Supplementary Materials

Table S1. Description of counties used in study

County	Species Richness	Species Present (# of observations)
Durham	8	Bufo americanus (130) Hyla versicolor (51) Pseudacris crucifer (135) Rana catesbeiana (3) Rana clamitans melanota (206) Rana pipiens (64) Rana septentrionalis (4) Rana sylvatica (100)
Halton & Hamilton-Wentworth	10	Bufo americanus (75) Hyla chrysoscelis (1) Hyla versicolor (93) Pseudacris crucifer (163) Rana catesbeiana (4) Rana clamitans melanota (173) Rana palustris (2) Rana pipiens (91) Rana septentrionalis (8) Rana sylvatica (33)
Lennox Addington	9	Bufo americanus (20) Hyla versicolor (42) Pseudacris crucifer (73) Rana catesbeiana (46) Rana clamitans melanota (63) Rana palustris (3) Rana pipiens (34) Rana septentrionalis (1) Rana sylvatica (13)
Niagara	8	Bufo americanus (81) Hyla chrysoscelis (2) Hyla versicolor (41) Pseudacris crucifer (126) Rana catesbeiana (59) Rana clamitans melanota (105) Rana pipiens (42) Rana sylvatica (8)
Northumberland	8	Bufo americanus (33)

		Hyla versicolor (61) Pseudacris crucifer (155) Rana catesbeiana (46) Rana clamitans melanota (68) Rana palustris (1) Rana pipiens (36) Rana sylvatica (29)
Peel	7	Bufo americanus (62) Hyla versicolor (70) Pseudacris crucifer (102) Rana catesbeiana (2) Rana clamitans melanota (90) Rana pipiens (36) Rana sylvatica (57)
Prince Edward	7	Bufo americanus (107) Hyla versicolor (155) Pseudacris crucifer (256) Rana catesbeiana (159) Rana clamitans melanota (196) Rana pipiens (139) Rana sylvatica (23)
Toronto	8	Bufo americanus (53) Hyla chrysoscelis (1) Hyla versicolor (26) Pseudacris crucifer (3) Rana catesbeiana (2) Rana clamitans melanota (76) Rana pipiens (56) Rana sylvatica (18)

Table S2. Exclusion of factors from the water quality dataset. Factors with over 15 NAs were excluded as well as factors that measured the same chemical as a retained factor. The only exceptions were nitrate and ammonia, which contained 29 NA values, but were retained due to biological relevance. The NA rows of these factors did not overlap with the NA rows of the removed factors.

Factors excluded	**Reason(s) for exclusion**
Temperature profile	All values are the same
Temperature EBT	1. Over 15 NAs 2. Not biologically relevant as the changes in temperature is not a good

	indicator of water quality
Dissolved inorganic/organic carbon	1. Over 15 NAs 2. Lots of variations between shoreside and tributary, therefore can not easily infer marsh values
Oxygen % saturation dissolved	1. Over 15 NAs 2. Oxygen is already measured by concentration
Chlorophyll A, uncorrected	Over 15 NAs
Chlorophyll A, corrected phaeophytin	Over 15 NAs
Alkalinity phenolphthalein, filtered	Another measure of alkalinity has more values
Phosphorus soluble reactive, filtered	Contains more NAs than total phosphorus
Total phosphorus, filtered	Contains more NAs than total phosphorus
Turbidity	Over 15 NAs
Residual non filtered	Over 15 NAs

