopt BMPs and other progressive agricultural strategies can lead to the development and persistence of harmful algal blooms across North America (Munawar and Fitzpatrick 2018; Graham et al. 2020). Stakeholder engagement remains crucial to successful water quality management, as the rates of BMP adoption by producers can influence the effectiveness of said practices at a watershed scale (Zammali et al. 2021).

Academics, federal, provincial and territorial governments, as well as watershed organizations have shown an increased interest in assessing the current state of freshwater ecosystems across the country (Pérez-Jvostov et al. 2020). To answer the need for a cohesive, national water-monitoring program, such as the National Lakes Assessment (NLA) in the United States, the NSERC (Natural Sciences and Engineering Research Council of Canada) Canadian Lake Pulse Network was founded in 2016 (Huot et al. 2019). Over three years, the Lake Pulse Network sampled 664 lakes in 12 ecozones spread across the country, following a standardized sampling protocol. Lakes were targeted according to a stratified random sampling design to account for a range of lake sizes and of LULC in watersheds, with land use serving as a proxy for human

impact (Huot et al. 2019). Very large lakes (>100 km^2) were excluded since most are already consistently monitored by both private and public interest groups.

Our overarching goal is to provide a more robust understanding of how freshwater ecosystems respond to varying levels of anthropogenic stress in a heterogeneous and expansive country. By using water physiochemistry and LULC data from the 664 Canadian LakePulse study lakes and watersheds, we quantified the spatial variation in water quality (mainly TP, TN, major ion, and Chl-*a* concentrations) among lakes as explained by watershed land use. We tested the hypothesis that the impact of LULC on lake water quality varies across lakes due to differences in lake and watershed physiography and geology. We applied both linear and nonlinear models to quantify the variation in water quality explained by various LULC categories including known stressors such as agriculture and urbanization, while also considering lake and watershed physiography (altitude, circularity, area, depth, and shoreline slope). Furthermore, we determined thresholds of agricultural and urban land use that were correlated with significantly poorer water quality parameters.

Materials and methods

Lake data

Over the span of three summers (June to September) from 2017 to 2019, teams from the NSERC Canadian Lake Pulse Network sampled 664 lakes across all ten Canadian provinces and two territories (Yukon and Northwest Territories). Approximately 600 lakes were in the nine southernmost ecozones (Pacific Maritime, Montane Cordillera, Semi-Arid Plateaux, Prairies, Boreal Plains, Boreal Shield, Mixedwood Plains, Atlantic Maritime, and Atlantic Highland; Fig.

1), while the remaining lakes were situated in the three northernmost, remote ecozones (Taiga Plains, Boreal Cordillera, and Taiga Cordillera). Target lakes were within 1 km from a road and selected using a random stratified design, controlling for three levels of human impact (low, medium, high; calculated from watershed LULC categories) and three lake sizes (0.1–0.5, 0.5–5, 5–100 km^2; see Huot et al. 2019). In a few cases, randomly selected lakes were replaced with lakes in the same size and human impact category because they had associated historical data or because they were important to certain LakePulse partners.

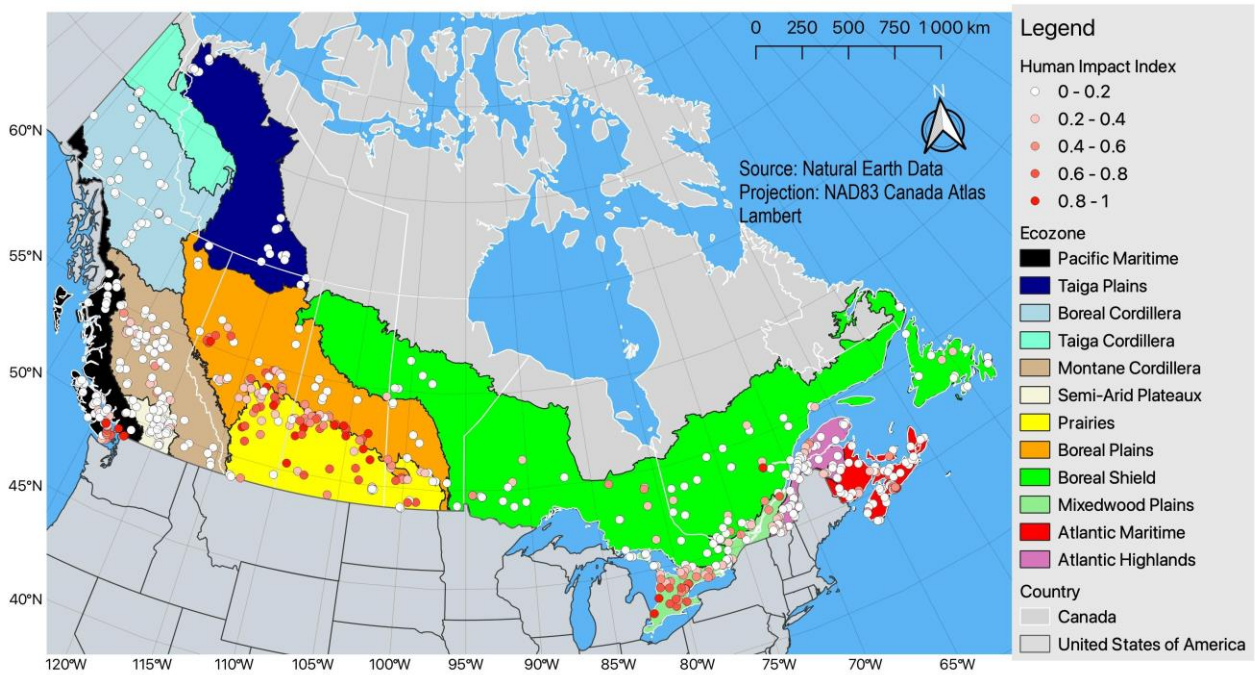


Figure 1. Ecozone map of the 664 sampled watersheds shaded by human impact index. The bulk of sampling was conducted in the nine southernmost main ecozones (Pacific Maritime, Montane Cordillera, Semi-Arid Plateaux, Prairies, Boreal Plains, Boreal Shield, Mixedwood Plains, Atlantic Maritime, and Atlantic Highland), while sampling was limited in the three northernmost and remote ecozones (Taiga Plains, Boreal Cordillera, Taiga Cordillera).

From the deepest point of each lake, survey teams recorded dozens of physical and chemical variables (Table 1). Water quality variables included Secchi depth, dissolved hypolimnetic oxygen, specific conductance, magnesium, potassium, sodium, sulfate, soluble reactive P, total P, total N, Chlorophyll *a*, dissolved organic carbon, and dissolved inorganic carbon. Samples for most water quality variables were taken from the photic zone (2x Secchi depth up to 2 meters), and those for additional variables (e.g., temperature, specific conductance) were taken from the full water column. In an effort to standardize lake assessment and variable collection protocols across North America, the sampling process was modelled after the NLA in the United States. For full details of the field protocol, see the NSERC Canadian Lake Pulse Network Field Guide (2021). For the purposes of this study, lakes were categorized by their trophic status (oligotrophic, mesotrophic, eutrophic and hypereutrophic) using well-established TP thresholds (Wetzel 2001).

Watershed LULC and physiographic data

LULC information for each lake's watershed was obtained from a variety of sources, including the Annual Space-Based Crop Inventory for Canada 2016 (GOC 2017) and the Land Use 2010 database (GOC 2015a), and analyzed using geomatic software (see Huot et al. 2019). Each lake's respective watershed was delineated to a 30 m^2 pixel size. Watershed delineation used flow directions calculated with the Canadian Digital elevation Model (GOC 2015b). LULC type (Agriculture, Pasture, Urban, Forestry, Mines, Natural landscapes, and Upland water) was extracted for each pixel from various public databases and an aggregate watershed human impact was derived based on these LULC data (see Huot et al. 2019 for details). A human impact index

was calculated by assigning relative scores to each form of LULC (see Supplemental Table S1) and represents the average value across all pixels in the watershed.

Various physiographic variables were measured for a subset of 636 lakes and watersheds, four of which we used as covariates in the variation partitioning analysis: lake altitude, circularity, dynamic ratio (estimated by dividing the square root of the lake area by the mean depth of the lake; Håkanson 1982; Ferland et al. 2012) and the slope of a 100 m land buffer from the shoreline into the watershed. The map of Canada and the sampled lakes was illustrated with QGIS 3.0 (QGIS.org 2022).

Statistical analyses

All statistical analyses were performed using R v4.0 software (R Development Core Team 2020). Since several of our analyses assumed a normal distribution of data, we examined the distribution of each response variable using the Shapiro-Wilks test combined with a visual inspection of the histogram. For all variables that did not meet the test's normality assumption ($p > 0.05$), we applied a Box Cox (Borcard et al. 2018), logarithmic, or square root transformation. The Box Cox transformations were computed with the R package *geoR* (Ribeiro et al. 2020). The transformed variables were then evaluated again with the Shapiro-Wilks test and a histogram inspection, and the transformation that was closest to being normal was kept (specific transformations, where appropriate, are included in Table 1). To improve model fit, the same test for normality and transformation exercise were applied to the LULC variables, for which we additionally attempted an arcsine transformation. For the principal components, redundancy, and variation partitioning analyses, LULC variables with zero values were offset by adding 0.00001

when Box Cox or logarithmic transformations were applied. However, the raw LULC values were used for the multivariate regression tree (MRT), cascade multivariate regression tree (cMRT), and generalized additive models (GAMs) since these methods are robust to nonlinear responses. We found no significant collinearity (>0.7 Pearson correlation or >10 VIF) among the predictors. However, watershed proportions of natural landscapes and mines were omitted as predictors to prevent redundancy and convergence issues in our models.

