

GEOSTATISTICAL MODELLING OF IN-STREAM CHLORIDE CONCENTRATIONS  
ACROSS SEASONAL FLOW STATES IN THREE URBANIZING WATERSHEDS

by

Colin Richard Ash, B.Sc, Memorial University of Newfoundland, 2015

A thesis presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Applied Science

in the program of

Environmental Applied Science and Management

Toronto, Ontario, Canada, 2018

© Colin Richard Ash, 2018

## Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

## Abstract

GEOSTATISTICAL MODELLING OF IN-STREAM CHLORIDE CONCENTRATIONS ACROSS SEASONAL FLOW

STATES IN THREE URBANIZING WATERSHEDS

Master of Applied Science

2018

Colin Richard Ash

Environmental Applied Science and Management

Ryerson University

Road salt causes increasing environmental chloride ( $\text{Cl}^-$ ) concentrations which threaten aquatic ecosystems. Environment Canada recommends that road authorities should develop salt management plans including identification of salt vulnerable areas. In this thesis, a Spatial Stream Network (SSN) geostatistical modelling approach was used with seasonal longitudinal field data to develop reach scale models of in-stream  $\text{Cl}^-$  concentrations in three Southern Ontario watersheds. Significance of potential drivers (lane length density (LLD), agricultural & undifferentiated rural land (AURL), and permeability of surficial geology) of stream  $\text{Cl}^-$  were assessed. Results suggest that SSN models are not consistently better than Euclidean models across watersheds. Unexpectedly, LLD was the most important predictor in the rural watershed, and AURL was most important in the urban watershed. Results also show that spatial structure in stream  $\text{Cl}^-$  concentrations was lost under higher flow conditions, which has important implications for when data should be collected to map salt vulnerable areas.

## Acknowledgements

The completion of this thesis was a daunting task and required the guidance and support of faculty, staff, family, friends, and peers to see it from conception to conclusion. I cannot possibly thank all those in my life who made me who I am today, but I hope that these acknowledgments can reflect my deep appreciation for their efforts.

I'd like to thank my supervisors, Dr. Stephanie Melles and Dr. Claire Oswald, for their infinite patience and guidance through this Master's Thesis. I am grateful that Dr. Melles found my application when I applied to the ENSCIMAN program which allowed me to work both with her and with Dr. Oswald on this fantastic project. I would also like to thank Dr. Carl Mitchell and his research team members, especially Aimé Kayembe, for providing much needed chloride and conductivity data for the Mimico Creek watershed.

I'd like to thank my peers, Dwayne Kier, Felix Chan, Evan Miller, Sammy Tangir, Niko Mcglashan, Jelena Grbic, Kyle Smith, and Anna Brooker for their assistance in the office, in the lab, or in the field. Their light-hearted nature, runs to A&W, and interesting music playlists made the long days a little easier to digest.

Thanks to Dr. Andrew Laursen, Karen Puddephatt, and the rest of the Ryerson University Biology Lab teaching staff for my crash course as a lab instructor and grader.

I'd also like to thank my friends and family for their help in guiding me through some tough moments during my studies. Living away from home was a difficult experience but I was able to make it through with the love and support of those around me.

## Table of Contents

Author’s Declaration .....	ii
Abstract.....	iii
Acknowledgements.....	iv
List of Figures .....	vii
List of Tables .....	x
1 Introduction .....	1
2 Literature Review.....	2
2.1 Properties and Sources of Chloride .....	2
2.2 Historical Investigations into Chloride Impacts .....	4
2.3 Chloride in Groundwaters and Streamwaters .....	5
2.4 Chloride Trends in Southern Ontario.....	6
2.5 Landscape Controls on Chloride Migration to Freshwater Ecosystems .....	7
2.6 Ecotoxicological Impact of Chloride in Freshwater Ecosystems.....	8
2.6.1 Density Stratification of Lakes.....	9
2.7 Modelling Stream Chloride Concentrations .....	10
2.7.1 Mass Balance Studies.....	10
2.8 Spatial Relationships and Stream Network Modelling .....	12
2.9 Semivariance and Semivariograms .....	17
2.10 Thesis Aims.....	21
3 Methods.....	22
3.1 Study Watersheds.....	22
3.1.1 Mimico Creek Watershed .....	22
3.1.2 East Holland River Watershed .....	25
3.1.3 Willow Creek Watershed .....	28
3.2 Field Data Collection .....	30
3.2.1 Developing Electrical Conductivity vs. In-Stream Chloride Relationships .....	30
3.2.2 Seasonal Longitudinal Surveys.....	34
3.3 Spatial Stream Network (SSN) Models and Residual Semivariogram Interpretation.....	35
3.4 Covariate Selection .....	41
3.4.1 Lane Length Density.....	43
3.4.2 Subsurface Permeability .....	43
3.4.3 Agriculture and Undifferentiated Rural Land Cover .....	44

3.5	Expected Covariate-Chloride Relationships.....	45
3.6	Model Construction and Investigation of Covariates .....	51
3.7	Metrics for Model Selection and Fit .....	56
3.8	Data Processing.....	56
4	Results.....	59
4.1	Survey Timing and Survey Flow Categorization .....	59
4.2	Electrical Conductivity vs. Chloride Relationships .....	62
4.2.1	Mimico Creek Watershed .....	62
4.2.2	East Holland River Watershed .....	64
4.2.3	Willow Creek Watershed .....	66
4.3	Statistical Summary and Spatial Patterns in Chloride Concentrations .....	68
4.4	Spatial Stream Network Modelling.....	74
4.4.1	Autocovariance Function Selection .....	74
4.5	Model Selection .....	76
4.6	Model Outputs and Covariate Significance.....	80
4.7	Statistical Stream Network Modeling vs. Euclidean Modeling.....	84
4.8	Spatial Stream Network Flow-Connected Residual Semivariograms .....	87
5	Discussion.....	93
5.1	Spatial Patterns in Stream Chloride Concentrations .....	93
5.2	SSN Chloride Models.....	94
5.2.1	Effect of spatial relationship (Euclidean vs. flow-connected) used in SSN models .....	94
5.2.2	Comparison of SSN Model Performance With Respect to Urbanization.....	95
5.2.3	Comparison of SSN model performance across flow states in Mimico Creek.....	97
5.3	Importance of landscape predictors of stream Cl <sup>-</sup> concentrations across seasons and watersheds.....	99
5.4	Do the spatial patterns of residual semivariograms from Spatial Stream Network models of in-stream Cl <sup>-</sup> concentrations vary with urbanization or flow state?.....	100
5.5	Study Limitations .....	103
5.6	Future Work.....	105
	Appendices.....	107
	Mimico Creek Winter Survey Corrections.....	107
	References .....	111

## List of Figures

Figure 1: Three different methods to measure separation distance between sample points. (A) depicts a Euclidean (straight line) method of measuring sample distance between points A and B or A and C. (B) depicts a “Flow Connected” method to measure between sample points whereby only the sample points that can be connected and measured using the flow paths of the network are considered. (C) depicts a “Flow Unconnected” method to measure between sample points whereby the sample points that are not connected via the flow pathways of the network are considered. ....	16
Figure 2: Example semivariogram with a labeled nugget, sill, partial sill, and range. ....	18
Figure 3: An example semivariogram produced using data collected along Mimico Creek watershed, found in the western Greater Toronto Area. Please refer to Figure 2 for a description of “Flow Connected”, “Flow Unconnected”, and Euclidean measurement methods. ....	19
Figure 4: Figure adapted from McGuire et al. (2014) that depicts the varying scale of effects using theoretical semivariograms over a cropped and simplified stream system for the Mimico Creek watershed, located on the western side of the Greater Toronto Area. The colors for each stream network depict a value for an arbitrary in-stream water chemistry value (e.g. Cl <sup>-</sup> concentration). (A) shows a flat, spatially independent semivariogram related to a stream network that is completely uniform or completely random. (A) shows that there is no small scale or large scale structure within the stream network. (B) shows a linear semivariogram associated with a large scale linear change to the in-stream chemistry variable. (C) shows a semivariogram depicting small scale patches of differing in-stream water chemistry values with no large scale pattern. Note that the (C) semivariogram has a nugget, range and sill, with the range being relatively close to the y-axis (short separation distance). (D) depicts a mixed semivariogram showing both small scale and large scale patterns in variance. The (D) semivariogram is associated with small scale patchy water chemistry values on top of a larger scale gradient in water chemistry values. ....	20
Figure 5: Maps delineating the Mimico Creek, East Holland River, and Willow Creek study watersheds with longitudinal sampling points marked along the stream network. The dotted lines represent the full East Holland River and Willow Creek watersheds that extend beyond the study areas. Land cover data was sourced from the Ontario Land Cover Compilation v.2 (OMNRF, 2016a); stream networks were sourced from the Ontario Integrated Hydrology Layer (OMNRF, 2015); watershed boundaries were sourced from the Ontario Flow Assessment Tool (OMNRF, 2013). Reproduced and adapted with permission from Nikolas McGlashan, Ryerson University. ....	23
Figure 6: Map of the surficial geology of the Mimico Creek watershed. Map was created from the Surficial Geology of Southern Ontario dataset (OGS, 2010). ....	24
Figure 7: A geological map of the East Holland River watershed study area. Map was created from the Surficial Geology of Southern Ontario dataset (OGS, 2010). ....	27

Figure 8: A geological map of the Willow Creek watershed study area. Map was created from the Surficial Geology of Southern Ontario dataset (OGS, 2010).....29

Figure 9: The green markers denote the locations of the in-stream electrical conductivity (EC) stations for A) Mimico Creek Watershed, B) East Holland Watershed, and C) Willow Creek Watershed. The EC stations in Mimico Creek were installed and maintained by the Miller Research Group at the University of Toronto. The EC stations in East Holland and Willow Creek were installed and maintained by the Oswald Research Group at Ryerson University. The Provincial Water Quality Monitoring Network (PWQMN) stations (marked in red) were installed and maintained by the Government of Ontario, with the datasets for these stations made available to the public. The watershed areas used in this thesis are delineated by green polygons. ....32

Figure 10: A) Custom made levellogger casing anchored in position and B) a Solinst LTC Levellogger from the UP\_404\_S1 site in East Holland. ....33

Figure 11: Depiction of the tail-up moving average function going from point n on a stream to point n+1. The overlap between the two average functions is where covariance of the two points is measured. ....38

Figure 12: Depiction of the tail-up and tail-down averaging models and how each type of model deals with confluences. The tail down model continues past the confluence, and no splitting of the averaging function is necessary. The tail-up method requires a split of the averaging function to account for the proportional influence of each tributary. ....39

Figure 13: A depiction of the flow exceedance curves for the A) Mimico Creek Watershed, B) East Holland River Watershed, and C) Willow Creek Watershed. This curve depicts the percentage of time that, over the full survey year (June 14, 2016 to June 14, 2017), the average daily flow within each respective watershed will equal or exceed some specified flow value. The colored squares represent the specific dates for each survey for each watershed. Data extracted from the Environment and Climate Change Canada Hydrometric Flow Data web site ([https://wateroffice.ec.gc.ca/mainmenu/historical\\_data\\_index\\_e.html](https://wateroffice.ec.gc.ca/mainmenu/historical_data_index_e.html)) .....60

Figure 14: Specific conductivity versus in-stream chloride plot for the Mimico Creek Study Area. Conductivity and chloride data were obtained from five in-situ water quality stations placed throughout Mimico Creek from the Mitchell Research Group at the University of Toronto Scarborough. Sample points used in this plot were selected to correspond with total length of the study period for Mimico Creek. The polynomial relationship determined from this dataset was used to convert specific conductivity to in-stream chloride for this study. 63

Figure 15: Specific conductivity versus in-stream chloride for East Holland watershed. The DOWN404\_S1, UP404\_S1, and MID\_EH\_S1 are sample sites, located in the headwaters of the watershed, from which in-stream water samples were taken and then measured for both SpCond and Cl<sup>-</sup> concentration. EH\_S1 is a water quality measurement site found outside of the East Holland study area. The Provincial Water Quality Monitoring Network (PWQMN) station (Aurora Creek; id# 03007700702) is found just downstream of MID\_EH\_S1 and this site provided three years of in-stream SpCond and in-stream Chloride data. The

combined headwater + PWQMN dataset was used to calculate the polynomial trend used in converting specific conductivity to in-stream chloride.....	65
Figure 16: Specific conductivity versus chloride graph for Willow Creek watershed. The MID_WIL_S1 site was selected as the downstream site as it is near the outlet for the Willow Creek study area. UP_WIL_S1 is found within the headwaters of the Willow Creek study area. Both LOW_WIL_S1 site and the Provincial Water Quality Monitor Network (PWQMN) site (id# 03005703002) are located outside the Willow Creek watershed study area. The MID_WIL_S1 SpCond-Cl <sup>-</sup> relationship was used for all survey sites downstream of Little Lake and the UP_WIL_S1 SpCond-Cl <sup>-</sup> relationship was used for all survey sites upstream of Little Lake.....	67
Figure 17: In-stream chloride maps for the Mimico Creek watershed showing the results of five spatially intensive longitudinal surveys. Chloride values were measured via specific conductivity measurements at each sample site, with the conversion equation from specific conductivity to chloride shown in Figure 14. ....	71
Figure 18: In-stream chloride maps for the East Holland River Watershed showing the results of four spatially intensive longitudinal surveys. Chloride values were measured via specific conductivity measurements at each sample site, with the conversion equation from specific conductivity to chloride shown in Figure 15. ...	72
Figure 19: In-stream chloride maps for the Willow Creek Watershed showing the results of four spatially intensive longitudinal surveys. Chloride values were measured via specific conductivity measurements at each sample site, with the conversion equation from specific conductivity to chloride shown in Figure 16. ....	73
Figure 20: Semivariograms for the Spring 1, Summer, Fall, Winter, and Spring 2 longitudinal surveys of the Mimico Creek watershed. ....	88
Figure 21: Semivariograms for the Summer, Fall, Winter, and Spring longitudinal surveys of the East Holland watershed. ....	90
Figure 22: Semivariograms for the Summer, Fall, Winter, and Spring longitudinal surveys of the Willow Creek watershed. ....	92
Figure 23: The R <sup>2</sup> , Akaike Information Criterion (AIC), and Root Mean Square Prediction Error (RMSPE), plotted against flow for all five seasonal surveys for the Mimico Creek watershed. ....	98

## List of Tables

Table 1: Semivariogram parameters for maximum site-pair distance and minimum number of site-pairs. Parameters set using Zimmerman & Ver Hoef (2017) .....	40
Table 2: The sample size and dates for the Mimico Creek, East Holland River, and Willow Creek longitudinal surveys of in-stream electrical conductivity. ....	42
Table 3: Expected importance of Lane Length Density, Weighted Permeability, and Agriculture & Undifferentiated Rural Land covariates in models of in-stream chloride within the Mimico Creek surveys. Expected covariate importance was based on watershed characteristics, weather, and potential Cl <sup>-</sup> inputs. ....	48
Table 4: Expected importance of Lane Length Density, Weighted Permeability, and Agriculture & Undifferentiated Rural Land covariates in models of in-stream chloride within the East Holland River surveys. Expected covariate importance was based on watershed characteristics, weather, and potential Cl <sup>-</sup> inputs. ....	49
Table 5: Expected importance of Lane Length Density, Weighted Permeability, and Agriculture & Undifferentiated Rural Land covariates in models of in-stream chloride within the Willow Creek surveys. Expected covariate importance was based on watershed characteristics, weather, and potential Cl <sup>-</sup> inputs. ....	50
Table 6: Tested models for each season within Mimico Creek, East Holland River, and Willow Creek watersheds. These models are based on all possible non-redundant combinations of model covariates. LLD is Lane Length Density as defined in Section 3.4.1, WPerm is Weighted Permeability as defined in Section 3.4.2, and AURL is the Agricultural and Undifferential Rural Land Cover variable as defined in Section 3.4.3.....	52
Table 7: Expected importance of the input covariates for all Mimico Creek Watershed models. Low importance means little improvement in model performance was expected if this variable were included, moderate impact means moderate improvement in expected model performance, and high impact means significant improvement in expected model performance. The expected best model for each survey is highlighted in green. Covariate impact estimations are based on the rationale seen in Table 3. The definitions for model acronyms are found in Table 6. ....	53
Table 8: Expected importance of the input covariates for all East Holland River Watershed models. Low importance means little improvement in model performance was expected if this variable were included, moderate impact means moderate improvement in expected model performance, and high impact means significant improvement in expected model performance. The expected best model for each survey is highlighted in green. Covariate impact estimations are based on the rationale found in Table 4. The definitions for model acronyms are found in Table 6. ....	54
Table 9: Expected importance of the input covariates for all Willow Creek Watershed models. Low importance means little improvement in model performance was expected if this variable were included, moderate impact means moderate improvement in expected model performance, and high impact means significant improvement in expected model performance. The expected best model for each survey is highlighted in	

green. Covariate impact estimations are based on the expected found in Table 5. The definitions for model acronyms are found in Table 6. ....55

Table 10: Transformations applied to observed chloride values and covariate values at each sample point. Transformations were applied with the aim of normalizing chloride and covariate distribution curves.....58

Table 11: A categorization of the flow states for each longitudinal watershed survey using data from Water Survey of Canada hydrometric gauges located within Mimico Creek, East Holland, and Willow Creek watersheds (refer to Figure 13 for flow exceedance curves).....61

Table 12: The dates, sample sizes, and basic statistics for calculated in-stream chloride concentrations in the Mimico Creek, East Holland River, and Willow Creek watersheds. The equations used to convert electrical conductivity to chloride concentration can be found in Figure 14 for Mimico Creek, Figure 15 for East Holland River, and Figure 16 for Willow Creek. ....70

Table 13: A comparison of statistical metrics when different tail-up autocovariance functions are applied to each watershed using Lane Length Density, Weighted Permeability, and Rural and Undifferentiated Land Cover covariates. The resulting values found in this table are an average of the outputs for each of the individual longitudinal surveys. All values within 5% of the highest  $R^2$ , the lowest RMSPE, the lowest AIC, and within  $\pm 5\%$  of the ideal values for cov.80 (0.80), cov.90 (0.90), and cov.95 (0.95) are bolded.....75

Table 14: The expected best models, model comparison metrics, and a sum of the number of metrics that fit within the “best” category for each model for the Mimico Creek dataset. All values within 5% of the highest  $R^2$ , the lowest RMSPE, the lowest AIC, and within  $\pm 5\%$  of the ideal values for cov.80 (0.80), cov.90 (0.90), and cov.95 (0.95) are bolded and categorized as the “best” result for comparison with other models. The expected best model is highlighted in green, while the model with the highest number of bolded metrics is highlighted in yellow.....77

Table 15: The expected best models, model comparison metrics, and a sum of the number of metrics that fit within the “best” category for each model for the East Holland dataset. All values within 5% of the highest  $R^2$ , the lowest RMSPE, the lowest AIC, and within  $\pm 5\%$  of the ideal values for cov.80 (0.80), cov.90 (0.90), and cov.95 (0.95) are bolded and categorized as the “best” result for comparison with other models. The expected best model is highlighted in green, while the model with the highest number of bolded metrics is highlighted in yellow.....78

Table 16: The expected best models, model comparison metrics, and a sum of the number of metrics that fit within the “best” category for each model for the Willow Creek dataset. All values within 5% of the highest  $R^2$ , the lowest RMSPE, the lowest AIC, and within  $\pm 5\%$  of the ideal values for cov.80 (0.80), cov.90 (0.90), and cov.95 (0.95) are bolded and categorized as the “best” result for comparison with other models. The expected best model is highlighted in green, while the model with the highest number of bolded metrics is highlighted in yellow.....79

Table 17: Spatial Stream Network model outputs for the Mimico Creek watershed. All predictor variables with  $p < 0.05$  are considered significant and are bolded. Model  $R^2$  and proportion of variation in  $Cl^-$  concentration explained by covariates and spatial component is included beneath the outputs for each model. ....81

Table 18: Spatial Stream Network model outputs for the East Holland River watershed. All predictor variables that have a significance value of  $p < 0.05$  are considered significant and are bolded. Model  $R^2$  and proportion of variation in  $Cl^-$  concentration explained by covariates and spatial component is included beneath the outputs for each model. ....82

Table 19: Spatial Stream Network model outputs for the Willow Creek watershed. All predictor variables that have a significance value of  $p < 0.05$  are considered significant and are bolded. Model  $R^2$  and proportion of variation in  $Cl^-$  concentration explained by covariates and spatial component is included beneath the outputs for each model. ....83

Table 20: A comparison of the Spatial Stream Network models used in this thesis to a Euclidean counterpart. All Euclidean models in this thesis used the exponential autocovariance function and uses the direct distance between points instead of stream distances. ....86

# 1 Introduction

Canadian winters are severe enough to warrant the application of over 5 million tonnes of road de-icing salt across the country to ensure the safety of automobiles on roadways (Perera et al., 2013). While road salt is useful to protect public health, there are a number of short- and long-term negative freshwater environmental implications associated with its use (USEPA, 1971; Murray & Ernst, 1976; Cain et al., 2000; CCME, 2011; Perera et al., 2013). Rainfall and snowmelt events can quickly transport road salt off of roadways via surface runoff and into the environment, resulting in short pulses of relatively high salinity in soils, streams and rivers (Crowther & Hynes, 1977; Daley et al., 2009; CCME, 2011). The quick onset of these pulses does not allow time for aquatic biota to adapt or retreat from the new environmental conditions that meet or exceed toxic salinity thresholds for many species (CCME, 2011). Continuous use of road salt over many years can increase background concentrations of the constituents of road salt, typically sodium (Na) and chloride (Cl<sup>-</sup>), in surface and subsurface water (Howard & Haynes, 1993; Howard & Maier, 2007; Perera et al., 2013). Elevated export of salts into receiving water bodies (e.g. ponds and lakes) can result in density stratification, loss of seasonal mixing, a reduced oxygen state, and the promotion of eutrophication processes (Hoffman et al., 1981; Judd, 1970; Koretsky et al., 2012; MacLeod et al., 2011; Novotny et al., 2008).

Increasing background subsurface Cl<sup>-</sup> concentrations is an area of concern as the accumulation of Cl<sup>-</sup> within groundwater aquifers can take many years to be flushed out due to the long water residence time of these systems (Kaushal et al., 2005; Morgan et al., 2012). Subsurface Cl<sup>-</sup> concentrations near urban areas and adjacent to major roadways in Northeastern United States and Southern Ontario are reaching levels higher than what is recommended by Health Canada, the United States Environmental Protection Agency, and the Canadian Council of Ministers for the Environment (CCME, 2011; Kaushal et

al., 2005; Perera et al., 2013; Winter et al., 2011). Without corrective action, for example by reducing road salt application rates, it is expected that soil-water and groundwater  $\text{Cl}^-$  concentrations throughout rural regions of Southern Ontario and Northeast United States will potentially exceed Health Canada's guidelines within the next century (Howard & Maier, 2007; Kaushal et al., 2005).

The overall goal of this study was to improve our understanding of stream  $\text{Cl}^-$  concentrations at a resolution more relevant for mapping salt vulnerable areas. The following research questions were addressed: 1) How does the Spatial Stream Network (SSN)  $\text{Cl}^-$  model performance vary with respect to urbanization and flow state?, 2) Does the importance of landscape predictors of in-stream  $\text{Cl}^-$  concentrations vary across seasons and watersheds?, 3) Do the spatial patterns of residual semivariograms from SSN models of in-stream  $\text{Cl}^-$  concentrations vary with urbanization or flow state?

## 2 Literature Review

### 2.1 Properties and Sources of Chloride

Chloride is a component of common salts like sodium chloride ( $\text{NaCl}$ ), potassium chloride ( $\text{KCl}$ ) and magnesium chloride ( $\text{MgCl}_2$ ), and all  $\text{Cl}^-$ -containing salts easily dissociate in water into their constituent ions (CCME, 2011). The  $\text{Cl}^-$  ion is generally considered to be conservative, meaning it does not undergo significant physical or chemical interactions in the environment which would lead its degradation, bioaccumulation, sorption to soils, precipitation, or volatilization (CCME, 2011). These conservative properties of  $\text{Cl}^-$  make it a useful environmental tracer in areas that do not receive anthropogenic inputs, allowing scientists to characterize such things as solute transport properties through soils (Mallants, 2014), groundwater recharge properties (Derby & Knighton, 2001; Edmunds & Gaye, 1994; Eriksson & Khunakasem, 1969), and tracing sewage pathways through a watershed from the source (Katz et al., 2011; Kelly et al., 2010). However, recent studies show that  $\text{Cl}^-$ -soil interactions are more complex than previously thought and that short-term  $\text{Cl}^-$  retention may occur in areas with low  $\text{Cl}^-$  inputs due to

temporary binding to soil organic matter (Bastviken et al., 2006; Bastviken et al., 2007; Svensson et al., 2012).

The use of Cl<sup>-</sup> bearing road salt on impervious surfaces (e.g. roadways & parking lots) can result in Cl<sup>-</sup> being mobilized directly from these surfaces during precipitation or melt events (Cooper et al., 2014). Road salt use can also result in Cl<sup>-</sup> being stored in snowbanks and in shallow soils that are found near road salt deposition sites, and this stored Cl<sup>-</sup> can be mobilized from melting snowbanks and from the flushing of soil waters during precipitation and/or melt events (Cooper et al., 2014). The mobilization of a large quantity of Cl<sup>-</sup> from an individual precipitation-runoff event, known as a *chloride pulse*, can result in extremely high short-term in-stream Cl<sup>-</sup> concentrations of over 6000 mg/L in some areas (Cooper et al., 2014; Gardner & Royer, 2010; Perera et al., 2010).

Chloride has various anthropogenic sources aside from road salt. The United States Geological Service (Bolen et al., 2016) reports that 55 megatons (MT) of Cl<sup>-</sup> is reportedly consumed in the United States, with approximately 1.6 MT used for agriculture and approximately 1 MT in water treatment. Agricultural-sourced Cl<sup>-</sup> can stem from animal waste which can have Cl<sup>-</sup> concentrations as high as 847 mg L<sup>-1</sup> as found in Illinois (Mullaney et al., 2009), or Cl<sup>-</sup> can come from potassium Cl<sup>-</sup> fertilizers which contain 47.5% Cl<sup>-</sup> by weight (Chapra et al., 2009). Landfills can contain high levels of Cl<sup>-</sup> as a by-product of Cl<sup>-</sup>-bearing waste products. For example, seven landfills across Illinois had an average Cl<sup>-</sup> leachate concentration of 2310 mg L<sup>-1</sup> (Panno et al., 2005), and a landfill in Niagara Falls, Canada, had leachate with a Cl<sup>-</sup> concentration of 1704 mg L<sup>-1</sup> (Howard & Beck, 1993). Wastewater treatment facilities may also fail to eliminate Cl<sup>-</sup> from incoming waste products due to chlorides' conservative properties (Mullaney et al., 2009). More broadly, the atmosphere can act as a source for Cl<sup>-</sup> as volcanic emissions and marine spray can deposit Cl<sup>-</sup> in a watershed (Kelly et al., 2010; CCME, 2011), such that above ground or subsurface Cl<sup>-</sup> infiltration can occur in close proximity to marine shorelines (CCME, 2011). Natural background concentrations of Cl<sup>-</sup> can be heavily dependent on location of measurement with Southern

Ontario lakes and streams having a natural background concentration of 10 to 25 mg L<sup>-1</sup> (Howard & Beck, 1993; CCME, 2011).

## 2.2 Historical Investigations into Chloride Impacts

Interest in the environmental impacts of Cl<sup>-</sup> has followed the fast growth of urban and roadway infrastructure with initial studies being published in the 1960s and the 1970s. In 1954, the United States Department of Agriculture released a manual on how to diagnose and treat overly saline soils but made no mention of road salts being a possible pollutant (Allison et al., 1954). The United States Environmental Protection Agency (USEPA) released a comprehensive report on the impacts of road salt in 1971, which suggests that the primary concerns at the time were the deleterious effects of road salt on roadway infrastructure, roadside plant health, soil chemistry and fertility, and the dispersal and infiltration of road salts onto and into soils. The relationship between road salt and increased environmental concentrations in streams and lakes seemed to be an incidental concern (USEPA, 1971). Many studies were driven by state government and highway research boards to help determine the indirect economic costs of road salt. These economic concerns culminated in a report released by the USEPA in 1976 that estimated the negative environmental and infrastructure-related impacts of applying road salt as a de-icing material cost individual state governments a minimum of \$2.91 Billion per year (Murray & Ernst, 1976).

Studies on effects of road salt as a de-icer from the mid-20<sup>th</sup> century still had a predominate focus on economic factors that affected infrastructure and agriculture. For example, researchers investigated elevated ion concentrations and dispersal patterns in proximity to roadways (Davison, 1971; Muethel, 1976), Cl<sup>-</sup> migration through soils (Davison, 1971), the negative impact of road salts on crop fertility, road side vegetation health, and salt damage to grasses and trees (Prior & Berthouex, 1967; USDI, 1969; Davison, 1971; Murray & Ernst, 1976; Townsend, 1980). Increasing salt concentrations

led materials scientists and engineers to create best roadway and bridge construction practices to mitigate and reduce damage caused by road salt (Murray & Ernst, 1976)

### 2.3 Chloride in Groundwaters and Streamwaters

As  $\text{Cl}^-$  is highly dissolvable in water and has conservative properties, it has the ability to percolate through soils via surface water infiltration and enter shallow groundwater (Howard & Beck, 1993; Meriano et al., 2009; Betts et al., 2014). Continued road salt inputs into a watershed can lead to a net influx of  $\text{Cl}^-$  into that watershed's groundwater reservoirs via overland mobilization of road salt from a precipitation or melt event through a groundwater recharge zone (Howard & Beck, 1993; Meriano et al., 2009; Perera et al., 2013). Several watershed studies in Southern Ontario and Finland found that surface water flow only removed 35% to 50% of applied road salts from a watershed, with the rest accumulating within the shallow subsurface (Howard & Beck, 1993; Ruth, 2003; Meriano et al., 2009). Due to the relatively slow rate of groundwater flow out of an aquifer, large inputs of  $\text{Cl}^-$  through groundwater recharge zones can lead to high  $\text{Cl}^-$  concentrations in groundwater aquifers (Bester et al., 2006; Kaushal et al., 2005; Morgan et al., 2012).

Assuming a steady input of  $\text{Cl}^-$  into a groundwater aquifer, groundwater  $\text{Cl}^-$  concentrations will increase until the  $\text{Cl}^-$  inputs into a reservoir equal the  $\text{Cl}^-$  outputs via groundwater outflow from the reservoir, resulting in a steady state equilibrium (Meriano et al., 2009). However, if there is a consistent increase of  $\text{Cl}^-$  input into an aquifer, then a steady state equilibrium may not be reached and groundwater  $\text{Cl}^-$  concentrations may continue to increase (Novotny et al., 2009). There also exists the potential for a lagged response between  $\text{Cl}^-$  inputs into a reservoir and subsequent outputs as it takes time for  $\text{Cl}^-$  to percolate through the subsurface (Perera et al., 2013). This slow response may result in increasing baseflow  $\text{Cl}^-$  concentrations even if all road salt inputs are stopped (Perera et al., 2013). It can

be very difficult and expensive to remove the groundwater  $\text{Cl}^-$  contamination and highlights the need to deal with  $\text{Cl}^-$  contamination by reducing or eliminating the use of  $\text{Cl}^-$ -based road salt (Perera et al., 2013).

## 2.4 Chloride Trends in Southern Ontario

In the mid-20<sup>th</sup> century wastewater effluent was considered to be the primary source of increasing  $\text{Cl}^-$  concentrations in Ontario freshwater systems (MOECC, 2016). Chloride concentrations in Lake Ontario peaked in the early 1970s at approximately 30 mg/L, but the implementation of Ontario's *Clean Water Act* (1972) regulated wastewater effluent and led to a subsequent decrease in  $\text{Cl}^-$  concentrations to approximately 22 mg/L by the mid-1990s (Chapra et al., 2009). Since then,  $\text{Cl}^-$  concentrations have been rising in both Lake Erie and Lake Ontario due to increasing use of road salt associated with significant urban growth in Southern Ontario (Chapra et al., 2009; CCME, 2011).

Studies conducted in the Greater Toronto Area (GTA) and Lake Simcoe watershed highlight increasing  $\text{Cl}^-$  concentrations in streams, rivers, and receiving water bodies over the past several decades. Perera et al. (2013) measured groundwater  $\text{Cl}^-$  concentrations as high as 250 mg L<sup>-1</sup> in the Highland Creek watershed of Toronto, Ontario, and these authors predicted that concentrations could increase to 505 mg L<sup>-1</sup> over the next century based on current trends of road salt application within the City. Howard & Beck (1993) found elevated levels of  $\text{Cl}^-$  in over half of 400 groundwater wells across Southern Ontario, with domestic wells ranging up to 700 mg/L and some pore waters reaching a concentration of 13,700 mg/L. Williams et al. (2000) measured  $\text{Cl}^-$  concentrations in 23 freshwater springs across the GTA and found concentrations as high as 1,200 mg/L.

During winter, in-stream  $\text{Cl}^-$  pulses are reported as high as 6,000 mg/L in the Highland Creek watershed of Toronto, Ontario, (Perera et al., 2010) and 5,000 mg/L in some regions of Northeastern United States (Kaushal et al., 2005). Winter et al. (2011) showed that  $\text{Cl}^-$  concentrations in Lake Simcoe outflow increased over three times their initial values over a span of 36 years, predicting that the

outflow will have a concentration of  $150 \text{ mg L}^{-1}$  in approximately 140 years which will exceed the CCME standard for chronic ecotoxicological  $\text{Cl}^-$  concentrations of  $120 \text{ mg L}^{-1}$  (CCME, 2011).

## 2.5 Landscape Controls on Chloride Migration to Freshwater Ecosystems

One of the strongest drivers of in-stream and groundwater  $\text{Cl}^-$  concentrations is total impervious surface area (e.g. roadways, parking lots, driveways) within a watershed (Kaushal et al., 2005; Kelting et al., 2012; Morgan et al., 2012). Not only do impervious surfaces need to be salted for safety reasons, but impervious surfaces reduce the amount of permeable soils that can capture  $\text{Cl}^-$ -rich runoff, instead allowing this runoff to directly enter the stream (Kelting et al., 2012). Total impervious area and road density within a watershed are thus strong factors that influence stream  $\text{Cl}^-$  concentrations (Kelting et al., 2012).

Other aspects of urbanization, such as channelization of streams with associated loss of connection between the stream and groundwater, can alter the movement of water through the groundwater/ $\text{Cl}^-$  interface, thus potentially altering the pathway that dissolved  $\text{Cl}^-$  may take through a watershed (Walsh et al., 2005). Landfill leachate and wastewater treatment plant discharge may also result in  $\text{Cl}^-$  being introduced into a stream network (Panno et al., 2006).

Agricultural activities may impact  $\text{Cl}^-$  concentrations as salts are found in animal food, animal waste products, fertilizers, and irrigation waters sourced from  $\text{Cl}^-$  laden groundwater (Mullaney et al., 2009). Panno et al. (2006) noted that type of fertilizer is important when considering potential environmental  $\text{Cl}^-$  inputs, and they estimated  $\text{Cl}^-$  concentrations of  $5.7$  to  $36.5 \text{ mg L}^{-1}$  in their study of tile drainage areas in Illinois, assuming  $\sim 90 \text{ kg/acre}$  of potassium chloride fertilizer loading rates. Panno et al. (2006) also observed animal wastes with  $\text{Cl}^-$  concentrations as high as  $1980 \text{ mg L}^{-1}$ .