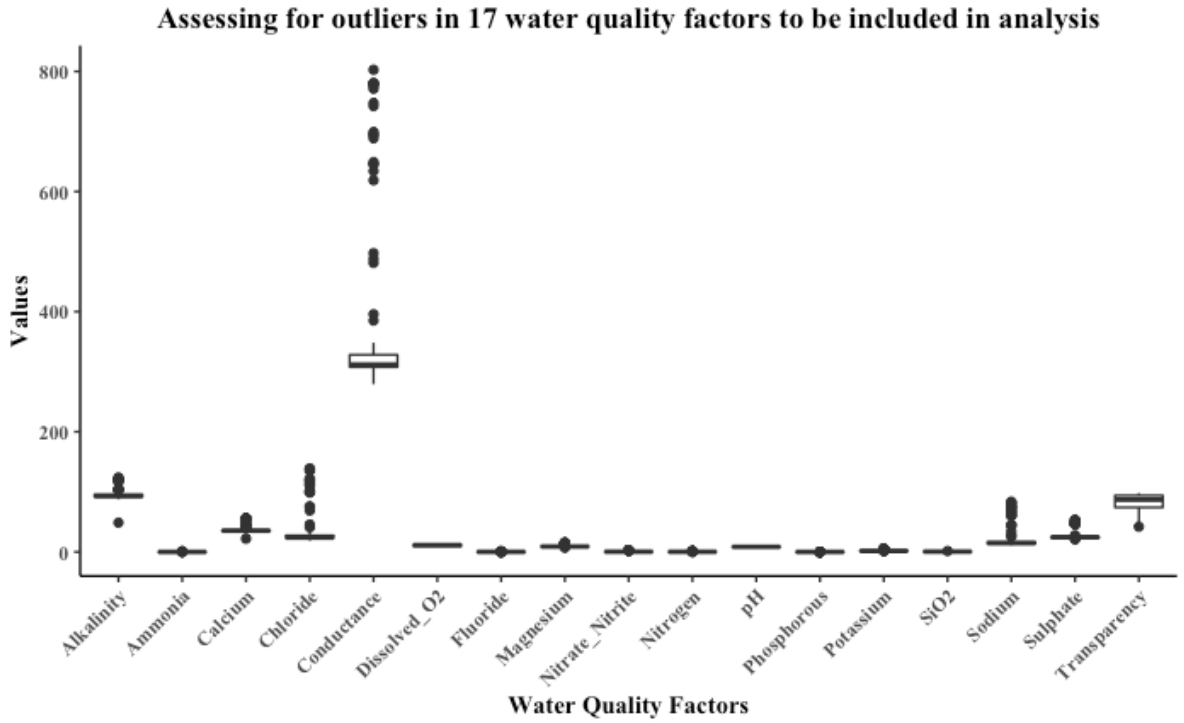

Figure S1. Boxplot of water quality factors used in PCA. Despite the presence of outliers in most water quality factors, we did not remove any outliers as this would drastically reduce our sample size and we have no indication from the dataset whether these represent measurement errors, sampling errors, or natural variations.

Table S3. Assumptions of ANOVA for water chemistry between counties

Shapiro-Wilk normality test	
W	p
0.56465	2.626e-14***

* p < 0.05, ** p < 0.01, ***p < 0.001

Levene's test for homogeneity of variance		
df	F	p
7	12.766	1.115e-10***

* p < 0.05, ** p < 0.01, ***p < 0.001

Table S4. Assumptions of ANOVA for pH between counties

Shapiro-Wilk normality test	
W	p
0.94936	0.00254**

* p < 0.05, ** p < 0.01, ***p < 0.001

Levene's test for homogeneity of variance		
df	F	p
7	0.7932	0.5953

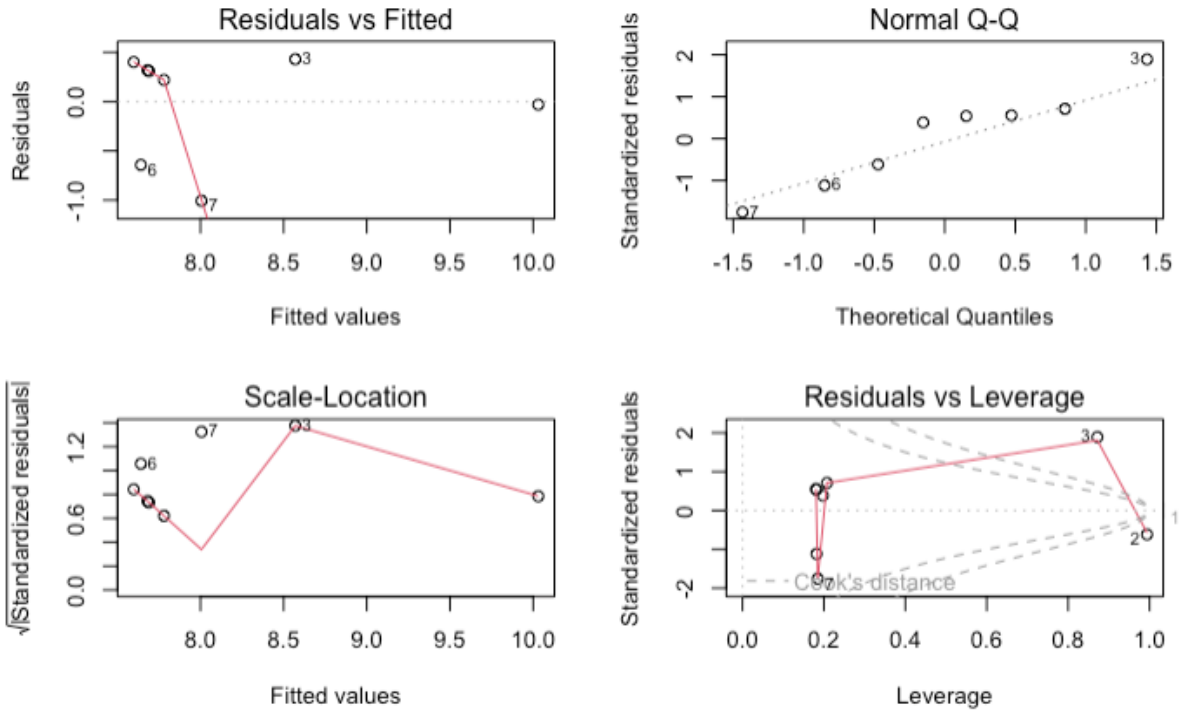

Figure S2. Checking assumptions of multiple linear regression. By visually assessing the plots, the assumption of normality seems to be met by the Q-Q plot, however, the assumption of homogeneity of variances seems to be violated by the residuals vs. fitted plot. The assumption of homogeneity of variances was still not met when transformations were performed on the data.