Table 1. Distribution and transformation coefficients of land use/land cover, water quality, and lake physiography variables for use in the principal component analysis, the redundancy analysis, and the variation partitioning (N = 664 lakes).

	Variable (symbol)	**Unit**	**Min/Median /Max**	**Transformation**	**Shapiro-Wilks p-value pre-transformation**	**Shapiro-Wilks p-value post-transformation**
Land Use/Land Cover	Percent agriculture (Agriculture)	%	0/0/86	None	< 2.2e-16	< 2.2e-16
	Percent pasture (Pasture)	%	0/0/55	None	< 2.2e-16	< 2.2e-16
	Percent urbanization (Urban)	%	0/2/99	Box Cox	< 2.2e-16	1.2e-10
	Percent forestry (Forestry)	%	0/0/59	Box Cox	< 2.2e-16	5.2e-14
	Percent mines (Mines)	%	0/0/28	None	< 2.2e-16	< 2.2e-16

	Percent natural landscapes (NatLand)	%	2/80/100	Arcsine	< 2.2e-16	4.8e-11
	Percent upland water (Water)	%	0/1/39	None	< 2.2e-16	< 2.2e-16
Water quality	Secchi depth (Secchi)	m	0.08/2.45/17.75	Box Cox	< 2.2e-16	0.05
	Epilimnetic temperature (Temp)	°C	11.80/20.44/29.73	Box Cox	9.5e-08	1.7e-04
	Dissolved hypolimnetic oxygen (DO)	mg/L	0/7.02/15.19	None	< 2.2e-16	< 2.2e-16
	Specific conductance (SC)	µS/cm at 25 °C	7.4/186.2/37 392.1	Box Cox	< 2.2e-16	4.7e-04
	Calcium (Ca)	mg/L	0.01/19.33/535.49	Box Cox	< 2.2e-16	3.7e-08
	Magnesium (Mg)	mg/L	0.01/4.96/2248.99	Box Cox	< 2.2e-16	7.9e-04
	Potassium (K)	mg/L	0.01/0.99/279.17	Box Cox	< 2.2e-16	2.0e-09
	Sodium (Na)	mg/L	0.01/5.63/14804.91	Box Cox	< 2.2e-16	1.5e-07
	Sulfate (SO_4)	mg/L	0.02/4.81/16971.23	Box Cox	< 2.2e-16	4.9e-12

	Soluble reactive phosphorus (SRP)	µg/L	0.5/4.7/4066.1	Box Cox	< 2.2e-16	5.7e-11
	Total phosphorus (TP)	µg/L	1.4/18.2/10053.0	Box Cox	< 2.2e-16	1.6e-07
	Total nitrogen (TN)	µg/L	0.03/0.2/4.4	Box Cox	< 2.2e-16	1.1e-06
	Chlorophyll *a* (Chl *a*)	mg/L	0.05/2.89/381.49	Box Cox	< 2.2e-16	1.4e-06
	Dissolved organic carbon (DOC)	mg/L	0.12/8.17/220.22	Box Cox	< 2.2e-16	6.2e-05
	Dissolved inorganic carbon (DIC)	mg/L	0.07/15.92/879.03	Box Cox	< 2.2e-16	2.3e-05
Lake physiography	Lake altitude	m	2.0/338.5/1555	Square root	< 2.2e-16	9.5e-07
	Lake circularity	-	0.003/0.31/0.93	Box Cox	3.5e-13	1.1e-03
	Lake dynamic ratio	-	0.03/0.17/6.65	Box Cox	< 2.2e-16	6.0e-03
	Shoreline slope	°	0.09/2.94/25.10	Box Cox	< 2.2e-16	0.07

To visualize the correlations among lake water quality variables and similarities among ecozones, we conducted a principal component analysis (PCA) using the *rda* function of the R package *vegan* (Oksanen et al. 2019). All response variables were transformed (Table 1), then standardized (scaled to zero mean and unit variance) prior to performing the PCA. LULC variables were fitted passively on the PCA to illustrate their association with water quality parameters. We used the broken-stick method (Jackson 1993) to determine the PC axes that accounted for a significant portion of the variation observed. Since most of the variation was explained by PC axis 1, we then used an analysis of variance (ANOVA) approach to quantify the observed variability in axis scores between ecozones and post-hoc Tukey's tests to determine significant differences.

With the goal of exploring the relationships between different LULC and water quality parameters, we then performed a multivariate canonical ordination (i.e., a redundancy analysis [RDA]) using the *rda* function of the R package *vegan* (Oksanen et al. 2019). We set the LULC proportions as predictors and water physiochemistry measurements as the response variables. RDAs assume linear relationships between explanatory and response variables. Therefore, the RDA was performed using the transformed variables (Table 1). All response variables were then standardized (scaled to zero mean and unit variance) prior to performing RDA and significant predictors were selected using forward selection (Legendre and Legendre 2012). The omitted LULC variables (Natural landscapes and Mines) were also fitted passively on the RDA to illustrate their association with water quality parameters. Since most of the variation was explained by RDA axis 1, we used the same ANOVA and Tukey's tests approach as for the PCA to quantify variation among ecozones.

To disentangle the added influence of lake and watershed physiography on these relationships, we performed a variation partitioning using the *varpart* function of the R package *vegan* (Oksanen et al. 2019). That is, we considered the unique vs. combined proportion of variation in lake water quality that land use as well as lake and watershed physiography explained, where altitude, lake circularity, shoreline slope, and the dynamic ratio were bundled into the physiography metric.

Finally, to examine potential nonlinear relationships that were not captured by the linear RDA models, we first investigated the direct relationships between the raw proportion of agriculture in the watershed and three commonly used indicators of water quality (Chl-*a*, TP, and TN; all log transformed for interpretability) using generalized additive models (GAMs). We chose GAMs due to the hypothesized non-linearity of these relationships in ecozones with either low or high human impact. We also performed a multivariate regression tree (MRT) using the R package *mvpart* (Therneau and Atkinson 2014) to test for nonlinear relationships and threshold responses (Ouelette and Legendre 2013) using all response variables (transformed and standardized environmental data). We used the cross-validation relative error (CVRE), the ratio of the variation unexplained by the model to the total variation in the response, to select the most parsimonious tree (i.e., the tree with the least number of nodes whose CVRE value was within one standard error of the tree with the lowest CVRE; De'Ath 2009). As an alternative way to calculate specific threshold LULC proportions, we also applied a cascade MRT (Ouellette et al. 2012), which permitted us to impose a nested structure on the analysis and partition the explanatory power of two independent datasets, a 664 × 7 matrix of raw LULC percentages and a 664 × 1 ecozone vector, to identify regional variability in the relationship between LULC and

water quality (see Supplementary Fig. S2). Essentially, the model computes three trees in the following order: (1) it initially separates all the sites according to ecozone, highlighting significant geographic discrepancies in lake water quality; (2) it then uses the environmental data from only one branch (i.e., group of ecozones) of the first tree to separate those sites with respect to LULC proportions; and (3) it finally uses the environmental data from the other branch to separate those sites with respect to LULC proportions.

Results

Local to regional variation in water quality, lake physiography, and LULC

We found significant variation in lake water quality, lake physiography, and LULC among study lakes and ecozones. A PCA of the lake water chemistry data ordinated the various sites into a single significant PC axis (PC1) explaining 51.2% of the variance, with non-significant PC axis 2 (PC2) explaining an additional 12.9% of the variance.

Ions (Na^+, K^+, Ca^{2+}, Mg^{2+}, and SO_4^{2-}), specific conductance (SC), and dissolved inorganic carbon (DIC) primarily loaded on PC1 and were associated with watersheds with more urban development and agriculture (Fig. 2a). TN, TP, DOC, soluble reactive phosphorus (SRP), and Chl-*a* also loaded strongly on PC1, but were associated with higher proportions of agriculture and pasture (Fig. 2a). PC2 distinguished sites with a higher proportion of natural landscapes that had lower TP, TN, DOC and DIC but greater Secchi disk depths (Fig. 2a). Sites in the Prairies and Boreal Plains (negative PC1 loadings), tended to differ from sites in the Mixedwood Plains or the Atlantic and Pacific Maritimes (tendency for positive PC1 loadings; Fig. 2b). This difference was supported by a one-way ANOVA which showed that lakes within the Prairies and

Boreal Plains ecozones had significantly different PC1 loadings than all other ecozones, distinguishing them from lakes across the rest of Canada ($F_{(11, 652)} = 103.5$, $p < 0.001$; Tukey HSD: $p < 0.05$; Fig. 2c).