Geology can have considerable influence on in-stream  $\text{Cl}^-$  concentrations in two primary ways. First,  $\text{Cl}^-$  can readily enter subsurface aquifers through aquifer recharge zones, so position of the aquifer

in the watershed, along with the level of connection of the aquifer to the stream network, can dictate amount of Cl<sup>-</sup> laden groundwater discharge entering a stream system (Derby & Knighton, 2001; Meriano et al., 2011). Second, the size and location of Cl<sup>-</sup> rich geological deposits (e.g. halite deposits) can result in the formation of Cl<sup>-</sup> rich brines, and the movement of Cl<sup>-</sup> from these geological deposits to a stream (Godwin et al., 2003; Gutchess et al., 2016; Panno et al., 2002).

## 2.6 Ecotoxicological Impact of Chloride in Freshwater Ecosystems

Increasing the concentration of Cl<sup>-</sup> within watersheds can result in many negative impacts on freshwater ecosystems. An increase in environmental salinity, caused by increasing Cl<sup>-</sup> concentrations, can exert osmotic pressure on aquatic biota such that there may be a net flow of water out of the cells of an organism potentially leading to a disruption in homeostasis (CCME, 2011; Holland et al., 2011). If the environmental concentration of Cl<sup>-</sup> rises above a maximum threshold, or rises too quickly, such as in an intense Cl<sup>-</sup> pulse, then an organism may not be able to compensate for the homeostatic shift and may suffer deleterious effects (e.g. stress, illness, reproductive and developmental problems, or death) (Davison, 1971; Richburg et al., 2001; Sanzo & Hecnar, 2006; Collins & Russell, 2009; Brand et al., 2010; CCME, 2011; Karraker & Gibbs, 2011). The Northern Riffleshell Mussel (*Epioblasma torulosa rangiana*) is an endangered in-stream species in Ontario and these bivalves are particularly vulnerable to acute Cl<sup>-</sup> pulses (Gillis, 2011). Riffleshell mussel larvae have a 50% probability to lose developmental viability when exposed to a Cl<sup>-</sup> concentrations of about 244 mg/L over a span of 24 hours (Gillis, 2011). This is important as acute Cl<sup>-</sup> concentrations in riffleshell mussel habitats can reach as high as 1300 mg/L (Gillis, 2011). Certain amphibian species, such as the spotted salamander (*Ambystoma maculatum*), are extremely sensitive to small changes in background Cl<sup>-</sup> concentrations, with developmental abnormalities observed at Cl<sup>-</sup> concentrations as low as 77.5 mg/L over a period of 90 days (Sanzo & Hecnar, 2006). Road salt is hypothesized to promote proliferation of salt tolerant species at the expense of salt sensitive species (Green et al., 2008; Novotny et al., 2008; Prasser & Zedler, 2010). Areas

alongside roadways can act as vectors for salt-tolerant species, like the invasive reed *Phragmites australis*, to spread throughout a landscape (Jodoin et al., 2008). Salinization of sedge meadow marshes in Wisconsin is also an explanation for the replacement of native tussock (*Carex stricta*) for the invasive reed canarygrass (*Phalaris arundinacea*) (Prasser & Zedler, 2010). Increased Cl<sup>-</sup> ion concentrations are linked to elevated environmental concentrations of heavy metals (e.g. copper, zinc and iron) through the formation of Cl-metal compounds that can enter a dissolved state in surface and groundwaters (Bäckström et al., 2004; Warren & Zimmerman, 1994).

Recognition of the environmental impacts of Cl<sup>-</sup> prompted Environment Canada and Health Canada to state that exposure to chronic Cl<sup>-</sup> concentrations  $\geq 230$  mg/L over a period of weeks can have detrimental impacts on aquatic life, whereas Cl<sup>-</sup> concentrations  $\geq 800$  to 1000 mg/L can pose the same threat after exposure for a few days (Perera et al., 2010). The CCME (2011) has released more conservative Cl<sup>-</sup> ecotoxicological guidelines, stating that Cl<sup>-</sup> concentrations  $\geq 120$  mg/L may have chronic toxicological effects and  $\geq 640$  mg/L may have acute toxicological effects. The CCME (2011) Cl<sup>-</sup> concentration guidelines are presented with some caveats: they are based on a limited number of toxicology studies, they may not be accurate for an ecosystem that has adapted to an elevated level of background Cl<sup>-</sup> concentrations, and these generalized guidelines may not have strong relevance to any one particular site (*i.e.* lack of site specificity).

### 2.6.1 Density Stratification of Lakes

Salinization of lakes may create a density gradient as saltwater will sink in freshwater, creating or reinforcing a stable vertical stratification of lake water (Boehrer & Schultze, 2008; Judd, 1970; Koretsky et al., 2012). Stable stratification of a lake reduces the ability for wind stress and temperature differences to thoroughly mix the different layers within a lake, thereby decreasing the cycling of nutrients throughout the water column (Boehrer & Schultze, 2008). This decreased mixing may promote

the formation of an oxygen-poor zone in the lower layers of a lake, which may cause death within this zone and subsequent decomposition of lake organisms (Boehrer & Schultze, 2008; Koretsky et al., 2012; MacLeod et al., 2011). Continuous reduction in oxygen can decrease the habitability of a lake leading to a “dead lake” with little biotic activity (Koretsky et al., 2012; MacLeod et al., 2011). Examples of significant vertical stratification caused by road salt is tied to formation of suburbs near a lake in California (Judd, 1970), urban and rural lakes in Michigan (Judd, 1970; Koretsky et al., 2012; MacLeod et al., 2011) and rural lakes in Norway (Kjensmo, 1997).

## 2.7 Modelling Stream Chloride Concentrations

Environment Canada’s *Code of Practice for the Environmental Management of Road Salt* (2004) states that one of the primary issues facing implementation of best practices for road salt is the lack of identification of salt vulnerable areas (SVAs) within watersheds. Identification of SVAs, or areas that have environmental and/or ecotoxicological vulnerabilities to the presence of road salt, are important as this would allow road authorities to strategically reduce or eliminate the use of road salt within SVAs while preserving public safety (ECCC, 2004). Conducting fine scale, spatially intensive physical surveys with the aim of studying water chemistry would allow for identification of  $\text{Cl}^-$  concentration hotspots and better delineation of SVAs, but fine scale watershed studies are expensive and previous work to study and understand  $\text{Cl}^-$  have mostly been conducted using mass balance studies. Mass balance studies, further discussed below (Section 2.7.1), are conducted at a watershed scale using simple and relatively inexpensive data collection methods.

### 2.7.1 Mass Balance Studies

Mass balance approaches commonly use the  $\text{Cl}^-$  ion to further understand watershed hydrologic properties due to the ion’s conservative nature. Mass balance studies rely on the principle of mass conservation which states that a chemical can only do three things within a watershed (Hemond &

Fechner, 2015): 1) be added to the system via transformation or transportation processes, 2) be transported within the watershed, and 3) be eliminated from the watershed via transformation or transportation processes. When the inputs and outputs for a system like a watershed can be characterized, mass balance principles can be used to better understand the underlying processes that govern the fate and transport of chemicals within a watershed.

Mass balance methods are used to study several catchment processes: the pathways that  $\text{Cl}^-$  takes within a watershed (Meriano et al., 2009; Novotny et al., 2009; Perera et al., 2010, 2013); the accumulation of  $\text{Cl}^-$  within groundwater (Meriano et al., 2009; Perera et al., 2010) and soils (Svensson et al., 2012); the estimations of in-stream  $\text{Cl}^-$  concentrations in stream systems (Betts et al., 2014); and in studying hydrological properties like groundwater recharge rates and surface water percolation depths (Bazuhair & Wood, 1996; Fouty, 1989; Guan et al., 2010; Ma et al., 2012; Sharma & Hughes, 1985).

In-stream  $\text{Cl}^-$  models can aid in the detection of  $\text{Cl}^-$  hotspots and salt vulnerable areas, but studies to date typically lack the spatially intense, longitudinal, stream data that would allow one to study watershed processes at fine scales (Hrachowitz et al., 2013; Peterson et al., 2006). Attempts to understand watershed biological or physical processes are typically conducted at the watershed outlet via mass balance studies due to their ease of implementation, but attempting to understand how these processes operate within the watershed are difficult due to the coarseness of these measures (Levin, 1992; Pickett & Cadenasso, 1995).

For example, Betts et al. (2014) used a mass balance equation to estimate in-stream  $\text{Cl}^-$  concentrations: they estimated in-stream  $\text{Cl}^-$  concentrations in six urban watersheds in the Greater Toronto Area, and combined this information with ecotoxicological data to identify areas in which biota may be vulnerable to salting activities. Their mass balance equation model was used to estimate in-stream  $\text{Cl}^-$  concentrations at arbitrary stream points based on salt application rates, land cover data, and groundwater input estimates. However, their method did not account for reach level spatial variability

of in-stream  $\text{Cl}^-$  concentrations and they simply model in-stream  $\text{Cl}^-$  using salt inputs and readily available GIS data.

## 2.8 Spatial Relationships and Stream Network Modelling

Most data collected in a geographical, environmental or ecological context contains some form of spatial structure whereby points with shorter separation distances tend to correlate (Dale & Fortin, 2014; Peterson et al., 2013). This trend of increasing similarity with decreasing spatial distance was recognized by Tobler (1970) and codified as the first law of geography: *“Everything is related to everything else, but near things are more related than distant things.”*

The complexity of spatial relationships found within an ecological dataset often depends on the type of system under investigation. For example, it may be expected that water quality measurements from the surface of a well-mixed lake would exhibit a relatively simple relationship between location of points on the surface and observed water chemistry values. More complex spatial relationships can form from ecological sampling of points along a stream network due to the interaction between stream topology, independent variables like landscape characteristics, and the characteristics of the dependant variable being studied. For example, the ability for fish to swim upstream or downstream requires researchers to examine both upstream and downstream factors of any arbitrary point along a stream network, and not accounting for stream directionality can negatively impact the power of predictive models (Lois, 2016).

Traditional terrestrial geospatial studies have used Euclidean methods to determine spatial relationships in environmental, geographical, or ecological datasets, however stream networks benefit from non-Euclidean methodologies as stream point locations are confined to a potentially complex stream network and stream points are also influenced by directionality of stream flow (Peterson et al., 2013; Ver Hoef et al., 2006). Information may be lost if one does not account for the unique

characteristics that differentiate stream networks from other landscape environments (Peterson et al., 2013; Ver Hoef et al., 2006). In the case of riverine systems, tools like Spatial Stream Network (SSN) Modelling were created to examine physicochemical and biological riverine processes within the context of a stream network, from the regional scale to the reach scale (McGuire et al., 2014; Peterson et al., 2013; Ver Hoef et al., 2006). Euclidean methods are less appropriate for dendritic ecological networks as they fail to capture the influences of complex branch structures, flow connectivity, and flow directionality (Peterson et al., 2013; Ver Hoef et al., 2006), particularly at finer spatial scales. These stream network influences can be better captured by measuring separation distances along the network between points that are flow connected, flow unconnected, or both flow connected and unconnected (Figure 1; Peterson et al., 2013; Ver Hoef et al., 2006). Measuring distances along the network may also improve model accuracy as the chemical or bio-aquatic variable being measured in the stream may be realistically confined to the stream network (Isaak et al., 2014; Lois, 2016; Peterson et al., 2013; Ver Hoef et al., 2006).

A Spatial Stream Network (SSN) model is a spatially explicit model that focuses on understanding spatial structures found within riverine Dendritic Ecological Networks (DENs) (Peterson et al., 2013; Ver Hoef et al., 2006; Ver Hoef & Peterson, 2010). Riverine systems, like cave networks and plant structures, have a network structure that becomes more complex as you move up from the terminus of the network, but terrestrial geospatial methods do not fully capture the complexities of riverine DEN structure and hydrologic flow connectivity (Isaak et al., 2014; Peterson et al., 2013; Ver Hoef et al., 2006; Ver Hoef & Peterson, 2010). SSN models account for stream topography, which allows these models to use proportional weighting of predictor variables (e.g. land cover), based on the relative size of upstream catchment areas and based on the distance of predictor variables to points on a stream network (Peterson et al., 2013). Previous studies suggest that a network aware model can improve model quality and predictive power, allowing for more accurate predictions of in-stream response

variables relative to a Euclidean model (Isaak et al., 2009; Lois, 2016; Peterson et al., 2006; Peterson & Urquhart, 2006; Steel et al., 2016; Turschwell et al., 2016). These findings explain the rise in popularity of the SSN methodology.

The use of Spatial Stream Network models has increased over the past 10 years, particularly in the areas of water chemistry (Gardner & McGlynn, 2009, McGuire et al., 2014), biotic abundance (Lois, 2016), and in-stream temperature (Isaak et al., 2009; Steel et al., 2016; Turschwell et al., 2016). Peterson & Urquhart (2006) and Peterson et al. (2006) compared SSN methodologies to non-spatial and other geospatial techniques and found a general improvement in in-stream water quality predictions so long as the density of the sampling points is high enough relative to the size of the watershed. Gardner & McGlynn (2009) utilized SSN methodology to determine that the relationship between water nitrate concentrations and land cover can be variable depending on season of observation: biological variables were the key predictor in the growing season and anthropogenic variables were dominant in the dormant season. Isaak et al. (2009) found a significant improvement in the ability to accurately predict stream temperature when using SSN methodologies, and also found improvements in spatial inferential power using SSN methods as compared to previous studies that evaluated the impacts of climate change on stream temperature. Lois (2016) utilized SSN methodology to assess both upstream and downstream contributions to mussel abundance and found that more variation in her data was explained by considering a downstream-to-upstream perspective which is enabled by SSN methods. Both Turschwell et al. (2016) and Steel et al. (2016) used SSN methods to determine the implications of aggregating near-continuous temperature data into specific time-based stream temperature metrics (e.g. average daily temperature & average monthly temperature). Turschwell et al. (2016) found that SSN models based on aggregate data were better than both random forest models and non-spatial models in determining thermal magnitude and duration, and in determining the frequency of heating events. Steel et al. (2016) found that relationships between landscape predictors and stream temperature varied through the

study period and that different spatiotemporal complexities in the thermal regime of a stream network can be revealed depending on what temperature data metrics are used. McGuire et al. (2014) used SSN models to show that significant multiscale differences exist in the spatial structure of in-stream anion (including Cl<sup>-</sup>), cation and heavy metal compositions, along with other water quality attributes, in a single dendritic riverine network. Their intense study of the Hubbard Brook watershed showed that spatial variation was more pronounced when comparing Euclidean distance methods with methods that use distances measured along the stream network. McGuire et al. (2014) also noted that better delineation of spatial structure via SSN methodologies can allow for more accurate predictions of stream chemistry in unsampled areas.

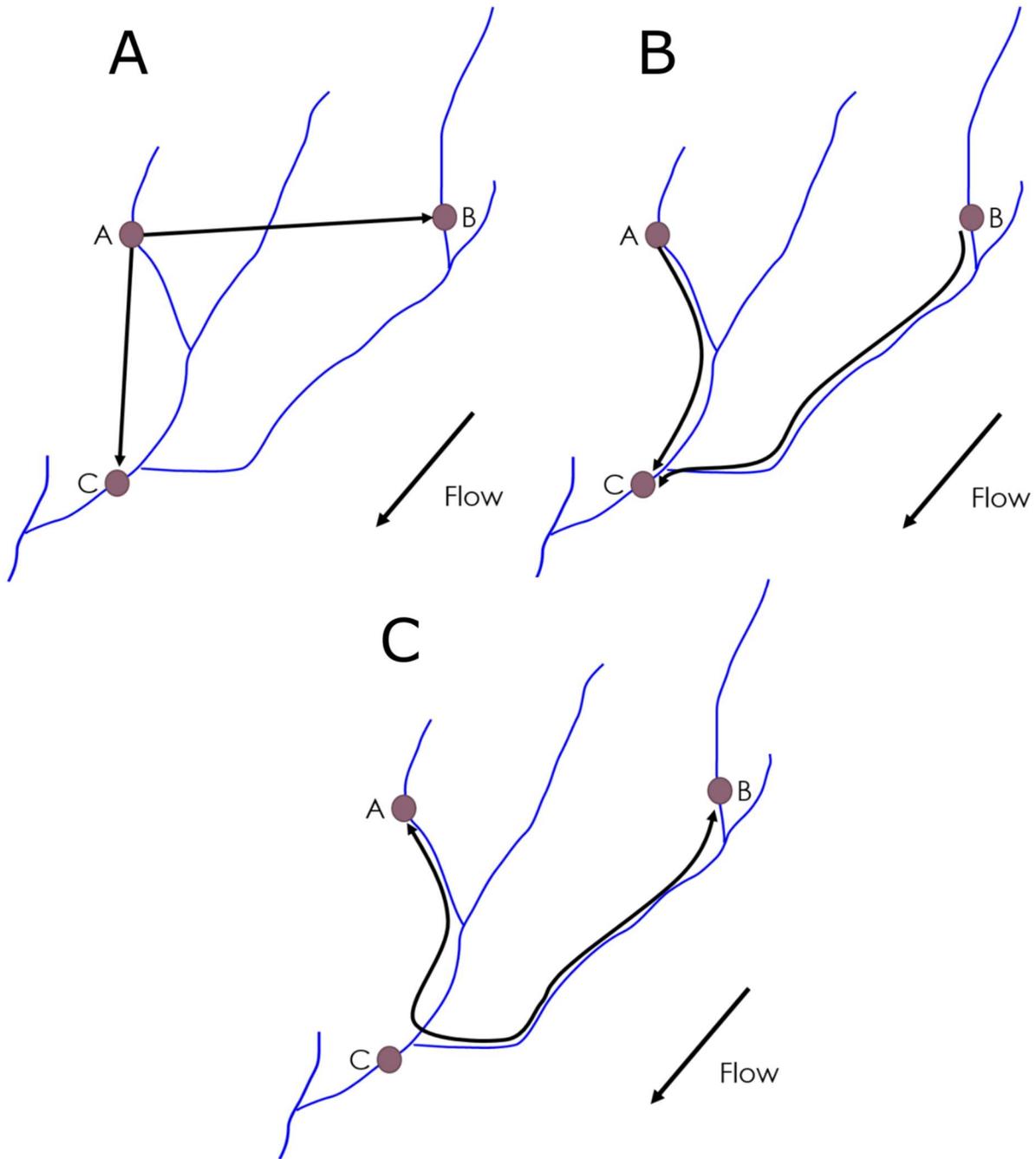


Figure 1: Three different methods to measure separation distance between sample points. (A) depicts a Euclidean (straight line) method of measuring sample distance between points A and B or A and C. (B) depicts a “Flow Connected” method to measure between sample points whereby only the sample points that can be connected and measured using the flow paths of the network are considered. (C) depicts a “Flow Unconnected” method to measure between sample points whereby the sample points that are not connected via the flow pathways of the network are considered.

## 2.9 Semivariance and Semivariograms

Semivariance is the average variance between any two observation points as a function of separation distance, and a plot of semivariance versus separation distance is known as a semivariogram (McGuire et al., 2014; Zimmerman & Ver Hoef, 2017). A semivariogram is an empirical tool that allows for study of multiscale geospatial structure within a stream network when using a longitudinal, synoptic data collection method (McGuire et al., 2014). The nugget, sill and range are depicted on the example semivariogram in Figure 2 where both the empirical semivariogram (observed semivariance shown in red) and the model semivariogram (black solid line) are shown. The y-intercept of the model semivariogram defines the *nugget*, which is an estimate of the average difference of a) the values of points at smaller-than-sampled separation distances, and b) sampling error (Isaak et al., 2014; Peterson et al., 2013; Ver Hoef et al., 2006). The *sill* represents the semivariance at which you reach spatial independence (*i.e.* where the curve flattens out), and the *range* is the separation distance required to reach spatial independence (Isaak et al., 2014; Peterson et al., 2013; Ver Hoef et al., 2006).

Differentiating between the different types of separation distances can have a noticeable impact on the structure of a semivariogram as shown in Figure 3. McGuire et al. (2014) reasoned that semivariograms created using only flow connected points highlight upstream spatial structure that is influenced by hydrologic connectivity, whereas Euclidean and flow unconnected methods can reveal the spatial structure related to landscape influences (e.g. land cover & geology). McGuire et al. (2014) also showed that multiscale patterns of in-stream water chemistry that appear on a semivariogram can be due to the spatial distribution of observed chemistry values within a stream network, and this relationship between semivariance and geospatial distribution is idealized in Figure 4 as adapted from McGuire et al. (2014).

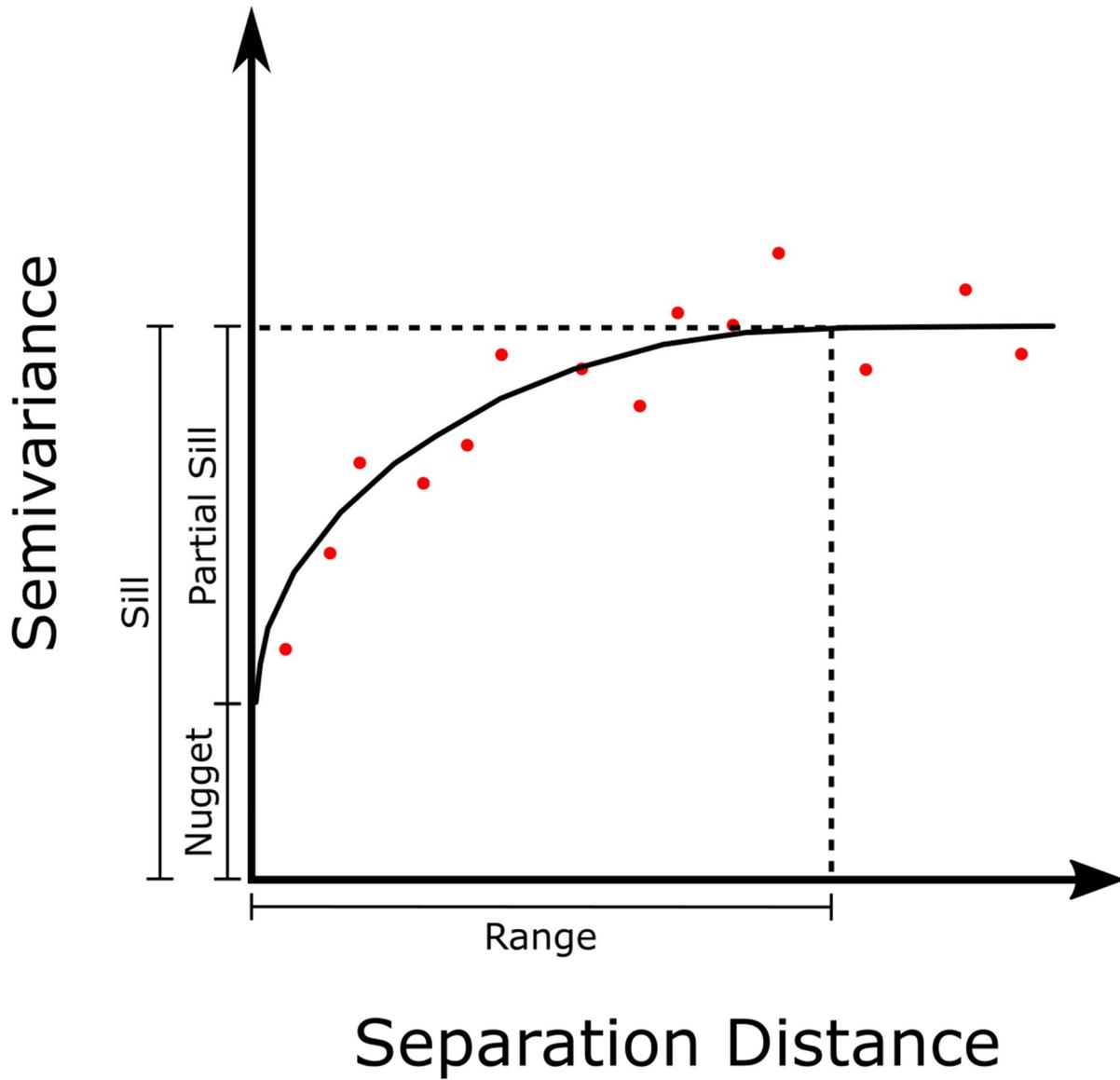


Figure 2: Example semivariogram with a labeled nugget, sill, partial sill, and range.

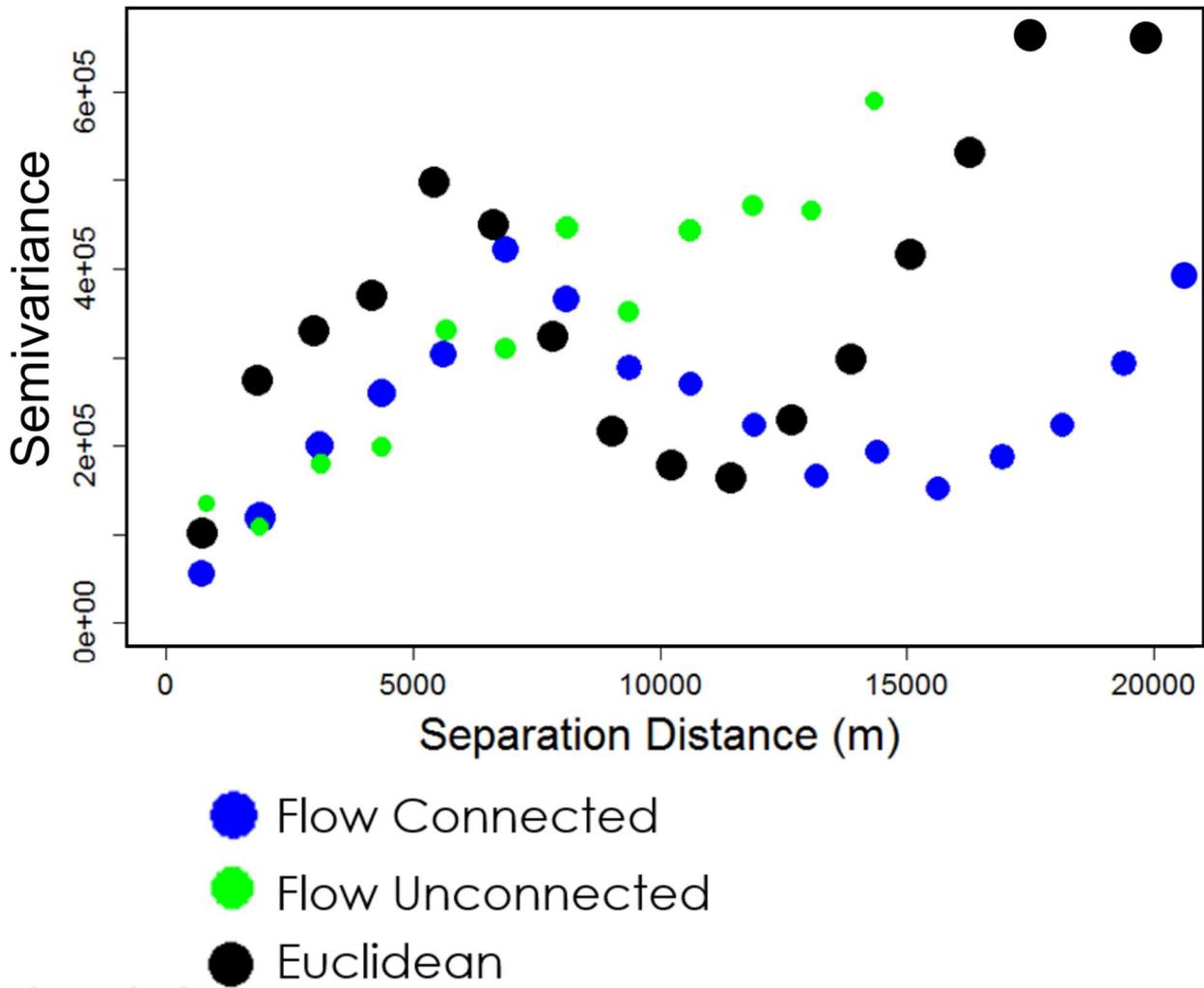


Figure 3: An example semivariogram produced using data collected along Mimico Creek watershed, found in the western Greater Toronto Area. Please refer to Figure 2 for a description of “Flow Connected”, “Flow Unconnected”, and Euclidean measurement methods.

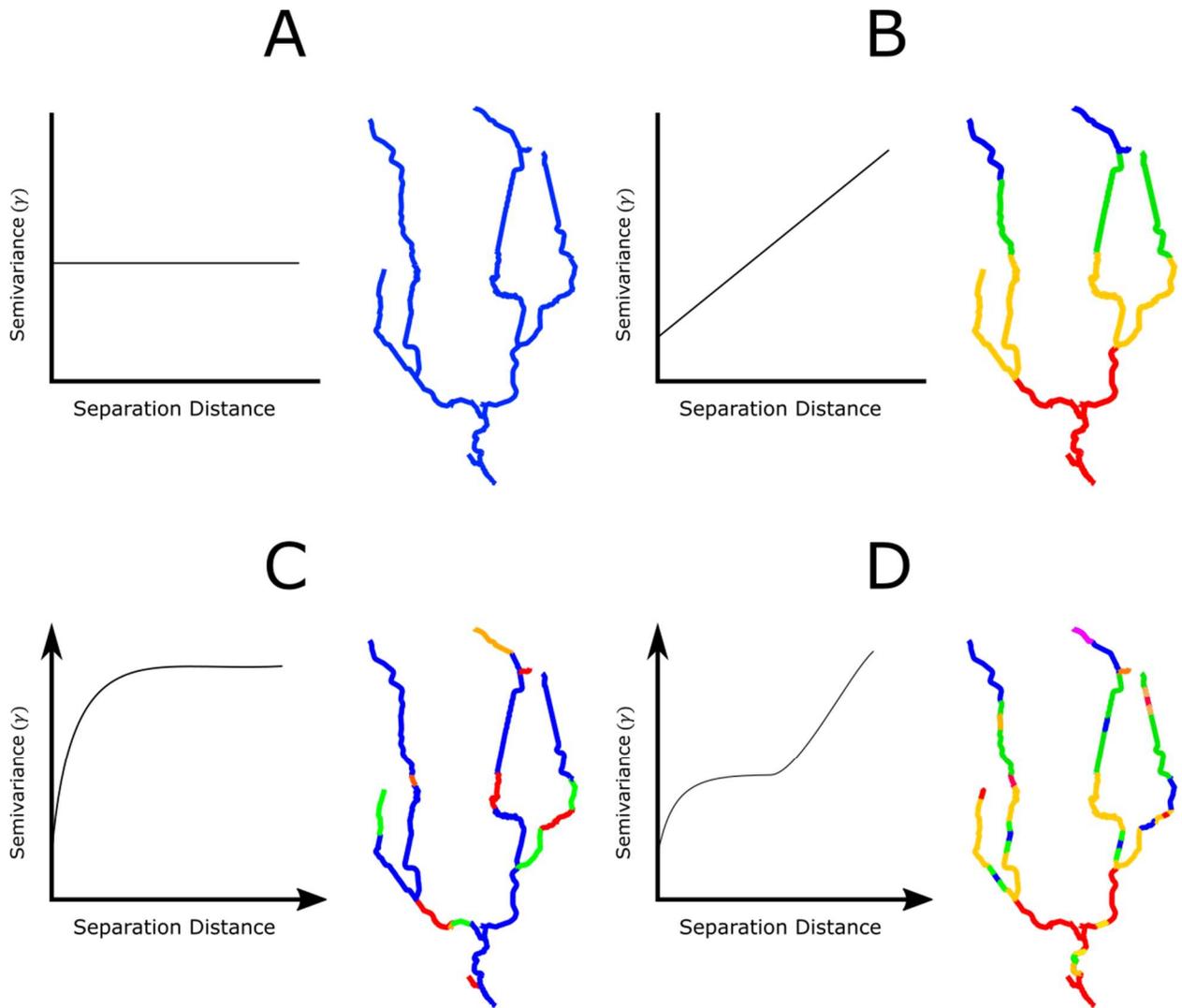


Figure 4: Figure adapted from McGuire et al. (2014) that depicts the varying scale of effects using theoretical semivariograms over a cropped and simplified stream system for the Mimico Creek watershed, located on the western side of the Greater Toronto Area. The colors for each stream network depict a value for an arbitrary in-stream water chemistry value (e.g.  $\text{Cl}^-$  concentration). (A) shows a flat, spatially independent semivariogram related to a stream network that is completely uniform or completely random. (A) shows that there is no small scale or large scale structure within the stream network. (B) shows a linear semivariogram associated with a large scale linear change to the in-stream chemistry variable. (C) shows a semivariogram depicting small scale patches of differing in-stream water chemistry values with no large scale pattern. Note that the (C) semivariogram has a nugget, range and sill, with the range being relatively close to the y-axis (short separation distance). (D) depicts a mixed semivariogram showing both small scale and large scale patterns in variance. The (D) semivariogram is associated with small scale patchy water chemistry values on top of a larger scale gradient in water chemistry values.

## 2.10 Thesis Aims

Given the potential for negative environmental impacts from anthropogenic  $\text{Cl}^-$  inputs to watersheds there is a need to identify localized areas of high  $\text{Cl}^-$  concentrations so that road authorities can minimize the use of road salt in salt vulnerable areas, thereby preventing environmental damage and the costs associated with environmental remediation. Previous studies have examined  $\text{Cl}^-$  dynamics at regional or watershed scales to identify key landscape drivers of elevated stream  $\text{Cl}^-$  concentrations (Kaushal et al., 2005; Kelting et al., 2012; Morgan et al., 2012) resulting in large-scale correlative models. Only one study by McGuire *et. al.* (2014) utilized spatially intensive water sampling to develop predictive maps of in-stream  $\text{Cl}^-$  concentrations at the reach scale (i.e. the scale of individual stream lengths).

The overall goals of this study were to: i) explore the spatial structure of in-stream  $\text{Cl}^-$  concentrations in three watersheds that span a gradient in urbanization across seasons and flow states and ii) assess the utility of potential landscape predictors of salt vulnerable areas. The Spatial Stream Network (SSN) model, a novel geospatial technique developed by Ver Hoef et al. (2006), was used to develop high resolution models of in-stream  $\text{Cl}^-$  concentrations across three study watersheds in Southern Ontario. Spatially intensive in-stream field data were collected in four seasons and used to develop the models. The specific objectives of the study were to: 1) evaluate variability in spatial patterns of stream  $\text{Cl}^-$  concentrations across seasons, 2) explore whether or not consideration of flow connectivity could improve our models of the influence of landscape characteristics on stream  $\text{Cl}^-$  concentrations, 3) compare importance of landscape predictors of stream  $\text{Cl}^-$  concentrations across seasons and watersheds, and 4) assess if residual semivariograms provide additional insight into how spatial patterns of stream  $\text{Cl}^-$  concentrations change with season.

## 3 Methods

### 3.1 Study Watersheds

#### 3.1.1 Mimico Creek Watershed

Mimico Creek Watershed has an area of 77 km<sup>2</sup> and extends north from Lake Ontario along the borders of Toronto, Brampton, and Mississauga (Figure 5). The watershed is dominated (96 %) by urban land cover with the remaining 4% considered to be in some stage of urbanization (TRCA, 2013). The mean annual flow of Mimico Creek is 0.766 m<sup>3</sup>/s and the average annual precipitation is 819 mm (OMNRF, 2013). The watershed received failing grades in the Toronto and Region Conservation Authority (2010) report card for surface water quality, forest conditions, and stormwater management. Mimico Creek has elevated levels of total phosphorus, elevated *E.coli* bacterial levels from human or animal waste, limited natural surface cover (e.g., forests and wetlands), and channelized stream beds contributing to “flashy” hydrologic responses to precipitation events (TRCA, 2013).

