Estimating Flow, Hydraulic Geometry, and Hydrokinetic Power at Ungauged Locations in Canada

James Bomhof

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in the partial fulfillment of the requirements for the Master of Applied Science degree in Civil Engineering

The Ottawa-Carleton Institute for Civil Engineering

Department of Civil Engineering

Faculty of Engineering

University of Ottawa

Acknowledgements

Thank you to Dr. Colin Rennie, my supervisor at the University of Ottawa. I admire his enthusiasm for river research, dedication to his students and their success, and strong work ethic while having a good work/life balance. I would also like to thank Dr. Wayne Jenkinson, my supervisor at the National Research Council, who challenged me to think big and taught me how to program. I was very very fortunate to have them as my supervisors and am thankful for all of their help and guidance.

Thank you to all of my friends and colleagues who have made Ottawa home for me. Thank you to my family for all of their support. And finally thank you to my wife Alanna who has journeyed through this process with me. It's time for a new adventure!

Abstract

A resource assessment of the hydrokinetic potential in Canada's rivers was completed. The main objectives of the study were to (1) quantify the potential hydrokinetic energy available for development both nationally and regionally, and (2) develop geospatial datasets identifying streams or areas of high hydrokinetic potential. Flow estimates at ungauged locations were found using multiple linear regression coupled with Canonical correlation analysis (MLR-CCA). Total theoretical hydro power, equivalent to total theoretical hydrokinetic power was calculated using these flow estimates and hydraulic head estimates from DEMs. It is estimated that there are 710 GW of potential power in Canadian rivers, with 97.5% confidence that there is at least 433 GW. Downstream hydraulic geometry (DHG) relations were applied to flow estimates to find cross section velocity and power at ungauged locations. Further testing was done on DHG relations, and were found to be most accurate when characterized by soil drainage characteristics.

Contents

Acknowledgements **I**

Abstract **II**

List of Figures **VII**

List of Figures **VIII**

List of Tables **IX**

List of Tables **IX**

Nomenclature **X**

1 Introduction **1**

1.1 Background 2

1.2 Significance of Study 3

1.3 Study Objectives and Approach 5

1.4 Organization of Thesis 6

2 Literature Review **8**

2.1 Hydrokinetic Potential in Rivers 8

2.2 Regionalization Techniques 11

2.3 Hydraulic Geometry 12
2.3.1 Introduction 12
2.3.2 Defining Channel Forming Discharge 16
Bankfull Discharge 17
Effective Discharge 18
Recurrence Interval 19
2.3.3 Proposed Hydraulic Geometry Theory 20
Regime Theory 21
Tractive Force Theory 21
Hydrodynamic Theory 22
Thermodynamic Entropy Theory 22
Minimal Extremal Theories 23
Maximum Extremal Theories 24
Summary 26
2.3.4 Downstream Hydraulic Geometry 26
Bed and Bank Material 29
Vegetation 30
2.3.5 Regional Curves 32
2.4 Conclusions 35

3 Hydrokinetic Estimation **38**

3.1 Methodology 38
3.1.1 Data sources 38
HYDAT 39
Water Survey of Canada - Measurement Database (WSC-MD) 39
Digital stream network 41

Voronoi-delineated Catchments 42
Canadian digital elevation data 44
Physiographic and climatic datasets 45
Calibration dataset 47
3.1.2 Flow prediction 51
3.1.3 Influence of neighbourhood definition through CCA 53
3.1.4 Augmenting predicted flows with measured flows 58
3.1.5 Propagation of uncertainty in flow estimates 61
3.1.6 Predicting power at a section from flow estimates 64
3.1.7 Total available power 67
Total power using elevation head 67
Total power using velocity head 68
3.2 Results 71
3.2.1 Hydrokinetic power and stream velocity estimates 71
3.2.2 Regional power estimates 74

4 Hydraulic Geometry **80**
4.1 Abstract 80
4.2 Introduction 81
4.3 Dataset 83
4.4 Analysis and Results 85
4.5 Discussion 92
4.6 Conclusions 94

5 Discussion **98**
5.1 Quantifying Uncertainty 98
5.2 Potential in Reducing Uncertainty 100

5.3 Validation of Results . . . 103
5.4 Impact and Recommendations . . . 105

6 Conclusions 109

6.1 Conclusions and Recommendations . . . 109

7 Bibliography 114

Appendices 130

A R-Scripts 131

A.1 Script to predict aggregate flows using MLR-CCA . . . 131
A.2 Script to predict increment flows using MLR-CCA . . . 142
A.3 Script to combine the aggregrate and increment flows . . . 152
A.4 Script to apply hydraulic geometry to flow results and estimate hydrokinetic power . . . 157