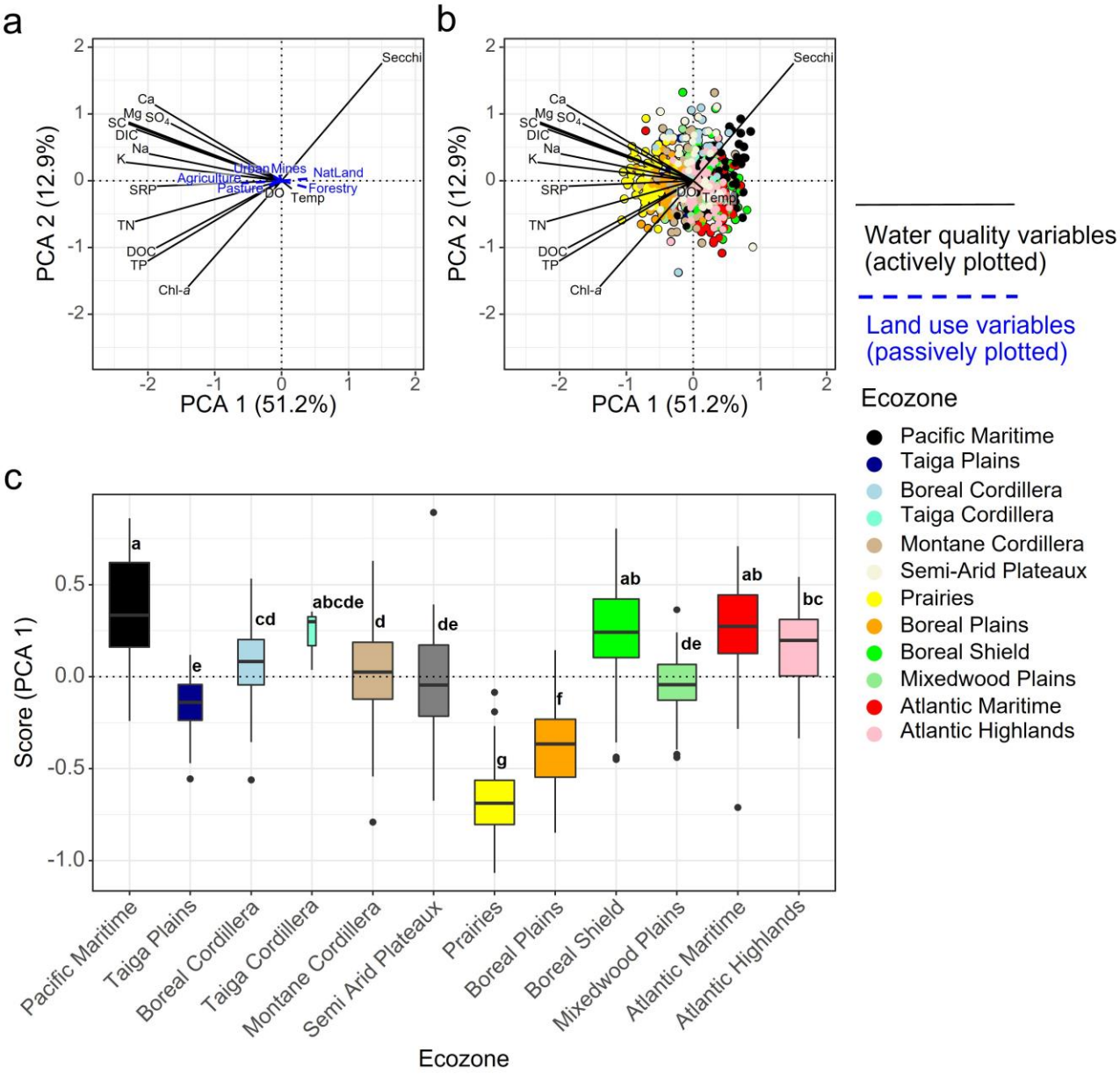


Figure 2. Principal component analysis of lake variables from the LakePulse dataset of Canadian lakes (N = 664). The PCA are plotted twice to highlight the distribution of a) vectors of LULC types plotted passively (colored in blue, dashed lines) and b) site scores, color-coded according to ecozone. c) Boxplot of site scores from PC axis 1 grouped and color-coded by ecozone, where the width of the boxes reflects sample size ranging from 3 to 91 lakes. Letter superscripts denote significantly different groups determined via a post-hoc Tukey's test.

Agriculture and pasture explain the most variation in water quality

The redundancy analyses with LULC proportions as the explanatory variables revealed similar gradients as the PCA, with an important difference: the resulting vectors for agriculture and urbanization were orthogonal in the RDA. The analysis ordinated the various sites into three significant axes that explain 23.7% of the variance. Axes 1, 2 and 3 explained 21.1%, 1.9% and 0.7% of the variance, respectively. The proportion of agriculture or pasture were found to have strong loadings on RDA axis 1, and were correlated with TN, TP, DIC, DOC, SRP, Chl-*a*, all ion concentrations, and hypolimnetic oxygen (DO), whereas Secchi disk depth had a negative RDA axis 1 loading and was associated with forestry activity and natural landscapes (Fig. 3a). In contrast, urbanization was positively associated with RDA axis 2 and correlated with higher epilimnion temperatures and sodium, but lower hypolimnetic oxygen concentrations (Fig. 3a). The one-way ANOVA of RDA 1 scores revealed that sites in the Prairies and Boreal Plains, strongly associated with agriculture and pasture, differed significantly from other ecozones ($F_{(11, 652)} = 99.4$, $p < 0.001$; Tukey HSD: $p < 0.05$; Fig. 3b).

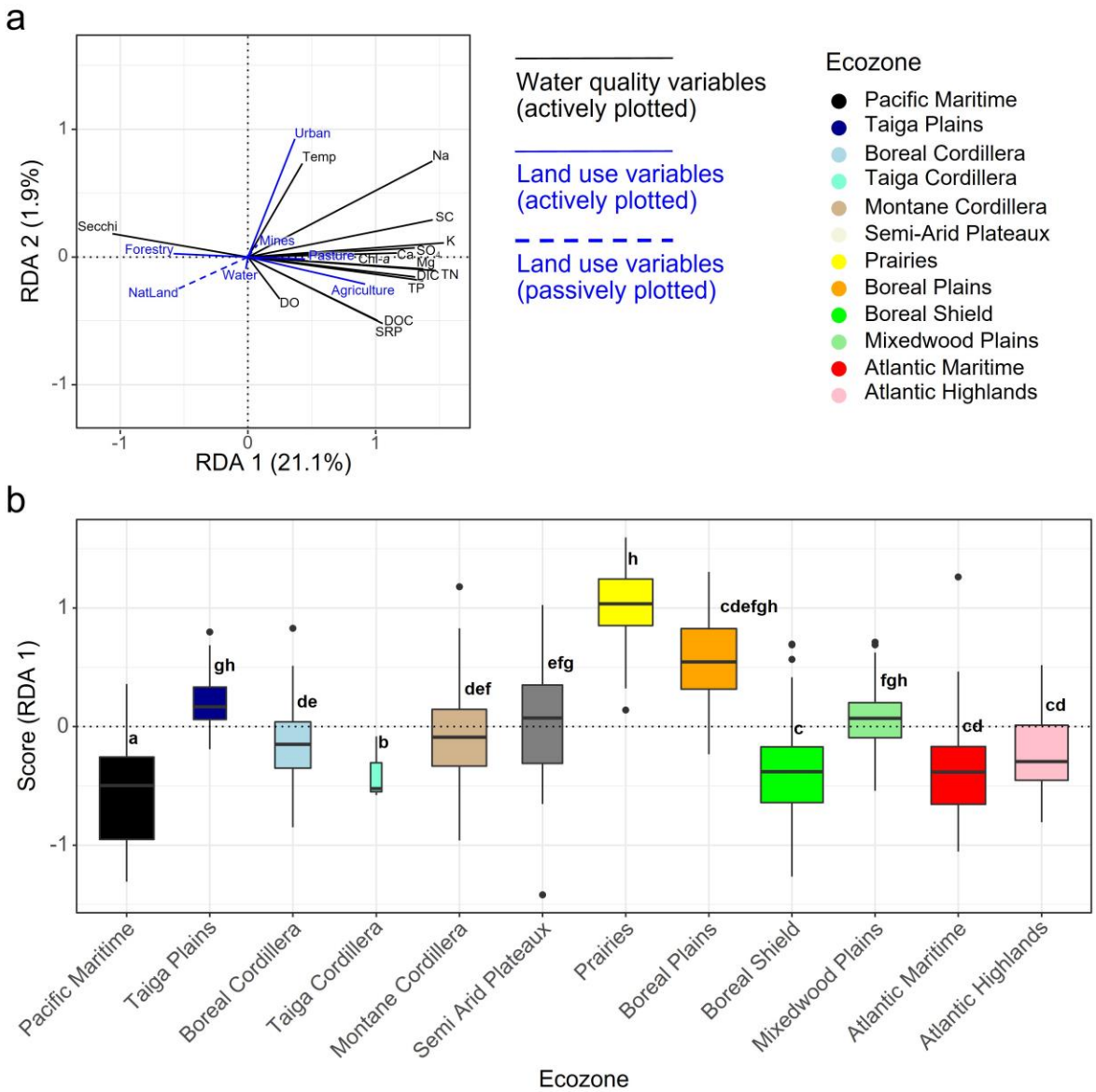


Figure 3. Redundancy analysis examining the relationships between water quality variables (response variables) and LULC metrics (predictors; N = 664). Response variables were transformed and standardized prior to analysis. The analyses are divided into two plots to highlight the distribution of a) vectors of LULC variables (colored in blue), and b) boxplot of site scores for RDA axis 1 grouped by ecozone, where the width of the boxes reflects sample size ranging from 3 to 91 lakes. Letter superscripts denote significantly different groups determined via a post-hoc Tukey's test.