The surficial geology of Mimico Creek is composed of four major units (Figure 6; OGS, 2010): An alluvial gravel, sand, and silt layer with variable permeability comprises much of the Mimico Creek streambed. A low permeability silt and clay layer forms the northern third of the watershed, followed by a clay-silt to silt layer that continues until a highly permeable gravelly sand to silty layer forms near the outlet of the watershed.

This system was of interest for this thesis as 1) the watershed is almost completely urbanized and was therefore considered as an urban extreme watershed, 2) the simple topography of the Mimico Creek stream network lent itself to sampling the entire stream network over the course of each field survey, and 3) water quality monitoring was already being carried out in Mimico Creek by researchers at the University of Toronto Scarborough and these measurements were used as a baseline reference to develop electrical conductivity – chloride relationships.

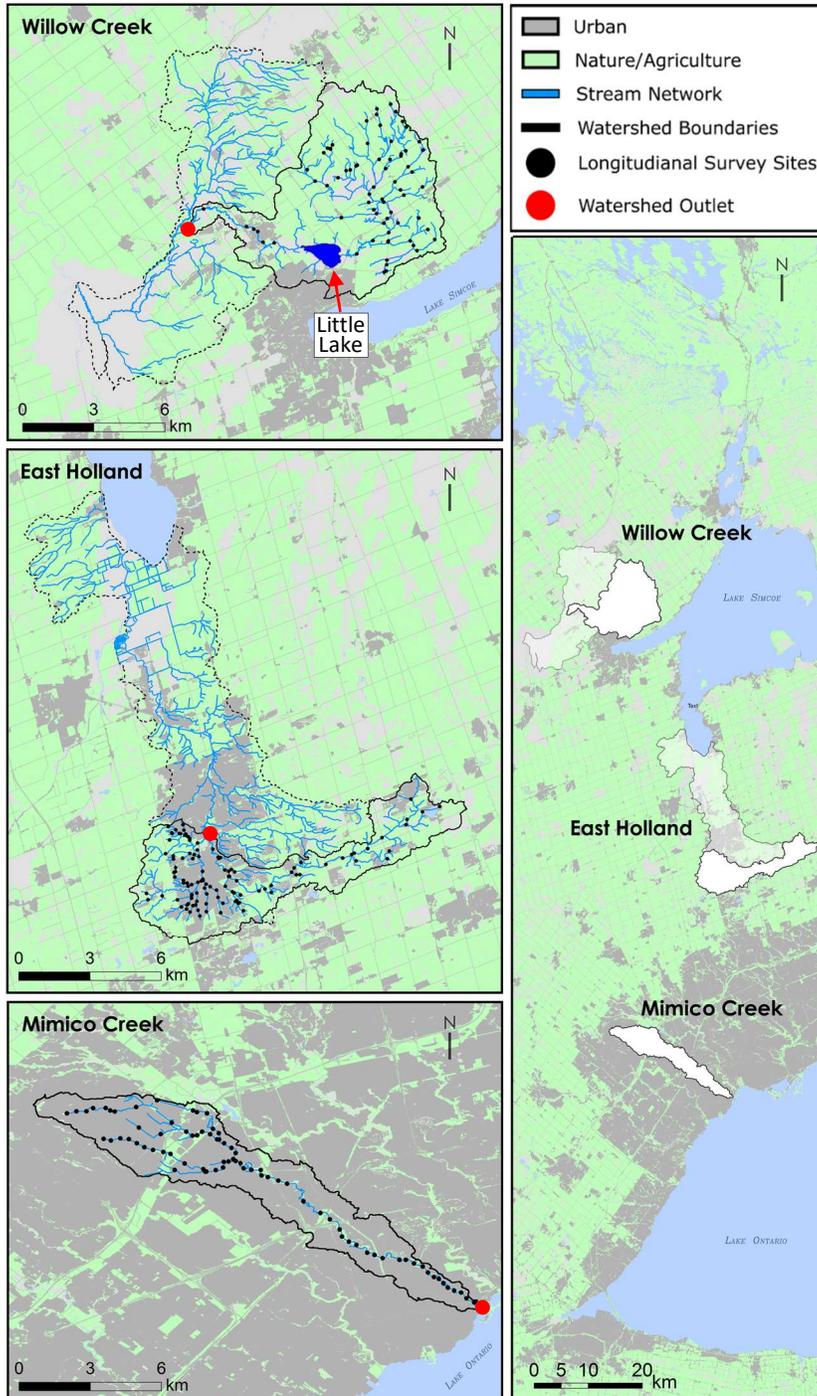


Figure 5: Maps delineating the Mimico Creek, East Holland River, and Willow Creek study watersheds with longitudinal sampling points marked along the stream network. The dotted lines represent the full East Holland River and Willow Creek watersheds that extend beyond the study areas. Land cover data was sourced from the Ontario Land Cover Compilation v.2 (OMNRF, 2016a); stream networks were sourced from the Ontario Integrated Hydrology Layer (OMNRF, 2015); watershed boundaries were sourced from the Ontario Flow Assessment Tool (OMNRF, 2013). Reproduced and adapted with permission from Nikolas McGlashan, Ryerson University.

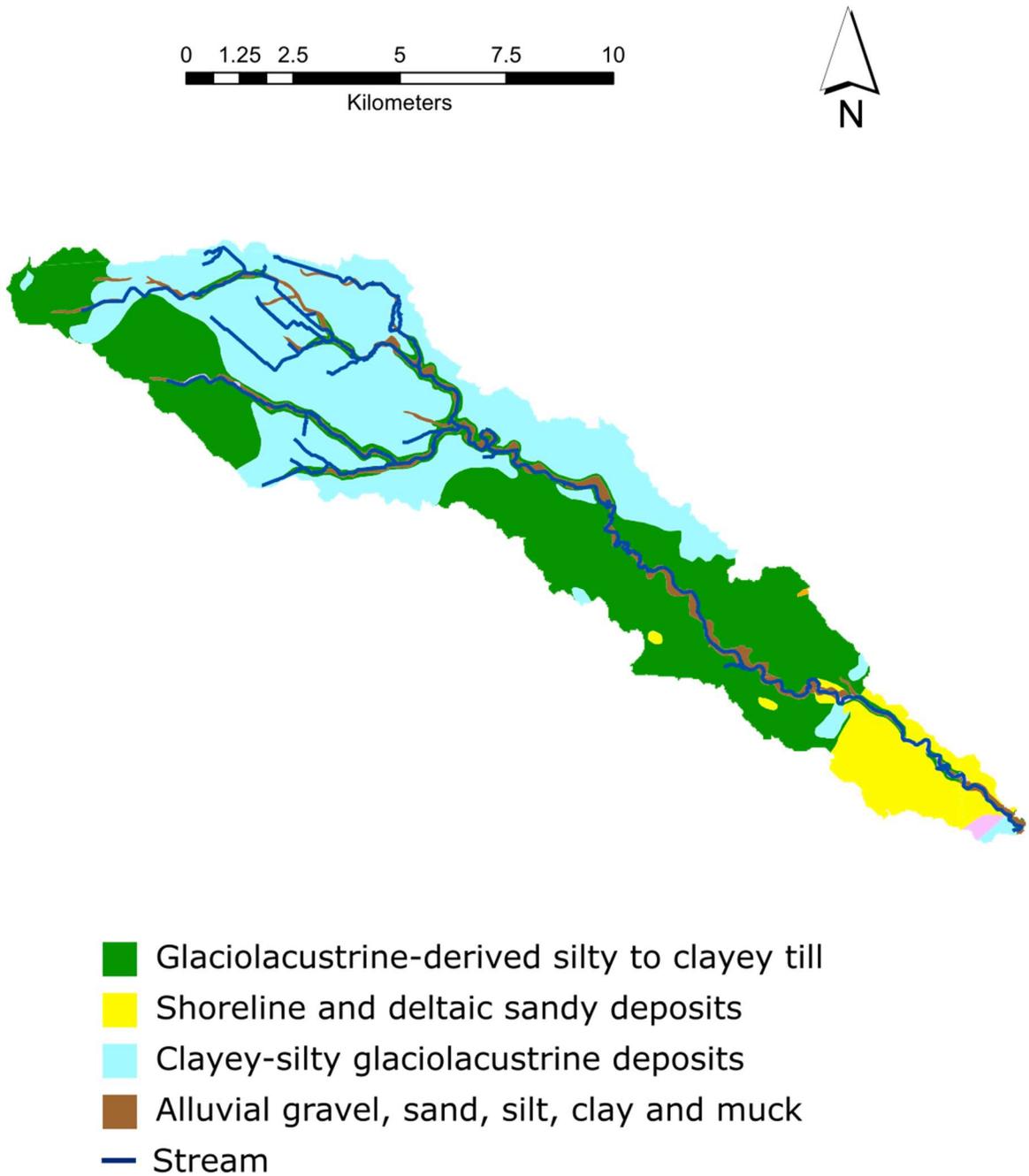


Figure 6: Map of the surficial geology of the Mimico Creek watershed. Map was created from the Surficial Geology of Southern Ontario dataset (OGS, 2010).

### 3.1.2 East Holland River Watershed

East Holland River Watershed is a 247 km<sup>2</sup> L-shaped watershed that drains into Lake Simcoe and is located almost entirely within the Regional Municipality of York (Figure 5). This watershed is approximately 17 % urban with approximately 20% of the total watershed area being impervious (LSRCA, 2010). This level of imperviousness has been linked to reduced groundwater volume and base flow due to reduced surface water infiltration, erosion of stream banks, the addition of contaminants from impervious surface run off (e.g. road salt & heavy metals), and negative impacts on habitat and biodiversity (LSRCA, 2010). Agricultural practices that mainly occur on the northern and southeastern parts of the watershed have also resulted in environmental impacts like riverbank erosion and excess nutrient loading (LSRCA, 2010). The combination of urban expansion and agricultural activities has made the East Holland River Watershed the largest contributor of phosphorous and Cl<sup>-</sup> to Lake Simcoe (LSRCA, 2010; Winter et al., 2011).

Due to the size of the watershed and need to conduct longitudinal surveys within a consecutive day period with no precipitation inputs, a sub-watershed of the East Holland River Watershed was focused on in this study. The East Holland River sub-watershed is located in the southern portion of the East Holland River Watershed and has an area of 85 km<sup>2</sup> with relatively easy access to the stream network, and has a mean annual flow of 0.741 m<sup>3</sup>/s and the average annual precipitation is 834 mm (OMNRF, 2013). The East Holland River also features a near equal mix between urban (54.4%) and non-urban (45.6%) areas (OMNRF, 2016), which allows for a classification as the “semi-urban” watershed for the purposes of this thesis as it has a relatively equal urban to non-urban area.

The East Holland River watershed has variable geology (Figure 7; OGS, 2010) with the eastern half of the watershed primarily composed of permeable sandy deposits on the northern side of the stream network less-permeable clayey silt to sandy silt on the southern side, the western side of the

watershed is centrally composed of glaciolacustrine silt and clay with sandy deposits on the western and southern periphery, and the streambed is composed of a mix of sand, clay, silt, and gravel.

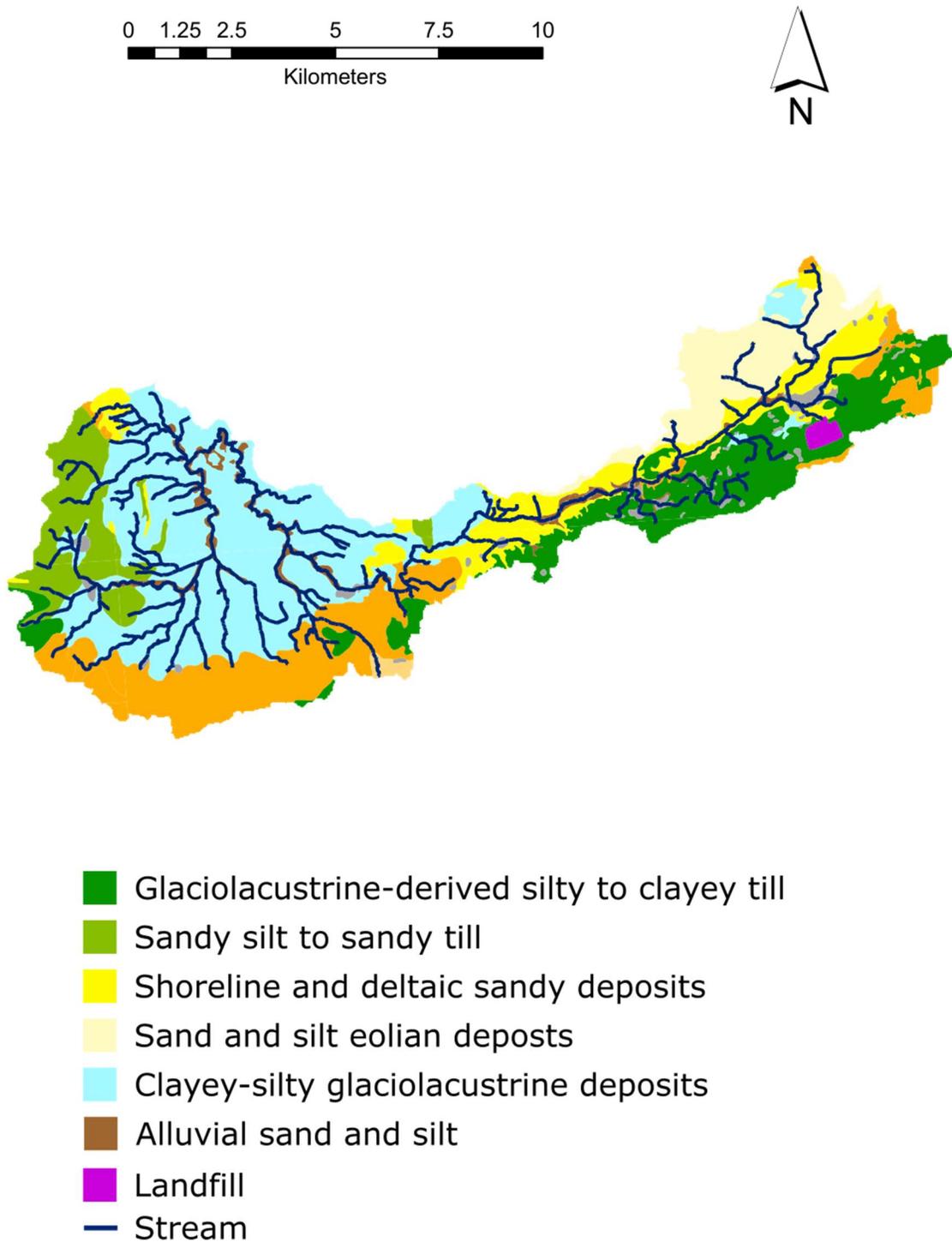


Figure 7: A geological map of the East Holland River watershed study area. Map was created from the Surficial Geology of Southern Ontario dataset (OGS, 2010).

### 3.1.3 Willow Creek Watershed

Willow Creek watershed (Figure 5) has an area of 338 km<sup>2</sup> and is part of the larger Nottawasaga Valley watershed which drains into the Georgian Bay of Lake Huron (NVCA, 2013). This watershed is approximately 8% urban and 43% Agricultural/Rural, and it contains the small community of Midhurst and only a northern section of the City of Barrie (OMNRF, 2016).

A smaller sub-watershed was selected within the Willow Creek watershed for the same reasons outlined above for East Holland River. This sub-watershed is 137 km<sup>2</sup> and comprises one of the two major tributary networks within the full watershed, and has a mean annual flow of 0.960 m<sup>3</sup>/s and the average annual precipitation is 941 mm (OMNRF, 2013). The Willow Creek sub-watershed area is 15.4% urban and 84.6% non-urban, contains a seven km<sup>2</sup> lake, known as Little Lake, towards the sub-watershed outlet, and was largely accessed at stream roadway crossings (OMNRF, 2016). Due to the relatively low amount of urban area, the Willow Creek sub-watershed was classified as “non-urban/rural” for this study.

The geology of Willow Creek (Figure 8) is largely comprised of stony to sandy/till silt with low-medium permeability, with highly permeable shoreline and deltaic sandy deposits running parallel to much of the central downstream stem of the stream network (OGS, 2010). The headwaters of the stream network overlays with sporadic glaciolacustrine clays and silts, Organic deposits, and with variably grained sand/gravel with silt/clay tills (OGS, 2010).

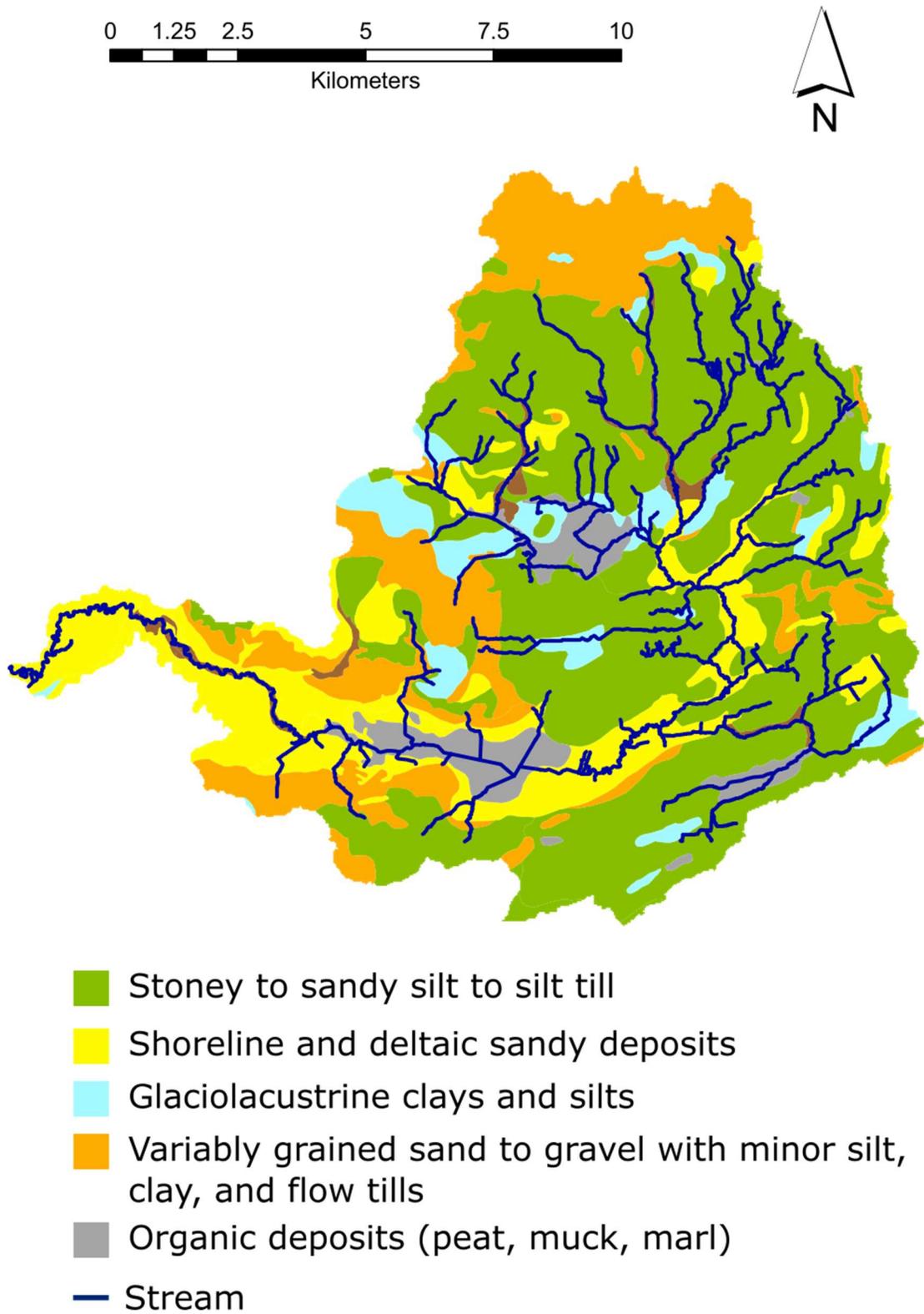


Figure 8: A geological map of the Willow Creek watershed study area. Map was created from the Surficial Geology of Southern Ontario dataset (OGS, 2010).

## 3.2 Field Data Collection

### 3.2.1 Developing Electrical Conductivity vs. In-Stream Chloride Relationships

In lieu of an instrument that would allow for a direct measurement of in-stream  $\text{Cl}^-$  concentrations, an electrical conductivity (EC) vs. in-stream chloride ( $\text{Cl}^-$ ) relationship was derived for each study watershed to allow conversion of EC to  $\text{Cl}^-$  with a reasonable degree of certainty. Although other ions and dissolved materials can influence electrical conductivity, ion composition studies can determine the exact  $\text{Cl}^-$  concentration within the water and previous studies have developed an EC vs  $\text{Cl}^-$  relationship with success (Howard & Haynes, 1993; Perera et al., 2010). All absolute electrical conductivity values collected in this thesis were converted to specific conductivity (conductivity at  $25^\circ\text{C}$ ). This conversion to specific conductivity was a feature in both the Solinst LTC Levelloggers and the YSI EXO water quality Sonde EC Probe that were used to collect data for this study.

Continuous in-stream EC information were collected and paired with water samples that were analyzed for  $\text{Cl}^-$  ion concentrations within the East Holland River and Willow Creek watersheds. Solinst LTC Levelloggers (Figure 10b) were installed in four locations within the East Holland River watershed (Figure 9b) and three locations within the Willow Creek Watershed (Figure 9c) over the 2016 Winter season. All levelloggers were initially set to capture hourly in-stream EC, but were later reprogrammed in the 2016 Fall season to measure EC values in 15 minute intervals. The levellogger sites in both watersheds were selected based on potential for round-year access, likelihood for continuous flow in the summer and winter, and on permissions obtained for continuous access and levellogger installation. These sites were also distributed so headwater and downstream water quality data could be obtained. Physical water (grab) samples were collected on a bi-weekly (2-week) basis at the location of each levellogger to determine the exact stream ion composition via a Dionex ICS-3000 Ion Chromatography System with a Dionex IonPack AS18 RFIC 4x 250 mm analytical column, and a Dionex IonPack AG18 RFIC

4 x 50mm guard column. All ion composition analysis were conducted by the Department of Geography and Environmental Management at the University of Waterloo. Chloride and EC data from a Provincial Water Quality Monitoring Network (PWQMN) station, located within the East Holland River Watershed (Figure 9b), was paired with ion composition and Solinst LTC data to create a more robust dataset. No PWQMN station was found in the Willow Creek study area.

In-stream levelloggers were installed either in the deepest part of a shallow stream or approximately 1 to 3 meters away from deeper river banks as measured during low-flow periods. This proximity to the shore during the low-flow season permits the levelloggers to remain at a wadeable depth during high-flow periods. In-stream levelloggers were housed in a custom PVC casing with perforations that permitted unimpeded waterflow and the casings were secured onto a metal t-post with u-bolts and wire (Figure 10a). The casings were examined and cleaned during each biweekly field trip.

Five *in-situ* Electrical Conductivity stations were installed and maintained in the Mimico Creek watershed by the Mitchell Research Group at the University of Toronto Scarborough (Figure 9a). These water quality stations used the Hydrolab DS5X instrument to collect conductivity values on an hourly basis. Physical water samples were collected over a period of two years in all four seasons and across high and low flows. Ion composition of the water samples were determined using a Perkin Elmer Series 200 High Performance Liquid Chromatography system. Stations G1 and G2 monitor the primary tributaries that feed into the Mimico Creek main stem, while stations G3, G4, and G6 monitor the main stem from the major upstream confluences to the downstream outlet.

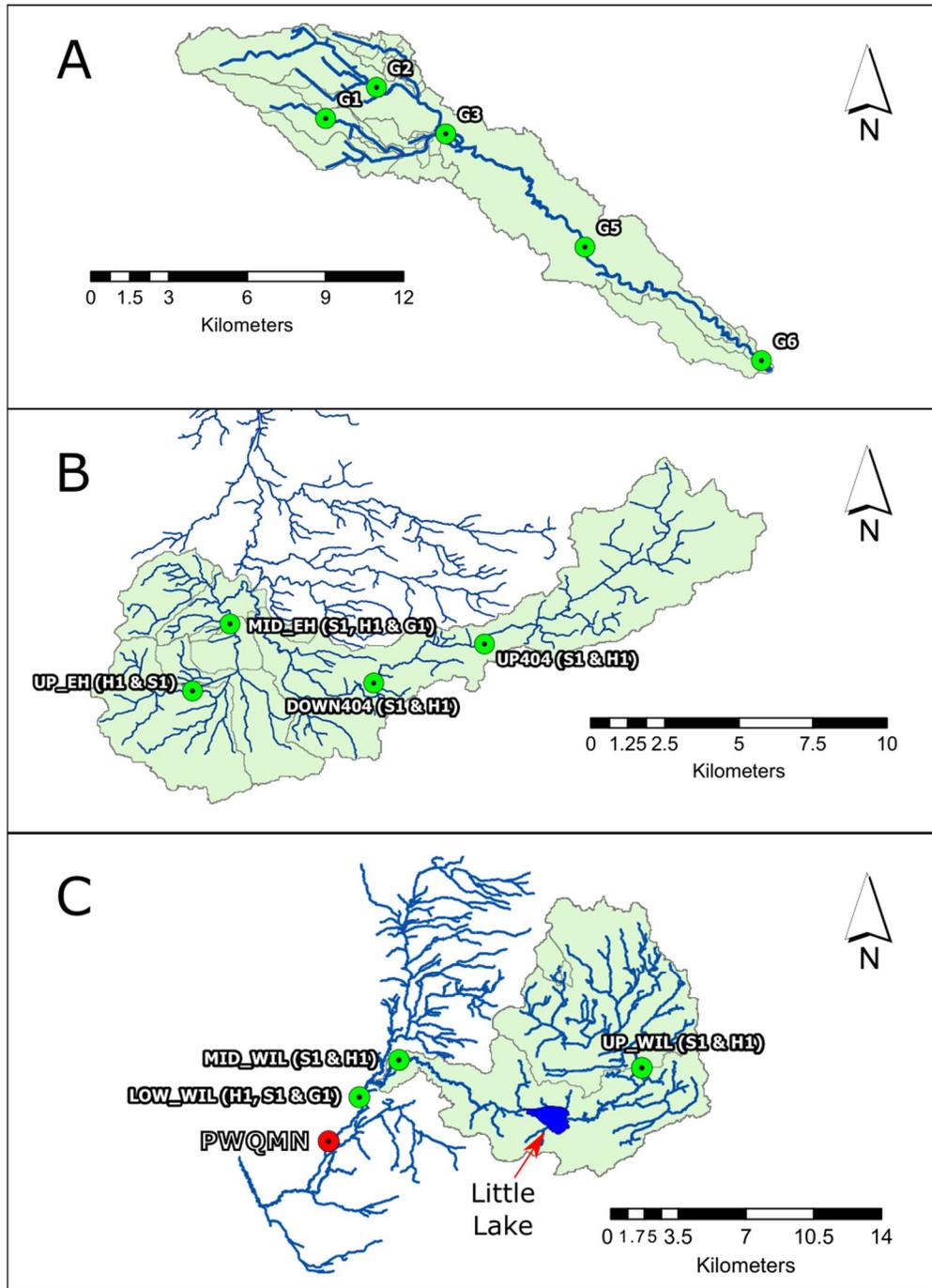


Figure 9: The green markers denote the locations of the in-stream electrical conductivity (EC) stations for A) Mimico Creek Watershed, B) East Holland Watershed, and C) Willow Creek Watershed. The EC stations in Mimico Creek were installed and maintained by the Miller Research Group at the University of Toronto. The EC stations in East Holland and Willow Creek were installed and maintained by the Oswald Research Group at Ryerson University. The Provincial Water Quality Monitoring Network (PWQMN) stations (marked in red) were installed and maintained by the Government of Ontario, with the datasets for these stations made available to the public. The watershed areas used in this thesis are delineated by green polygons.



Figure 10: A) Custom made levellogger casing anchored in position and B) a Solinst LTC Levellogger from the UP\_404\_S1 site in East Holland.

### 3.2.2 Seasonal Longitudinal Surveys

High intensity longitudinal surveys within each watershed stream network were conducted to capture a snapshot of Cl<sup>-</sup> concentrations across the entire study area within a span of two to five days. This snapshot provided a high geospatial density of electrical conductivity (EC) measurements within relatively stable periods of weather with no precipitation. An effort was made to keep the distance between sampling points 500m or less, but sample size was constrained by ease of access and by permission to access non-public lands. A YSI EXO water quality sonde with an attached EC sensor, calibrated before each field season, was used to collect in-stream conductivity data. The EC sensor was thermoequilibrated in the stream for 30 to 60 seconds before the EC measurements were recorded. If the stream was too shallow, a sample of the water was carefully collected in a cup which was used to collect EC measurements.

The flow state for each seasonal survey was determined using a flow exceedance curve, which is a commonly used technique when interpreting data that is categorized over time intervals (Risley et al., 2009). All daily flow data for the Mimico Creek, East Holland, and Willow Creek watersheds were obtained from the Environment and Climate Change Canada Historical Hydrometric Data (ECCC, 2017). Guided by Risley et al. (2009), the following equation was used to calculate the flow exceedance curves for each watershed:

$$P = 100 * \frac{m}{n + 1} \quad (1)$$

where P is the probability (calculated as a percentage) that a given flow will be equal or exceeded within the bounds of the dataset, m is the ranking of a given flow value amongst the dataset, and n is the total number of flow values within the dataset (Risley et al., 2009).

For the purposes of this thesis, a low flow state was defined when the majority of a survey's daily flow values had a  $P < 25\%$ ; the low-medium flow state was defined when the majority of a survey's daily

flow values had a  $P$  that was  $\geq 25\%$  and  $< 50\%$ ; the medium-high flow state was defined when the majority of a survey's daily flow values had a  $P$  that was  $\geq 50\%$  and  $< 75\%$ ; and the high flow state was defined when the majority of a survey's daily flow values had a  $P \geq 75\%$ .

### 3.3 Spatial Stream Network (SSN) Models and Residual Semivariogram Interpretation

Spatial statistical network (SSN) modelling, as introduced by Ver Hoef et al. (2006), is a relatively new methodology that incorporates network topography and stream directionality to improve analysis of riverine datasets. Using the work by Ver Hoef et al. (2006), Peterson et al. (2013) outlines how SSN modelling utilizes well-known multiple linear regression which has the following general form in matrix notation:

$$y = \beta X + z + \varepsilon, \quad (2)$$

where  $y$  is a  $1 \times n$  vector containing the values of the response variable and  $n$  is the number of observations. Matrix  $X$  is the  $n \times p$  design matrix of environmental predictor variables (also referred to herein as covariates), and  $\beta$  is the  $1 \times p$  vector of regression coefficients. The spatially structured component of error,  $z$ , is a vector of random spatially correlated variables, and  $\varepsilon$  represents a  $1 \times n$  vector of random errors that are considered to be spatially independent (Peterson et al., 2013). The variance of  $z$  gives a covariance matrix,  $\Sigma$ , which provides  $n(n+1)/2$  parameters to estimate, with the elements on the diagonal representing the variance of each observation point, and the off-diagonal elements representing the covariances between different observation points (Isaak et al., 2014; Peterson et al., 2013). The number of parameters can be reduced by using an autocovariance function to measure the covariance of the values within  $z$  as a function of separation distance (Ver Hoef et al., 2006; Peterson et al., 2013; Isaak et al., 2014). This provides 3 parameters: the nugget effect, the partial-sill, and the range, all of which

describe the spatial structure of the values from the spatial stream model, which were also previously described in Section 2.9 (Ver Hoef et al., 2006; Peterson et al., 2013; Isaak et al., 2014).

To clarify how SSN modelling accounts for spatial structure in a flow-connected hydrologic context, the description by Ver Hoef et al. (2006, 2014) and Ver Hoef & Peterson (2010) defines how SSN utilizes a weighted moving average function that operates along the stream length in the up-stream and/or down-stream direction. A moving average function calculates the average of a set of points as a function of distance from a point  $n$ , and after this average is calculated, the function moves to the next point,  $n+1$ , and calculates another average for this second subset of points (Figure 11; Ver Hoef et al., 2006; Ver Hoef & Peterson, 2010). This moving average function continues until it has generated an average for all possible subsets of points under consideration within the system. The moving average function also utilizes a weighting function, usually with the points closest to  $n$  having the most weight in determining the average, whereas points further away from point  $n$  have less influence on the calculated average. The moving average function can be applied in either a “tail-up” fashion, where the moving average function only considers points that are up-stream of the current location (Figure 12), or they can be applied in a “tail-down” fashion, where the averaging function considers only points downstream of the current location (Figure 12). This distinction in directionality is important as the tail-up method looks at potential influences that are upstream of a point on a stream, while the tail-down method looks for downstream influences of a point. When a tail-up autocovariance function comes across a confluence in a stream network, a weighting strategy is used to determine the proportional influence that each contributing stream segment has on any downstream point. This proportional weighting scheme is additive in that it incorporates the upstream influence of any stream segment and not just the immediate segments that contribute to any confluence, allowing for a more nuanced consideration of upstream variables with respect to hydrological distances and complexity. This additive function is further described by Ver Hoef & Peterson (2010). Catchment area was the watershed

characteristic used in this thesis to derive the proportional weighting of stream segments that are upstream of a given stream point.

The overlapping area between moving average functions at points  $n$  and  $n+1$  is the area within which the covariance between the two points can be measured. The moving average functions described by Ver Hoef et al. (2014) outlines all of the weighted average functions utilized by the Spatial Stream Network package in R (R Core Team, 2017). The general linear model used for the purposes of this research is one that can incorporate tail-up, tail-down, and Euclidean models, and is further described by Ver Hoef and Peterson (2010).

The shape of a semivariogram, as described in Section 2.9, can be influenced by the method in which point pair separation distances are measured. This thesis used separation distance between points that share a common flow pathway (i.e. are flow-connected), as it is assumed that stream water chemistry at any given point is influenced by the characteristics of the watershed that is upstream of that point. Therefore, points that are flow-connected are likely to be more similar than points that do not belong to the same flow pathway (i.e. are flow-disconnected). Euclidean models were generated to provide a comparative reference for the SSN models, and flow-connected semivariograms created herein follow recommendations from Zimmerman & Ver Hoef (2017; Table 1) when setting the maximum distance between pairs of sample sites (site pairs) and the minimum number of site pairs required for a separation distance bin. All semivariograms were produced using the SSN package (Ver Hoef et al., 2014) for the R statistical software (R Core Team, 2017).

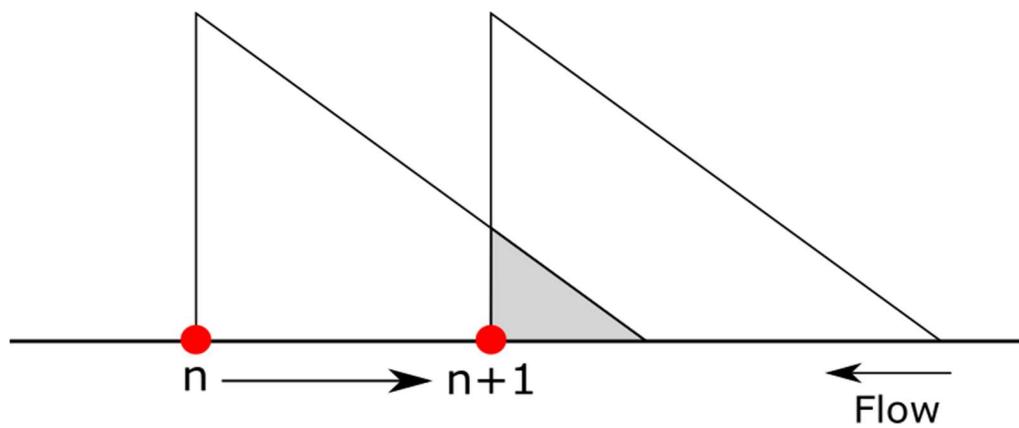


Figure 11: Depiction of the tail-up moving average function going from point  $n$  on a stream to point  $n+1$ . The overlap between the two average functions is where covariance of the two points is measured.