B Attribute Descriptions 162

B.1 Soil Landscapes of Canada . . . 162

List of Figures

2.1 Hydraulic Geometry Parameters, (Leopold and Maddock, 1953) 17

3.1 National Atlas of Canada - stream segments by major drainage basin. 41

3.2 Comparison of a DEM and Voronoi Delineated Catchment 42

3.3 Voronoi Delineated Catchments along Coastline 44

3.4 Voronoi Delineated Catchment along Regional Boundary 45

3.5 Voronoi Delineated Catchments in the Maritimes Region 46

3.6 HYDAT stations used in calibration dataset 50

3.7 Estimate confidence with no regionalization, FDC-50 55

3.8 Estimate confidence with CCA regionalization ($\alpha = 0.0001$), FDC-50 56

3.9 Estimate confidence with CCA regionalization ($\alpha = 0.0075$), FDC-50 57

3.10 Example of watershed data along Canada-USA border 59

3.11 Comparison of Confidence and Prediction Intervals 62

3.12 Estimate confidence of FDC50 flows including HYDAT flow data 63

3.13 National velocity map - 20% exceedence probability 72

3.14 National velocity map - 50% exceedence probability 73

3.15 Hydropower in Canada by Major Drainage Basin 75

3.16 Hydropower in Canada by Province 76

3.17 Elevations of Canada, Source: NRCan 77

3.18 Hydrokinetic power in Canada by Major Drainage Basin 78
3.19 Hydrokinetic power in Canada by Province 79

4.1 Robust regression for at-a-station hydraulic geometry 86
4.2 Map of HYDAT stations in WSC-MD 87
4.3 Downstream Hydraulic Geometry at Bankfull 88
4.4 Downstream Hydraulic Geometry with Regionalization by Drainage Characterization 90

5.1 Uncertainty in Hydro and HydroKinetic Estimates 100
5.2 Validation of Flow Estimates using the HYDAT and Calibration Dataset: RRMSE 105
5.3 Validation of Power Estimates using the Calibration Dataset: RRMSE . . . 106
5.4 Validation of Area Estimates using the Calibration Dataset: RRMSE 107
5.5 Validation of Velocity Estimates using the Calibration Dataset: RRMSE . . 108

List of Tables

3.1 Summary of datasets used 40

3.2 FDC Exceedence Probability Index 48

3.3 Attributes mapped to watersheds 49

3.4 Regional calibrated neighbourhood definitions α 53

4.1 NASH Statistics on Regional Downstream Hydraulic Geometry 91

4.2 NASH Statistics on Multiple Regression Downstream Hydraulic Geometry . 92

4.3 Downstream width parameters grouped by drainage conditions. 93

4.4 Downstream width parameters grouped by mode of deposition. 94

5.1 Uncertainty in Hydro and HydroKinetic Estimates 99

Nomenclature

γ Specific Weight of Water (kN/m^3)

ρ Density of Water (kg/m^3)

A Cross Sectional Area (m^3)

C_R Relative Confidence Interval

d River Depth (m)

H Hydraulic Head (m)

L Length of the Given Stream Reach (m)

L Suspended Sediment Load (kg/m^3)

n Manning's Roughness Coefficient $(Unitless)$

P Wetted Perimeter of Cross Section (m)

P_H HydroKinetic Power $(Watts)$

P_K HydroKinetic Power $(Watts)$

P_{KS} Cross Sectional HydroKinetic Power $(Watts)$

P_{SS} Specific Power $(Watts/m^3)$

Q Flow (m^3/s)

R Hydraulic Radius ($Unitless$)

S Slope ($Unitless$)

v Velocity (m/s)

w River Width, (m)

Chapter 1

Introduction

Canada has a vast network of rivers that have been harnessed for over 100 years to generate electricity from hydropower. These rivers also contain significant potential for power production from hydrokinetic resources. Studies undertaken in the past to quantify national hydropower and hydrokinetic resources have only been completed for specific rivers or regions. The National Research Council of Canada - Oceans, Coasts and Rivers Engineering division (NRC-OCRE) was contracted by Natural Resources Canada - CanmetENERGY to perform a three-phase study to quantify the hydrokinetic resources in Canada. The phases were defined as follows (Jenkinson, 2010):

- Phase I Methodology Review and Data Review: a review of the methodologies, techniques and available data for conducting a regional hydrokinetic power assessment as well as a selection of proposed methodologies and validation datasets and locations.
- Phase II Methodology Validation: the implementation of a recommended set of methodologies against the validation datasets including sensitivity/uncertainty analysis.
- Phase III Assessment Determination: the application of the recommended methodologies to conduct a nation-wide assessment on the hydrokinetic potential for theoretical

energy extraction.