The variation partitioning including altitude, lake circularity, shoreline slope, and the dynamic ratio as covariables describing lake and watershed physiography further emphasized the importance of LULC as a predictor of lake water quality. LULC uniquely explained a total 15% of the variation in water quality, physiography uniquely explained 11%, while 9% of the variation was co-explained by both groups of variables (Fig. S1).

Variable thresholds and plateau responses in the Canadian Prairies

GAMs showed that agricultural land use had varying effects on lake water quality indicators (Chl-*a*, TP, and TN concentrations) depending on ecozone (i.e., local bedrock geology, lake and watershed physiography). Mean values for TP and TN in watersheds with little to no agriculture (intercepts) were relatively higher for the Prairies than for the Mixedwood Plains and Boreal Shield ecozones (Fig. 4), and the response to greater proportions of agriculture in the watershed tended to plateau in the Prairies and Boreal Plains relative to trends observed in other ecozones.

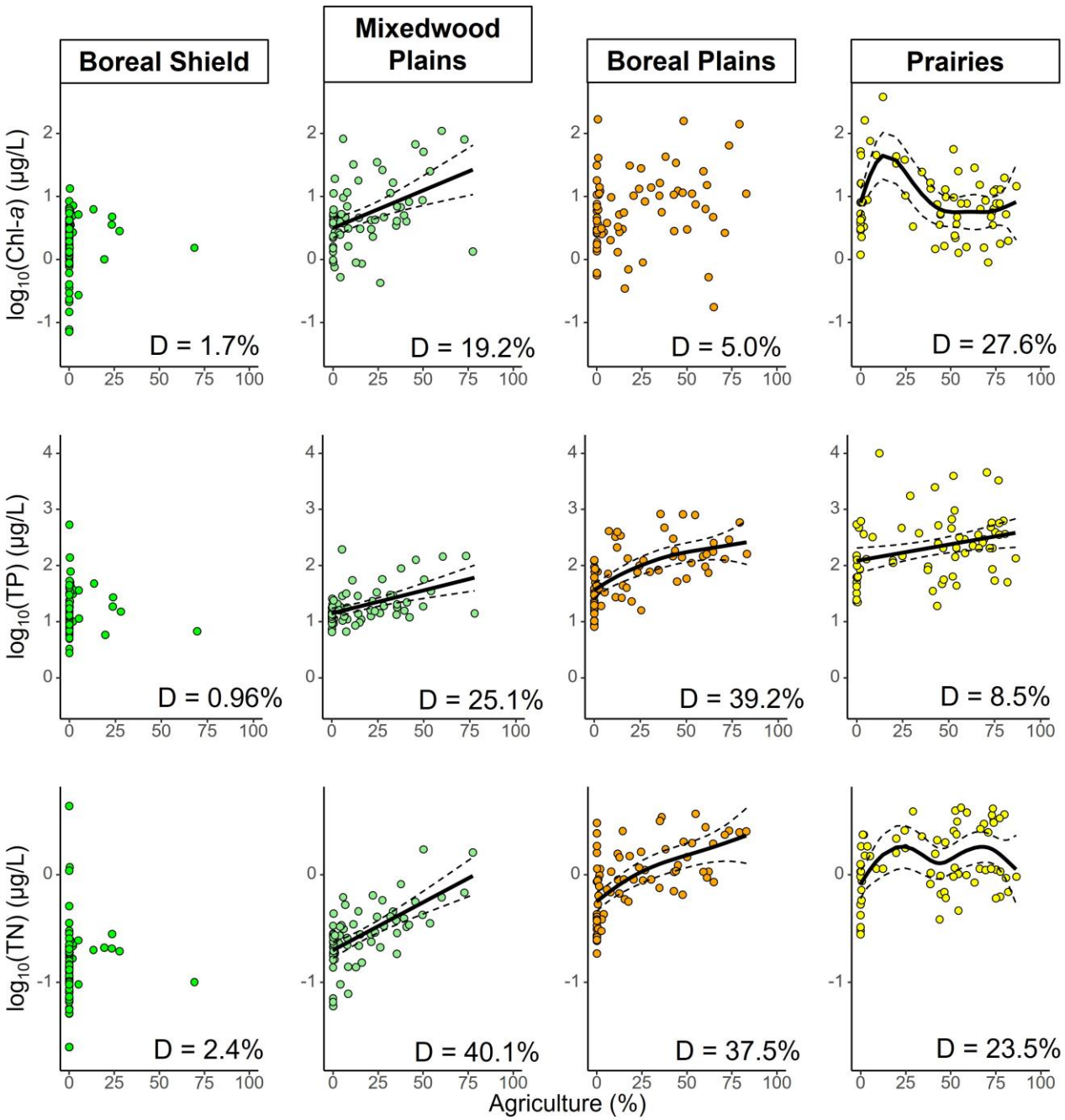


Figure 4. GAMs of the relationships between % Agriculture and total phosphorous (TP), total nitrogen (TP), and chlorophyll a (Chl-*a*) concentration in four populous ecozones along an increasing human impact gradient. Dashed lines represent the 95% confidence intervals. Sites are colored according to ecozone. Deviation explained by each model (D) is shown at the bottom right of each figure.

The multivariate regression tree analysis identified threshold values for agriculture and urbanization where we could expect pronounced changes in lake water quality across the country (Fig. 5). Across Canada, the presence of agriculture (>0.1% of the watershed used) was identified as the most important predictor of water quality, whereby sites with minimal agriculture and urbanization had higher than average (i.e., the baseline within each branch of the MRT) Secchi depths, lower than average concentrations of ions, nutrients, and lower epilimnetic temperatures (Fig. 5). Among these sites, those with some urban land use (i.e., >1.3%) showed ions, nutrients and water temperatures that were closer to the national average (Fig. 5). Water quality in sites with moderate agriculture (i.e., between 0.2% and 36%) was similar to that of sites with some urban land use, albeit with slightly higher than average values as opposed to slightly lower than average (Fig. 5). Sites with substantial agriculture (i.e., >36%) had significantly higher than average ion and nutrient concentrations while Secchi depths were considerably lower than average (Fig. 5).

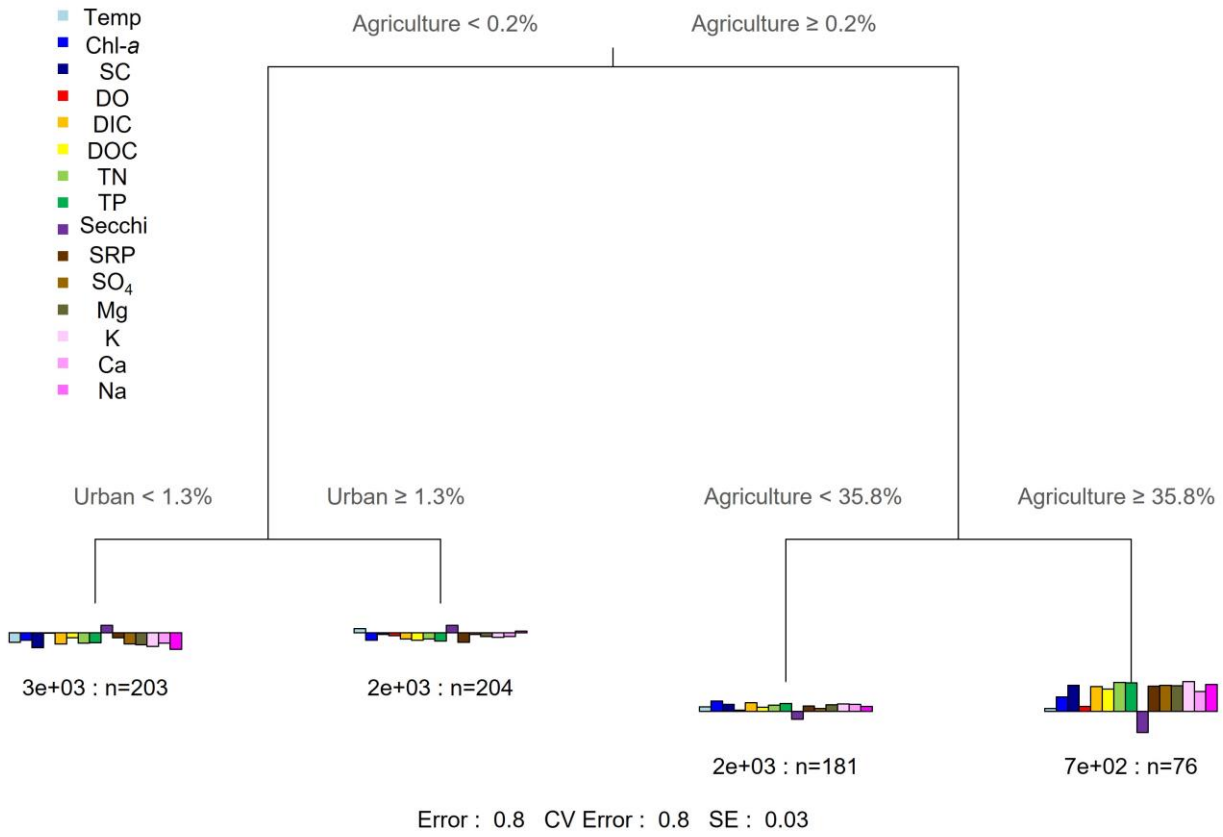


Figure 5. Multivariate regression tree (MRT) of the transformed and standardized water quality variables, constrained by the LULC variables (R^2 = 0.20; N = 664).