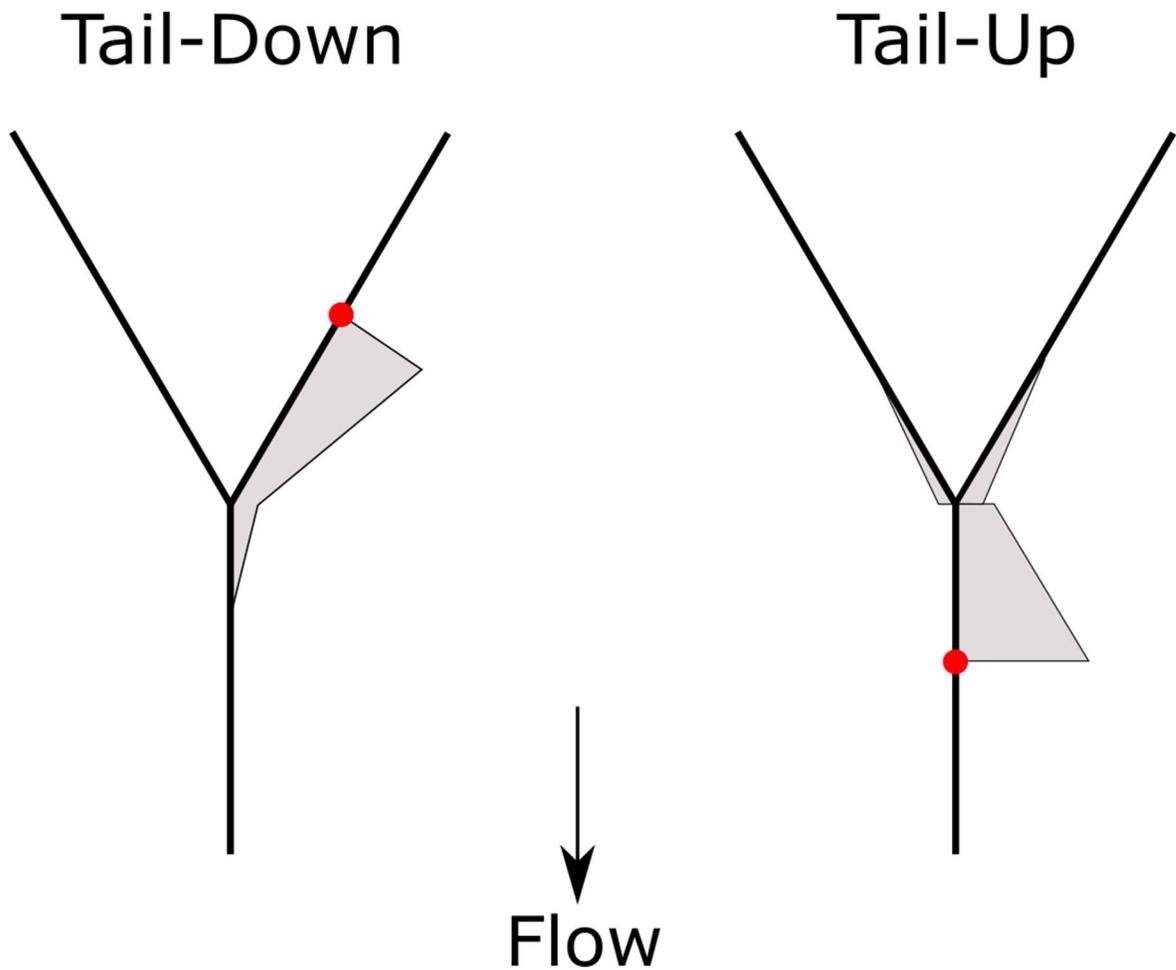


Figure 12: Depiction of the tail-up and tail-down averaging models and how each type of model deals with confluences. The tail down model continues past the confluence, and no splitting of the averaging function is necessary. The tail-up method requires a split of the averaging function to account for the proportional influence of each tributary.

Table 1: Semivariogram parameters for maximum site-pair distance and minimum number of site-pairs. Parameters set using Zimmerman & Ver Hoef (2017) .

<b>Flow Connected Semivariograms</b>	
Maximum Site-Pair Distance	$\frac{1}{2}$ Maximum Site-Pair Distance
Minimum Number of Site-Pairs	5*
<p>* Zimmerman &amp; Ver Hoef (2017) added the caveat that a minimum 5-10 site-pair flow connected semivariogram may not allow for the same level of confidence as that of the minimum site pair number in Euclidean semivariograms due to the decreased sample size.</p> <p>** Willow Creek S1 &amp; S2 does not meet this minimum requirement</p>	

### 3.4 Covariate Selection

To avoid the issue of overfitting in linear regression a 10 datapoint-per-covariate rule of thumb was applied to this study (Babyak, 2004). Applying this rule of thumb resulted in a maximum of three covariates for all surveys based on the lowest survey sample size (Table 2). This limit was applied to all surveys for consistency when comparing model results between Mimico Creek, East Holland River, and Willow Creek study watersheds.

Table 2: The sample size and dates for the Mimico Creek, East Holland River, and Willow Creek longitudinal surveys of in-stream electrical conductivity.

<b>Survey (S)</b>	<b>Season</b>	<b>Dates</b>	<b>Sample Size</b>
Mimico S1	Spring 1	Jun 14-16, 2016	73
Mimico S2	Summer	Aug 03-05, 2016	73
Mimico S3	Fall	Oct 11-13, 2016	68
Mimico S4	Winter	Feb 20-23, 2017	68
Mimico S5	Spring 2	Apr 21-23, 2017	69
East Holland S1	Summer	Jul 21-25, 2016	41
East Holland S2	Fall	Oct 24-26, 2016	54
East Holland S3	Winter	Mar 02-05, 2017	110
East Holland S4	Spring	May 15-17, 2017	109
Willow S1	Summer	Jul 28-29, 2016	35
Willow S2	Fall	Oct 18-19, 2016	49
Willow S3	Winter	Feb 06-08, 2017	54
Willow S4	Spring	Apr 17-18, 2017	68

### 3.4.1 Lane Length Density

Road salt on roadways is a significant contributor of  $\text{Cl}^-$  in watersheds that experience seasonal freezing temperatures as previously discussed in Section 1 and Section 2.1. To determine the magnitude of the relationship between roadways and in-stream  $\text{Cl}^-$  concentrations in different seasons, Lane Length Density (LLD) was calculated as a metric for the density of roadways within a given watershed catchment area. Lane Length Density was calculated as:

$$\text{Lane Length Density (Km}^{-1}\text{)} = \frac{\text{Length of Roadway (Km)} \times \text{\# of Lanes in the Roadway}}{\text{Catchment Area (n) (Km}^2\text{)}} \quad (3)$$

This simple metric accounted for the length and width of a road as measured by road length times the number of lanes a segment of road contains. This metric does not account for other road characteristics, like surface type (e.g., gravel or paved), roadside infrastructure (e.g. rural or urban), etc. The upstream LLD of any sample point was calculated using the catchment area for each sample point provided by the Ontario Flow Assessment Tool (OMNRF, 2013), along with the length of roadways and the number of lanes in a roadway as provided by the Ontario Road Network: Road Net Element dataset via Land Information Ontario (OMNRF, 2016b).

### 3.4.2 Subsurface Permeability

Surficial geology plays a role in the transport, storage, and release of  $\text{Cl}^-$  within a watershed as previously discussed in Section 2.3. As there were many geological units to consider within the study watersheds, and as there was a three covariate limit for all linear regression models in this thesis, the permeability of the rock types found in Southern Ontario was used to form a single weighted permeability metric. This weighted permeability metric was calculated using the geological data and the ordinal permeability scale provided in the Surficial Geology of Southern Ontario Dataset created by the Ontario Geological Survey (2010). Geological units are labeled as having “Low”, “Low-Medium”, “Medium”, “Medium-High”, “High”, or “Variable” permeability depending on the geological unit in

question. The weighted permeability function for any point,  $n$  on a stream,  $WeightedPermeability(n)$ , was calculated using the following equation,

$$\begin{aligned}
 WeightedPermeability(n) = & \\
 & Low * AreaLowPermeability(n) + \\
 & LowMedium * AreaLowMediumPermeability(n) + \\
 & Medium * AreaMediumPermeability(n) + \\
 & MediumHigh * AreaMediumHighPermeability(n) + \\
 & High * AreaHighPermeability(n)
 \end{aligned}
 \tag{4}$$

Where:

$$\begin{aligned}
 Low &= 1 \\
 LowMedium &= 2 \\
 Medium &= 3 \\
 MediumHigh &= 4 \\
 High &= 5
 \end{aligned}$$

where the constants  $Low, \dots, High$  were set in order according to the ordinal scale provided by the Surficial Geology of Southern Ontario dataset (OGS, 2010). The “Variable” permeability metric was not used due to the ambiguity of placement on an ordinal scale. The factors  $AreaLowPermeability(n), \dots, AreaHighPermeability(n)$  represent the proportion of area containing geological rock units of a given permeability within the entire upstream catchment area for any sampled point,  $n$ , on a stream.

### 3.4.3 Agriculture and Undifferentiated Rural Land Cover

Chloride-based fertilizers may act as a source for  $Cl^-$  within the agricultural areas of a watershed (Section 2.1). The *Agriculture and Undifferentiated Rural Land Use* (AURL) land cover layer, obtained from the Ontario Land Cover Compilation V.2 (OMNRF, 2016a), was used to delineate areas of pastures, abandoned fields, and lands that were previously used for industrial or commercial activity (i.e. urban

brown fields). No distinction was made based on type of agricultural land due to lack of information and to limit the number of model covariates.

The proportion of AURL within a given catchment area was calculated using the following equation:

$$\%AURL(n) = \frac{\text{Area of AURL}}{\text{Catchment Area ( } n \text{)}} \quad (5)$$

where  $\%AURL(n)$  is the proportion of AURL area within an upstream catchment area for any given point  $n$  on a stream network. The area for AURL along with the catchment area for all sample points were calculated using the STARS package v2.0.3 (Peterson & Ver Hoef, 2014) in ESRI's ArcGIS 10.2.2.

### 3.5 Expected Covariate-Chloride Relationships

As road salt was expected to be the dominant source of  $\text{Cl}^-$  (Kaushal et al., 2005; Kelting et al., 2012; Morgan et al., 2012) in Mimico Creek, East Holland River, and Willow Creek watersheds, an increase in watershed lane length density was expected to result in an increase in  $\text{Cl}^-$  inputs. The major processes for watershed  $\text{Cl}^-$  inputs are as follows: 1)  $\text{Cl}^-$ -containing road salts are used on Canadian roadways during freezing weather periods (Perera et al., 2013), 2)  $\text{Cl}^-$  containing salts disassociate easily in water and can be mobilized during precipitation or melt events (CCME, 2011), 3) this precipitation or melt can flow away from roadways into the surrounding environment, and 4) this  $\text{Cl}^-$ -laden flow can directly enter a stream, or it can percolate through the ground and enter the shallow subsurface, contaminating groundwater reservoirs (Howard & Beck, 1993; Meriano et al., 2009; Perera et al., 2013). Therefore, it was expected that winter surveys would see the highest in-stream  $\text{Cl}^-$  concentration across all watersheds (Table 3, Table 4, and Table 5); and, as  $LLD_{\text{MIMICO}} > LLD_{\text{EAST HOLLAND}} > LLD_{\text{WILLOW}}$ , winter in-stream  $\text{Cl}^-$  was expected to be highest in Mimico Creek, followed by East Holland River, and finally Willow Creek.

With the absence of road salt in the summer, and as a lesser amount of road salt was likely used in the fall and spring seasons, it was expected that the magnitude of the relationship between LLD and in-stream  $\text{Cl}^-$  would be reduced. At the same time,  $\text{Cl}^-$  inputs from other sources, like groundwater reservoirs or  $\text{Cl}^-$ -containing fertilizers, was expected to become relatively more important.

Groundwater aquifers are known reservoirs of  $\text{Cl}^-$  (Howard & Beck, 1993; Meriano et al., 2009; Perera et al., 2013), and the weighted permeability covariate was designed to estimate the potential flow between groundwater and streamwater, with higher weighted permeability values signifying a higher relative proportion of permeable rock for any given point, and thus a higher flow from groundwater to streamwater. Thus, it was expected that weighted permeability would become a relatively influential covariate in periods of low road salt application (e.g. non-winter months) as the subsurface may continuously release  $\text{Cl}^-$  into a stream. It is also recognized that agricultural activities can result in a release of  $\text{Cl}^-$  into the watershed (Chapra et al., 2009; Mullaney et al., 2009), and the *Agriculture and Undifferentiated Land* (AURL) covariate was chosen to estimate the proportion of agricultural land within Mimico Creek, East Holland River, and Willow Creek watersheds. Thus, it was expected that areas with relatively higher AURL values would see higher  $\text{Cl}^-$  inputs from agricultural activities.

It was expected that the LLD covariate would have a strong and significant relationship with in-stream  $\text{Cl}^-$  in the Winter season and that the weighted permeability and AURL covariates would have larger and significant relationships with in-stream  $\text{Cl}^-$  in the Summer, Fall, and Spring surveys. However, due to the low presence of AURL within all surveys of Mimico Creek, and given the low AURL presence in Summer and Fall surveys of East Holland, it was not expected that AURL would have a significant relationship with in-stream  $\text{Cl}^-$  within these surveys.

Freezing weather events are also an important factor to consider when surmising expected covariate-Cl<sup>-</sup> relationships as freezing weather in the period leading up to and/or during a survey likely results in the presence of road salt and consequently a stronger LLD-Cl<sup>-</sup> relationship. The Mimico Creek Spring 1 survey took place in the late Spring of 2016 and had no freezing events leading up to or during the survey and had no snow on the ground, so the roads were expected to have a low Cl<sup>-</sup> source potential. The Mimico Creek Spring 2 survey had freezing weather in the period leading up to and during the survey and had small patches of snow on the ground, so road salt would expected to have been used on the roads and thus would act as a source for Cl<sup>-</sup>. Weather for all surveys was considered, along with watershed characteristics and potential Cl<sup>-</sup> inputs in determining expected covariate importance (Table 3 for Mimico Creek, Table 4 for East Holland River, and Table 5 for Willow Creek).

Table 3: Expected importance of Lane Length Density, Weighted Permeability, and Agriculture & Undifferentiated Rural Land covariates in models of in-stream chloride within the Mimico Creek surveys. Expected covariate importance was based on watershed characteristics, weather, and potential Cl<sup>-</sup> inputs.

Mimico Creek Watershed				
Spring 1 (June 14-16, 2016)				
Covariates	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Highly Urbanized -Relative lack of AURL	-Warm, no freezing events -No significant precipitation -No snow on ground	-Little Cl <sup>-</sup> on roads	Low
Weighted Permeability			-Cl <sup>-</sup> is stored in subsurface from winter inputs	High
Agriculture & Undifferentiated Rural Land (AURL)			-Low agricultural Cl <sup>-</sup> inputs	Low
Summer (August 03-05, 2016)				
Covariates	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Highly Urbanized -Relative lack of AURL	-Warm, no freezing events -No significant precipitation -No snow on ground	-Little Cl <sup>-</sup> on roads	Low
Weighted Permeability			-Cl <sup>-</sup> is stored in subsurface from winter inputs	High
Agriculture & Undifferentiated Rural Land (AURL)			-Low agricultural Cl <sup>-</sup> inputs	Low
Fall (October 11-13, 2016)				
Covariates	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Highly Urbanized -Relative lack of AURL	-Cool, No freezing events -No significant precipitation -No snow on ground	-Little Cl <sup>-</sup> on roads	Low
Weighted Permeability			-Cl <sup>-</sup> is stored in subsurface from winter inputs	High
Agriculture & Undifferentiated Rural Land (AURL)			-Low agricultural Cl <sup>-</sup> inputs	Low
Winter (February 20-23, 2017)				
Covariates	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Highly Urbanized -Relative lack of AURL	-Freezing temperatures -Significant rain event during survey -Snow on ground -Stream partially Iced	-High Cl <sup>-</sup> mobilization from roads and snowbanks	High
Weighted Permeability			-Cl <sup>-</sup> input is relatively negligible	Low
Agriculture & Undifferentiated Rural Land (AURL)			-Low agricultural Cl <sup>-</sup> inputs	Low
Spring 2 (April 21-23, 2017)				
Covariates	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Highly Urbanized -Relative lack of AURL -Snow on ground -Stream partially Iced	-Freezing events in week leading up to survey -Freezing event during survey -Small precipitation event	-Low to Moderate Cl <sup>-</sup> mobilization from roads and snowbanks	High
Weighted Permeability			-Cl <sup>-</sup> input is relatively negligible	Low
Agriculture & Undifferentiated Rural Land (AURL)			-Low agricultural Cl <sup>-</sup> inputs	Low

Table 4: Expected importance of Lane Length Density, Weighted Permeability, and Agriculture & Undifferentiated Rural Land covariates in models of in-stream chloride within the East Holland River surveys. Expected covariate importance was based on watershed characteristics, weather, and potential Cl<sup>-</sup> inputs.

East Holland River Watershed				
Summer (July 21-25, 2016)				
Covariates	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Half of survey area surveyed -Highly Urbanized -Relatively low AURL	-Warm, no freezing events	-Little Cl <sup>-</sup> on roads	Low
Weighted Permeability		-Premature end of survey due to precipitation	-Cl <sup>-</sup> is stored in subsurface from winter inputs	High
Agriculture & Undifferentiated Rural Land (AURL)		-No snow on ground	-Low agricultural Cl <sup>-</sup> inputs	Low
Fall (October 24-26, 2016)				
Covariates	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Half of survey area surveyed -Highly Urbanized -Relatively low AURL	-One freezing event on Oct 26.	-Little Cl <sup>-</sup> on roads	Moderate
Weighted Permeability		-No freezing events in week before survey	-Cl <sup>-</sup> is stored in subsurface from winter inputs	High
Agriculture & Undifferentiated Rural Land (AURL)		-Premature end of survey due to precipitation -Little to no snow on ground	-Low agricultural Cl <sup>-</sup> inputs	Low
Winter (March 02-05, 2017)				
Covariates	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Full study area surveyed -Roughly Half urban, half rural/agricultural	-Extended freeze event during survey period	-Low to Moderate Cl <sup>-</sup> mobilization from roads, roadsides, and from snowbanks	High
Weighted Permeability		-One freeze period in week before survey	-Cl <sup>-</sup> input is negligible	Low
Agriculture & Undifferentiated Rural Land (AURL)		-Light precipitation event during survey -Snow on ground	-Low agricultural Cl <sup>-</sup> inputs	Low
Spring (May 15-17, 2017)				
Covariate	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Full study area surveyed -Roughly Half urban, half rural/agricultural	-2 freezing events in week before survey	-Low to Moderate Cl <sup>-</sup> mobilization from roads, roadsides, and from snowbanks	Moderate
Weighted Permeability		-No freezing events during survey	-Cl <sup>-</sup> is stored in subsurface from winter inputs	Moderate
Agriculture & Undifferentiated Rural Land (AURL)		-No significant precipitation during survey -Snow on ground	-Higher agricultural Cl <sup>-</sup> inputs due to potential agricultural activity	Moderate

Table 5: Expected importance of Lane Length Density, Weighted Permeability, and Agriculture & Undifferentiated Rural Land covariates in models of in-stream chloride within the Willow Creek surveys. Expected covariate importance was based on watershed characteristics, weather, and potential Cl<sup>-</sup> inputs.

Willow Creek Watershed				
Summer (July 28-29, 2016)				
Covariates	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Mostly unurbanized -Significant agricultural land	-No freezing events -No precipitation during survey period -No snow on the ground	-Little Cl <sup>-</sup> on roads	Low
Weighted Permeability			-Cl <sup>-</sup> is stored in subsurface from winter inputs	High
Agriculture & Undifferentiated Rural Land (AURL)			-Higher agricultural Cl <sup>-</sup> inputs due to potential agricultural activity	Moderate
Fall (October 18-19, 2016)				
Covariates	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Mostly unurbanized -Significant agricultural land	-No freezing events -No precipitation events -No snow on the ground	- Little Cl <sup>-</sup> on roads	Low
Weighted Permeability			-Cl <sup>-</sup> is stored in subsurface from winter inputs	Moderate
Agriculture & Undifferentiated Rural Land (AURL)			-Higher agricultural Cl <sup>-</sup> inputs due to potential agricultural activity	Moderate
Winter (February 06-08, 2017)				
Covariates	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Mostly unurbanized -Significant agricultural land	-Extended freeze event during survey period -Freeze events in week before survey -Minor freezing precipitation event during survey -Snow on the ground	-Moderate to high Cl <sup>-</sup> mobilization from roads, roadsides, and from snowbanks	High
Weighted Permeability			-Cl <sup>-</sup> input is negligible	Low
Agriculture & Undifferentiated Rural Land (AURL)			-Low agricultural Cl <sup>-</sup> inputs	Low
Spring (April 17-18, 2017)				
Covariate	Characteristics	Weather	Potential Cl <sup>-</sup> Inputs	Estimated Importance
Lane Length Density	-Mostly unurbanized -Significant agricultural land	-Freezing period halfway through survey -Freezing events in week leading up to survey -No precipitation during survey -Snow on the ground	-Moderate Cl <sup>-</sup> mobilization from roads, roadsides, and from snowbanks	Moderate
Weighted Permeability			-Cl <sup>-</sup> is stored in subsurface from winter inputs	Moderate
Agriculture & Undifferentiated Rural Land (AURL)			-Higher agricultural Cl <sup>-</sup> inputs due to potential agricultural activity	Moderate

### 3.6 Model Construction and Investigation of Covariates

To maintain consistency across watersheds, the same three variables, and all possible (non-redundant) variable subsets, were explored for each survey resulting in seven related models per season (Table 6).

The expected importance of each covariate, as discussed in Section 3.5, were summarized for each model in Table 7 for Mimico Creek, Table 8 for East Holland, and Table 9 for Willow Creek.

Table 6: Tested models for each season within Mimico Creek, East Holland River, and Willow Creek watersheds. These models are based on all possible non-redundant combinations of model covariates. LLD is Lane Length Density as defined in Section 3.4.1, WPerm is Weighted Permeability as defined in Section 3.4.2, and AURL is the Agricultural and Undifferential Rural Land Cover variable as defined in Section 3.4.3.

Global Model 1 (GMI)	LLD + WPerm + AURL
Model 2 (M2)	LLD
Model 3 (M3)	AURL
Model 4 (M4)	WPerm
Model 5 (M5)	LLD + AURL
Model 6 (M6)	LLD + WPerm
Model 7 (M7)	WPerm + AURL

Table 7: Expected importance of the input covariates for all Mimico Creek Watershed models. Low importance means little improvement in model performance was expected if this variable were included, moderate impact means moderate improvement in expected model performance, and high impact means significant improvement in expected model performance. The expected best model for each survey is highlighted in green. Covariate impact estimations are based on the rationale seen in Table 3. The definitions for model acronyms are found in Table 6.

Mimico Creek Watershed	Model	Covariates		
		LLD	WePerm	Ag & RL
Spring 1 (June 14-16, 2016)	GMI	Low Impact	High Impact	Low Impact
	M2	Low Impact	Not Applicable	Not Applicable
	M3	Not Applicable	Not Applicable	Low Impact
	M4	High Impact	High Impact	Not Applicable
	M5	Low Impact	Not Applicable	Low Impact
	M6	Low Impact	High Impact	Not Applicable
	M7	Not Applicable	High Impact	Low Impact
Summer (August 03-05, 2016)	GMI	Low Impact	High Impact	Low Impact
	M2	Low Impact	Not Applicable	Not Applicable
	M3	Not Applicable	Not Applicable	Low Impact
	M4	High Impact	High Impact	Not Applicable
	M5	Low Impact	Not Applicable	Low Impact
	M6	Low Impact	High Impact	Not Applicable
	M7	Not Applicable	High Impact	Low Impact
Fall (October 11-13, 2016)	GMI	Low Impact	High Impact	Low Impact
	M2	Low Impact	Not Applicable	Not Applicable
	M3	Not Applicable	Not Applicable	Low Impact
	M4	High Impact	High Impact	Not Applicable
	M5	Low Impact	Not Applicable	Low Impact
	M6	Low Impact	High Impact	Not Applicable
	M7	Not Applicable	High Impact	Low Impact
Winter (February 20-23, 2017)	GMI	High Impact	Low Impact	Low Impact
	M2	High Impact	Not Applicable	Not Applicable
	M3	Not Applicable	Not Applicable	Low Impact
	M4	Not Applicable	Low Impact	Not Applicable
	M5	High Impact	Not Applicable	Low Impact
	M6	High Impact	Low Impact	Not Applicable
	M7	Not Applicable	Low Impact	Low Impact
Spring 2 (April 21-23, 2017)	GMI	High Impact	Low Impact	Low Impact
	M2	High Impact	Not Applicable	Not Applicable
	M3	Not Applicable	Not Applicable	Low Impact
	M4	Not Applicable	Low Impact	Not Applicable
	M5	High Impact	Not Applicable	Low Impact
	M6	High Impact	Low Impact	Not Applicable
	M7	Not Applicable	Low Impact	Low Impact
Low Impact		Low Impact		
Medium Impact		Low Impact		
High Impact		High Impact		
Not Applicable		Not Applicable		

Table 8: Expected importance of the input covariates for all East Holland River Watershed models. Low importance means little improvement in model performance was expected if this variable were included, moderate impact means moderate improvement in expected model performance, and high impact means significant improvement in expected model performance. The expected best model for each survey is highlighted in green. Covariate impact estimations are based on the rationale found in Table 4. The definitions for model acronyms are found in Table 6.

East Holland River Watershed	Model	Covariates		
		LLD	WePerm	Ag & URL
Summer (July 21-25, 2016)	GMI	Low	High	Low
	M2	Low	Not Applicable	Not Applicable
	M3	Not Applicable	Not Applicable	Low
	M4	Not Applicable	High	Not Applicable
	M5	Low	Not Applicable	Low
	M6	Low	High	Not Applicable
	M7	Not Applicable	High	Low
Fall (October 24-26, 2016)	GMI	Low	High	Low
	M2	Low	Not Applicable	Not Applicable
	M3	Not Applicable	Not Applicable	Low
	M4	Not Applicable	High	Not Applicable
	M5	Low	Not Applicable	Low
	M6	Low	High	Not Applicable
	M7	Not Applicable	High	Low
Winter (March 11-13, 2016)	GMI	High	Low	Low
	M2	High	Not Applicable	Not Applicable
	M3	Not Applicable	Not Applicable	Low
	M4	Not Applicable	Low	Not Applicable
	M5	High	Not Applicable	Low
	M6	High	Low	Not Applicable
	M7	Not Applicable	Low	Low
Spring (May 15-17, 2017)	GMI	Low	High	High
	M2	Low	Not Applicable	Not Applicable
	M3	Not Applicable	Not Applicable	Low
	M4	Not Applicable	Low	Not Applicable
	M5	Low	Not Applicable	Low
	M6	Low	Low	Not Applicable
	M7	Not Applicable	Low	Low
Low Impact		Low		
Medium Impact		Low		
High Impact		High		
Not Applicable		Not Applicable		

Table 9: Expected importance of the input covariates for all Willow Creek Watershed models. Low importance means little improvement in model performance was expected if this variable were included, moderate impact means moderate improvement in expected model performance, and high impact means significant improvement in expected model performance. The expected best model for each survey is highlighted in green. Covariate impact estimations are based on the expected found in Table 5. The definitions for model acronyms are found in Table 6.

Willow Creek Watershed	Model	Covariates		
		LLD	WePerm	Ag & URL
Summer (June 14-16, 2016)	GMI	Light Blue	Dark Blue	Light Blue
	M2	Light Blue	Grey	Grey
	M3	Grey	Grey	Light Blue
	M4	Grey	Dark Blue	Grey
	M5	Light Blue	Grey	Light Blue
	M6	Light Blue	Dark Blue	Grey
	M7	Green	Grey	Dark Blue
Fall (August 03-05, 2016)	GMI	Light Blue	Light Blue	Light Blue
	M2	Light Blue	Grey	Grey
	M3	Grey	Grey	Light Blue
	M4	Grey	Light Blue	Grey
	M5	Light Blue	Grey	Light Blue
	M6	Light Blue	Light Blue	Grey
	M7	Green	Grey	Light Blue
Winter (October 11-13, 2016)	GMI	Dark Blue	Light Blue	Light Blue
	M2	Green	Dark Blue	Grey
	M3	Grey	Grey	Light Blue
	M4	Grey	Light Blue	Grey
	M5	Dark Blue	Grey	Light Blue
	M6	Dark Blue	Light Blue	Grey
	M7	Grey	Light Blue	Light Blue
Spring (February 20-23, 2017)	GMI	Green	Light Blue	Light Blue
	M2	Light Blue	Grey	Grey
	M3	Grey	Grey	Light Blue
	M4	Grey	Light Blue	Grey
	M5	Light Blue	Grey	Light Blue
	M6	Light Blue	Light Blue	Grey
	M7	Grey	Light Blue	Light Blue
Low Impact		Light Blue		
Medium Impact		Light Blue		
High Impact		Dark Blue		
Not Applicable		Grey		

### 3.7 Metrics for Model Selection and Fit

The statistical metrics used for model selection included Root Mean Square Prediction Error (RMSPE), Akaike Information Criterion (AIC), and confidence interval coverage (COV). RMSPE is the root square of the average of the calculated error for the model with error defined as the difference between the regression line and the actual data point. Smaller RMSPE values indicate a lower calculated error for a given model. AIC, as first presented in Akaike (1973), allows for the relative comparison of models with different variables based on how well the model fits the data. AIC does not test the absolute quality of a model, instead it allows you to pick a model that minimizes deviation from the sampled entity, and the goal is to minimize this value when selecting a model. The COV value captures the proportion of (cross-validation) model predictions that fall within a specific confidence interval. Three different confidence intervals, 80% (cov.80,) 90% (cov.90), and 95% (cov.95), were used and the COV value that was as close as possible to the respective ideal confidence interval was selected (e.g. COV.80 = 0.81 is better than COV.80 = 0.90). All these metrics of model fit were used to compare models and those models with fit metrics within +/-5% of the respective 'best' values were considered comparable. The best values for each metric include the lowest RMSPE value + 5%, the lowest AIC value + 5%, and all values found within ideal confidence intervals  $\pm 5\%$  (e.g. ideal cov.80 values fall within the range of  $0.80 \pm 5\%$ ). Additionally, models that made the most ecological sense in accordance with our expectations, and that explained higher amounts of variability (as indicated by  $R^2$  values) were preferred in accordance with the principle of model parsimony.

### 3.8 Data Processing

Calculated CI' values were checked for normality and outliers. Twelve outliers were identified (in the full dataset) as sample locations with erroneous and unreasonably small catchment areas. These GIS catchment delineation errors were likely introduced by pseudo-nodes (arbitrary breaks) found within

the stream networks for each of the study watersheds as extracted from the Integrated Hydrology Layer dataset (OMNR, 2015). The observed  $\text{Cl}^-$  values and the covariate data were assessed for distribution normality by plotting the data as histograms and then applying logarithmic transformations where necessary (Table 10).

Table 10: Transformations applied to observed chloride values and covariate values at each sample point. Transformations were applied with the aim of normalizing chloride and covariate distribution curves.

<b>Observed Chloride Transformations</b>	
Mimico Creek	No Transformation Applied
East Holland	$\text{Log}(\text{Cl}^- + 1)$
Willow Creek	$\text{Log}(\text{Cl}^- + 1)$
<b>Covariate Transformations</b>	
<b>Mimico Creek</b>	
Lane Length Density	No Transformation Applied
Weighted Permeability	No Transformation Applied
Agriculture and Undiff. Rural Land	$\text{Log}(x) - \text{Log}(1-x)$
<b>East Holland</b>	
Lane Length Density	$\text{Log}(x)$
Weighted Permeability	$\text{Log}(x)$
Agriculture and Undiff. Rural Land	$\text{Log}(x) - \text{Log}(1-x)$
<b>Willow Creek</b>	
Lane Length Density	$\text{Log}(x)$
Weighted Permeability	No Transformation Applied
Agriculture and Undiff. Rural Land	$\text{Log}(x) - \text{Log}(1-x)$

## 4 Results

### 4.1 Survey Timing and Survey Flow Categorization

A total of 13 longitudinal surveys were carried out across seasons in the three study watersheds generating a total sample size of 871, with an average of 70 sample points for Mimico Creek, 79 sample points for the East Holland River, and 52 sampling points for Willow Creek. The Mimico Creek Spring 1 survey was the only survey conducted in the Spring 2016 season and was used for field training and for setting field operational procedures.

The flow state for each survey (Table 11) was categorized for descriptive purposes into low, low-medium, medium-high, and high flow categories based on flow exceedance criteria described in Section 3.2.2 and depicted in Figure 13.