The second and third phases of the study were done in collaboration with the University of Ottawa and form the basis of this thesis. Downstream hydraulic geometry equations were a critical portion of the study for the purposes of estimating velocities at ungauged stations. Hence an in-depth investigation into these equations also form a major portion of this thesis.

1.1 Background

There is significant interest in hydrokinetic or in-stream river potential in Canada and internationally for power production using in-stream or zero-head turbines. Hydrokinetic (or simply kinetic) turbines differ from traditional hydro turbines in that they extract energy from the moving water, whereas hydro turbines extract energy from the high pressure of water due to an elevation drop. There are many advantages to using kinetic turbines. Because a dam or barrage is not required, many issues relating to the flooding of upstream reaches, including the loss of land and environment concerns, are avoided. Additionally, the large up-front construction costs of building dam and access infrastructure are avoided. Instead, kinetic turbines can be deployed incrementally, lowering the investment risks. Kinetic turbines do have some disadvantages compared to classical hydro. They are less efficient due to friction losses in the wake of the turbine (Lalander and Leijon, 2011). Because they intrinsically have a run-of-the-river operation, as opposed to large dams which store water in reservoirs, the power output is less consistent and dependant on natural flow regimes[1]. It is also unlikely that hydrokinetic turbines will generate power on the same scale as traditional hydropower. For these reasons hydrokinetic power installations will not replace hydropower dams but instead supplement the electrical grid where they are deemed most advantageous. Hydrokinetic installations have shown great potential in rural communities that currently

[1]Hydrokinetic turbines installed in areas influenced by tidal currents do have predictable power outputs.

rely on diesel generators for power production (Vermaak et al., 2013).

Currently there are at least five private companies developing and deploying riverine kinetic turbines in Canada (described below in alphabetical order). 1) Clean Current, based in Vancouver, was the first company to install a ducted horizontal bi-directional turbine for tidal currents in 2006. They have recently deployed a turbine in the Winnipeg River in Manitoba. 2) Instream Energy Systems, also based in Vancouver, has installed vertical kinetic turbines rated to produce 25 KW of power in streams in Duncan, BC (tailrace of dam), Yakima, Washington, and has received approval for a project in Bordeaux, France. 3) New Energy Corporation, based in Calgary, has deployed vertical riverine turbines in Alberta, Manitoba and Alaska, and is planning projects in British Columbia, Ontario, Northwest Territories and the Yukon. 4) Renewable Energy Research (formerly RSW)is based in, and has deployed a horizontal kinetic turbine near Montreal. 5) Verdant Power is based in the USA and has installed horizontal kinetic turbines in the St. Lawrence River near Cornwall, as well as in the East River in New York City. The Canadian government has also committed to help develop riverine kinetic power, with a goal of supplying "50 % of water-to-wire river-current power systems to the international market"(Marine Renewable Energy Technology Roadmap Steering Committee, 2011). As part of the government's support of the burgeoning industry, it has also initiated research into assessing potential riverine sites. This study details some of the outcomes of that initiative.

1.2 Significance of Study

This research is significant from both a practical and technical perspective. First, a nation-wide hydrokinetic or hydropower resource assessment has never been completed in Canada. The results of this research can be used by the Canadian government and by industry to quantify national and regional potential of riverine power resources as well as provide a pre-reconnaissance estimation of power potential on specific reaches. This author hopes that the information will assist in the responsible development of Canadian rivers from an environmental and social perspective.

From a technical perspective this research is novel because of the scale to which flow regionalization was done. Typically flow regionalization has been applied at a single ungauged station or a series of ungauged stations within a specified region. Because flows in all the rivers in Canada needed to be estimated, it required the amalgamation of national datasets into a form that had never been done before, and even resulted in the creation of a new national Voronoi watershed dataset. In addition to new datasets, a regionalization scheme using Canonical Correlation analysis (CCA) was automated and implemented on a national scale for the first time. Previously CCA had only been implemented on a regional scale for proof-of-concept purposes.