The cascade regression tree (cMRT), which was performed to consider potential regional differences, clearly identified the difference in lakes of the Prairie ecozone from other regions. The cMRT also increased the proportion of variation explained slightly (R^2 = 0.29; Fig. S2), relative to the MRT, but identified similar thresholds in land use proportions. In the Prairies, water quality varied slightly by the presence of agriculture in the watershed, since catchments with even negligible agriculture (i.e., >0.2%) had ion and nutrient concentrations above the national average (Fig. S2). In other ecozones, water quality parameters only reached above

average values when the proportion of agricultural land surpassed 30%, with lakes maintaining relatively low ion and nutrient levels below that threshold (Fig. S2). Additionally, water quality in catchments with less than 0.7% agriculture varied mainly by their proportion of urbanization, with sites over 6.6% urban land showing higher epilimnetic temperatures and specific conductance, but lower hypolimnetic oxygen compared to the national average as well as their less urbanized counterparts (Fig. S2).

Discussion

Our analyses of 664 lakes from twelve ecozones, spread across a total study area of 412 052 km^2, revealed considerable variation in lake water quality parameters across the nation, with the Canadian Prairies showing significantly different lake water quality compared to the rest of Canada. The higher proportions of agriculture and pasture in the Prairies suggest an important link (Figs. 2 and 3), although it is also clear that other factors (see below) are at play given the heightened baseline nutrient values for this ecozone (Fig. 4). Indeed, both linear and nonlinear multivariate regression methods identified agriculture and urbanization as the main land use drivers of lake water quality across a wide geographic gradient (Figs. 3, 5 and S2). Variation partitioning with lake and watershed physiography factors as covariates demonstrated that this relationship was modulated by geological and topographical factors inherently correlated with LULC (Fig. S1). Multivariate regression tree models also identified the thresholds at which agricultural and urban land use are associated with poorer water quality and highlighted differences between the Prairies and other ecozones (Figs. 5 and S2). Furthermore, we found that nonlinear models explain more variation in water quality than linear models, which underscores

the importance of further examining ecological causes for nonlinearity in the relationship between LULC and water quality metrics.

LULC, physiography and lake water quality vary among Canadian ecozones

We observed that most lakes in the Prairies, and at least some in the Boreal Plains, had significantly different water quality values relative to the rest of the country. Specifically, lakes in the Prairies and Boreal Plains had higher than average concentrations of TP, TN, major ions, dissolved carbon and algal biomass (inferred from Chl-*a*), and lower than average Secchi depths and hypolimnetic oxygen, all of which contributed to substantial variation in the national dataset (Figs. 2b and 2c). These marked discrepancies in water quality for the Prairies and Boreal Plains ecozones are likely the result of a higher local prevalence of agriculture and a difference in baseline nutrient conditions (Prairies TP_{median} = 204.1 µg/L, Canada TP_{median} = 18.2 µg/L; Fig. 4; Taranu and Gregory-Eaves 2008). There are multiple factors that might contribute to this regional discrepancy in the response of lakes to LULC development. For one, the region's flat catchment topography (Bharath and Elshorbagy 2018), naturally nutrient-rich geology (Taranu and Gregory-Eaves 2008; Griffiths et al. 2022) and its changing hydrology (i.e., declining river flows due to a reduction in yearly persistence and maximum depth of local snowpacks; Schindler and Donahue 2006) may further modulate the effect of agricultural development (Bunting et al. 2016) and urban expansion (Leavitt et al. 2006). Indeed, the combined effects of increasing agricultural and urban development and the Prairie region's flat, nutrient-rich landscape on lake water quality were supported by our variation partitioning analysis, where LULC explained a considerable amount of the variation (15%) in water quality alongside lake physiography (11%) and their interaction (9%; Fig. S1). There is also a well-known tendency for eutrophication to

occur in shallow lakes set in flat catchments, and this coupled with the natural productivity of the Prairies likely modulated the relationship between LULC and lake water quality. Consistent with these findings, we found that TP, TN, and DOC were higher than the national average, and that DO was lower than average in the LakePulse Prairie sites. In addition, climatic variability (e.g., Chinook winds locally and droughts regionally; see Lemmen et al. 1997) has been shown to profoundly affect continental lakes (Schindler and Donahue 2006), leading to variations in water quality across different spatial and temporal scales (Pham et al. 2009). However, several studies have also shown that internal lake characteristics (i.e., limiting nutrients, oxygen content, DOC) can have an even larger influence on the water quality of lakes in this region relative to that of climate and meteorology (Bunting et al. 2016; Vogt et al. 2018).

There is a growing body of evidence that has identified that long-term anthropogenic impacts may have disproportionately altered the productivity and water physiochemistry of Prairie lakes relative to other ecozones, leading to the present-day dichotomy in baseline nutrient levels and water quality. For example, time-series analyses revealed widespread increases in algal biomass since pre-industrial times (Taranu et al. 2015; Ho et al. 2019; Griffiths et al. 2022), and indicate that in highly developed and economically important ecozones, such as the Prairies and the Mixedwood Plains, these trends have accentuated. Furthermore, analyses of bioindicators preserved in lake sediment records have shown that the highest degree of taxonomic turnover since pre-industrial conditions have been localized in ecozones that currently have the more intense land uses, with agriculture and urbanization being key stressors to those lakes (Griffiths et al. 2021; Griffiths et al. 2022).

Land use thresholds for lake water quality changes

With nonlinear modelling methods, we identified land use thresholds that lead to significant changes in water quality. The initial branch of both the MRT and the cMRT separated sites according to whether agriculture was present in the watershed, suggesting that agricultural land use was the main determinant of water quality across all sites. Interestingly, nested analyses suggested that subsequent inflection points differed between the Prairies sites and the rest of Canada. Outside of the Prairies, the regression trees identified an agricultural threshold of 30–35% after which a large decrease in water quality was observed (i.e., increase in limiting nutrients, ions, carbon, and algal biomass (Figs. 4 and S2). Water quality across all sites was poorer when even a minute proportion of the catchment was devoted to agricultural activities (0.2% in the Prairies, 0.8% in the rest of Canada; Fig. S2). However, the differences in water quality were much more pronounced after surpassing the 30–35% agriculture threshold across Canada than when exceeding the 0.2% threshold in the Prairies ecozone (Fig. S2), suggesting a lessened effect of additional nutrient input on algal biomass in eutrophic and hypereutrophic Prairie lakes. Many lakes in this region were historically productive, being at least mesotrophic (Lemmen et al. 1997; Bunting et al. 2016), which might partially explain why the change in water quality is notably less severe in the Prairies, where baseline nutrients are already high.

We also identified a national urbanization threshold of 1–7%, after which we observed altered water quality though increases in sodium concentrations and epilimnion temperatures (Figs. 4 and S2). Since roads were classified as urban land use (Supplemental Table S1), the observed increases in sodium concentrations might be due to various factors such as the widespread practice of road salt application, leaky septic fields, deteriorating infrastructure (Gagnon et al.

2008; CIRC 2019), or a combination of these factors. Increases in epilimnetic temperature are consistent with past studies on examining the effect of urbanization on lake water temperature (LeBlanc et al. 1997; Yang et al. 2020), although climate change is also impacting lake temperatures globally (O'Reilly et al. 2015; Woolway et al. 2020). Additionally, the association between urbanization and epiliminetic temperatures could be related to geography as Canadian municipalities are concentrated in the south (and we observed a significantly negative correlation between latitude and percent urban land use in our dataset).

Across the body of literature that quantified land use threshold values in association with lake water quality, there is substantial heterogeneity. For example, Qiu and Turner (2015) found an agricultural threshold of 13% using data from a Wisconsin watershed, Baker (2006) reported an evident reduction in water quality at a threshold of 5% urban land use when using a global array of studies. It is possible that the climatic and geomorphological context modifies the influence of land use globally and could help reconcile these seemingly disparate findings, as we have seen in our Canadian dataset. Further refinement of these models could be performed with a closer examination of the type of agriculture (e.g., traditional vs. organic farming, type of crop) and urbanization across different watersheds, to help inform policy development and to provide more data and incentives for farmers.

Conclusions

Through the application of regression models to an expansive dataset of limnology and LULC proportions, we have shown that lake water quality responds differently to land use across regions, which can be explained by geology, physiography, land use history, and climate.

Nonetheless, with further expansion in the amount of land converted to agricultural activities due to population growth and increasing uncertainty with regards to global food security, we would predict declines in lake water quality. Looking forward, initiatives such as organic and vertical farming might pave the way for more space-efficient agriculture that results in less run-off of nutrients to downstream water bodies (Viana et al. 2022). Since organic farms incur a considerable risk of soil P depletion (Carr et al. 2019), recycling anthropogenic P waste is also being considered as a possible alternative source of nutrients (Nicksy and Entz 2021). Land management strategies must take these particularities into account if we are to preserve crucial ecosystem services such as food production and potable water access worldwide and beyond the foreseeable future.