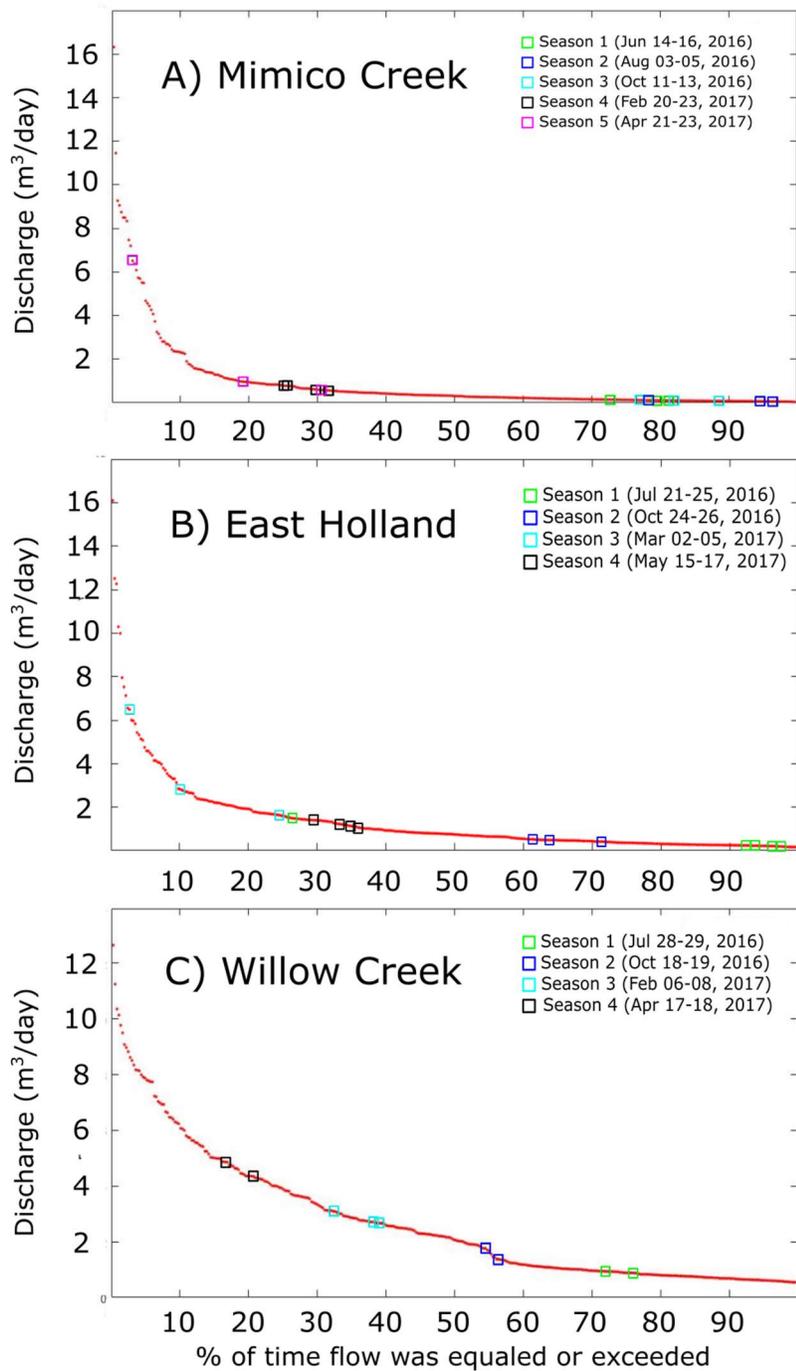


Figure 13: A depiction of the flow exceedance curves for the A) Mimico Creek Watershed, B) East Holland River Watershed, and C) Willow Creek Watershed. This curve depicts the percentage of time that, over the full survey year (June 14, 2016 to June 14, 2017), the average daily flow within each respective watershed will equal or exceed some specified flow value. The colored squares represent the specific dates for each survey for each watershed. Data extracted from the Environment and Climate Change Canada Hydrometric Flow Data web site ([https://wateroffice.ec.gc.ca/mainmenu/historical\\_data\\_index\\_e.html](https://wateroffice.ec.gc.ca/mainmenu/historical_data_index_e.html))

Table 11: A categorization of the flow states for each longitudinal watershed survey using data from Water Survey of Canada hydrometric gauges located within Mimico Creek, East Holland, and Willow Creek watersheds (refer to Figure 13 for flow exceedance curves)

<b>Mimico Creek</b>	
<b>Survey</b>	<b>Flow State</b>
Spring 1 (Jun 14-16, 2016)	Low Flow
Summer (Aug 3-5, 2016)	Low Flow
Fall (Oct 11-13, 2016)	Low Flow
Winter (Feb 20-23, 2017)	Medium-High Flow
Spring 2 (Apr 21-23, 2017)	High Flow
<b>East Holland</b>	
<b>Survey</b>	<b>Flow State</b>
Summer (Jul 21-25, 2016)	Low Flow
Fall (Oct 24-26, 2016)	Low-Medium Flow
Winter (Mar 02-05, 2017)	High Flow
Spring (May 15-17, 2017)	Medium-High Flow
<b>Willow Creek</b>	
<b>Survey</b>	<b>Flow State</b>
Summer (Jul 28-29, 2016)	Low Flow
Fall (Oct 18-19, 2016)	Low-Medium Flow
Winter (Feb 06-08, 2017)	Medium-High Flow
Spring (Apr 17-18, 2017)	High Flow
<ul style="list-style-type: none"> <li>• <u>Low Flow</u> if majority of survey points are &lt;25% on the flow exceedance curve</li> <li>• <u>Low-Medium Flow</u> if majority of survey points are ≥25% and &lt;50% on the flow exceedance curve</li> <li>• <u>Medium-High Flow</u> if majority of survey points are ≥50% and &lt;75% on the flow exceedance curve</li> <li>• <u>High flow</u> if majority of survey points fall are ≥75% on the flow exceedance curve</li> </ul>	

## 4.2 Electrical Conductivity vs. Chloride Relationships

### 4.2.1 Mimico Creek Watershed

In-stream chloride ( $\text{Cl}^-$ ) data points ( $n = 143$ ) were collected from five stations across the study area by the Mitchell Research Group at the University of Toronto Scarborough and these values were paired with electrical conductivity (EC) values to establish the relationship between EC and  $\text{Cl}^-$ . This EC and  $\text{Cl}^-$  data provided a very strong SpCond- $\text{Cl}^-$  correlation using both a linear trendline ( $R^2 = 0.97$ ) and a polynomial trend line ( $R^2 = 0.98$ , Figure 14), and this strong relationship provides evidence that  $\text{Cl}^-$  may be the dominant factor controlling electrical conductivity within the Mimico Creek watershed. The polynomial relationship was applied to all five Mimico Creek longitudinal surveys because both linear and polynomial terms were significant and given its higher  $R^2$  value.

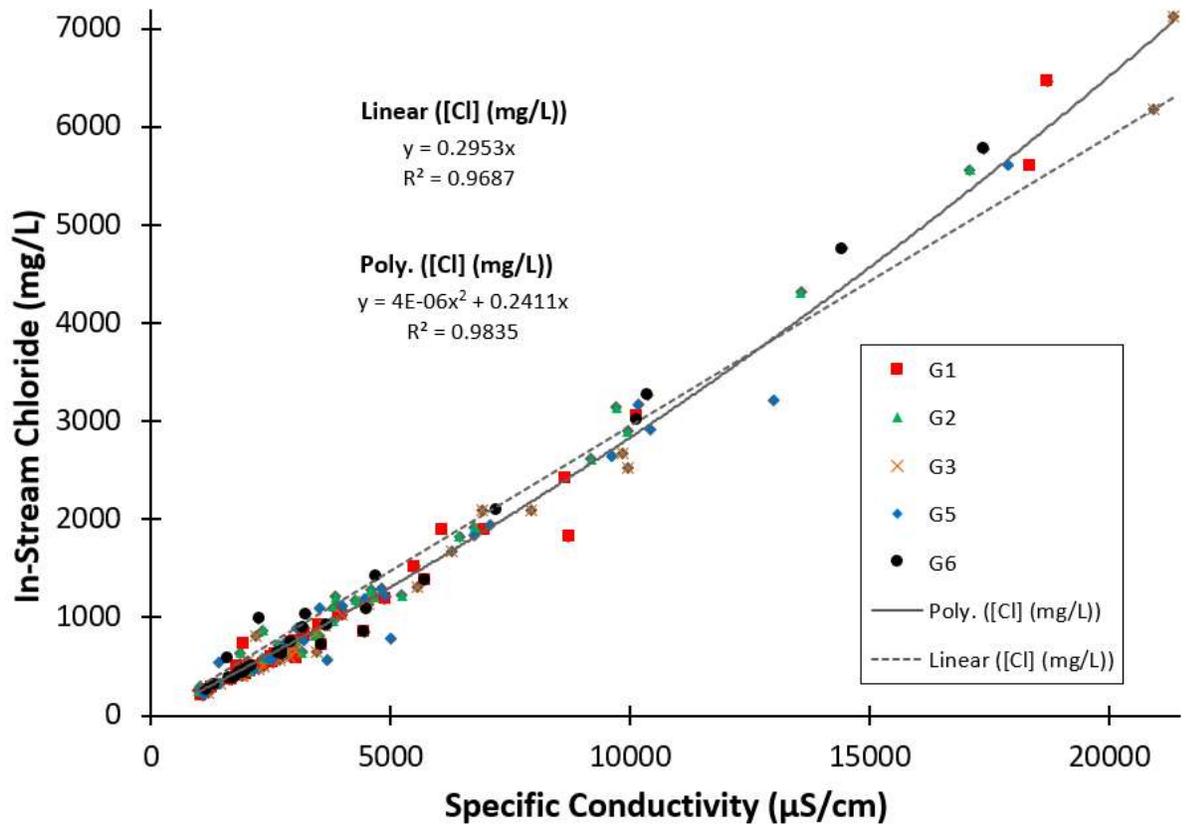


Figure 14: Specific conductivity versus in-stream chloride plot for the Mimico Creek Study Area. Conductivity and chloride data were obtained from five in-situ water quality stations placed throughout Mimico Creek from the Mitchell Research Group at the University of Toronto Scarborough. Sample points used in this plot were selected to correspond with total length of the study period for Mimico Creek. The polynomial relationship determined from this dataset was used to convert specific conductivity to in-stream chloride for this study.

#### 4.2.2 East Holland River Watershed

There were two EC-Cl<sup>-</sup> trends for East Holland watershed depending on geographic context of the sampled data (Figure 15). Stream water samples collected at the main downstream confluence, EH\_S1, showed a weaker correlation with a higher average Cl<sup>-</sup> concentration than samples from all headwater sources. The point data taken from the western headwaters (MID\_EH\_S1 & UP\_EH\_S1) and eastern headwaters (UP404\_S1 & DOWN404\_S1) showed similarities to the relationships derived from the PWQMN data. It is unknown if the difference between the downstream trend versus the upstream trend is due to contributions from unmeasured upstream sections of the East Holland Watershed. The specific conductance-Cl<sup>-</sup> polynomial relationship was derived from 25 data points: 4 from the headwater levellogger stations and 21 from the PWQMN station.

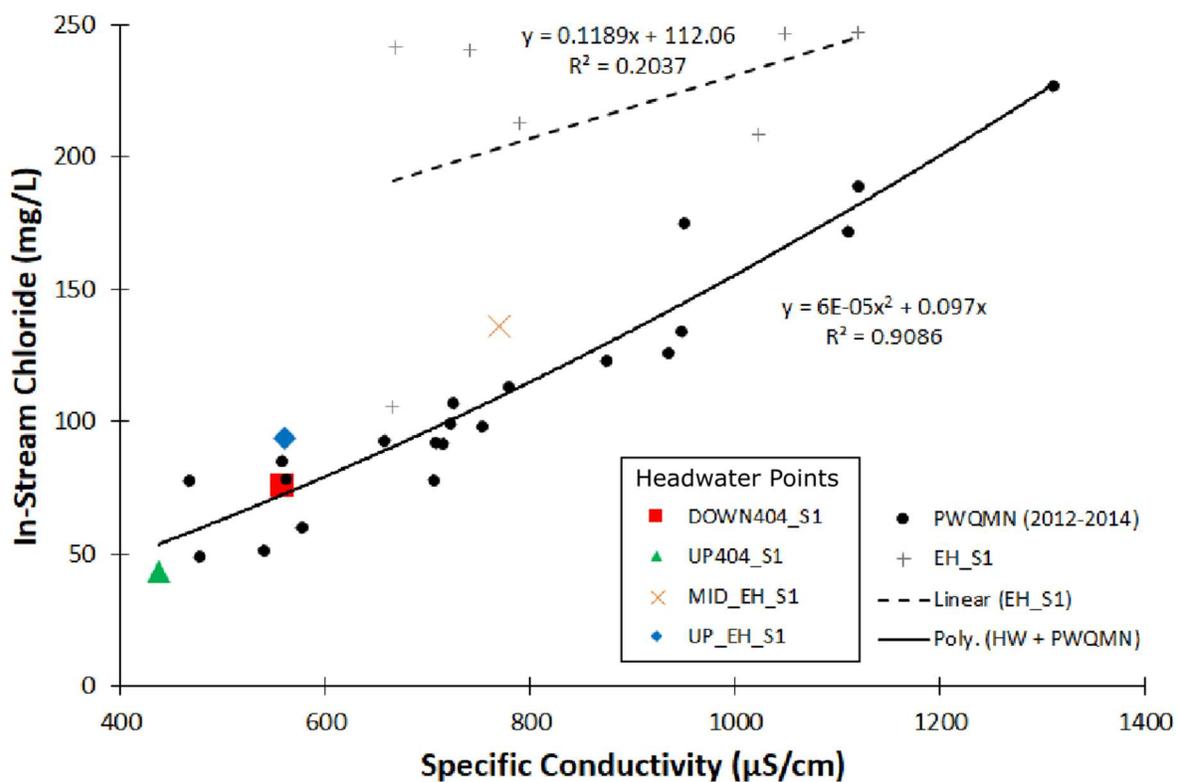


Figure 15: Specific conductivity versus in-stream chloride for East Holland watershed. The DOWN404\_S1, UP404\_S1, and MID\_EH\_S1 are sample sites, located in the headwaters of the watershed, from which in-stream water samples were taken and then measured for both SpCond and Cl<sup>-</sup> concentration. EH\_S1 is a water quality measurement site found outside of the East Holland study area. The Provincial Water Quality Monitoring Network (PWQMN) station (Aurora Creek; id# 03007700702) is found just downstream of MID\_EH\_S1 and this site provided three years of in-stream SpCond and in-stream Chloride data. The combined headwater + PWQMN dataset was used to calculate the polynomial trend used in converting specific conductivity to in-stream chloride.

#### 4.2.3 Willow Creek Watershed

The Willow Creek watershed did not have a strong EC-Cl<sup>-</sup> relationship potentially due to variation based on geographic context. Data points from the headwater sampling site (UP\_WIL\_S1) tended to have a lower SpCond-Cl<sup>-</sup> relationship compared to watershed outlet data points (MID\_WIL\_S1). Note that the lower watershed outlet location (LOW\_WIL\_S1) and the PWQMN station were not considered as these sampling sites were located just outside of the selected study area for Willow Creek. A headwater curve ( $n = 3$ ) and a downstream curve ( $n = 3$ ) were applied to all Willow Creek surveys, and as Little Lake separates the upstream headwaters from the downstream outlet (Figure 5), and as this lake was not sampled for Cl, the upstream curve was applied to all points upstream of Little Lake and the downstream curve was applied to all points downstream of Little Lake.

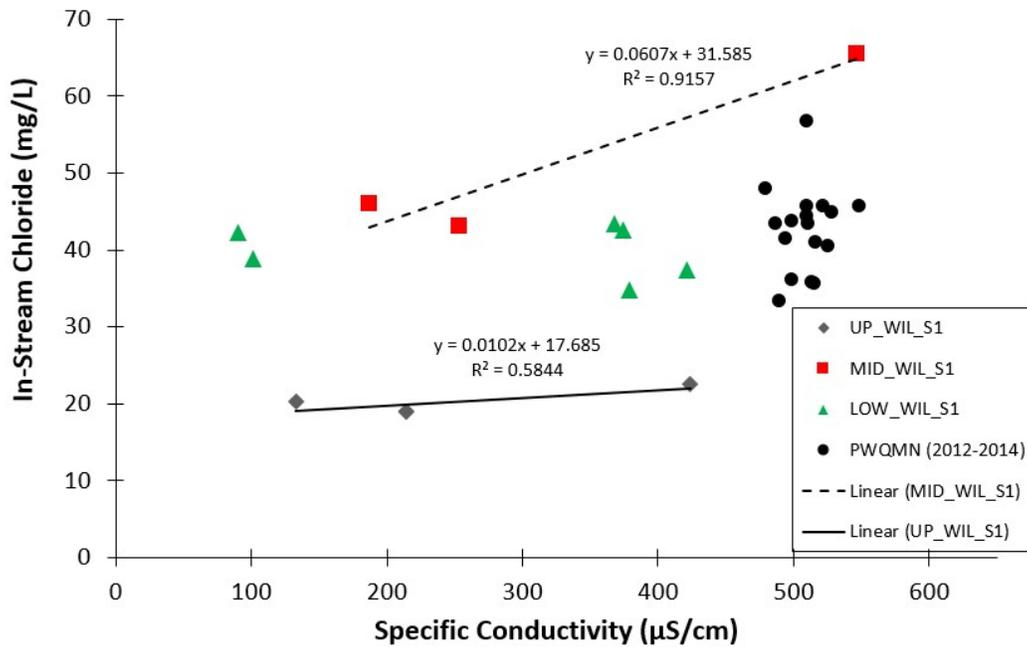


Figure 16: Specific conductivity versus chloride graph for Willow Creek watershed. The MID\_WIL\_S1 site was selected as the downstream site as it is near the outlet for the Willow Creek study area. UP\_WIL\_S1 is found within the headwaters of the Willow Creek study area. Both LOW\_WIL\_S1 site and the Provincial Water Quality Monitor Network (PWQMN) site (id# 03005703002) are located outside the Willow Creek watershed study area. The MID\_WIL\_S1 SpCond-Cl<sup>-</sup> relationship was used for all survey sites downstream of Little Lake and the UP\_WIL\_S1 SpCond-Cl<sup>-</sup> relationship was used for all survey sites upstream of Little Lake.

### 4.3 Statistical Summary and Spatial Patterns in Chloride Concentrations

Observed  $\text{Cl}^-$  values for all surveys are shown in Table 12, with the equations used to convert electrical conductivity to  $\text{Cl}^-$  displayed in Figure 14 for Mimico Creek, Figure 15 for East Holland River, and Figure 16 for Willow Creek. Mean  $\text{Cl}^-$  values tended to positively correlate with urbanization, with the highly urbanized Mimico Creek having a mean  $\text{Cl}^-$  concentration of 812.4 mg/L, the moderately urbanized East Holland having a mean  $\text{Cl}^-$  concentration of 188.8 mg/L, and the relatively un-urbanized Willow Creek showing a mean  $\text{Cl}^-$  concentration of 30.9 mg/L. There was considerable variation in seasonal patterns across the three watersheds. The winter season had mean  $\text{Cl}^-$  concentration peaks in Mimico Creek and East Holland, but no such peak existed in Willow Creek. Summer baseflow provided the lowest mean concentrations in East Holland, and the second lowest in Mimico Creek, but interestingly provided the highest mean  $\text{Cl}^-$  concentration in Willow Creek.

Across all Mimico Creek surveys, the highest  $\text{Cl}^-$  concentrations were seen (Figure 17) in the central headwaters with a dilution after the main confluences which form the main stem. Chloride concentrations tended to increase as the main stem approaches the watershed outlet into Lake Ontario. The variability in the Mimico Creek Winter survey tended to be opposite of that in all other Mimico Creek surveys as in-stream  $\text{Cl}^-$  concentrations tended to increase after the main confluences, though a decrease in concentration was seen closer to the outlet. The in-stream  $\text{Cl}^-$  concentrations were also much higher in survey 4 when compared to any other survey.

The limitation of sample size in the Summer and Fall East Holland surveys was clearly apparent in the East Holland  $\text{Cl}^-$  concentration map (Figure 18). The western urbanized headwaters tended to have higher in-stream  $\text{Cl}^-$  concentrations than samples taken from the eastern half of the watershed. The urbanized headwaters in East Holland also tended to have small scale patchiness in  $\text{Cl}^-$  concentrations whereas the eastern rural area tended to have relatively homogenized  $\text{Cl}^-$  values.

Willow Creek could not be fully sampled as the stream tended to become more ephemeral in lower flow states. Sample sizes ( $n = 35$ ; Table 12) tended to be lowest in the low flow state (Table 11) of the Summer survey, but sample sizes increased relative to the flow state of the watershed: low-medium flow (Fall season) resulted in 49 sample points, medium-high flow (Winter season) resulted in 54 sample points, and high flow (Spring season) resulted in 68 sample points. The Willow Creek watershed (Figure 19) headwaters had consistently lower  $\text{Cl}^-$  concentrations across all seasons when compared to the main stem below Little Lake, and the  $\text{Cl}^-$  concentration values were relatively homogenous above Little Lake.

Table 12: The dates, sample sizes, and basic statistics for calculated in-stream chloride concentrations in the Mimico Creek, East Holland River, and Willow Creek watersheds. The equations used to convert electrical conductivity to chloride concentration can be found in Figure 14 for Mimico Creek, Figure 15 for East Holland River, and Figure 16 for Willow Creek.

Survey (S)	Season	Sample Size	Minimum (Cl <sup>-</sup> [mg/L])	Maximum (Cl <sup>-</sup> [mg/L])	Mean (Cl <sup>-</sup> [mg/L])	Std. Dev. (Cl <sup>-</sup> [mg/L])
Mimico S1	Spring 1	73	399.7	1053.1	726.4	134.0
Mimico S2	Summer	73	201.3	888.9	484.5	162.1
Mimico S3	Fall	68	197.4	764.8	421.1	128.3
Mimico S4	Winter	68	740.2	3052.7	1902.4	444.6
Mimico S5	Spring 2	69	243.4	697.6	527.6	121.5
East Holland S1	Summer	41	58.9	907.8	161.5	126.1
East Holland S2	Fall	54	46.4	613.9	168.0	100.8
East Holland S3	Winter	110	37.1	1832.4	246.4	206.1
East Holland S4	Spring	109	35.9	505.6	179.2	91.7
Willow S1	Summer	35	21.4	82.9	39.6	25.5
Willow S2	Fall	49	21.7	78.5	28.9	14.1
Willow S3	Winter	54	20.9	73.7	27.9	14.3
Willow S4	Spring	68	20.1	67.9	27.0	13.9

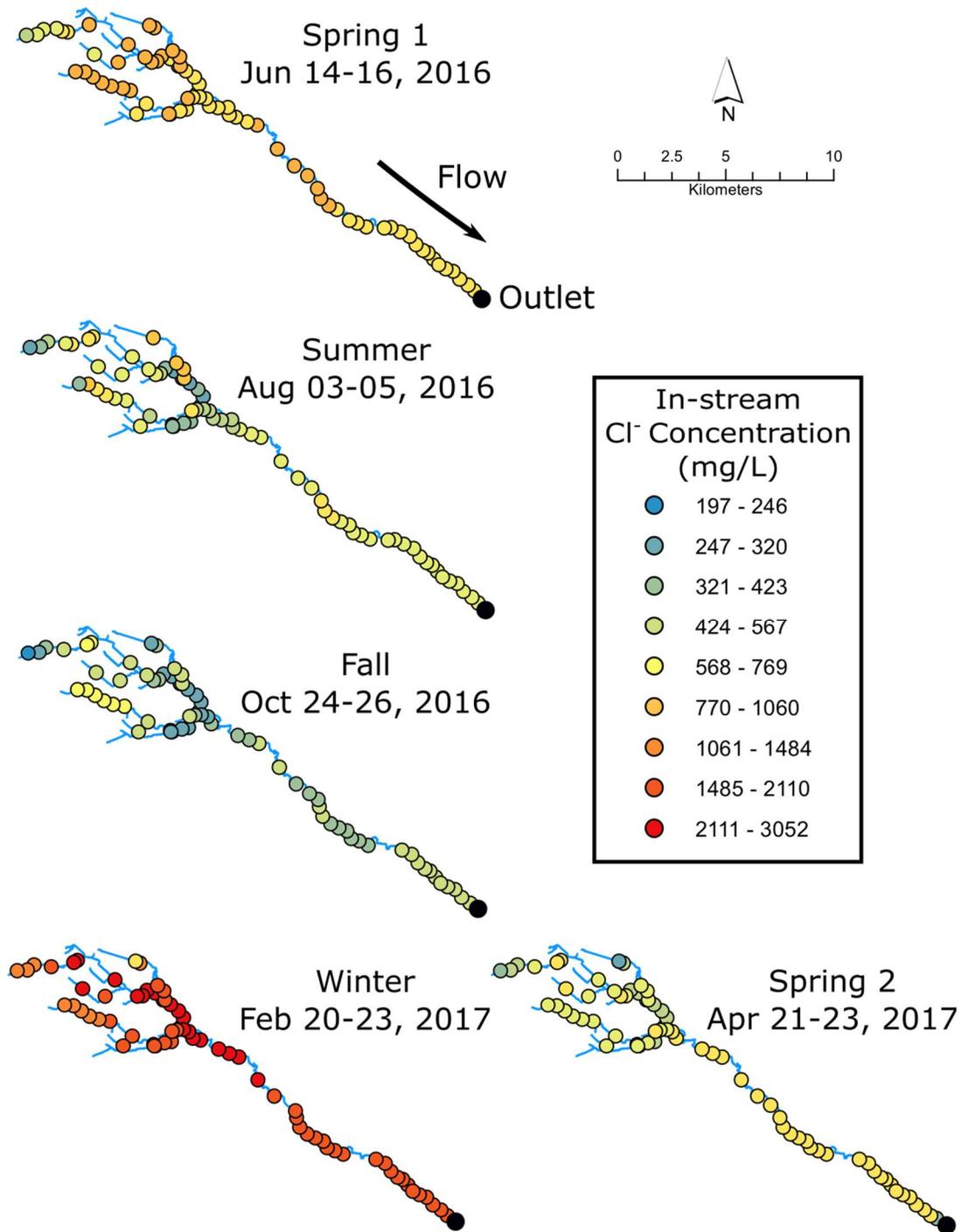


Figure 17: In-stream chloride maps for the Mimico Creek watershed showing the results of five spatially intensive longitudinal surveys. Chloride values were measured via specific conductivity measurements at each sample site, with the conversion equation from specific conductivity to chloride shown in Figure 14.

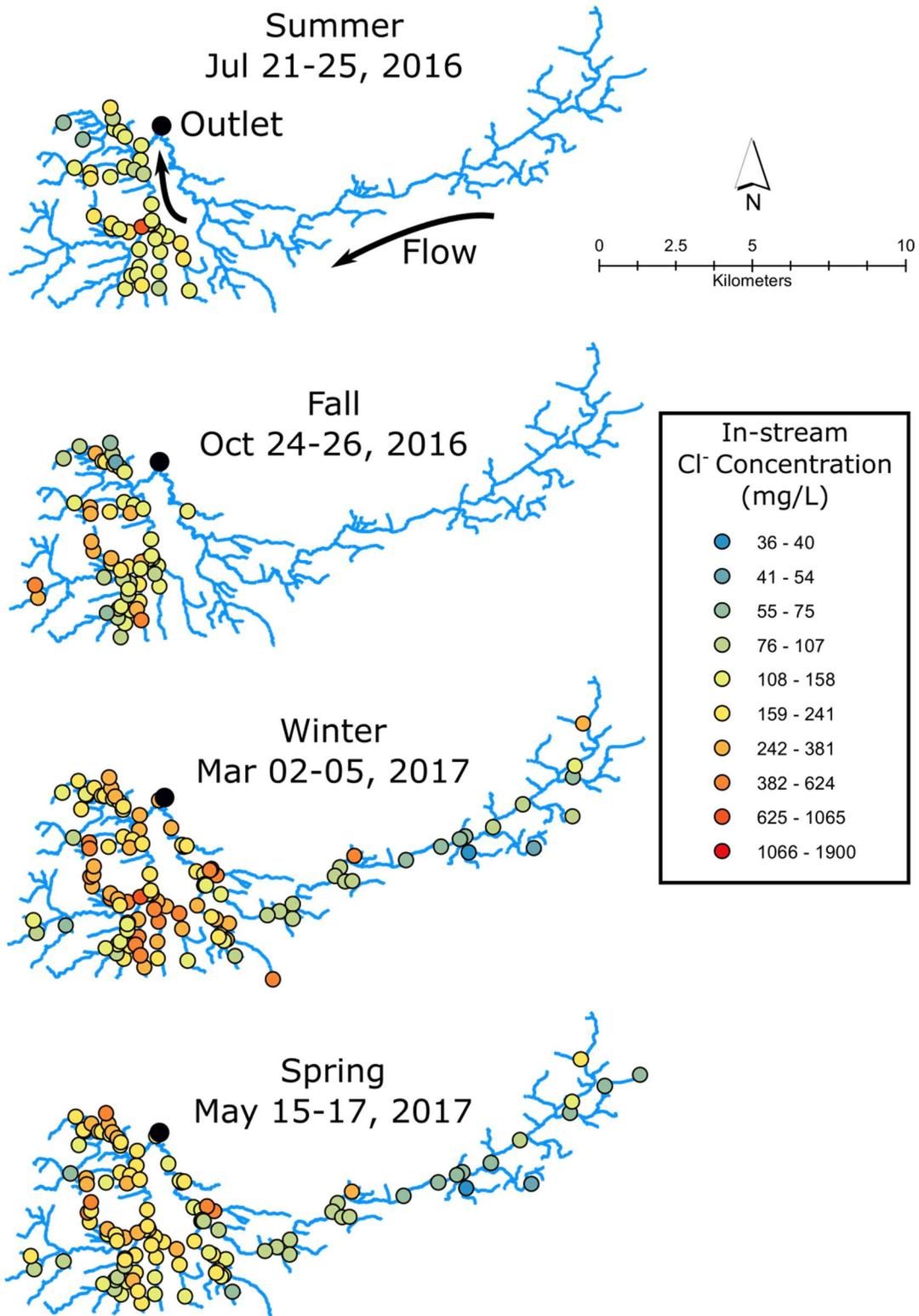


Figure 18: In-stream chloride maps for the East Holland River Watershed showing the results of four spatially intensive longitudinal surveys. Chloride values were measured via specific conductivity measurements at each sample site, with the conversion equation from specific conductivity to chloride shown in Figure 15.

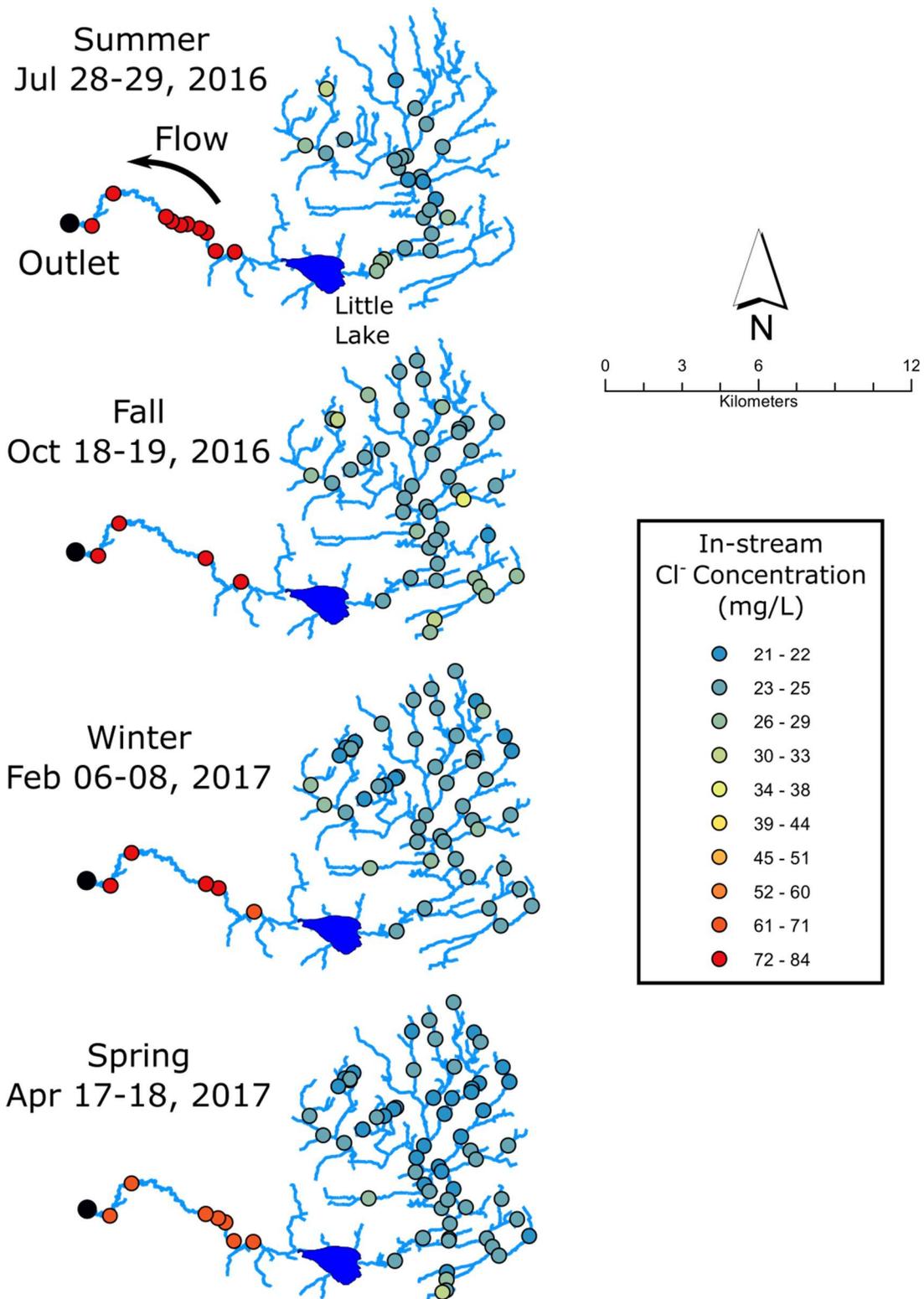


Figure 19: In-stream chloride maps for the Willow Creek Watershed showing the results of four spatially intensive longitudinal surveys. Chloride values were measured via specific conductivity measurements at each sample site, with the conversion equation from specific conductivity to chloride shown in Figure 16.

## 4.4 Spatial Stream Network Modelling

### 4.4.1 Autocovariance Function Selection

A statistical comparison of tail-up autocovariance functions appropriate for modelling the residual spatial structure of in-stream Cl<sup>-</sup> in Mimico Creek, East Holland River, and Willow Creek, is shown in Table 13 and this table revealed that there was no clear “best” spatial model appropriate either for a given watershed, or across different watersheds. Model selection metrics (e.g., R<sup>2</sup>, RMSPE, AIC, Cov.80, etc) were highly similar and where small improvements in R<sup>2</sup> were noted, these improvements seemed to come at the expense of other model fit metrics (e.g., AIC or RMSPE). For simplicity, a Linear-With-Sill model was initially preferred for all watersheds and this model of network spatial structure was selected for Mimico Creek given the linear nature of this watershed (Figure 5). Exponential and spherical variogram models are well known in terrestrial applications and these models are typically used with Euclidean distances (Dale & Fortin, 2014). Either of these two models seemed ecologically appropriate for East Holland and Willow Creek watersheds given the non-linear shape of their stream networks. The final autocovariance models were Linear-With-Sill for Mimico Creek, and the Exponential function for both East Holland River and Willow Creek. A small improvement in R<sup>2</sup> model fit may be found if the Linear-With-Sill, Spherical, or Epanechnikov functions were used instead of the exponential autocovariance function, but the difference was not significant enough to warrant rerunning all models.

Table 13: A comparison of statistical metrics when different tail-up autocovariance functions are applied to each watershed using Lane Length Density, Weighted Permeability, and Rural and Undifferentiated Land Cover covariates. The resulting values found in this table are an average of the outputs for each of the individual longitudinal surveys. All values within 5% of the highest  $R^2$ , the lowest RMSPE, the lowest AIC, and within  $\pm 5\%$  of the ideal values for cov.80 (0.80), cov.90 (0.90), and cov.95 (0.95) are bolded.

Mimico Creek Watershed						
Autocovariance Function (Tail Up)	$R^2$	RMSPE	AIC	cov.80	cov.90	cov.95
Linear-With-Sill	0.11	<b>802.67</b>	<b>94.60</b>	0.8432	<b>0.9084</b>	<b>0.9372</b>
Spherical	0.11	<b>804.36</b>	<b>94.72</b>	0.8428	<b>0.8943</b>	<b>0.9425</b>
Epanechnikov	0.10	<b>803.34</b>	<b>94.23</b>	0.8431	<b>0.8886</b>	<b>0.9341</b>
Mariah	<b>0.13</b>	<b>812.58</b>	<b>97.50</b>	0.8597	<b>0.9230</b>	<b>0.9515</b>
Exponential	0.12	<b>806.13</b>	<b>95.56</b>	0.8431	<b>0.8917</b>	<b>0.9454</b>
East Holland Watershed						
Autocovariance Function (Tail Up)	$R^2$	RMSPE	AIC	cov.80	cov.90	cov.95
Linear-With-Sill	0.17	<b>68.87</b>	<b>0.40</b>	<b>0.8281</b>	<b>0.9007</b>	<b>0.9335</b>
Spherical	0.17	<b>69.07</b>	<b>0.40</b>	<b>0.8281</b>	<b>0.8961</b>	<b>0.9274</b>
Epanechnikov	<b>0.18</b>	<b>68.55</b>	<b>0.39</b>	<b>0.8281</b>	<b>0.8984</b>	<b>0.9335</b>
Mariah	<b>0.18</b>	72.88	<b>0.40</b>	<b>0.8334</b>	<b>0.9021</b>	<b>0.9320</b>
Exponential	<b>0.17</b>	<b>69.11</b>	<b>0.40</b>	<b>0.8281</b>	<b>0.8999</b>	<b>0.9274</b>
Willow Creek Watershed						
Autocovariance Function (Tail Up)	$R^2$	RMSPE	AIC	cov.80	cov.90	cov.95
Linear-With-Sill	<b>0.21</b>	<b>-49.26</b>	<b>0.12</b>	0.9050	<b>0.9394</b>	<b>0.9614</b>
Spherical	<b>0.21</b>	<b>-47.77</b>	<b>0.13</b>	0.9050	<b>0.9445</b>	<b>0.9614</b>
Epanechnikov	<b>0.22</b>	<b>-49.73</b>	<b>0.12</b>	0.8942	<b>0.9343</b>	<b>0.9639</b>
Mariah	<b>0.21</b>	-16.80	0.16	0.9127	0.9603	<b>0.9723</b>
Exponential	0.20	-44.37	<b>0.13</b>	0.9198	<b>0.9445</b>	<b>0.9711</b>

## 4.5 Model Selection

The Spatial Stream Network models used to study the Mimico Creek, East Holland River, and Willow Creek watersheds were based on lane length density, weighted permeability, and agriculture and undifferentiated rural land cover as covariates. All possible models based on combinations of these covariates (displayed in Table 6) were tested using model selection metrics, as discussed in Section 3.7, to create a model configuration that aimed to represent the watershed survey being studied. The expected model, the final selected model, and model selection metrics were organized and presented in Table 14 for Mimico Creek, Table 15 for East Holland River, and Table 16 for Willow Creek.