The inroads made in the topic of downstream hydraulic geometry are also novel and significant because of the data and methods used. Previous studies on downstream hydraulic geometry have been limited to the testing of correlations within a specified physiographic or hydro-geologic region, likely due to dataset limitations. This is the first known study to define regions in downstream hydraulic geometry by attributes that are not necessarily confined within a contiguous region.

The literature shows a significant interest in hydrokinetic resource assessment in the 1980s (UMA Group, 1980; Miller et al., 1986). These previous resource assessments have been limited in scope because of the limitations in data and computational power. This study (and a more recent study in the USA by Jacobson (2012)) is the first to fully utilize GIS capabilities and automated data processing algorithms to calculate a truly national hydrokinetic resource assessment.

1.3 Study Objectives and Approach

This research is part of a series of studies designed to quantify the national potential of riverine hydrokinetic resources in Canada. The first phase included a review of the methodologies, techniques and available data for conducting a regional hydrokinetic power assessment as well as a selection of proposed methodologies and validation datasets and locations. The results of this study are outlined in Jenkinson (2010). The second phase tested the implementation of a recommended set of methodologies against validation datasets including sensitivity/uncertainty analysis while the third phase implemented these methodologies nationwide.

The final objectives of these studies are to:

1. Quantify the national potential in Canada for hydrokinetic power production; and

2. Develop a geospatial dataset to identify regions of high hydrokinetic potential where no measurements were otherwise available (regionalization).

This report details the third phase of the investigation: the application of hydrokinetic prediction techniques developed by NRC. The approach uses multiple-linear regression (MLR) with Canonical correlation analysis (CCA) to provide flow regionalization. Hydropower equations are applied to these flows in order to quantify the regional and national energy potential from Canadian rivers. Finally, hydraulic geometry equations are used to extract velocities, widths and depths at a river cross section in order to produce maps showing river reaches with the highest hydrokinetic potential. Because the accurate prediction of geometry at an ungauged cross section is a critical component in determining hydrokinetic potential (Jenkinson, 2010), objectives were also set to analyze the uncertainty in these equations and find ways to improve their accuracy. The large majority of the research and written report that this thesis includes was completed by this author under the supervision of Dr. Colin D. Rennie (University of Ottawa), and Dr. Wayne Jenkinson (National Research

Council). Some small portions written by Dr. Jenkinson included in this thesis are stated at the beginning of the pertinent sections and the text is indented.

1.4 Organization of Thesis

This document is outlined as follows:

Chapter 2 reviews previous hydrokinetic and hydropower resource assessments done in North America. It also provides an extensive literature review related to hydraulic geometry, including channel forming discharge, hydraulic geometry theory, the use of regional curves, and the influence of bank materials and riparian vegetation.

Chapter 3 outlines the methodology and data employed to accomplish the hydrokinetic assessment, including data collection, implementation and analysis of flow regionalization approaches, and hydropower and hydrokinetic resource estimation. It summarizes and presents the results of the hydropower resource assessment as aggregated regional and national datasets. Finally it presents the results of hydrokinetic resource assessment, showing river reaches with the highest resource potential.

Chapter 4 utilizes data sets collected in the hydrokinetic study to further investigate influences of geology and vegetation on downstream hydraulic geometry equations. These equations can be applied to the flow estimates determined in Chapter 3 to determine hydrokinetic potential at a reach. This chapter is being submitted as a stand-alone journal article.

Chapter 5 analyses the uncertainty in the hydro and hydrokinetic predictions and discusses potential methods and datasets for improving the accuracy of these estimates. This is followed by a validation analysis of the flow predictions and finally a review of the study's findings is provided and recommendations for the employment of the data products developed.

Chapter 6 summarizes the results of the thesis and recommends areas of future research that the author believes show the greatest potential.