Acknowledgements

The authors thank the LakePulse data collection and data curating teams for their provision of the data, Maxime Fradette at Université de Sherbrooke for his invaluable GIS expertise, as well as Katherine Griffiths, Alexandre Baud, Candice Aulard, David Zilkey, Rebecca Garner, and all the Gregory-Eaves lab for their continued support and insightful advice. This research was funded by McGill University's BIPOC Research Award (J.R.S.S.), the NSERC National Lake Pulse Program (funded to Y.H & I.G-E., among others) and a Canada Research Chair grant (I.G-E.).

Declaration of interest statement

The authors report there are no competing interests to declare.

References

Abirhire O, North RL, Hunter K, Vandergucht D, Hudson J. 2016. Do human activities affect phytoplankton biomass and composition in embayments on Lake Diefenbaker? J Freshw Ecology. 31(4):597–608.

Almanza V, Pedreros P, Laughinghouse HD IV, Félez J, Parra O, Azócar M, Urrutia R. 2019. Association between trophic state, watershed use, and blooms of cyanobacteria in south-central Chile. Limnologica. 75:30–41.

Baker A. 2006. Encyclopedia of Hydrological Sciences. New York (NY): Wiley. 188, Land Use and Water Quality; p. 1–3.

Beusen AH, Bouwman AF, Van Beek LP, Mogollón JM, Middelburg JJ. 2016. Global riverine N and P transport to ocean increased during the 20th century despite increased retention along the aquatic continuum. Biogeosciences. 13(8):2441–2451.

Bharath R, Elshorbagy A. 2018. Flood mapping under uncertainty: a case study in the Canadian prairies. Nat Hazards. 94(2):537–60.

Bunting L, Leavitt PR, Simpson GL, Wissel B, Laird KR, Cumming BF, St. Amand A, Engstrom DR. 2016. Increased variability and sudden ecosystem state change in Lake Winnipeg, Canada, caused by 20th century agriculture. Limnol Oceanogr. 61(6):2090–2107.

[CIRC] Canadian Infrastructure Report Card. 2019. Monitoring the State of Canada's Core Public Infrastructure: The Canadian Infrastructure Report Card 2019. [accessed 2022 Mar 13]. http://canadainfrastructure.ca/.

Carpenter SR, Caraco NF, Correll DL, Howarth RW, Sharpley AN, Smith VH. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecol Appl. 8(3):559–568.

Carr PM, Cavigelli MA, Darby H, Delate K, Eberly JO, Gramig GG, Heckman JR, Mallory EB, Reeve JR, Silva EM, Suchoff DH. 2019. Nutrient cycling in organic field crops in Canada and the United States. Agron J. 111(6):2769–2785.

Cassman KG, Dobermann A. 2022. Nitrogen and the future of agriculture: 20 years on. Ambio. 51(1):17–24.

Collins SM, Yuan S, Tan PN, Oliver SK, Lapierre JF, Cheruvelil KS, Fergus CE, Skaff NK, Stachelek J, Wagner T, Soranno PA. 2019. Winter precipitation and summer temperature predict lake water quality at macroscales. Water Resour Res. 55(4):2708–2721.

De'Ath G. 2009. Multivariate regression trees: A new technique for modeling species–environment relationships. Ecology. 83:1105–1117.

de Mello K, Taniwaki RH, de Paula FR, Valente RA, Randhir TO, Macedo DR, Leal CG, Rodrigues CB, Hughes RM. 2020. Multiscale land use impacts on water quality: assessment, planning, and future perspectives in Brazil. J Environ Manag. 270:110879.

Dillon PJ, Kirchner WB. 1975. The effects of geology and land use on the export of phosphorus from watersheds. Water Res. 9(2):135–148.

Doubek JP, Carey CC, Cardinale BJ. 2015. Anthropogenic land use is associated with N-fixing cyanobacterial dominance in lakes across the continental United States. Aquat Sci. 77(4):681–694.

Faust DR, Kumar S, Archer DW, Hendrickson JR, Kronberg SL, Liebig MA. 2018. Integrated crop-livestock systems and water quality in the northern Great Plains: Review of current practices and future research needs. J Environ Qual. 47(1):1–15.

Ferland ME, del Giorgio PA, Teodoru CR, Prairie YT. 2012. Long-term C accumulation and total C stocks in boreal lakes in northern Québec. Glob Biogeochem Cycles. 26(4).

Gagnon M, Gaudreault V, Overton D. 2008. Age of public infrastructure: A provincial perspective. Ottawa (ON): Statistics Canada.

[GOC] Government of Canada. 2015a Jan 1. Data from: Land Use 2010 [dataset]. Open Government Portal. [accessed 2018 Sep 4]. https://open.canada.ca/data/en/dataset/9e1efe92-e5a3-4f70-b313-68fb1283eadf.

[GOC] Government of Canada. 2015b Jan 1. Data from: Canadian Digital Elevation Model [dataset]. Open Government Portal. [accessed: 2018 Feb 5]. https://open.canada.ca/data/en/dataset/7f245e4d-76c2-4caa-951a-45d1d2051333.

[GOC] Government of Canada. 2016 Jul 11. Data from: Agri-Environmental Indicator (AEI) - Risk of P release in agricultural land (P-Source) - 2011. Open Government Portal. [accessed: 2022 June 1]. https://open.canada.ca/data/en/dataset/fc9e5c73-1c1a-47c1-9de4-612569b718fd.

[GOC] Government of Canada. 2017 Feb 14. Data from: Annual Crop Inventory 2016. Open Government Portal. [accessed: 2018 April 9]. https://open.canada.ca/data/en/dataset/b8e4da73-fb5f-4e6e-93a4-8b1f40d95b51.

[GOC] Government of Canada. 2019. Canadian agricultural partnership. [accessed 2022 April 10]. https://agriculture.canada.ca/en/about-our-department/key-departmental-initiatives/canadian-agricultural-partnership.

[GOC] Government of Canada. 2021. Overview of Canada's agriculture and agri-food sector. [accessed 2022 July 15]. https://agriculture.canada.ca/en/canadas-agriculture-sectors/overview-canadas-agriculture-and-agri-food-sector.

Graham JL, Dubrovsky NM, Foster GM, King LR, Loftin KA, Rosen BH, Stelzer EA. 2020. Cyanotoxin occurrence in large rivers of the United States. Inland Waters. 10(1):109–117.

Griffith MB. 2014. Natural variation and current reference for specific conductivity and major ions in wadeable streams of the conterminous USA. Freshw Sci. 33(1):1–7.

Griffiths K, Jeziorski A, Paquette C, Taranu ZE, Baud A, Antoniades D, Beisner B, Hamilton PB, Smol JP, Gregory-Eaves I. 2021. Multi-trophic level responses to environmental stressors over the past~ 150 years: Insights from a lake-rich region of the world. Ecol Indic. 127:107700.

Griffiths K, Jeziorski A, Antoniades D, Beaulieu M, Smol JP, Gregory-Eaves I. 2022. Pervasive changes in algal indicators since pre-industrial times: a paleolimnological study of changes in primary production and diatom assemblages from ~200 Canadian lakes. Sci Total Environ. 838:155938.

Grimvall A, Stålnacke P, Tonderski A. 2000. Time scales of nutrient losses from land to sea—a European perspective. Ecol Eng. 14(4):363–71.

Håkanson L. 1982. Lake bottom dynamics and morphometry: the dynamic ratio. Water Resour Res. 18(5):1444–1450.

Hassanzadeh E, Strickert G, Morales-Marin L, Noble B, Baulch H, Shupena-Soulodre E, Lindenschmidt KE. 2019. A framework for engaging stakeholders in water quality modeling and management: Application to the Qu'Appelle River Basin, Canada. J Environ Manag. 231:1117–1126.

Ho JC, Michalak AM, Pahlevan N. 2019. Widespread global increase in intense lake phytoplankton blooms since the 1980s. Nature. 574(7780):667–670.

Huot Y, Brown CA, Potvin G, Antoniades D, Baulch HM, Beisner BE, Bélanger S, Brazeau S, Cabana H, Cardille JA et al. 2019. The NSERC Canadian Lake Pulse Network: a national assessment of lake health providing science for water management in a changing climate. Sci Total Environ. 695:133668.

Jackson DA. 1993. Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. Ecology. 74:2204–2214.

Jeppesen E, Søndergaard M, Jensen JP, Havens KE, Anneville O, Carvalho L, Coveney MF, Deneke R, Dokulil MT, Foy BOB, Gerdeaux D. 2005. Lake responses to reduced nutrient loading–an analysis of contemporary long-term data from 35 case studies. Freshw Biol. 50(10):1747–1771.

Kronvang B, Jeppesen E, Conley DJ, Søndergaard M, Larsen SE, Ovesen NB, Carstensen J. 2005. Nutrient pressures and ecological responses to nutrient loading reductions in Danish streams, lakes and coastal waters. J Hydrol. 304(1–4):274–288.

Leavitt PR, Brock CS, Ebel C, Patoine A. 2006. Landscape-scale effects of urban nitrogen on a chain of freshwater lakes in central North America. Limnol Oceanogr. 51(5):2262–2277.

LeBlanc RT, Brown RD, FitzGibbon JE. 1997. Modeling the effects of land use change on the water temperature in unregulated urban streams. J Environ Manag. 49(4):445–469.

Legendre P, Legendre L. 2012. Numerical ecology. Amsterdam (NE): Elsevier.