For the Mimico Creek watershed, the only season where the final selected model matched the expected best model was the Winter survey. The Global Model One (GMI; LLD + WPerm + AURL) was selected for Mimico Creek Spring 1, Fall, and Winter surveys as they provided the best  $R^2$  fit with the other selection metrics being comparable. Model 7 (M7; WPerm + AURL) was used in the Mimico Creek Spring 2 survey as all M7 metrics were comparable with GMI, but M7 was the simpler model and was thus chosen.

The expected best models for the East Holland River Summer and Winter surveys matched the models that were finally selected. The GMI was chosen for the Fall and Spring surveys as they had the best  $R^2$  fit, with all other metrics being comparable to the other “best” models.

In Willow Creek, no selected model matched the expected best models. Model six (M6; LLD + WPerm) was chosen over GMI in the Summer and Spring surveys as they both had comparable selection metrics, but M6 was more parsimonious than GMI. GMI was chosen for the Fall survey as it had the best fit along with the other metrics being comparable to the other “best” models. The model for the Willow Creek Spring survey was difficult to select as all model fits were poor, so It was decided to use GMI for this survey as it has the best fit with a tradeoff of higher potential error.

Table 14: The expected best models, model comparison metrics, and a sum of the number of metrics that fit within the “best” category for each model for the Mimico Creek dataset. All values within 5% of the highest R<sup>2</sup>, the lowest RMSPE, the lowest AIC, and within ±%5 of the ideal values for cov.80 (0.80), cov.90 (0.90), and cov.95 (0.95) are bolded and categorized as the “best” result for comparison with other models. The expected best model is highlighted in green, while the model with the highest number of bolded metrics is highlighted in yellow.

Mimico Creek	Expected Best Model	Selected Model	R <sup>2</sup>	AIC	RMSPE	cov.80	cov.90	cov.95	“Best” Metrics
Spring 1 Jun 14-16, 2016	GMI	<b>X</b>	<b>0.148</b>	<b>791.71</b>	68.88	<b>0.7945</b>	<b>0.9041</b>	<b>0.9178</b>	<b>5</b>
	M2		0.015	<b>818.24</b>	72.316	<b>0.8356</b>	<b>0.9041</b>	<b>0.9178</b>	<b>4</b>
	M3		0.013	<b>814.43</b>	75.002	<b>0.8356</b>	<b>0.8904</b>	<b>0.9589</b>	<b>4</b>
	M4		0.038	<b>811.65</b>	<b>64.644</b>	<b>0.8082</b>	<b>0.8904</b>	<b>0.9589</b>	<b>5</b>
	M5		0.029	<b>810.17</b>	77.134	<b>0.8356</b>	<b>0.9041</b>	<b>0.9315</b>	<b>4</b>
	M6		0.123	<b>800.72</b>	68.807	<b>0.7945</b>	<b>0.8904</b>	<b>0.9178</b>	<b>4</b>
	M7		0.114	<b>796.88</b>	<b>66.609</b>	<b>0.8082</b>	<b>0.8904</b>	<b>0.9315</b>	<b>5</b>
Summer Aug 03-05, 2016	GMI		0.041	<b>860.09</b>	<b>100.05</b>	0.8904	<b>0.9452</b>	<b>0.9589</b>	<b>4</b>
	M2		0.000	<b>884.64</b>	<b>97.706</b>	0.9041	0.9726	<b>0.9726</b>	<b>3</b>
	M3		0.001	<b>878.43</b>	<b>95.460</b>	0.9178	<b>0.9452</b>	<b>0.9726</b>	<b>4</b>
	M4	<b>X</b>	<b>0.060</b>	<b>873.33</b>	<b>99.522</b>	0.9041	<b>0.9315</b>	<b>0.9726</b>	<b>5</b>
	M5		0.002	<b>877.32</b>	<b>98.068</b>	0.9178	0.9589	<b>0.9726</b>	<b>3</b>
	M6		0.041	<b>867.65</b>	<b>99.364</b>	0.8904	<b>0.9452</b>	<b>0.9589</b>	<b>4</b>
	M7		0.041	<b>864.60</b>	<b>99.516</b>	0.8904	<b>0.9452</b>	<b>0.9589</b>	<b>4</b>
Winter Oct 11-13, 2016	GMI	<b>X</b>	<b>0.110</b>	<b>720.48</b>	54.894	<b>0.8235</b>	<b>0.8971</b>	<b>0.9265</b>	<b>5</b>
	M2		0.003	<b>743.49</b>	<b>50.864</b>	0.8529	<b>0.8971</b>	<b>0.9118</b>	<b>4</b>
	M3		0.089	<b>735.78</b>	<b>48.918</b>	0.8529	<b>0.8971</b>	<b>0.9118</b>	<b>4</b>
	M4		0.003	<b>737.83</b>	<b>50.348</b>	0.8529	<b>0.9118</b>	<b>0.9118</b>	<b>4</b>
	M5		0.099	<b>731.79</b>	<b>51.061</b>	<b>0.8235</b>	<b>0.8971</b>	<b>0.9118</b>	<b>5</b>
	M6		0.006	<b>732.75</b>	51.705	<b>0.8382</b>	<b>0.8971</b>	<b>0.9118</b>	<b>4</b>
	M7		0.100	<b>724.44</b>	52.896	<b>0.8382</b>	<b>0.8971</b>	<b>0.9118</b>	<b>4</b>
Fall Feb 20-23, 2017	GMI	<b>X</b>	<b>0.137</b>	<b>865.24</b>	177.58	<b>0.8235</b>	<b>0.8824</b>	<b>0.9265</b>	<b>5</b>
	M2		0.021	<b>891.25</b>	<b>149.49</b>	<b>0.8382</b>	<b>0.9265</b>	<b>0.9412</b>	<b>5</b>
	M3		0.043	<b>887.59</b>	<b>153.69</b>	0.8824	<b>0.8824</b>	<b>0.9265</b>	<b>4</b>
	M4		0.023	<b>883.74</b>	160.74	0.8529	<b>0.8824</b>	<b>0.9118</b>	<b>3</b>
	M5		0.051	<b>880.26</b>	<b>153.41</b>	<b>0.8382</b>	<b>0.9265</b>	<b>0.9412</b>	<b>5</b>
	M6		0.048	<b>877.66</b>	167.04	<b>0.8235</b>	<b>0.8676</b>	<b>0.9265</b>	<b>4</b>
	M7		0.086	<b>872.12</b>	171.61	<b>0.8235</b>	<b>0.8824</b>	<b>0.9118</b>	<b>4</b>
Spring 2 Apr 21-23, 2017	GMI		<b>0.110</b>	<b>775.86</b>	<b>71.615</b>	0.8841	<b>0.9130</b>	<b>0.9565</b>	<b>5</b>
	M2		0.008	<b>800.28</b>	<b>71.403</b>	0.8696	<b>0.9275</b>	<b>0.9565</b>	<b>4</b>
	M3		0.095	<b>791.93</b>	<b>68.543</b>	0.8841	<b>0.9130</b>	<b>0.9565</b>	<b>4</b>
	M4		0.004	<b>793.72</b>	74.451	<b>0.8551</b>	<b>0.9130</b>	<b>0.9565</b>	<b>4</b>
	M5		0.094	<b>787.58</b>	<b>68.943</b>	0.8841	<b>0.9130</b>	<b>0.9565</b>	<b>4</b>
	M6		0.012	<b>789.00</b>	74.579	<b>0.8551</b>	<b>0.9275</b>	<b>0.9565</b>	<b>4</b>
	M7	<b>X</b>	<b>0.105</b>	<b>780.25</b>	<b>71.096</b>	0.8841	<b>0.9130</b>	<b>0.9565</b>	<b>5</b>

Table 15: The expected best models, model comparison metrics, and a sum of the number of metrics that fit within the “best” category for each model for the East Holland dataset. All values within 5% of the highest R<sup>2</sup>, the lowest RMSPE, the lowest AIC, and within ±%5 of the ideal values for cov.80 (0.80), cov.90 (0.90), and cov.95 (0.95) are bolded and categorized as the “best” result for comparison with other models. The expected best model is highlighted in green, while the model with the highest number of bolded metrics is highlighted in yellow.

East Holland	Expected Best Model	Selected Model	R <sup>2</sup>	AIC	RMSPE	cov.80	cov.90	cov.95	“Best” Covariates
Summer Jul 21-25, 2016	GMI		<b>0.084</b>	55.544	0.5536	0.8537	<b>0.9024</b>	0.9024	<b>2</b>
	M2		0.034	<b>53.654</b>	0.531	<b>0.8293</b>	<b>0.8780</b>	0.9024	<b>3</b>
	M3		0.013	<b>54.407</b>	<b>0.3957</b>	<b>0.8293</b>	<b>0.9024</b>	<b>0.9268</b>	<b>5</b>
	M4	<b>X</b>	0.043	<b>52.388</b>	<b>0.3873</b>	<b>0.8293</b>	<b>0.9268</b>	<b>0.9512</b>	<b>5</b>
	M5		0.035	55.677	0.5792	<b>0.8293</b>	<b>0.8780</b>	0.9024	<b>2</b>
	M6		0.067	<b>54.357</b>	0.5122	0.8537	<b>0.9024</b>	<b>0.9512</b>	<b>3</b>
	M7		0.045	55.021	<b>0.3899</b>	0.8537	<b>0.9268</b>	<b>0.9512</b>	<b>3</b>
Fall Oct 24-26, 2016	GMI	<b>X</b>	<b>0.077</b>	<b>66.595</b>	0.4175	<b>0.8148</b>	<b>0.8889</b>	<b>0.9259</b>	<b>5</b>
	M2		0.030	<b>64.607</b>	<b>0.3956</b>	0.8704	<b>0.9074</b>	<b>0.9259</b>	<b>4</b>
	M3		0.003	<b>65.967</b>	<b>0.4015</b>	0.8519	<b>0.8889</b>	<b>0.9259</b>	<b>4</b>
	M4		0.014	<b>64.201</b>	<b>0.3962</b>	0.8519	<b>0.8889</b>	<b>0.9259</b>	<b>4</b>
	M5		0.068	<b>65.129</b>	<b>0.4097</b>	<b>0.8333</b>	<b>0.8889</b>	<b>0.9259</b>	<b>5</b>
	M6		0.051	<b>65.423</b>	<b>0.404</b>	0.8519	<b>0.9074</b>	<b>0.9074</b>	<b>4</b>
	M7		0.014	<b>67.318</b>	<b>0.4082</b>	0.8519	<b>0.8889</b>	<b>0.9259</b>	<b>4</b>
Winter Mar 02-05, 2017	GMI		<b>0.282</b>	<b>110.65</b>	<b>0.3459</b>	<b>0.8273</b>	<b>0.9182</b>	<b>0.9455</b>	<b>6</b>
	M2	<b>X</b>	<b>0.273</b>	<b>105.86</b>	<b>0.342</b>	<b>0.8273</b>	<b>0.9000</b>	<b>0.9455</b>	<b>6</b>
	M3		0.123	123.6	0.3783	<b>0.8364</b>	<b>0.9182</b>	<b>0.9455</b>	<b>3</b>
	M4		0.023	133.9	0.3983	0.8545	<b>0.9000</b>	<b>0.9273</b>	<b>2</b>
	M5		<b>0.279</b>	<b>108.23</b>	<b>0.3442</b>	<b>0.8364</b>	<b>0.9091</b>	<b>0.9455</b>	<b>6</b>
	M6		<b>0.277</b>	<b>108.08</b>	<b>0.3438</b>	<b>0.8273</b>	<b>0.9000</b>	<b>0.9545</b>	<b>6</b>
	M7		0.128	125.56	0.379	<b>0.8273</b>	<b>0.9273</b>	<b>0.9455</b>	<b>3</b>
Spring May 15-17, 2017	GMI	<b>X</b>	<b>0.253</b>	<b>43.639</b>	<b>0.2798</b>	<b>0.8165</b>	<b>0.8899</b>	<b>0.9358</b>	<b>6</b>
	M2		0.191	<b>44.682</b>	<b>0.2798</b>	<b>0.8257</b>	<b>0.8807</b>	<b>0.9266</b>	<b>5</b>
	M3		0.134	50.424	<b>0.2807</b>	0.8532	<b>0.9083</b>	<b>0.9541</b>	<b>3</b>
	M4		0.048	58.400	0.3322	<b>0.8257</b>	<b>0.8899</b>	<b>0.9266</b>	<b>3</b>
	M5		0.228	<b>43.947</b>	<b>0.2706</b>	<b>0.8165</b>	<b>0.8899</b>	<b>0.9541</b>	<b>5</b>
	M6		0.223	<b>43.281</b>	0.2908	<b>0.8073</b>	<b>0.8899</b>	<b>0.9266</b>	<b>4</b>
	M7		0.158	50.249	0.2871	0.8440	<b>0.9174</b>	<b>0.9450</b>	<b>2</b>

Table 16: The expected best models, model comparison metrics, and a sum of the number of metrics that fit within the “best” category for each model for the Willow Creek dataset. All values within 5% of the highest R<sup>2</sup>, the lowest RMSPE, the lowest AIC, and within ±%5 of the ideal values for cov.80 (0.80), cov.90 (0.90), and cov.95 (0.95) are bolded and categorized as the “best” result for comparison with other models. The expected best model is highlighted in green, while the model with the highest number of bolded metrics is highlighted in yellow.

Willow Creek	Expected Best Model	Selected Model	R <sup>2</sup>	AIC	RMSPE	cov.80	cov.90	cov.95	“Best” Covariates
Summer Jul 28-29, 2016	GMI		<b>0.357</b>	<b>-49.085</b>	0.209	0.9143	<b>0.9429</b>	<b>0.9714</b>	<b>4</b>
	M2		0.309	<b>-51.103</b>	0.1762	0.9143	<b>0.9429</b>	<b>0.9714</b>	<b>3</b>
	M3		0.004	-37.134	<b>0.1227</b>	0.8857	<b>0.9429</b>	1.0000	<b>2</b>
	M4		0.066	-40.316	0.141	<b>0.8286</b>	<b>0.9429</b>	1.0000	<b>2</b>
	M5		0.315	-48.711	0.1924	0.9143	<b>0.9429</b>	<b>0.9714</b>	<b>2</b>
	M6	<b>X</b>	<b>0.353</b>	<b>-51.579</b>	0.1889	0.8857	<b>0.9429</b>	<b>0.9714</b>	<b>4</b>
	M7		0.068	-38.014	0.1672	0.8571	<b>0.9143</b>	1.0000	<b>1</b>
Fall Oct 18-19, 2016	GMI	<b>X</b>	<b>0.117</b>	-22.8	<b>0.1216</b>	<b>0.8980</b>	<b>0.9388</b>	<b>0.9796</b>	<b>5</b>
	M2		0.073	<b>-28.524</b>	<b>0.1218</b>	<b>0.8980</b>	<b>0.9388</b>	<b>0.9796</b>	<b>5</b>
	M3		0.002	-23.693	<b>0.1241</b>	0.9184	<b>0.9184</b>	<b>0.9388</b>	<b>3</b>
	M4		0.007	-25.466	<b>0.1252</b>	<b>0.8980</b>	<b>0.9184</b>	<b>0.9388</b>	<b>4</b>
	M5		0.078	-23.972	<b>0.1228</b>	<b>0.8980</b>	<b>0.9388</b>	<b>0.9796</b>	<b>4</b>
	M6		0.103	-26.8	<b>0.1219</b>	<b>0.8980</b>	<b>0.9388</b>	<b>0.9796</b>	<b>4</b>
	M7		0.012	-21.006	<b>0.1263</b>	<b>0.8980</b>	<b>0.9184</b>	<b>0.9388</b>	<b>4</b>
Winter Feb 06-08, 2017	GMI		<b>0.207</b>	-33.098	0.0957	<b>0.9259</b>	<b>0.9259</b>	<b>0.9630</b>	<b>4</b>
	M2		0.118	<b>-36.541</b>	0.093	<b>0.9259</b>	<b>0.9444</b>	<b>0.9444</b>	<b>4</b>
	M3		0.034	-32.222	<b>0.0849</b>	<b>0.9259</b>	<b>0.9259</b>	<b>0.9259</b>	<b>4</b>
	M4		0.030	-32.822	<b>0.0891</b>	<b>0.9259</b>	<b>0.9259</b>	<b>0.9259</b>	<b>4</b>
	M5		0.135	-32.774	0.098	<b>0.9259</b>	<b>0.9259</b>	<b>0.9259</b>	<b>3</b>
	M6	<b>X</b>	<b>0.198</b>	<b>-37.285</b>	0.0937	0.9444	<b>0.9444</b>	<b>0.9630</b>	<b>4</b>
	M7		0.061	-29.738	0.091	<b>0.9259</b>	<b>0.9259</b>	<b>0.9259</b>	<b>3</b>
Spring Apr 17-18, 2017	GMI	<b>X</b>	<b>0.123</b>	-72.502	0.074	0.9412	0.9706	<b>0.9706</b>	<b>2</b>
	M2		0.066	<b>-78.136</b>	<b>0.0687</b>	0.9118	0.9706	<b>0.9706</b>	<b>3</b>
	M3		0.045	<b>-77.371</b>	<b>0.0667</b>	0.9265	<b>0.9559</b>	<b>0.9706</b>	<b>4</b>
	M4		0.024	<b>-77.155</b>	<b>0.0675</b>	<b>0.8824</b>	0.9706	<b>0.9706</b>	<b>4</b>
	M5		0.105	<b>-75.452</b>	0.0739	0.9412	0.9706	<b>0.9706</b>	<b>2</b>
	M6		0.083	<b>-75.127</b>	0.0712	0.9118	0.9706	<b>0.9706</b>	<b>2</b>
	M7		0.070	<b>-74.874</b>	<b>0.0699</b>	0.8971	0.9706	<b>0.9706</b>	<b>3</b>

#### 4.6 Model Outputs and Covariate Significance

Final selected model outputs along with covariate significance were organized in Table 17 for Mimico Creek, Table 18 for East Holland, and Table 19 for Willow Creek. The model outputs for Mimico Creek showed that Weighted Permeability (WPerm) and Agricultural and Undifferentiated Rural Land (AURL) were the only significant predictor variables, with WPerm having a positive relationship with in-stream  $\text{Cl}^-$  and AURL having a negative relationship with in-stream  $\text{Cl}^-$ . The Summer and Fall East Holland models had a more limited sampling area and these models had no significant covariates, whereas Lane Length Density was a significant covariate for both Winter and Spring surveys and this variable had a positive relationship with in-stream  $\text{Cl}^-$ . Lane Length Density was significant in all seasons for the Willow Creek model outputs while Weighted Permeability was significant only in the Winter survey.

Table 17: Spatial Stream Network model outputs for the Mimico Creek watershed. All predictor variables with  $p < 0.05$  are considered significant and are bolded. Model  $R^2$  and proportion of variation in Cl<sup>-</sup> concentration explained by covariates and spatial component is included beneath the outputs for each model.

Mimico Creek Watershed				
Spring 1 (GMI) Jun 14-16, 2016				
Predictor	b	Std. Error	t	p
<b>Intercept</b>	434.26	99.12	4.38	<b>p &lt;&lt; 0.001</b>
Lane Length Density	2.67	1.76	1.52	p > 0.05
<b>Weighted Permeability</b>	272.10	91.93	2.96	<b>p &lt; 0.01</b>
AURL	-18.88	13.34	-1.42	p > 0.05
R <sup>2</sup> = 0.15 (Covariates: 14.8%; Linear-with-Sill Tail-Up: 85.2%; Nugget: <0.01%)				
Summer (M4) Aug 03-05, 2016				
Predictor	b	Std. Error	t	P
<b>Intercept</b>	291.50	102.00	2.86	<b>p &lt; 0.01</b>
<b>Weighted Permeability</b>	259.40	121.80	2.13	<b>p &lt; 0.05</b>
R <sup>2</sup> = 0.06 (Covariates: 6.0%; Linear-with-Sill Tail-Up: 84.1%; Nugget: 9.9%)				
Fall (GMI) Oct 11-13, 2016				
Predictor	b	Std. Error	T	p
<b>Intercept</b>	211.51	88.946	2.378	<b>p &lt; 0.05</b>
Lane Length Density	1.75	2.068	0.845	p > 0.05
Weighted Permeability	71.80	77.958	0.921	p > 0.05
<b>AURL</b>	-29.046	11.034	-2.632	<b>p &lt; 0.05</b>
R <sup>2</sup> = 0.11 (Covariates: 11.0%; Linear-with-Sill Tail-Up: 88.8%; Nugget: 0.2%)				
Winter (GMI) Feb 20-23, 2017				
Predictor	b	Std. Error	t	p
<b>Intercept</b>	1183.53	283.02	4.182	<b>p &lt;&lt; 0.001</b>
Lane Length Density	-9.035	6.331	-1.427	p > 0.05
<b>Weighted Permeability</b>	457.967	223.987	2.045	<b>p &lt; 0.05</b>
<b>AURL</b>	-77.337	35.651	-2.169	<b>p &lt; 0.05</b>
R <sup>2</sup> = 0.14 (Covariates: 13.7%; Linear-with-Sill Tail-Up: 86.2%; Nugget: <0.01%)				
Spring 2 (M7) Apr 21-23, 2017				
Predictor	b	Std. Error	t	p
<b>Intercept</b>	333.08	87.41	3.811	<b>p &lt; 0.001</b>
Weighted Permeability	84.64	90.14	0.939	p > 0.05
<b>AURL</b>	-33.79	12.55	-2.693	<b>p &lt; 0.001</b>
R <sup>2</sup> = 0.11 (Covariates: 10.5%; Linear-with-Sill Tail-Up: 78.4%; Nugget: 11.1%)				

Table 18: Spatial Stream Network model outputs for the East Holland River watershed. All predictor variables that have a significance value of  $p < 0.05$  are considered significant and are bolded. Model  $R^2$  and proportion of variation in  $Cl^-$  concentration explained by covariates and spatial component is included beneath the outputs for each model.

East Holland River Watershed				
Summer (M4) Jul 21-25, 2016				
Predictor	b	Std. Error	t	p
<b>Intercept</b>	5.0008	0.1361	36.749	<b>p &lt;&lt; 0.001</b>
Weighted Permeability	-0.1883	0.1423	-1.323	p > 0.05
$R^2 = 0.04$ (Covariates: 4.3%; Exponential Tail-Up: 95.7%; Nugget: <0.01%)				
Fall (GMI) Oct 24-26, 2016				
Predictor	b	Std. Error	t	p
<b>Intercept</b>	4.7251	0.2249	21.014	<b>p &lt;&lt; 0.001</b>
Lane Length Density	0.2056	0.111	1.851	p > 0.05
Weighted Permeability	0.1027	0.1539	0.667	p > 0.05
AURL	0.1346	0.1096	1.229	p > 0.05
$R^2 = 0.08$ (Covariates: 7.7%; Exponential Tail-Up: 84.9%; Nugget: 7.4%)				
Winter (M2) Mar 02-05, 2017				
Predictor	b	Std. Error	t	p
<b>Intercept</b>	4.6488	0.1141	40.742	<b>p &lt;&lt; 0.001</b>
<b>Lane Length Density</b>	0.34536	0.05419	6.374	<b>p &lt;&lt; 0.001</b>
$R^2 = 0.27$ (Covariates: 27.3%; Exponential Tail-Up: 55.8%; Nugget: 16.9%)				
Spring (GMI) May 15-17, 2017				
Predictor	b	Std. Error	t	p
<b>Intercept</b>	4.86176	0.11256	43.191	<b>p &lt;&lt; 0.001</b>
<b>Lane Length Density</b>	0.17018	0.04947	3.44	<b>p &lt; 0.001</b>
Weighted Permeability	-0.16333	0.08859	-1.844	p > 0.05
AURL	-0.10169	0.05255	-1.935	p > 0.05
$R^2 = 0.25$ (Covariates: 25.3%; Exponential Tail-Up: 64.4%; Nugget: 10.3%)				

Table 19: Spatial Stream Network model outputs for the Willow Creek watershed. All predictor variables that have a significance value of  $p < 0.05$  are considered significant and are bolded. Model  $R^2$  and proportion of variation in  $Cl^-$  concentration explained by covariates and spatial component is included beneath the outputs for each model.

Willow Creek Watershed				
Summer (M6) Jul 21-25, 2016				
Predictor	b	Std. Error	t	p
<b>Intercept</b>	1.6835	0.4699	3.582	<b>p &lt; 0.001</b>
<b>Lane Length Density</b>	1.0016	0.2653	3.775	<b>p &lt; 0.001</b>
Weighted Permeability	0.2566	0.173	1.484	p > 0.05
$R^2 = 0.35$ (Covariates: 35.3%; Exponential Tail-Up: 64.7%; Nugget: <0.01%)				
Fall (GMI) Oct 24-26, 2016				
Predictor	b	Std. Error	t	p
<b>Intercept</b>	2.79377	0.25743	10.852	<b>p &lt;&lt; 0.001</b>
<b>Lane Length Density</b>	0.17598	0.07603	2.315	<b>p &lt; 0.05</b>
Weighted Permeability	0.11725	0.08367	1.401	p > 0.05
AURL	0.03229	0.03733	0.865	p > 0.05
$R^2 = 0.12$ (Covariates: 11.7%; Exponential Tail-Up: 88.3%; Nugget: <0.01%)				
Winter (M6) Mar 02-05, 2017				
Predictor	b	Std. Error	t	p
<b>Intercept</b>	2.73299	0.15173	18.012	<b>p &lt; 0.001</b>
<b>Lane Length Density</b>	0.1417	0.04344	3.262	<b>p &lt; 0.01</b>
<b>Weighted Permeability</b>	0.12121	0.05499	2.204	<b>p &lt; 0.05</b>
$R^2 = 0.20$ (Covariates: 19.8%; Exponential Tail-Up: 80.2%; Nugget: <0.01%)				
Spring (GMI) May 15-17, 2017				
Predictor	b	Std. Error	t	p
<b>Intercept</b>	3.04969	0.11681	26.109	<b>p &lt;&lt; 0.001</b>
<b>Lane Length Density</b>	0.03629	0.01854	1.957	p > 0.05
Weighted Permeability	0.0548	0.04762	1.151	p > 0.05
AURL	-0.0441	0.0261	-1.69	p > 0.05
$R^2 = 0.12$ (Covariates: 12.2%; Exponential Tail-Up: 87.8%; Nugget: <0.01%)				

#### 4.7 Statistical Stream Network Modeling vs. Euclidean Modeling

Euclidean models were generated using the same selected covariates for the Spatial Stream Network (SSN) models as seen in Table 14 for the Mimico Creek watershed, for Table 15 East Holland River watershed, and Table 16 for Willow Creek watershed. The Euclidean models measure straight line separation distances and ignore stream topology and stream flow directionality, so these models provide a comparative reference for SSN models that measured separation distance only along the stream network, and only measured separation distances between points that were flow connected. All Euclidean models used an exponential autocovariance function to match the exponential autocovariance functions used for East Holland River and for Willow Creek, and because there is not an equivalent Linear-With-Sill autocovariance function available for use in Mimico Creek. The Euclidean model outputs are paired with their comparative SSN model outputs in Table 20.

Comparing the entire three watershed dataset, the SSN models resulted in  $R^2$  (-0.4%) and AIC (-5.2%) values that are comparable to their Euclidean counterparts. The SSN models had a large decrease in RMSPE of -29.4% when compared to the Euclidean models. The SSN models were not consistently better than Euclidean models which was disappointing. Only the SSN models for the Willow Creek watershed were an improvement over Euclidean models with an 83% increase in average  $R^2$  values. The SSN models for Mimico Creek and East Holland River showed  $R^2$  values that were 17% and 28% lower than the average value of the comparative Euclidean Model  $R^2$  fits. The AIC values were lower (and therefore better) in the SSN models across all three watersheds, and the RMSPE values were only lower in Mimico Creek (-30%) and in the East Holland (-10%) watersheds. The Willow Creek SSN models had an RMSPE increase of 27% compared to the Willow Creek Euclidean models.

Comparing SSN and Euclidean surveys within a watershed also yielded highly variable results. The SSN models for Mimico Creek Spring 1 and Summer surveys showed significant increases in  $R^2$ , slight

decreases in AIC, and significant decreases in RMSPE compared to the Spring 1 and Summer Euclidean surveys. Model performance improvement reverses in the SSN Mimico Creek Winter and Spring 2 surveys, where the models had significantly lower  $R^2$  fits relative to the Euclidean models. Increasing sample sizes in East Holland River Winter and Spring surveys resulted in significantly better SSN  $R^2$  values when compared to the SSN Summer and Fall surveys, but the SSN methodologies did not outperform Euclidean  $R^2$  fits apart from a 5% improvement in  $R^2$  fit in the East Holland River Winter survey. SSN methodologies did have modest improvements in AIC and RMSE for East Holland River and Willow Creek surveys - except for the East Holland River Fall Survey and Willow Creek Summer and Winter surveys. In contrast to the East Holland River, Willow Creek Summer and Winter SSN models had an 6960% increase and 450% increase in model fit  $R^2$  values compared to their respective Euclidean models.

Table 20: A comparison of the Spatial Stream Network models used in this thesis to a Euclidean counterpart. All Euclidean models in this thesis used the exponential autocovariance function and uses the direct distance between points instead of stream distances.

Mimico Creek						
Survey (S)	R <sup>2</sup>	AIC	RMSPE	cov.80	cov.90	cov.95
S1 (M4: Tail Up)	<b>0.15</b>	791.71	68.88	0.7945	0.9041	0.9178
S1 (M4: Euclidean)	<b>0.17</b>	843.64	91.19	0.8767	<b>0.9178</b>	0.9315
S2 (M4: Tail Up)	<b>0.06</b>	<b>873.33</b>	<b>99.52</b>	0.9041	0.9315	0.9726
S2 (M4: Euclidean)	0.02	924.23	144.98	<b>0.8356</b>	<b>0.8493</b>	<b>0.9178</b>
S3 (GMI: Tail Up)	<b>0.11</b>	<b>720.48</b>	<b>54.89</b>	0.8235	<b>0.8971</b>	<b>0.9265</b>
S3 (GMI: Euclidean)	<b>0.11</b>	760.86	68.25	<b>0.8088</b>	0.8824	<b>0.9265</b>
S4 (GMI: Tail Up)	0.14	<b>865.24</b>	<b>177.58</b>	0.8235	<b>0.8824</b>	<b>0.9265</b>
S4 (GMI: Euclidean)	<b>0.18</b>	934.54	283.55	<b>0.8088</b>	0.8676	<b>0.9265</b>
S5 (M7: Tail Up)	0.11	<b>780.25</b>	<b>71.10</b>	0.8841	0.9130	<b>0.9565</b>
S5 (M7: Euclidean)	<b>0.20</b>	788.70	81.68	<b>0.8261</b>	<b>0.8841</b>	<b>0.9565</b>
East Holland						
S1 (M4: Tail Up)	0.043	<b>52.39</b>	<b>0.3873</b>	<b>0.8293</b>	<b>0.9268</b>	<b>0.9512</b>
S1 (M4: Euclidean)	<b>0.096</b>	54.49	0.4147	0.9024	0.9512	<b>0.9512</b>
S2 (GMI: Tail Up)	0.077	<b>66.60</b>	0.4175	<b>0.8148</b>	0.8889	<b>0.9259</b>
S2 (GMI: Euclidean)	<b>0.19</b>	74.72	<b>0.4142</b>	0.8333	<b>0.9074</b>	0.9210
S3 (M2: Tail Up)	<b>0.27</b>	<b>105.86</b>	<b>0.3420</b>	0.8273	<b>0.9000</b>	<b>0.9455</b>
S3 (M2: Euclidean)	0.26	149.19	0.4245	<b>0.8727</b>	0.9364	0.9727
S4 (GMI: Tail Up)	0.25	<b>43.64</b>	<b>0.2798</b>	<b>0.8165</b>	<b>0.8899</b>	0.9358
S4 (GMI: Euclidean)	<b>0.35</b>	90.69	0.3331	0.8807	0.9266	<b>0.9450</b>
Willow Creek						
S1 (M6: Tail Up)	<b>0.35</b>	<b>-51.58</b>	0.1889	0.8857	0.9429	<b>0.9714</b>
S1 (M6: Euclidean)	0.005	-47.09	<b>0.1187</b>	<b>0.8286</b>	<b>0.9143</b>	<b>0.9714</b>
S2 (GMI: Tail Up)	0.12	-22.80	<b>0.1216</b>	0.8980	0.9388	0.9796
S2 (GMI: Euclidean)	<b>0.12</b>	<b>-32.41</b>	0.1245	<b>0.8571</b>	<b>0.9184</b>	<b>0.9592</b>
S3 (M6: Tail Up)	<b>0.20</b>	-37.29	0.0937	0.9444	0.9444	0.9630
S3 (M6: Euclidean)	0.036	<b>-81.63</b>	<b>0.0666</b>	<b>0.8704</b>	<b>0.9074</b>	<b>0.9444</b>
S4 (GMI: Tail Up)	0.12	-72.50	0.0740	0.9412	0.9706	0.9706
S4 (GMI: Euclidean)	<b>0.27</b>	<b>-121.52</b>	<b>0.0653</b>	<b>0.8382</b>	<b>0.8676</b>	<b>0.9559</b>

#### 4.8 Spatial Stream Network Flow-Connected Residual Semivariograms

Flow-connected residual model semivariograms are depicted in Figures 20-22 for Mimico, East Holland, and Willow Creek respectively, and the spatial structure in these results are expanded on in the paragraphs below. The following section will focus on interpreting residual spatial structure using semivariograms.

The Mimico Creek Watershed Spring 1, Summer, and Fall model residual semivariograms depicted a linear variance increase with a peak, then a subsequent linear variance decrease, corresponding to a repeating small scale spatial pattern centered at a 6000 meter sample point separation distance. This small scale semivariance pattern was replaced by a larger scale linear-with-sill pattern in the medium-high flow Mimico Creek Winter survey. The high flow regime of the Mimico Creek Spring 2 survey resulted in a near-flat semivariogram, indicating a loss of spatial dependence at all separation distances.