Chapter 2

Literature Review

2.1 Hydrokinetic Potential in Rivers[1]

The only known Canada-wide study of current-power potential was conducted by the UMA Group (1980) for the National Research Council (NRC). The study was restricted to a small number of rivers (the largest in Canada) and river reach inclusion was limited by threshold flow (>450 m^3/s), velocity (>1.5 m/s), width (>50 m) and depth (>3m) values. Flow was based on HYDAT data (HYDAT is described in section 3.1.1) records having a minimum 20 years of data. The UMA Group study did not examine a large number of small to medium rivers that, although possibly providing less available energy than those examined, could still offer hydrokinetic power potential. As the study was limited to techniques and data available circa 1980, it limited the spatial resolution of the analysis to very large reaches and manual techniques for channel property estimation. Advances in the types and quality of the data as well as computational power available today would have allowed for a much more thorough and detailed investigation.

[1] indented portions of the hydrokinetic literature review were used with permission from the Phase I report of this study(Jenkinson, 2010)

Miller et al. (1986) conducted a study for the US Department of Energy to estimate the hydrokinetic resource potential in the United States. The methodology involved first dividing up the continental US into 16 physiographic/hydrologic regions. Hydrometric data from the USGS were examined and analysis was restricted to river reaches with greater than 113 m^3/s mean flow rate and 1.3 m/s flow velocity. Flow rates were determined for each river developing a relationship between the channel distance to the mouth of the river and the mean flow rate. This was done by developing a linear relationship between channel distance to the river mouth and the drainage area, and then developing a relationship between drainage area and flow rate and combining the two to provide the desired channel distance flow rate relationship. Channel slope and geometry were estimated from USGS topographic maps and a rectangular cross section was assumed. The flow velocities were estimated using the Manning equation as per the UMA Group study (UMA Group, 1980). This approach was considered to provide a conservative resource estimate and the total US hydrokinetic energy resource was thus estimated to be 12.5 GW, with the Western, Northwest and Alaskan areas exhibiting the greatest overall resource contributions.

Acres conducted a study for Environment Canada, Inland Waters Directorate, in 1984 to develop a pre-feasibility methodology for assessing hydropower resources in Canada (Acres Consulting Services Limited, 1984). The study identified 12 hydrologic regions in Canada of which 11 were included in analysis (the arctic islands region was excluded). For each hydrologic region a representative flow duration curve was created, which was normalized to a median flood flow rate. The Index Flood Method was employed to apply the regional FDC to a local site, although an integration of the FDC was used to produce a "turbinable" flow curve for each ungauged site. A hydrokinetic power generation was addressed in this report, but it required data to be collected in the field to characterise the location to be investigated. This method was validated against a simulated flow analysis and showed an average of 5% absolute error in turbinable flow estimations across all 11 regions. The standard deviation of the percent error was found to vary from 5 to 20%.

The U.S. Department of Energy (USDOE) conducted a feasibility study to examine the hydropower energy resources across the United States for small- and low-power hydro (USDOE, 2004, 2006). This study focused on employing geographic information tools and digital elevation model data to determine the assessment, and the focus was on low-head and low-power resources. This US national study was preceded by a number of regional studies that examined the feasibility of the technique in Arkansas, the Pacific Northwest and the North Atlantic regions USDOE (2002a,b, 2003). The principal determination of available hydropower was conducted by determining a mean annual flow at a location and estimating the potential hydraulic head that could be developed at that location. It used GIS data to determine stream locations, delineate stream segments and power potential, or possible penstock height. Calculations of hydraulic head were based on elevation drop using DEM data. The study suggested some uncertainty in the estimation of the power potential, but the error sources were not identified or quantified.

In 2007, Kerr Wood Leidel Associates (KWL) developed a GIS assisted hydropower assessment system for BC Hydro and the BC Transmission Corporation called the Rapid Hydropower Assessment Model (RHAM) (Kerr Wood Leidal Associates, 2008; Monk et al., 2009). RHAM was used to conduct an inventory of potential run-of-river hydroelectric sites across British Columbia. In this study run-of-river referred to low-head hydropower with a constructed penstock. The study identified over 8,000 potential hydroelectric opportunities with over 12 GW of power generation potential in total.

The Electric Power Research Institute (EPRI) completed a riverine hydrokinetic study in the United States that included the 48 contiguous states and Alaska (Jacobson, 2012). They estimated the total theoretical power in river reaches without any hydroelectric dams and having a mean annual flow of at least 1000 cubic feet per second (28.3 m^3/s). Theoretical power was estimated using mean annual flow and the elevation drop over a reach, as estimated by the NHDPlus database (USGS and USEPA, 2006). They applied a recovery factor to these estimates that accounted for minimum required water velocity of 0.5 m/s and a depth of 2 m during low flow conditions (5^{th} percentile), device deployment and efficiency, ranges of flows and backwater effects. The recovery factor was modeled using HEC-RAS and developed as primarily a function of flow and slope. They estimated a total theoretical power of 1,381 TWh/y (158 GW) and a recoverable power of 120 TWh/y (14 GW).