Lemmen DS, Vance RE, Wolfe SA, Last WM. 1997. Impacts of future climate change on the southern Canadian Prairies: a paleoenvironmental perspective. Geosci Can. 24(3):121–133.

Lencha SM, Ulsido MD, Tränckner J. 2022. Estimating Point and Nonpoint Source Pollutant Flux by Integrating Various Models, a Case Study of the Lake Hawassa Watershed in Ethiopia's Rift Valley Basin. Water. 14(10):1569.

Liu H, Brouwer R. 2022. Incentivizing the future adoption of best management practices on agricultural land to protect water resources: The role of past participation and experiences. Ecol Econ. 196:107389.

MacDonald GK, Bennett EM, Potter PA, Ramankutty N. 2011. Agronomic phosphorus imbalances across the world's croplands. Proc Natl Acad Sci. 108(7):3086–3091.

Mbonde AS, Sitoki L, Kurmayer R. 2015. Phytoplankton composition and microcystin concentrations in open and closed bays of Lake Victoria, Tanzania. Aquat Ecosyst Health Manag. 18(2):212–220.

McCrackin ML, Jones HP, Jones PC, Moreno-Mateos D. 2017. Recovery of lakes and coastal marine ecosystems from eutrophication: A global meta-analysis. Limnol Oceanogr. 62(2):507–18.

Medalie L, Hirsch RM, Archfield SA. 2012. Use of flow-normalization to evaluate nutrient concentration and flux changes in Lake Champlain tributaries, 1990–2009. J Gt Lakes Res. 38:58–67.

Munawar M, Fitzpatrick MAJ. 2018. Eutrophication in three Canadian Areas of Concern: Phytoplankton and major nutrient interactions. Aquat Ecosyst Health Manag. 21(4):421–437.

Natural Earth. 2022 May 12. Data from: 1:50m cultural vectors: Admin 1 – States, provinces [dataset]. Free Vector and Raster Database. [accessed 2022 Jul 3]. https://www.naturalearthdata.com/downloads/50m-cultural-vectors/50m-admin-1-states-provinces/.

Nicksy J, Entz MH. 2021. Recycled nutrients as a phosphorus source for Canadian organic agriculture: a perspective. Can J Soil Sci. 101(4):571–580.

Nielsen A, Trolle D, Søndergaard M, Lauridsen TL, Bjerring R, Olesen JE, Jeppesen E. 2012. Watershed land use effects on lake water quality in Denmark. Ecol Appl. 22(4):1187–1200.

Oksanen J, Guillaume Blanchet F, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, et al. 2020. vegan: community ecology package [software]. R package Ver. 2.5–7; [accessed 2021 May 20]. https://CRAN.R-project.org/package=vegan.

O'Reilly CM, Sharma S, Gray DK, Hampton SE, Read JS, Rowley RJ, Schneider P, Lenters JD, McIntyre PB, Kraemer BM, Weyhenmeyer GA. 2015. Rapid and highly variable warming of lake surface waters around the globe. Geophys Res Lett. 42(24):10–773.

Ouellette MH, Legendre P. 2013. MVPARTwrap: Additional features for package mvpart [software]. R package Ver. 0.1–9.2; [accessed 2021 Aug 19]. https://cran.r-project.org/package=MVPARTwrap.

Ouellette MH, Legendre P, Borcard D. 2012. Cascade multivariate regression tree: a novel approach for modelling nested explanatory sets. Methods Ecol Evol. 3(2):234244.

Özkundakci D, Hamilton DP, Trolle D. 2011. Modelling the response of a highly eutrophic lake to reductions in external and internal nutrient loading. N Z J Marine Freshw Res. 45(2):165–85.

Pham SV, Leavitt PR, McGowan S, Wissel B, Wassenaar LI. 2009. Spatial and temporal variability of prairie lake hydrology as revealed using stable isotopes of hydrogen and oxygen. Limnol Oceanogr. 54(1):101–118.

Pérez-Jvostov F, Sutherland WJ, Barrett RD, Brown CA, Cardille JA, Cooke SJ, Cristescu ME, St-Gelais NF, Fussmann GF, Griffiths K et al. 2020. Horizon scan of conservation issues for inland waters in Canada. Can J Fish Aquat Sci. 77(5):869–81.

QGIS.org. 2022. QGIS Geographic Information System. QGIS Association [software]. Ver. 3.0; [accessed 2022 Jun 20]. http://www.qgis.org/.

Qiu J, Turner MG. 2015. Importance of landscape heterogeneity in sustaining hydrologic ecosystem services in an agricultural watershed. Ecosphere. 6(11):1–19.

R Development Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing [software]. Ver. 4.0.0; [accessed 2021 May 20]. https://www.r-project.org/.

Read EK, Patil VP, Oliver SK, Hetherington AL, Brentrup JA, Zwart JA, Winters KM, Corman JR, Nodine ER, Woolway RI et al. 2015. The importance of lake-specific characteristics for water quality across the continental United States. Ecol Appl. 25(4):943–955.

Reid AJ, Carlson AK, Creed IF, Eliason EJ, Gell PA, Johnson PT, Kidd KA, MacCormack TJ, Olden JD, Ormerod SJ et al. 2019. Emerging threats and persistent conservation challenges for freshwater biodiversity. Biol Rev. 94(3):849–873.

Ribeiro PJ Jr, Diggle PJ, Schlather M, Bivand R, Ripley B. 2020. geoR: Analysis of geostatistical data [software]. R package Ver. 1.8–1; [accessed 2021 May 20]. https://cran.r-project.org/package=geoR.

Schindler DW, Donahue WF. 2006. An impending water crisis in Canada's western prairie provinces. Proc Natl Acad Sci. 103(19):7210–7216.

Spiess E. 2011. Nitrogen, phosphorus and potassium balances and cycles of Swiss agriculture from 1975 to 2008. Nutr Cycl Agroecosys. 91(3):351–65.

Taranu ZE, Gregory-Eaves I. 2008. Quantifying relationships among phosphorus, agriculture, and lake depth at an inter-regional scale. Ecosystems. 11(5):715–725.

Taranu ZE, Gregory-Eaves I, Leavitt PR, Bunting L, Buchaca T, Catalan J, Domaizon I, Guilizzoni P, Lami A, McGowan S. 2015. Acceleration of cyanobacterial dominance in north temperate- subarctic lakes during the Anthropocene. Ecol Lett. 18(4):375–384.

Therneau TM, Atkinson B. 2014. mvpart: Multivariate partitioning [software]. R package Ver. 1.6–2; [accessed 19 August 2021]. https://CRAN.R-project.org/package=mvpart.

Van Meter KJ, Basu NB, Veenstra JJ, Burras CL. 2016. The nitrogen legacy: emerging evidence of nitrogen accumulation in anthropogenic landscapes. Environ Res Lett. 11(3):035014.

Viana LR, Dessureault PL, Marty C, Loubet P, Levasseur A, Boucher JF, Paré MC. 2022. Would transitioning from conventional to organic oat grains production reduce environmental impacts? A LCA case study in North-East Canada. J Clean Prod. 349:131344.

Vogt RJ, Sharma S, Leavitt PR. 2018. Direct and interactive effects of climate, meteorology, river hydrology, and lake characteristics on water quality in productive lakes of the Canadian Prairies. Can J Fish Aquat Sci. 75(1):47–59.

Wetzel RG. 2001. Limnology: lake and river ecosystems. San Diego (CA): Elsevier.

Woolway RI, Kraemer BM, Lenters JD, Merchant CJ, O'Reilly CM, Sharma S. 2020. Global lake responses to climate change. Nat Rev Earth Environ. 1(8):388–403.

Wunderlich SM, Martinez NM. 2018. Conserving natural resources through food loss reduction: Production and consumption stages of the food supply chain. Int Soil Water Conserv Res. 6(4):331–339.

Yang K, Yu Z, Luo Y. 2020. Analysis on driving factors of lake surface water temperature for major lakes in Yunnan-Guizhou Plateau. Water Res. 184:116018.

Zammali MK, Hassanzadeh E, Shupena-Soulodre E, Lindenschmidt KE. 2021. A generic approach to evaluate costs and effectiveness of agricultural Beneficial Management Practices to improve water quality management. J Environ Manag. 287:112336.