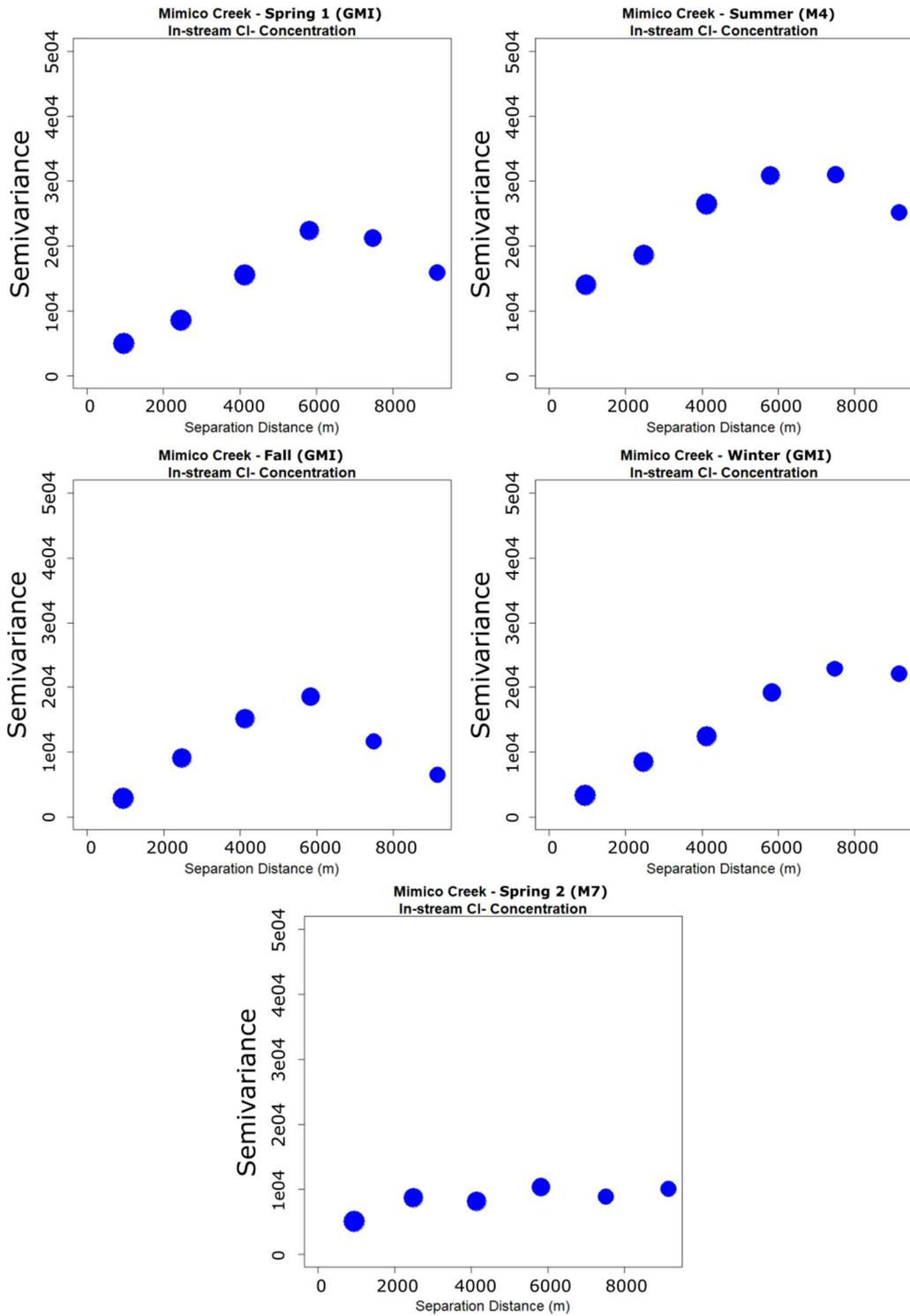


Figure 20: Semivariograms for the Spring 1, Summer, Fall, Winter, and Spring 2 longitudinal surveys of the Mimico Creek watershed.

The Summer survey for the East Holland River depicted a similar linear variance increase with a sudden linear variance decrease centered around 1000 meter separation distance. This small scale signal in the semivariogram was lost in the low-medium flow regime of the East Holland River Fall survey as the curve depicts a simple linear increase in semivariance with no sill. The East Holland River Winter and Spring surveys both exhibit a flat semivariogram, corresponding to a loss of spatial dependence at all separation distances.

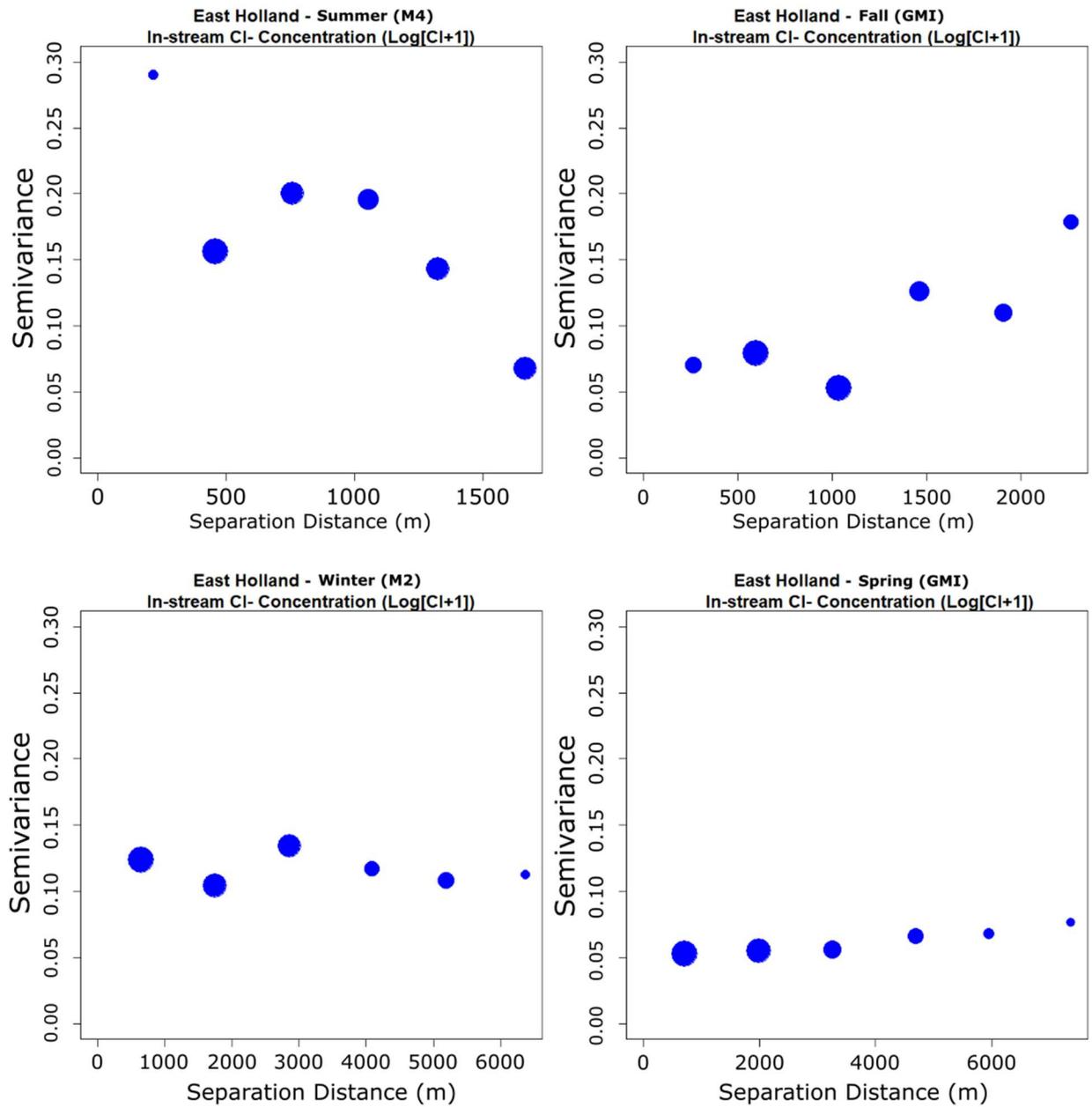


Figure 21: Semivariograms for the Summer, Fall, Winter, and Spring longitudinal surveys of the East Holland watershed.

A clear positive linear semivariance trend is evident in the low flow regime of the Willow Creek Summer survey, which changes to an exponentially increasing curve in the low-medium flow of the Willow Creek Fall survey. This exponential rise in semivariance is subdued with a sill in the medium-high flow regime of the Willow Creek Winter survey. A near-flat linear trend is seen in the high flow regime of the Willow Creek Spring survey, albeit with an extreme increase in semivariance in the highest separation distances.

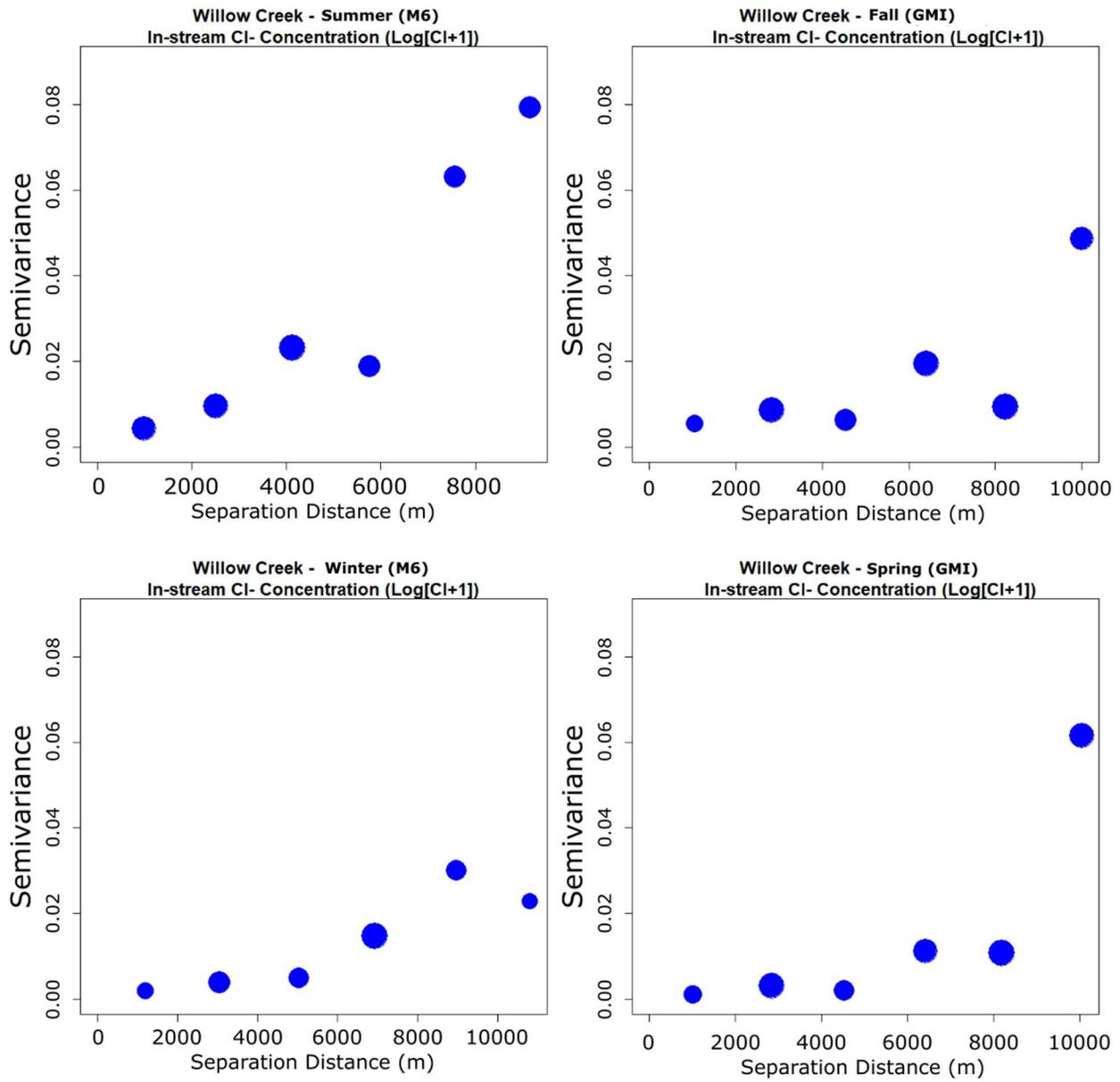


Figure 22: Semivariograms for the Summer, Fall, Winter, and Spring longitudinal surveys of the Willow Creek watershed.

## 5 Discussion

Previous attempts to map salt vulnerable areas have relied on data from large scale correlative models that relate salt concentrations at a catchment outlet to average landscape characteristics and mean annual flow. The goals of this study were to: i) explore the spatial structure of in-stream  $\text{Cl}^-$  concentrations in three watersheds that span a gradient in urbanization across seasons and flow states and ii) assess the utility of potential landscape predictors of salt vulnerable areas. The Spatial Stream Network (SSN) model, a novel geospatial technique developed by Ver Hoef et al. (2006), was used to develop high resolution models of in-stream  $\text{Cl}^-$  concentrations across three study watersheds in Southern Ontario. Spatially intensive in-stream field data were collected in four seasons and used to develop the models. The following research questions were addressed: 1) How does SSN  $\text{Cl}^-$  model performance vary with respect to Urbanization and Flow state?, 2) Does the importance of landscape predictors of in-stream  $\text{Cl}^-$  concentrations vary across seasons and watersheds?, 3) Do the spatial patterns of residual semivariograms from SSN models of in-stream  $\text{Cl}^-$  concentrations vary with urbanization or flow state?

### 5.1 Spatial Patterns in Stream Chloride Concentrations

The Mimico Creek, East Holland River, and Willow Creek watersheds all reveal spatial structure when viewing their observed in-stream  $\text{Cl}^-$  concentration maps (Figures 17-19). The field data from Mimico Creek displayed a small scale high-low-high-low  $\text{Cl}^-$  concentration pattern in the Spring 1, Summer, and Fall surveys from the headwaters to the outlet, while the Winter and Spring 2 surveys lost this small scale pattern for a homogenous, large scale  $\text{Cl}^-$  distribution with no repeating pattern. The Summer and Fall East Holland River surveys were confined to the western urban area in the watershed, had low separation distances, and had small scale patchiness that was limited to the headwaters. The Winter and Spring surveys for the East Holland River had small scale  $\text{Cl}^-$  concentration variation with no discernable

pattern in the western urban region of the watershed. This small scale patchiness was lost for a homogenous, low  $\text{Cl}^-$  concentration across the eastern agricultural section of East Holland River. The Willow Creek watershed exhibited homogenous in-stream  $\text{Cl}^-$  concentrations within the headwaters, however there was a sharp increase in  $\text{Cl}^-$  concentration for all points downstream of Little Lake and the town of Midhurst. This discontinuity, upstream and downstream of Little Lake, may be due to the influence of urbanized area both south and downstream of the lake, or it may be due to an unrepresentative EC- $\text{Cl}^-$  curve caused by a low sample count. In general, under low flow conditions, the urban areas (i.e., all of Mimico Creek, western portion of East Holland River, and downstream reaches of Willow Creek) had more small scale spatial patterns indicative of localized zones of elevated  $\text{Cl}^-$  concentrations.

The three study watersheds were simply categorized as follows: Mimico Creek exhibited a repeating small scale pattern in in-stream  $\text{Cl}^-$  concentrations from the headwaters to the outlet. The East Holland River exhibited small scale patchiness in the western urbanized section of the watershed and larger scale homogeneity in the eastern agricultural area. The Willow Creek watershed displayed a relatively stable two-step large scale homogeneity, with the first low concentration homogeneous section in the headwaters, and the second high concentration homogeneous section downstream of Little Lake.

## 5.2 SSN Chloride Models

### 5.2.1 Effect of spatial relationship (Euclidean vs. flow-connected) used in SSN models

Previous studies have shown that SSN methodologies improved in-stream model performance within the context of stream temperature (Isaak et al., 2009; Steel et al., 2016; Turschwell et al., 2016), stream chemistry (McGuire et al., 2014), and biotic abundance (Lois, 2016). In this thesis, SSN models were not necessarily better than comparable Euclidean models as there was a general tradeoff between  $R^2$  fits

and other model performance metrics (e.g. AIC, RMSPE, Cov.80, ...). For example, SSN models had a large improvement in  $R^2$  seen in Willow Creek (+83%) but was associated with a large increase in RMSPE (+27%) and a large decrease in AIC (-35%). The decrease in  $R^2$  fits for the SSN models for Mimico Creek (-18%) and East Holland River (-28%) was paired with lower (i.e. better) AIC (MC: -5%, EH:-27%) and RMSPE (MC: -30%, EH: -10%) values. Peterson & Urquhart (2006) and Peterson et al. (2006) found that if there was a lack of spatial neighbors found within the boundaries of a spatial neighborhood, then an SSN method may be no better than a non-spatial method to compare points. This may apply in the case of East Holland River and Willow Creek stream networks where sample densities were lower given that stream access was limited to points where broadly spaced roads intersected the stream in rural areas. Additionally, the lack of reach scale predictor variables, combined with a low sampling density, may have reduced SSN model performance when compared to their Euclidean counterparts. The Mimico Creek watershed had an extremely simple stream network with few tributaries and with an overall linear shape, so this watershed may not have been the best choice for the application of SSN methodologies.

### 5.2.2 Comparison of SSN Model Performance With Respect to Urbanization

The primary reason to collect sample data across Mimico Creek, East Holland River, and Willow Creek watersheds was to obtain stream data from environments in differing states of urbanization. Urban stormsewer infrastructure can alter flow pathways through a watershed and even alter the real catchment area beyond its natural boundaries (Kayembe & Mitchell, 2018). The addition of artificial and altered pathways may introduce variables that may not be present in natural watersheds, and it is of interest to see what proportion of the model fit is accounted by the tail-up SSN methodology within each watershed.

There was no clear relationship in SSN performance (Table 20) when looking at each watershed in terms of urbanization as described in Section 3.1. However, if the Summer and Fall East Holland River surveys are considered highly urbanized due to their survey area being limited to the urbanized western section of the study area, Mimico Creek (Spring 1, Summer, Winter, Fall, Spring 2) and East Holland River (Summer and Fall) can be considered as the urbanized watersheds, East Holland River (Winter and Spring) as the semi-urban surveys, and Willow Creek (Summer, Fall, Winter, Spring) as the rural surveys. Using this new urbanization classification, and comparing SSN performance to a Euclidean baseline, there existed more of an association between urbanization and SSN model performance with the urban surveys having a 30% lower  $R^2$  compared to baseline, the semi-urban SSN Models having the second lowest  $R^2$  at -13% below the baseline, and the rural SSN models having 83% higher  $R^2$  than the Euclidean baseline. SSN model performance improvements for AIC were less consistent with urbanization state, with urban, semi-urban, and non-urban SSN models showing lower AIC (i.e. better than baseline) values of -5%, -37%, and -34%, respectively. The RMSPE values had the opposite trend with the non-urban SSN models performing worse than the Euclidean models with a 27% higher RMSPE value, the semi-urban performing better with a RMSPE value that 17% below the baseline, and the urban SSN models performing the best with a RMSPE value that is 29% lower than the Euclidean baseline.

The SSN models for the urban (Mimico Creek; East Holland River S1 & S2) and semi-urban watersheds (East Holland River S3 & S4) tended to have lower (i.e. worse)  $R^2$  values, but lower (i.e. better) AIC and RMSPE values. The SSN models for the rural Willow Creek watershed had overall better  $R^2$ , AIC, and RMSPE values. The contributions to the urban and semi-urban  $R^2$  values by the nugget, 6.5% and 14% respectively, which represents the error in the dataset or variation in the dataset at separation distances smaller than what was sampled, may be evidence that there may be localized contributions to in-stream Cl<sup>-</sup> that are not resolved by the SSN models. The nugget contributes <0.1% to the  $R^2$  value for the rural Willow Creek watershed, thus any such resolution issues may not have

occurred. The covariates used in this thesis also seem to explain more of the variance seen in the semi-urban (26.3%) and rural watersheds (19.75%), and less so in the urban watershed (9.7%). It seems that much of the variation within the urban dataset was explained by the tail-up methodology (86%), but with the nugget at 6.5% and covariates only explaining 10%. This result shows that the consideration of spatial structure can be very important as it explains the clear majority of the variance in stream  $\text{Cl}^-$  concentration. As well, there is a need to refine predictor variables to better represent the processes that deliver  $\text{Cl}^-$  to the stream and increase sampling densities to better reflect reach-scale variations in predictor values.

### 5.2.3 Comparison of SSN model performance across flow states in Mimico Creek

Mimico Creek was the only one of the three watersheds where consistent sample numbers were achieved across all surveys; hence it is the only survey that SSN performance can be compared across flow states. Mimico Creek SSN model  $R^2$ , AIC, and RMSPE have no clear correlation with watershed flow (Figure 23), so the flow states in Mimico Creek watershed had no clear association with model performance in this thesis.

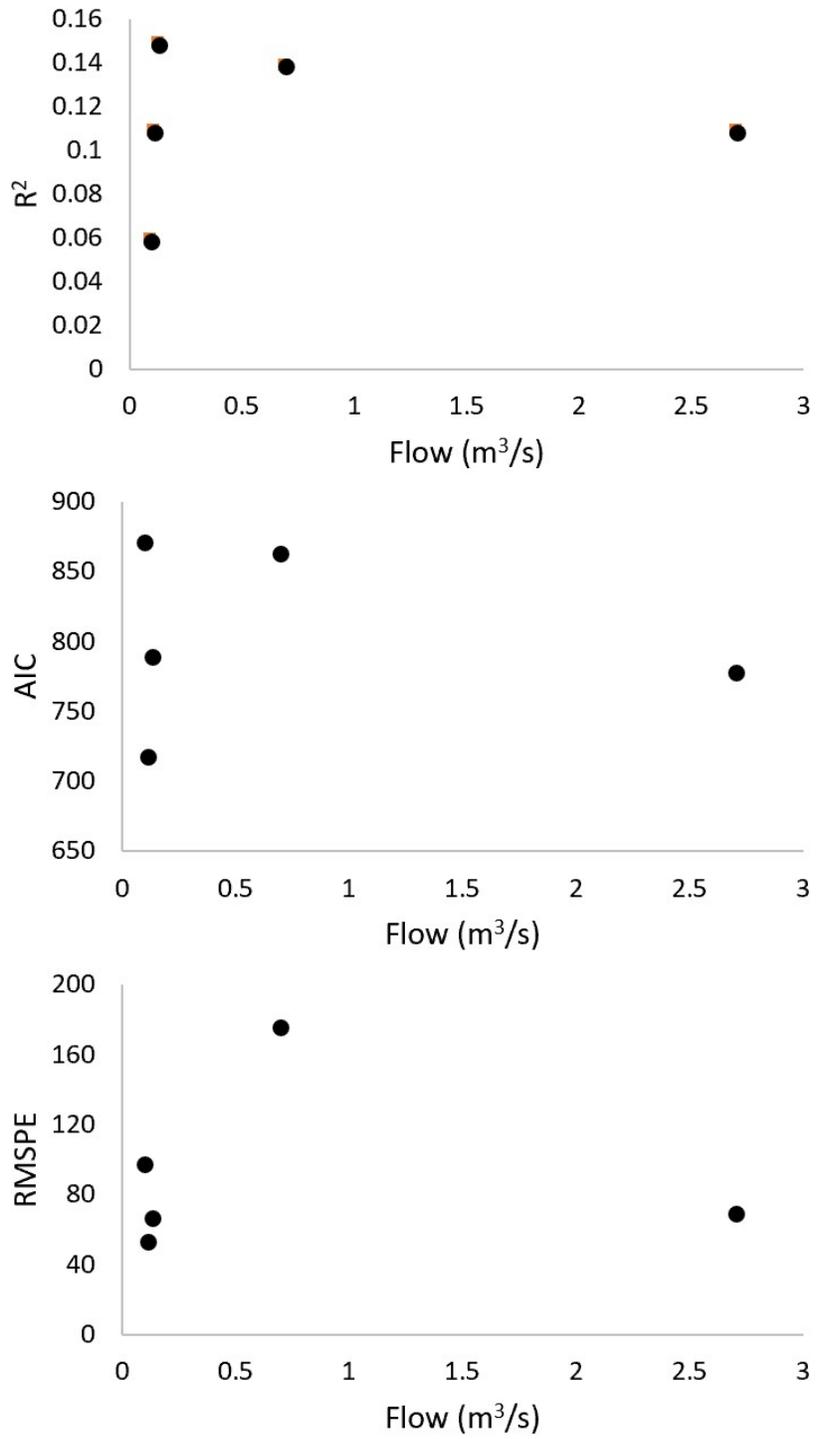


Figure 23: The  $R^2$ , Akaike Information Criterion (AIC), and Root Mean Square Prediction Error (RMSPE), plotted against flow for all five seasonal surveys for the Mimico Creek watershed.

### 5.3 Importance of landscape predictors of stream $\text{Cl}^-$ concentrations across seasons and watersheds

Reasonable expectations about the relative importance of variables, put forward in Section 3.5, were not supported in many of the selected best models. Indeed, Lane Length Density (LLD) was not a significant covariate in Mimico Creek in any season, but it was significant in all Willow Creek surveys. Also, of interest is the fact that Agricultural and Undifferentiated Rural Land (AURL) was significant in the Mimico Creek Fall, Winter, and Spring 2 models and that AURL correlated negatively with in-stream  $\text{Cl}^-$ . As expected, weighted permeability was significant in the low flow late-spring and summer models of Mimico Creek, but this variable was not significant in any other survey.

It is possible that the expected influence of  $\text{Cl}^-$  inputs on roads was not adequately captured by LLD in the Mimico Creek and East Holland River watersheds due to the coarse scale in which this covariate was generated. This geospatial coarseness resulted in an inability for the model to capture the relationship between impervious surface areas (e.g. roads and parking lots) within the catchment area for Mimico Creek and in the western half of the East Holland River subwatershed. The LLD covariate was significant in all seasons of Willow Creek when it was expected it would have relatively low importance in the watershed. This mismatch between the expected and actual importance of LLD in Willow Creek is likely due to an underestimation of the  $\text{Cl}^-$  source potential that roads possess in a rural watershed.

The importance of the AURL covariate was found to be less than expected in the Willow Creek watershed, likely due to an overestimation of the importance of the AURL and due to the watershed scale coarseness of AURL. The importance of the AURL covariate was likely underestimated in Mimico Creek as it was not expected that the low density of the AURL covariate in the Mimico Creek would have a meaningful impact on in-stream  $\text{Cl}^-$  concentrations. It is conceivable that, as the AURL covariate has a significant negative relationship with in-stream  $\text{Cl}^-$  concentration in Mimico Creek, AURL acts as a sink

for road salt inputs and thus prevents  $\text{Cl}^-$  from directly entering the stream network, or the AURL covariate could correlate with areas of reduced de-iceable surface area which would result in less road salt being applied in areas with higher AURL.

While weighted permeability was significant in two of the low flow periods of Mimico Creek, it was not significant in any other low flow periods for other watersheds and it was also unexpectedly significant in the medium flow period of Willow Creek. As the weighted permeability covariate was calculated on a watershed scale it may not be representative of the actual localized flow of  $\text{Cl}^-$  from subsurface reservoirs to the stream.

The unexpected significance of LLD in Willow Creek and AURL in Mimico Creek provides insight into the importance of categorizing the covariates as a means to compare across watersheds, but not within watersheds. Lane length density may be appropriate within Willow Creek because the gradient from urban to rural was more pronounced in this watershed, and the lack of urbanization within Willow Creek made changes to LLD, and resultant impacts on in-stream  $\text{Cl}^-$  easier to detect, whereas all of Mimico Creek is heavily urbanized, so the gradient in LLD was less pronounced and small changes in LLD may lead to undetectable changes in in-stream  $\text{Cl}^-$ . The converse idea applies to AURL where changes in AURL in the heavily urbanized Mimico Creek seem to lead to easily detectable in-stream chloride levels, but changes to AURL in Willow Creek were far less consequential. This result begs further delineation of contextually appropriate predictor variables for use in seasonal watershed studies.

#### 5.4 Do the spatial patterns of residual semivariograms from Spatial Stream Network models of in-stream $\text{Cl}^-$ concentrations vary with urbanization or flow state?

Semivariograms generated using the residuals from the SSN models can allow for the visualization of small and large scale spatial structure within a geostatistical dataset (Section 2.9). For all three study watersheds, a flattening of the residual semivariograms occurred for surveys conducted under high flow

conditions (Figures 20-22). This result suggests that the small scale spatial structure, or patchiness, caused by local landscape characteristics (e.g., road-stream intersections, storm sewer outfalls) were masked by high flows that led to more homogenous  $\text{Cl}^-$  concentrations throughout the stream network. Hence, if the SSN approach were applied again for mapping salt vulnerable areas, it would be important to collect  $\text{Cl}^-$  concentration measurements across a range of flow conditions to determine the relationship between flow regime and geospatial structure, and to determine if this relationship can be generalized to watersheds outside of this thesis. Also, as resources tend to be limited for environmental studies, it would be worthwhile to determine the efficacy of the intensive SSN approach with respect to flow regime. If small-to-large scale signals are masked during periods of higher flow it is likely that a less geospatially intensive study could still provide an accurate representation of a watershed.

A small scale linear rise-peak-fall pattern is seen in the semivariograms for the low flow Spring 1, Summer, and Fall surveys for Mimico Creek, and in the Summer survey for the East Holland River. In Mimico Creek, the low variance at low and high separation distances may be due to the increasing trend in  $\text{Cl}^-$  concentrations as one moves downstream in the headwater tributaries, followed by a dilution pattern as one moves downstream along the main stem, followed by another increasing pattern as one approaches the watershed outlet. The rise-peak-fall pattern seen in the semivariogram for the East Holland River Spring 1 survey has a low sample count with a very low maximum separation distance and it was difficult to make a meaningful interpretation of this very small scale pattern. The Willow Creek semivariograms show a dominant large scale spatial pattern with a linear-to-exponential rise with slope, and this slope decreases with increasing flow regime. A semivariance peak is seen at the end of the Summer, Fall, and Spring semivariograms for Willow Creek and this is likely due to the jump in conductivity between the headwaters upstream of Little Lake, and the main stem downstream of Little Lake. A lack of small scale spatial signals in the Willow Creek watershed may be due to the lack of localized  $\text{Cl}^-$  inputs that may be found in urbanized watersheds with significant stormwater and sewer

infrastructure, and it also may be due to a low sample point density which reduces the resolving power of the SSN model.

The urbanization state of a watershed may result in characteristic semivariogram shapes, but the transferability of these shapes to other watersheds is questionable. Mimico Creek is a topographically simple watershed with its landscape dominated by urban land cover and provided the most consistent semivariogram pattern across seasons relative to East Holland River and Willow Creek. Guided from the ideal semivariogram shapes from McGuire et al. (2014), it can be interpreted from the Mimico Creek semivariograms that a repeating small scale pattern (centered around 6000 m separation distance) gradually changed to a larger scale pattern with a sill forming at 7800 m in the Winter season. The Mimico Creek Spring 2 survey had a flat semivariogram with no interpretable large or small scale structure, which was likely caused by larger scale flows “washing out” smaller reach scale contributions, followed by larger scale landscape contributions.

It was harder to interpret spatial patterns in the urbanizing East Holland River due to erratic semivariance patterns relative to Mimico Creek or Willow Creek. The ability to see patterns in East Holland Summer and Spring surveys was extremely limited by the short separation distances, and the East Holland Winter and Spring surveys were generally flat which leads to the conclusion that the data from these surveys show spatial independence, however this independence is likely from the higher flow regimes of this survey and not due to a lack of relationship between watershed characteristics and in-stream CI.

A small scale repeating signal was seen in Mimico Creek watershed and this may be due to the simple topographical nature of the watershed, the single dominant urban land cover type, and/or localized inputs from urban infrastructure (e.g. storm sewers). The East Holland and Willow Creek surveys did not show a consistent small scale signal in their semivariograms and this could be due to an

increase in stream topographic complexity and/or a lack of adequate sampling density preventing the delineation of localized  $\text{Cl}^-$  sources or sinks. It was also impossible to determine if the small scale pattern seen in Mimico Creek is a trait of urbanized watersheds with simple network topography or if it is a pattern unique to the Mimico Creek watershed.

A link between watershed spatial structure and watershed flow regime necessitates further study to determine if this relationship is generalizable to other watersheds. Researchers could maximize available resources by determining the appropriate study methodology to use within the context of the hydrological state of a watershed. A geospatially intensive study during a period where relatively small scale geospatial signals are hidden may be an inefficient use of resources. Further study is also needed to determine the geospatial structure of differing response variables to flow regime. For instance in-stream chemistry and in-stream biotic abundance may have differing relationships with stream flow.

## 5.5 Study Limitations

The SSN approach used in this study was limited by the sample sizes obtained during the 13 different surveys, with the primary limiting factor for sample sizes being access to the stream networks. In parts of all three watersheds, but to a greater degree in the East Holland River and Willow Creek watersheds, the land adjacent to the stream is privately owned and hence stream access for measurements was only possible at road crossings. In less urbanized or rural areas, roads are widely spaced which reduces the density of stream-road intersections which can limit the extent to which a stream can be selectively measured.

$R^2$  fits for all three watersheds were generally low, in particular Mimico Creek and the Summer and Fall surveys for the East Holland River. The variance composition for all three watersheds showed that the SSN methodology, as opposed to the covariates, explained most of the variance in the selected models. This revealed that the covariates used in this thesis may not have been appropriate in the

context of each respective watershed: LLD in Mimico Creek may have been too crude to capture the complexity of differing road types, private infrastructure (e.g., commercial parking lots), and associated urban road infrastructure (e.g., storm drains). LLD was a more appropriate covariate in Willow Creek due to the paucity of urban infrastructure within the watershed, leading to a more detectable relationship between LLD and in-stream  $\text{Cl}^-$ . The relationship between weighted permeability and the actual connectivity between groundwater and stream waters is unknown. Watershed aquifer maps may provide additional information on groundwater contributions to in-stream  $\text{Cl}^-$ . AURL may have required a more nuanced approach in Willow Creek due to the dominance of agricultural land cover. For example, additional information on crop types, fertilizer constituents (particularly those containing  $\text{Cl}^-$ ), and timing of fertilizer application may be important to consider. Another reason why  $R^2$  may have been low was due to the low variability in observed in-stream  $\text{Cl}^-$  concentrations (e.g., Mimico Creek winter survey), so pseudoreplication may have occurred resulting in reduced model performance.

Another limiting factor in this study was weather conditions. All three watersheds are considered to be meso-scale and hence to capture a snapshot of stream  $\text{Cl}^-$  concentration patterns a single survey usually had to be conducted over a 2 to 5 day rain-free period. Weather constrained when surveys could take place and resulted in premature termination of the East Holland River Summer and Fall surveys due to precipitation. Mimico Creek was more consistently sampled without the issues found in East Holland or Willow Creek, however the Mimico Creek Winter survey in-stream  $\text{Cl}^-$  data had a correction factor applied due to a rain event, followed by a melt event occurring during the survey. This Mimico Creek correction factor (See Appendices) converted all pre-rain and post-melt values to a post-rain state and therefore the winter survey may not be totally representative of the watershed in a normal flow state in the winter months.

The East Holland Watershed is an example of an urbanizing watershed that still has significant rural and agricultural lands. However, incomplete sampling in East Holland Summer and Fall surveys

resulted in an inability to detect relationships between varying landscape characteristics and in-stream  $\text{Cl}^-$ . This lack of data also resulted in short separation distances between sampling points which made it difficult to resolve small scale patterns in their respective semivariograms. Willow Creek is a very rural watershed with significant agricultural lands, but the ephemerality of many tributaries along with difficulties with stream access made it difficult to consistently gather a high density of seasonal sample points which subsequently reduces the ability to resolve small scale geospatial structures within the watershed.