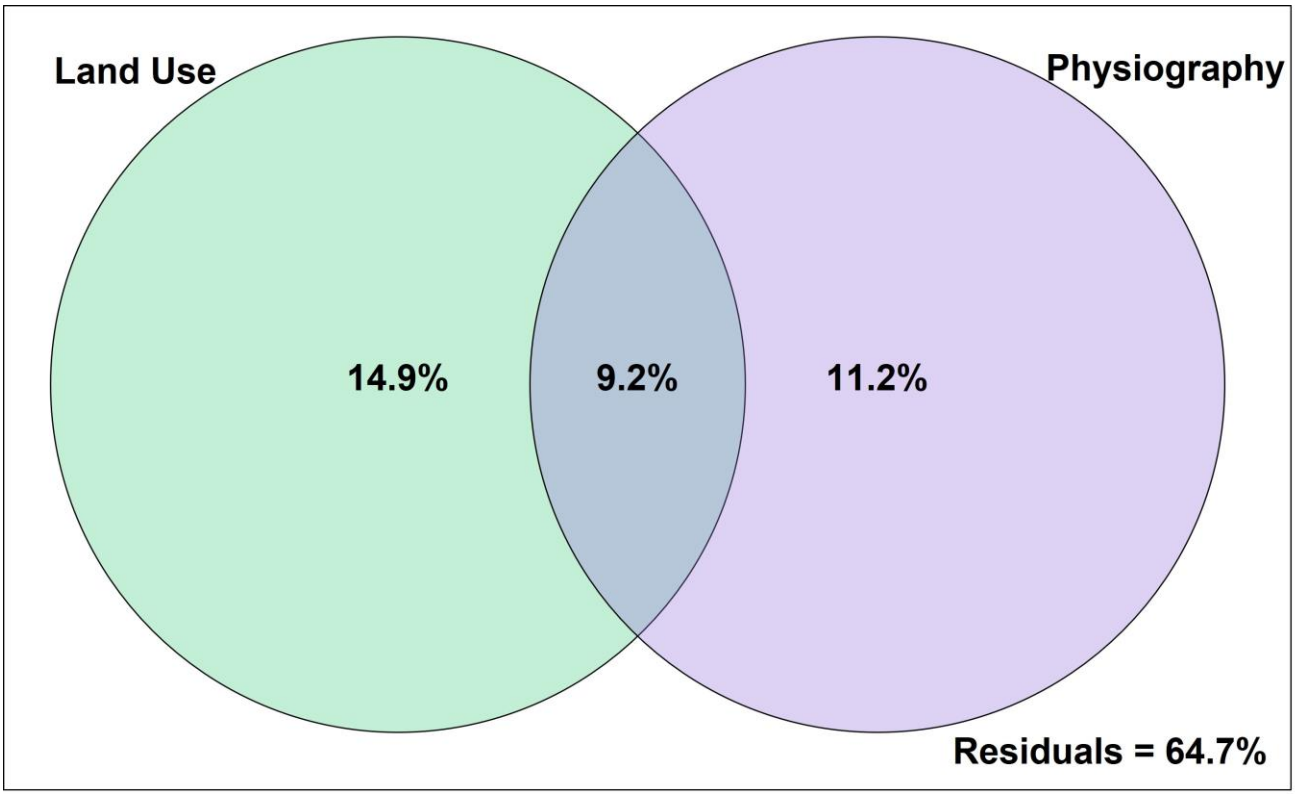
Appendix A: Supplemental Table

Supplemental Table S1. Annual Space-Based Crop Inventory for Canada 2016 and Land Use 2010 database categories and corresponding simplified categories of the various LULC types. Simplified categories were used in this article's analyses.

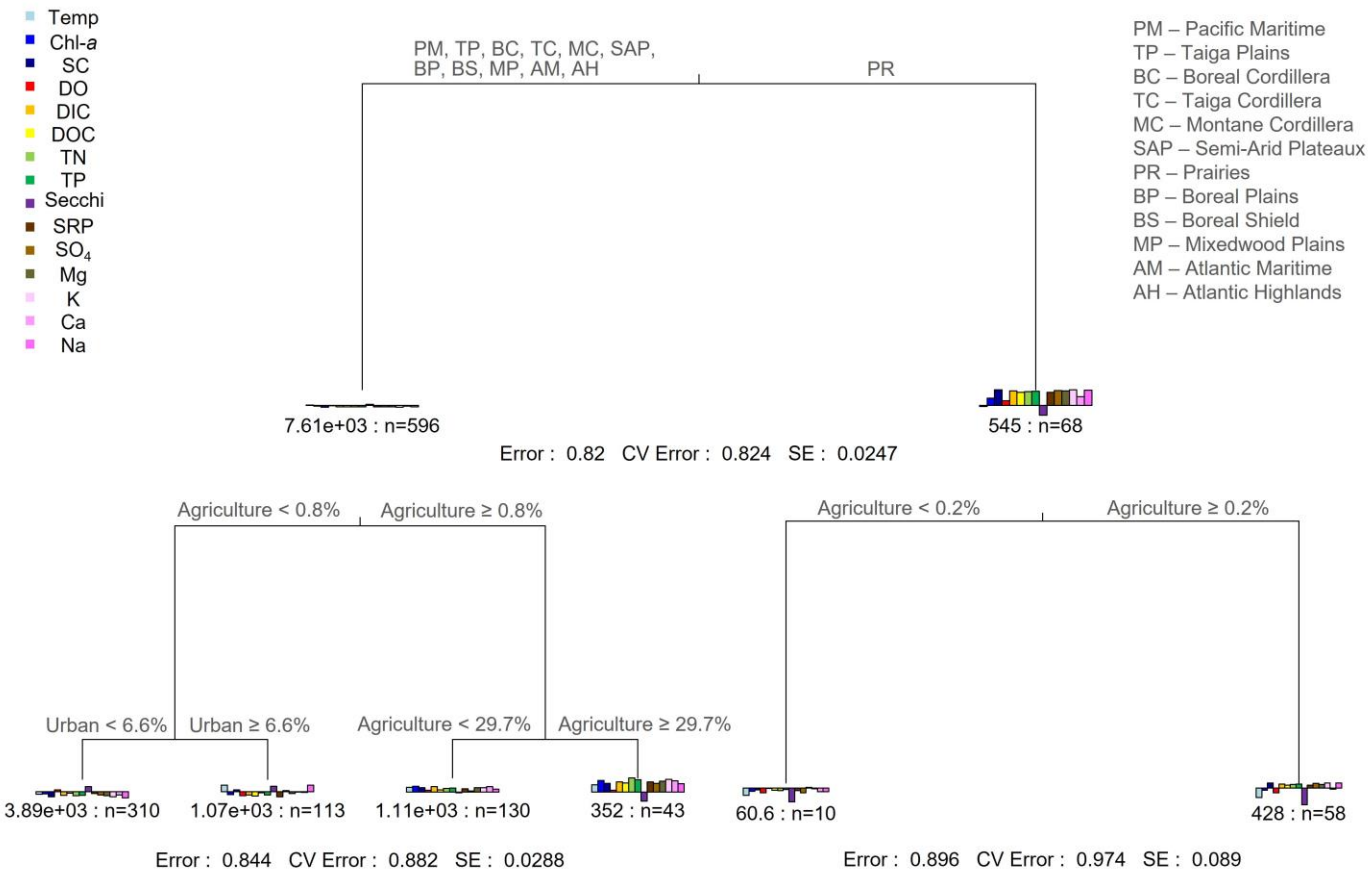
Definition	Simplified category
No Data	NA
Water	Water
Exposed Land / Barren / Rock / Snow / Ice	Natural Landscape
Natural landscape (unclassified)	Natural Landscape
Bryoids	Natural Landscape
Shrubland	Natural Landscape
Herb / Grassland	Natural Landscape
Coniferous forest	Natural Landscape
Broadleaf forest	Natural Landscape
Mixedwood forest	Natural Landscape
Forest (undifferentiated)	Natural Landscape
Wetland	Natural Landscape
Forest Loss (2012-2017)	Forestry
Urban / Developed	Urban
Sod	Urban
Greenhouses	Agriculture
Agriculture (undifferentiated)	Agriculture
Pasture / Forages	Pasture
Too Wet to be Seeded	Agriculture
Fallow	Agriculture
Cereals	Agriculture
Barley	Agriculture
Other Grains	Agriculture
Millet	Agriculture
Oats	Agriculture
Rye	Agriculture
Spelt	Agriculture
Triticale	Agriculture
Wheat	Agriculture
Switchgrass	Agriculture

Sorghum	Agriculture
Winter Wheat	Agriculture
Spring Wheat	Agriculture
Corn	Agriculture
Tobacco	Agriculture
Ginseng	Agriculture
Oilseeds	Agriculture
Borage	Agriculture
Camelina	Agriculture
Canola / Rapeseed	Agriculture
Flaxseed	Agriculture
Mustard	Agriculture
Safflower	Agriculture
Sunflower	Agriculture
Soybeans	Agriculture
Pulses	Agriculture
Peas	Agriculture
Beans	Agriculture
Lentils	Agriculture
Vegetables	Agriculture
Tomatoes	Agriculture
Potatoes	Agriculture
Sugar beets	Agriculture
Other Vegetables	Agriculture
Fruits	Agriculture
Berries	Agriculture
Blueberry	Agriculture
Cranberry	Agriculture
Other Berry	Agriculture
Orchards	Agriculture
Other Fruits	Agriculture
Vineyards	Agriculture
Hops	Agriculture
Herbs	Agriculture
Nursery	Agriculture
Buckwheat	Agriculture
Canary seed	Agriculture
Hemp	Agriculture
Vetch	Agriculture
Other Crops	Agriculture
Mines/excavation	Mines

Appendix B: Supplemental Figures



Supplemental Figure S1. Variation partitioning of water quality variables (response variables) among LULC proportions and lake physiography (explanatory variable groups; $R^2 = 0.35$; N = 636). Water quality variables were transformed and standardized prior to analysis. Percentages inside circles represent the total variation explained by the explanatory variables included in each group, with each group being distinguished by its own color.



Supplemental Figure S2. Cascade multivariate regression tree (cMRT) of the transformed and standardized water quality variables, constrained by the LULC proportions (R^2 = 0.29; N = 664). The model first splits the dataset geographically, then by LULC in response to water quality differences.

Total word count: 8 360