The Agricultural & Undifferentiated Rural Land cover layer, lane length density, and the weighted permeability variable are watershed scale metrics, so reach scale relationships between the covariates and in-stream  $\text{Cl}^-$  may be increasingly hidden as the overall point catchment size increases when moving downstream. The weighted permeability covariate was also created out of a need to simplify surficial geology land cover and it may not represent actual groundwater-surface water interactions along the stream network.

The limited sample size for each watershed, the fact that only one watershed was used for each urbanization classification (i.e. urbanized, semi-urban, and rural), and the coarseness of the covariates prevent the results of this thesis from being generalized to other watersheds. More data needs to be collected across more watersheds to determine if the multi-scale geospatial signals seen in this thesis, and the change of these signals with changing flow regime, are unique to each study watershed, the region, or are actually characteristic of watersheds with similar levels of urbanization.

## 5.6 Future Work

This study shows how the SSN modelling approach can be applied to watersheds with varying land use to map salt vulnerable areas at the stream reach scale. Here it is shown that flow state must be taken

into consideration when collecting input data for an SSN model and that predictor variables measured at the upstream catchment scale were able to capture local source areas that connect to the stream.

Future studies should focus on refining and validating the SSN models developed in this study. Additional field data should be collected to check the model and generate predictive maps of in-stream  $\text{Cl}^-$  concentrations in areas beyond where input data were collected. Larger sample sizes, combined with better predictor variables, such as public and private impervious surface area (e.g. roads, parking lots, buildings), real or estimated road salt application metrics, and/or aquifer maps, predictive in-stream  $\text{Cl}^-$  maps at (or near) reach scales could be generated to better constrain the location of in-stream  $\text{Cl}^-$  hotspots. Urban watershed catchment delineation could be refined by considering stormsewer infrastructure that either extends the catchment beyond the boundaries as commonly defined by elevation, or diverts water away from an area that would normally be included in a standard elevation based watershed boundary (Kayembe & Mitchell, 2018) . Better understanding of the inter-seasonal relationships between in-stream  $\text{Cl}^-$  and contextually appropriate predictor variables would help to inform future models when predicting the impacts of road salt on in-stream concentrations. Better delineation of  $\text{Cl}^-$  hotspots could be used in conjunction with critical habitat and/or species-at-risk maps to improve identification of salt vulnerable areas, which would make it easier for road authorities to efficiently apply road salt mitigation measures. Lastly, an SSN modelling approach could also be used to predict how urban growth and road salt mitigation measures influence the magnitude of stream  $\text{Cl}^-$  concentrations and the location of salt vulnerable areas.

## Appendices

### Mimico Creek Winter Survey Corrections

The following corrections were applied to the Mimico Creek Watershed Winter longitudinal survey. Site resampling occurred after a rain event and after a melt event to allow for a correction factor for measured conductivity values. The Mimico Creek watershed was organized into three hydrologic states: Pre-Rain, Post-Rain, and Post-Melt. Conversion factors between the pre-rain and post-rain states were calculated for the E-sites, C-sites, B-sites, and A71 to A15 sites. Conversion factors between pre-melt and post-melt states were calculated for A14 to A04 sites. As there is no data to allow for a conversion between the pre-rain watershed state and the post-melt watershed state, all sample points were converted to a post-rain state.

Site Name	Pre-Rain Measurement (µs/cm)	Post-Rain Measurement (µs/cm)	Post-Melt Measurement (µs/cm)	Value Difference	% Diff	FINAL VALUES	Change Type	Change?
E02	3410.8	8014.6				8014.6		
E01	10865.7	7385.8				7385.8		
C12	5183.8					5183.8	Rain	Expected change unclear
C08	2927.8					2927.8	Rain	Expected change unclear
C07	4230.1					4230.1	Rain	Expected change unclear
C06	5510.2					5510.2		
C03	6245.9	7110.8		864.9	12.16	7110.8		
C02	6169.7	6895.9		726.2	10.53	6895.9		
C01	5804.6				8.90	6321.2	Rain	Used trend from C3 to C2
BB02	6981.8					6981.8	Rain	No information that allows for justified change
BB01	7239					7239.0	Rain	No information that allows for justified change
BA01	7183.6					7183.6	Rain	No information that allows for justified change
B18	4250	5050.8		800.8	18.84	5050.8	Rain	
B17	4371.3				16.98	5113.5	Rain	Predicted % based on linear trend from B18 to B11
B16	4325.5				15.30	4987.3	Rain	Predicted % based on linear trend from B18 to B11
B15	4849.9	5477.2		627.3	12.93	5477.2	Rain	
B14	4997.3				11.94	5594.0	Rain	Predicted % based on linear trend from B18 to B11
B13	4942.8	5466.1		523.3	10.59	5466.1	Rain	
B12	5136.5				8.58	5577.2	Rain	Predicted % based on linear trend from B18 to B11
B11	6157.2				6.90	6582.0	Rain	Predicted % based on linear trend from B18 to B11
B08	4908.7	6136.5		1227.8	25.01	6136.5	Rain	
B06	5083				28.94	6554.0	Rain	Predicted % based on linear trend from B08 to B03
B05	5639.4				30.92	7383.1	Rain	Predicted % based on linear trend from B08 to B03
B04	5632.3				32.90	7485.3	Rain	Predicted % based on linear trend from B08 to B03
B03	5598.5	7552.1		1953.6	34.90	7552.1	Rain	
B02	5610.3				36.86	7678.3	Rain	Predicted % based on linear trend from B08 to B03
B01	5589.5				38.84	7760.5	Rain	Predicted % based on linear trend from B08 to B03
A71	5525.3	5041.5		-483.8		5041.5	Rain	
A70	5133.5				-5.69	4841.4	Rain	Predicted % based on linear trend from A71 to A66

A69	4724.3				-2.87	4588.7	Rain	Predicted % based on linear trend from A71 to A66
A68	4351				-0.05	4348.8	Rain	Predicted % based on linear trend from A71 to A66
A66	6082.8	6408.9		326.1	5.36	6408.9	Rain	
A64	7661.7				11.23	8522.1	Rain	Predicted % based on linear trend from A71 to A66
A63	7606				14.05	8674.6	Rain	Predicted % based on linear trend from A71 to A66
A58	6941.4	9937.2		2995.8	43.16	9937.2	Rain	
A55	5345.6				49.40	7986.3	Rain	Predicted % based on linear trend from A58 to A53
A54	6414.6				55.64	9983.7	Rain	Predicted % based on linear trend from A58 to A53
A53	6638.1	10745.8		4107.7	61.88	10745.8	Rain	
A52	6622.9				54.92	10260.0	Rain	Predicted % based on linear trend from A53 to A51
A51	6512.9	9636.1		3123.2	47.95	9636.1	Rain	
A50	6347.1				28.24	8139.5	Rain	Same LC throughout...applied change from A48
A49	6322.3				28.24	8107.7	Rain	Same LC throughout...applied change from A48
A48	6291.8	8068.6		1776.8	28.24	8068.6	Rain	
A47	6264.1				28.24	8033.1	Rain	Same LC throughout...applied change from A48
A46	5253.7					9125.0	Rain	Predicted values based on linear trend from A41 to A15
A45	5275.5					9015.3	Rain	Predicted values based on linear trend from A41 to A15
A44	5308.6					8905.5	Rain	Predicted values based on linear trend from A41 to A15
A41	10726.7	8166.1				8166.1		
A40		8419.8				8419.8		
A39		8656.2				8656.2		
A35		8469.9				8469.9		
A31		7089.4				7089.4		
A28		7115.8				7115.8		
A27		7007				7007		
A26		6968.3				6968.3		
A25		6839.3				6839.3		
A24		6667.2				6667.2		
A23		6525.8				6525.8		
A22		6526.2				6526.2		
A21		6423.9				6423.9		

A20		6349.8				6349.8		
A15		6130.8				6130.8		
A14			7455.6		15.23	6319.9	Melt	Used average change from A27, A23, A20 & A09. Did not find good linear pattern.
A13			7242.6		15.23	6139.4	Melt	Used average change from A27, A23, A20 & A09. Did not find good linear pattern.
A12			7199.5		15.23	6102.8	Melt	Used average change from A27, A23, A20 & A09. Did not find good linear pattern.
A11			7195.2		15.23	6099.2	Melt	Used average change from A27, A23, A20 & A09. Did not find good linear pattern.
A10			7200.7		15.23	6103.9	Melt	Used average change from A27, A23, A20 & A09. Did not find good linear pattern.
A09		6097.5	7199.8	1102.3	15.31	6097.5	Melt	
A08			7181.2		15.23	6087.3	Melt	Used average change from A27, A23, A20 & A09. Did not find good linear pattern.
A07			7016.1		15.23	5947.4	Melt	Used average change from A27, A23, A20 & A09. Did not find good linear pattern.
A06			7132.9		15.23	6046.4	Melt	Used average change from A27, A23, A20 & A09. Did not find good linear pattern.
A05			7168.7		15.23	6076.7	Melt	Used average change from A27, A23, A20 & A09. Did not find good linear pattern.
A04			7105.6		15.23	6023.2	Melt	Used average change from A27, A23, A20 & A09. Did not find good linear pattern.

## References

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory* (pp. 267–281). Budapest, Hungary.
- Allison, L. E., Bernstein, L., Bower, C. A., Brown, J. W., Fireman, M., Hatcher, J. T., ... Wilcox, L. V. (1954). *Agriculture handbook No. 60: Diagnosis and improvement of saline and alkali soils*. United States Department of Agriculture.
- Babayak, M. A. (2004). What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine*, 66(3).
- Bäckström, M., Karlsson, S., Bäckman, L., Folkesson, L., & Lind, B. (2004). Mobilisation of heavy metals by deicing salts in a roadside environment. *Water Research*, 38(3), 720–732. <https://doi.org/10.1016/j.watres.2003.11.006>
- Bastviken, D., Sandén, P., Svensson, T., Stahlberg, C., Magounakis, M., & Öberg, G. (2006). Chloride-retention and release in a boreal forest soil: Effects of soil water residence time and nitrogen and chloride loads. *Environmental Science & Technology*, 40(9), 2977–2982. <https://doi.org/10.1021/es0523237>
- Bastviken, D., Thomsen, F., Svensson, T., Karlsson, S., Sandén, P., Shaw, G., ... Öberg, G. (2007). Chloride retention in forest soil by microbial uptake and by natural chlorination of organic matter. *Geochimica et Cosmochimica Acta*, 71(13), 3182–3192. <https://doi.org/10.1016/j.gca.2007.04.028>
- Bazuhair, A., & Wood, W. (1996). Chloride mass-balance method for estimating ground water recharge in arid areas: examples from western Saudi Arabia. *Journal of Hydrology*, 186, 153–159.
- Bester, M. L., Frind, E. O., Molson, J. W., & Rudolph, D. L. (2006). Numerical investigation of road salt impact on an urban wellfield. *Ground Water*, 44(2), 165–175. <https://doi.org/10.1111/j.1745-6584.2005.00126.x>
- Betts, A. R. R., Gharabaghi, B., & McBean, E. A. A. (2014). Salt vulnerability assessment methodology for urban streams. *Journal of Hydrology*, 517, 877–888. <https://doi.org/10.1016/j.jhydrol.2014.06.005>
- Boehrer, B., & Schultze, M. (2008). Stratification of lakes. *Reviews of Geophysics*, 46(2), RG2005. <https://doi.org/10.1029/2006RG000210>
- Bolen, W. P., Barnes, L. M., & Wallace, G. J. (2016). 2014 Minerals yearbook - salt [advance release]. *United States Geological Survey*, (October). Retrieved from <https://minerals.usgs.gov/minerals/pubs/commodity/bismuth/myb1-2014-bismu.pdf>
- Brand, A. B., Snodgrass, J. W., Gallagher, M. T., Casey, R. E., & Van Meter, R. (2010). Lethal and sublethal effects of embryonic and larval exposure of *Hyla versicolor* to stormwater pond sediments. *Archives of Environmental Contamination and Toxicology*, 58(2), 325–331. <https://doi.org/10.1007/s00244-009-9373-0>
- Cain, N. P., Hale, B., Berkelaar, E., & Morin, D. (2000). *Review of effects of NaCl and other road salts on terrestrial vegetation in Canada*. Environment Canada. Ottawa, ON.
- Canadian Council of Ministers of the Environment. (2011). *Canadian water quality guidelines for the*

*protection of aquatic life - chloride. Canadian environmental quality guidelines. Winnipeg.*

- Chapra, S. C., Dove, A., & Rockwell, D. C. (2009). Great Lakes chloride trends: Long-term mass balance and loading analysis. *Journal of Great Lakes Research*, 35(2), 272–284. <https://doi.org/10.1016/j.jglr.2008.11.013>
- Collins, S. J., & Russell, R. W. (2009). Toxicity of road salt to Nova Scotia amphibians. *Environmental Pollution*, 157(1), 320–324. <https://doi.org/10.1016/j.envpol.2008.06.032>
- Cooper, C. A., Mayer, P. M., & Faulkner, B. R. (2014). Effects of road salts on groundwater and surface water dynamics of sodium and chloride in an urban restored stream. *Biogeochemistry*, 121(1), 149–166. <https://doi.org/10.1007/s10533-014-9968-z>
- Crowther, R. A., & Hynes, H. B. N. (1977). The effect of road deicing salt on the drift of stream benthos. *Environmental Pollution (1970)*, 14(2), 113–126. [https://doi.org/10.1016/0013-9327\(77\)90103-3](https://doi.org/10.1016/0013-9327(77)90103-3)
- Dale, M., & Fortin, M. (2014). *Spatial analysis: A guide for ecologists*. Cambridge University Press.
- Daley, M. L., Potter, J. D., & McDowell, W. H. (2009). Salinization of urbanizing New Hampshire streams and groundwater: effects of road salt and hydrologic variability. *Journal of the North American Benthological Society*, 28(4), 929–940. <https://doi.org/10.1899/09-052.1>
- Davison, A. W. (1971). The effects of de-icing salt on roadside verges. I. Soil and plant analysis. *Journal of Applied Ecology*, 8(2), 555–561.
- Derby, N. E., & Knighton, R. E. (2001). Field-scale preferential transport of water and chloride tracer by depression-focused recharge. *Journal of Environmental Quality*, 30(1), 194–199. <https://doi.org/10.1130/B30467.1>
- Edmunds, W. M., & Gaye, C. B. (1994). Estimating the spatial variability of groundwater recharge in the Sahel using chloride. *Journal of Hydrology*, 156(1–4), 47–59. [https://doi.org/10.1016/0022-1694\(94\)90070-1](https://doi.org/10.1016/0022-1694(94)90070-1)
- Environment and Climate Change Canada. (2004). Code of practice: The environmental management of road salts, <https://www.ec.gc.ca/sels-salts/default.asp?lang=E>. Retrieved from <http://www.ec.gc.ca/sels-salts/default.asp?lang=En&n=F37B47CE-1>
- Environment and Climate Change Canada. (2017). Historical Hydrometric Data. Retrieved from [https://wateroffice.ec.gc.ca/mainmenu/historical\\_data\\_index\\_e.html](https://wateroffice.ec.gc.ca/mainmenu/historical_data_index_e.html)
- Eriksson, E., & Khunakasem, V. (1969). Chloride concentration in groundwater, recharge rate and rate of deposition of chloride in the Israel Coastal Plain. *Journal of Hydrology*, 7(2), 178–197. [https://doi.org/10.1016/0022-1694\(69\)90055-9](https://doi.org/10.1016/0022-1694(69)90055-9)
- Fouty, S. C. (1989). *Chloride mass balance as a method for determining long-term groundwater recharge rates and geomorphic-surface stability in arid and semi-arid regions whisky flat and beatty, nevada*. University of Arizona.
- Gardner, K. K., & McGlynn, B. L. (2009). Seasonality in spatial variability and influence of land use/land cover and watershed characteristics on stream water nitrate concentrations in a developing watershed in the Rocky Mountain West. *Water Resources Research*, 45(8), 1–14. <https://doi.org/10.1029/2008WR007029>
- Gardner, K. M., & Royer, T. V. (2010). Effect of road salt application on seasonal chloride concentrations

- and toxicity in South-Central Indiana streams. *Journal of Environment Quality*, 39(3), 1036–1042. <https://doi.org/10.2134/jeq2009.0402>
- Gillis, P. L. (2011). Assessing the toxicity of sodium chloride to the glochidia of freshwater mussels: Implications for salinization of surface waters. *Environmental Pollution*, 159(6), 1702–1708. <https://doi.org/10.1016/j.envpol.2011.02.032>
- Godwin, K. S., Hafner, S. D., & Buff, M. F. (2003). Long-term trends in sodium and chloride in the Mohawk River, New York: the effect of fifty years of road-salt application. *Environmental Pollution*, 124(2), 273–281. [https://doi.org/10.1016/S0269-7491\(02\)00481-5](https://doi.org/10.1016/S0269-7491(02)00481-5)
- Green, S. M., Machin, R., & Cresser, M. S. (2008). Effect of long-term changes in soil chemistry induced by road salt applications on N-transformations in roadside soils. *Environmental Pollution*, 152(1), 20–31. <https://doi.org/10.1016/j.envpol.2007.06.005>
- Guan, H., Love, A. J., Simmons, C. T., Hutson, J., & Ding, Z. (2010). Catchment conceptualisation for examining applicability of chloride mass balance method in an area with historical forest clearance. *Hydrology and Earth System Sciences*, 14(7), 1233–1245. <https://doi.org/10.5194/hess-14-1233-2010>
- Gutchess, K., Jin, L., Lautz, L., Shaw, S. B., Zhou, X., & Lu, Z. (2016). Chloride sources in urban and rural headwater catchments, central New York. *Science of The Total Environment*, 565, 462–472. <https://doi.org/10.1016/j.scitotenv.2016.04.181>
- Hemond, H., & Fechner, E. (2015). *Chemical fate and transport in the environment* (Third Ed). Elsevier Inc.
- Hoef, J. M. Ver, Peterson, E. E., Clifford, D., & Shah, R. (2014). SSN: An R package for spatial statistical modeling on stream networks. *Journal of Statistical Software*, 56(3). <https://doi.org/10.18637/jss.v056.i03>
- Hoffman, R. W., Goldman, C. R., Paulson, S., & Winters, G. R. (1981). Aquatic impacts of deicing salts in the Central Sierra Nevada Mountains, California. *Journal of the American Water Resources Association*, 17(2), 280–285. <https://doi.org/10.1111/j.1752-1688.1981.tb03935.x>
- Holland, A. J., Gordon, A. K., & Muller, W. J. (2011). Osmoregulation in freshwater invertebrates in response to exposure to salt pollution. *Report to the Water Research Commission. Unilever Centre for Environmental Water Quality, Institute for Water Research, Rhodes University, Grahamstown, South Africa.*
- Howard, K. W. F., & Beck, P. J. (1993). Hydrogeochemical implications of groundwater contamination by road de-icing chemicals. *Journal of Contaminant Hydrology*, 12(1993), 245–268.
- Howard, K. W. F., & Haynes, J. (1993). Groundwater contamination due to road de-icing chemicals – salt balance implications. *Geoscience Canada*, 20(1), 1–8.
- Howard, K. W. F., & Maier, H. (2007). Road de-icing salt as a potential constraint on urban growth in the Greater Toronto Area, Canada. *Journal of Contaminant Hydrology*, 91(1–2), 146–170. <https://doi.org/10.1016/j.jconhyd.2006.10.005>
- Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., ... Cudennec, C. (2013). A decade of predictions in ungauged basins (PUB) — a review. *Hydrological Sciences Journal*, 58(6), 1198–1255. <https://doi.org/10.1080/02626667.2013.803183>

- Isaak, D. J., Peterson, E. E., Ver Hoef, J. M., Wenger, S. J., Falke, J. A., Torgersen, C. E., ... Monestiez, P. (2014). Applications of spatial statistical network models to stream data. *Wiley Interdisciplinary Reviews: Water*, 1(3), 277–294. <https://doi.org/10.1002/wat2.1023>
- Isaak, D., Luce, C., Rieman, B., Nagel, D., Peterson, E., Horan, D., ... Chandler, G. (2009). Effects of climate change and recent wildfires on stream temperature and thermal habitat for two salmonoids in a mountain river network. *Ecological Applications*, 20(5), 100319061507001. <https://doi.org/10.1890/09-0822>
- Jodoin, Y., Lavoie, C., Villeneuve, P., Theriault, M., Beaulieu, J., & Belzile, F. (2008). Highways as corridors and habitats for the invasive common reed *Phragmites australis* in Quebec, Canada. *Journal of Applied Ecology*, 45(2), 459–466. <https://doi.org/10.1111/j.1365-2664.2007.01362.x>
- Judd, J. (1970). Lake stratification caused by runoff from street deicing. *Water Research*, 4(8), 521–532. [https://doi.org/10.1016/0043-1354\(70\)90002-3](https://doi.org/10.1016/0043-1354(70)90002-3)
- Karraker, N. E., & Gibbs, J. P. (2011). Road deicing salt irreversibly disrupts osmoregulation of salamander egg clutches. *Environmental Pollution*, 159(3), 833–835. <https://doi.org/10.1016/j.envpol.2010.11.019>
- Katz, B. G., Eberts, S. M., & Kauffman, L. J. (2011). Using Cl/Br ratios and other indicators to assess potential impacts on groundwater quality from septic systems: A review and examples from principal aquifers in the United States. *Journal of Hydrology*, 397(3–4), 151–166. <https://doi.org/10.1016/j.jhydrol.2010.11.017>
- Kaushal, S. S., Groffman, P. M., Likens, G. E., Belt, K. T., Stack, W. P., Kelly, V. R., ... Fisher, G. T. (2005). Increased salinization of fresh water in the northeastern United States. *Proceedings of the National Academy of Sciences of the United States of America*, 102(38), 13517–13520. <https://doi.org/10.1073/pnas.0506414102>
- Kayembe, A., & Mitchell, C. P. J. (2018). Determination of subcatchment and watershed boundaries in a complex and highly urbanized landscape. *Hydrological Processes*. <https://doi.org/10.1002/hyp.13229>
- Kelly, W. R., Panno, S. V., Hackley, K. C., Hwang, H. H., Martinsek, A. T., & Markus, M. (2010). Using chloride and other ions to trace sewage and road salt in the Illinois Waterway. *Applied Geochemistry*, 25(5), 661–673. <https://doi.org/10.1016/j.apgeochem.2010.01.020>
- Kelting, D. L., Laxson, C. L., & Yerger, E. C. (2012). Regional analysis of the effect of paved roads on sodium and chloride in lakes. *Water Research*, 46(8), 2749–2758. <https://doi.org/10.1016/j.watres.2012.02.032>
- Kjensmo, J. (1997). The influence of road salts on the salinity and the meromictic stability of Lake Svinsjøen, southeastern Norway. *Hydrobiologia*, 347, 151–159. <https://doi.org/10.1023/A:1003035705729>
- Koretsky, C. M., MacLeod, A., Sibert, R. J., & Snyder, C. (2012). Redox stratification and salinization of Three Kettle Lakes in Southwest Michigan, USA. *Water, Air, & Soil Pollution*, 223(3), 1415–1427. <https://doi.org/10.1007/s11270-011-0954-y>
- Lake Simcoe Region Conservation Authority. (2010). *East Holland River subwatershed plan*. Retrieved from <https://www.lsrca.on.ca/Shared Documents/reports/east-holland-subwatershed-plan.pdf>

- Levin, S. A. (1992). The problem of pattern and scale in ecology. *Ecology*, 73(6), 1943–1967.  
<https://doi.org/10.2307/1941447>
- Lois, S. (2016). Modelling and predicting biogeographical patterns in river networks. *Frontiers of Biogeography*, 9(1), 1–5.
- Ma, J., Wang, Y., Zhao, Y., Jin, X., Ning, N., Edmunds, W. M., & Zhou, X. (2012). Spatial distribution of chloride and nitrate within an unsaturated dune sand of a cold-arid desert: Implications for paleoenvironmental records. *CATENA*, 96, 68–75. <https://doi.org/10.1016/j.catena.2012.04.012>
- MacLeod, A., Sibert, R., Snyder, C., & Koretsky, C. M. (2011). Eutrophication and salinization of urban and rural kettle lakes in Kalamazoo and Barry Counties, Michigan, USA. *Applied Geochemistry*, 26(SUPPL.), S214–S217. <https://doi.org/10.1016/j.apgeochem.2011.03.107>
- Mallants, D. (2014). Field-scale solute transport parameters derived from tracer tests in large undisturbed soil columns. *Soil Research*, 52(1), 13–26. <https://doi.org/10.1071/SR13143>
- McGuire, K. J., Torgersen, C. E., Likens, G. E., Buso, D. C., Lowe, W. H., & Bailey, S. W. (2014). Network analysis reveals multiscale controls on streamwater chemistry. *Proceedings of the National Academy of Sciences*, 111(19), 7030–7035. <https://doi.org/10.1073/pnas.1404820111>
- Meriano, M., Eyles, N., & Howard, K. W. F. (2009). Hydrogeological impacts of road salt from Canada's busiest highway on a Lake Ontario watershed (Frenchman's Bay) and lagoon, City of Pickering. *Journal of Contaminant Hydrology*, 107(1–2), 66–81.  
<https://doi.org/10.1016/j.jconhyd.2009.04.002>
- Meriano, M., Howard, K. W. F., & Eyles, N. (2011). The role of midsummer urban aquifer recharge in stormflow generation using isotopic and chemical hydrograph separation techniques. *Journal of Hydrology*, 396(1–2), 82–93. <https://doi.org/10.1016/j.jhydrol.2010.10.041>
- Morgan, R. P., Kline, K. M., Kline, M. J., Cushman, S. F., Sell, M. T., Weitzell, R. E., & Churchill, J. B. (2012). Stream conductivity: Relationships to land use, chloride, and fishes in Maryland streams. *North American Journal of Fisheries Management*, 32(5), 941–952.  
<https://doi.org/10.1080/02755947.2012.703159>
- Muethel, R. W. (1976). Effects of deicing salts on the chloride levels in water and soil adjacent to roadways. *Michigan Department of Transportation*, 16.
- Mullaney, J. R., Lorenz, D. L., & Arntson, A. D. (2009). *Chloride in groundwater and surface water in areas underlain by the glacial aquifer system, northern United States*. United States Geological Survey.
- Murray, D. M., & Ernst, U. F. W. (1976). An economic analysis of the environmental impact of highway deicing. *United States Environmental Protection Agency*, XXX, 352.
- Nottawasga Valley Conservation Authority. (2013). Willow Creek: 2013 subwatershed health check.
- Novotny, E. V., Murphy, D., & Stefan, H. G. (2008). Increase of urban lake salinity by road deicing salt. *Science of The Total Environment*, 406(1–2), 131–144.  
<https://doi.org/10.1016/j.scitotenv.2008.07.037>
- Novotny, E. V., Sander, A. R., Mohseni, O., & Stefan, H. G. (2009). Chloride ion transport and mass balance in a metropolitan area using road salt. *Water Resources Research*, 45(12), n/a-n/a.  
<https://doi.org/10.1029/2009WR008141>

- Ontario Geological Survey. (2010). Surficial geology of southern Ontario; Ontario Geological Survey, Miscellaneous Data 128 - Revised.
- Ontario Ministry of Natural Resources and Forestry. (2013). Ontario Flow Assessment Tool. Retrieved from <https://www.ontario.ca/page/watershed-flow-assessment-tool>
- Ontario Ministry of Natural Resources and Forestry. (2015). Ontario Integrated Hydrology Data.
- Ontario Ministry of Natural Resources and Forestry. (2016a). Ontario land cover compilation v.2.0.
- Ontario Ministry of Natural Resources and Forestry. (2016b). Ontario Road Network: Road Net Element.
- Ontario Ministry of the Environment and Climate Change. (2016). *Water Quality in Ontario 2014 Report*.
- Panno, S. V., Hackley, K. C., Hwang, H. H., Greenberg, S. E., Krapac, I. G., Landsberger, S., & O'Kelly, D. J. (2006). *Characterization and identification of Na-Cl sources in ground water*. *Ground Water* (Vol. 44). <https://doi.org/10.1111/j.1745-6584.2005.00127.x>
- Panno, S. V., Hackley, K. C. K., Hwang, H. H., Greenberg, S., Krapac, I. G., Landsberger, S., & O'Kelly, D. J. (2002). Source identification of sodium and chloride contamination in natural waters: Preliminary results. *Illinois State Geological Survey Open File Series 2005-1*. Retrieved from <http://www.water-research.net/Waterlibrary/privatewell/nacl.pdf>
- Perera, N., Gharabaghi, B., & Howard, K. (2013). Groundwater chloride response in the Highland Creek watershed due to road salt application: A re-assessment after 20 years. *Journal of Hydrology*, 479, 159–168. <https://doi.org/10.1016/j.jhydrol.2012.11.057>
- Perera, N., Gharabaghi, B., Noehammer, P., & Kilgour, B. (2010). Road salt application in Highland Creek watershed, Toronto, Ontario – Chloride mass balance. *Water Quality Research Journal of Canada*, 45(4), 451–461. <https://doi.org/https://doi.org/10.2166/wqrj.2010.044>
- Peterson, E. E., Merton, A. A., Theobald, D. M., & Urquhart, N. S. (2006). Patterns of spatial autocorrelation in stream water chemistry. *Environmental Monitoring and Assessment*, 121(1–3), 569–594. <https://doi.org/10.1007/s10661-005-9156-7>
- Peterson, E. E., & Urquhart, S. N. (2006). Predicting water quality impaired stream segments using landscape-scale data and a regional geostatistical model: A case study in Maryland. *Environmental Monitoring and Assessment*, 121(1–3), 613–636. <https://doi.org/10.1007/s10661-005-9163-8>
- Peterson, E. E., & Ver Hoef, J. M. (2014). STARS: An ArcGIS toolset used to calculate the spatial information needed to fit spatial statistical models to stream network data. *Journal of Statistical Software*, 56(2), 1–17. <https://doi.org/http://dx.doi.org/10.18637/jss.v056.i02>
- Peterson, E. E., Ver Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M. J., Jordan, C. E., ... Wenger, S. J. (2013). Modelling dendritic ecological networks in space: An integrated network perspective. *Ecology Letters*, 16(5), 707–719. <https://doi.org/10.1111/ele.12084>
- Pickett, S. . T. A., & Cadenasso, M. L. (1995). Landscape Ecology: Spatial Heterogeneity in Ecological Systems. *Science*, 269(5222), 331–334. <https://doi.org/10.1126/science.269.5222.331>
- Prasser, N., & Zedler, J. B. (2010). Salt tolerance of invasive *Phalaris arundinacea* exceeds that of native *Carex stricta* (Wisconsin). *Ecological Restoration*, 28(3), 238–240.
- Prior, G. A., & Berthouex, P. M. (1967). *A study of salt pollution of soil by highway salting*.

- R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.r-project.org/>
- Richburg, J. A., Patterson, W. A., & Lowenstein, F. (2001). Effects of road salt and *Phragmites australis* invasion on the vegetation of a Western Massachusetts calcareous lake-basin fen. *Wetlands*, *21*(2), 247–255. [https://doi.org/10.1672/0277-5212\(2001\)021\[0247:EORSAP\]2.0.CO;2](https://doi.org/10.1672/0277-5212(2001)021[0247:EORSAP]2.0.CO;2)
- Risley, J., Stonewall, A., & Haluska, T. (2009). *Estimating flow-duration and low-flow frequency statistics for unregulated streams in Oregon*. Retrieved from <https://pubs.usgs.gov/sir/2008/5126/index.html>
- Sanzo, D., & Hecnar, S. J. (2006). Effects of road de-icing salt (NaCl) on larval wood frogs (*Rana sylvatica*). *Environmental Pollution*, *140*(2), 247–256. <https://doi.org/10.1016/j.envpol.2005.07.013>
- Sharma, M. L., & Hughes, M. W. (1985). Groundwater recharge estimation using chloride, deuterium and oxygen-18 profiles in the deep coastal sands of Western Australia. *Journal of Hydrology*, *81*(1–2), 93–109. [https://doi.org/10.1016/0022-1694\(85\)90169-6](https://doi.org/10.1016/0022-1694(85)90169-6)
- Steel, A. E., Sowder, C., & Peterson, E. E. (2016). Spatial and temporal variation of water temperature regimes on the Snoqualmie River network. *Journal of the American Water Resources Association*, *52*(3), 769–787. <https://doi.org/10.1111/1752-1688.12423>
- Struzeski, E., & United States Environmental Protection Agency. (1971). Environmental impact of highway deicing. Retrieved from <http://nepis.epa.gov/Exe/ZyPDF.cgi/9100Y80B.PDF?Dockey=9100Y80B.PDF>
- Svensson, T., Lovett, G. M., & Likens, G. E. (2012). Is chloride a conservative ion in forest ecosystems? *Biogeochemistry*, *107*(1–3), 125–134. <https://doi.org/10.1007/s10533-010-9538-y>
- Toronto and Region Conservation Authority. (2013). Mimico Creek Watershed Report Card.
- Townsend, A. M. (1980). Identifying trees with tolerance to soil salts. *Proceedings of the Metropolitan Tree Improvement Alliance (METRIA)*, *3*, 24–32. <https://doi.org/10.1017/CBO9781107415324.004>
- Turschwell, M. P., Peterson, E. E., Balcombe, S. R., & Sheldon, F. (2016). To aggregate or not? Capturing the spatio-temporal complexity of the thermal regime. *Ecological Indicators*, *67*(April), 39–48. <https://doi.org/10.1016/j.ecolind.2016.02.014>
- United States Department of the Interior. (1969). Water pollution aspects of urban runoff.
- Ver Hoef, J. M., & Peterson, E. E. (2010). A moving average approach for spatial statistical models of stream networks. *Journal of the American Statistical Association*, *105*(489), 6–18. <https://doi.org/10.1198/jasa.2009.ap08248>
- Ver Hoef, J. M., Peterson, E., & Theobald, D. (2006). Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, *13*(4), 449–464. <https://doi.org/10.1007/s10651-006-0022-8>
- Walsh, C. J., Roy, A. H., Feminella, J. W., Cottingham, P. D., Groffman, P. M., Morgan II, R. P., & Morgan, R. P. (2005). The urban stream syndrome: current knowledge and the search for a cure. *Journal of the North American Benthological Society*, *24*(3), 706. [https://doi.org/10.1899/0887-3593\(2005\)024\[0706:TUSACK\]2.0.CO;2](https://doi.org/10.1899/0887-3593(2005)024[0706:TUSACK]2.0.CO;2)
- Warren, L. A., & Zimmerman, A. P. (1994). The influence of temperature and NaCl on cadmium, copper

and zinc partitioning among suspended particulate and dissolved phases in an urban river. *Water Research*, 28(9), 1921–1931. [https://doi.org/10.1016/0043-1354\(94\)90167-8](https://doi.org/10.1016/0043-1354(94)90167-8)

Williams, D. D., Williams, N. E., & Cao, Y. (2000). Road salt contamination of groundwater in a major metropolitan area and development of a biological index to monitor its impact. *Water Research*, 34(1), 127–138. [https://doi.org/10.1016/S0043-1354\(99\)00129-3](https://doi.org/10.1016/S0043-1354(99)00129-3)

Winter, J. G., Landre, A., Lembcke, D., O'Connor, E. M., & Young, J. D. (2011). Increasing chloride concentrations in Lake Simcoe and its tributaries. *Water Quality Research Journal of Canada*, 46(2), 157. <https://doi.org/10.2166/wqrjc.2011.124>

Zimmerman, D. L., & Ver Hoef, J. M. (2017). The torgegram for fluvial variography: Characterizing spatial dependence on stream networks. *Journal of Computational and Graphical Statistics*, 26(2), 253–264. <https://doi.org/10.1080/10618600.2016.1247006>