2.2 Regionalization Techniques

The basis of hydrokinetic power prediction in this study lies in the accurate estimation of flow at ungauged locations, or regionalization. Regionalization consists of two major components: 1) the identification of similar gauged watersheds which can be used as data sources (ie. defining a 'neighbourhood') and 2) the regional estimation method with transfers data from these gauged watersheds to the ungauged watersheds (Ouarda et al., 2001a). A literature review on the types of regionalization was completed in phase I, which can easily be found online (Jenkinson, 2010). The report recommended that multiple linear regression (MLR)be used as the regional estimation method (2) while Canonical correlation analysis (CCA) be used in the identification of similar gauged watersheds. While a full description of these methods are found in the phase I report, CCA will be described briefly.

CCA was first used in flow regionalization by Cavadias (1990). In essence, CCA finds correlations between two separate groups of variables. When used in flow regionalization, the first group of variables are a set of flow values and the second group of variables are independent watershed characteristics. Linear combinations of each of these groups are then tested to find statistical correlations between these groups in the newly defined canonical space. Cavadias (1990) and Ouarda et al. (2001a) demonstrate how a 'neighbourhood' of similar stations can be defined within this canonical space. Watersheds that are most statistically similar to one another will appear in close proximity to one another. A chi-squared distance can be defined and optimized to determine which stations should be included in the neighbourhood.

2.3 Hydraulic Geometry

2.3.1 Introduction

The basis for hydraulic geometry relationships (HGR) stems from observations of stable canals in India (Henderson, 1966). Engineers recognized that an understanding of the interaction a stream's variables was critical in design of irrigation and river systems. Regime was first defined in this period. A river in regime, or equilibrium, has adjusted its slope and dimensions to carry a balanced sediment load (ie. the sediment that is introduced to a stream equals the sediment that exits the stream). This early discussion on stable canals layed the groundwork for carryover into the study of river morphology.

Much work has since been completed to find out how a river moves and shapes itself. Though primarily driven by flow, many other factors influence the shape of a river. The complexity of sediment processes at river reach and water scales has necessitated the need to often describe influences in empirical and semi-empirical relationships. This literature review begins by introducing and defining hydraulic geometry and then discusses flow definitions most useful to the topic. An overview of theoretical equations that have been employed in determining hydraulic geometry is also given. The influences of bed and bank material, and riparian vegetation are discussed. Finally a summary of regional curves, denoting hydraulic geometry conditions as a function of drainage area rather than flow, is presented.

Kennedy and Lindley first proposed an exponent relationship between velocity and depth in 1895 (Henderson, 1966). Leopold and Maddock (1953) expanded on this work and cemented the name of the concept hydraulic geometry. Their landmark paper framed the discussion for years to come. They defined a set of equations in which hydraulic variables were primarily a function of river discharge. A subset of these variables were related to each other through continuity. More specifically:

$$w = aQ^b \tag{2.1}$$

$$d = cQ^f \tag{2.2}$$

$$v = kQ^m \tag{2.3}$$

Continuity shows that:

$$Q = wdv = aQ^b \cdot cQ^f \cdot kQ^m \tag{2.4}$$

and therefore:

$$b + f + m = 1 \tag{2.5}$$

$$a \cdot c \cdot k = 1 \tag{2.6}$$

where w is stream width, d is average depth, v is average velocity, Q is flow, a, c, k are empirical coefficients and b, f, m are empirical exponents. These relationships were defined for at-a-station and for the downstream direction, where at-a-station referred to the geometries found at a single cross section on a river with varying flow rates (Q), and downstream referred to geometries found at multiple cross sections on a river with an equal frequency flow. Leopold and Maddock plotted these relationships using data from gauge measuring stations in the American mid-west. The streams were assumed to be in an equilibrium or quasi-equilibrium state, where average channel dimensions are constant over time but lateral migrations of the channel can occur within the floodplain due to natural erosion and deposition processes (Watson et al., 1999).

At-a-station relationships plotted were very similar to the stage rating curves; the novelty was in transforming them on a log-log scale and comparing the slope of the lines. Some of the scatter among the plots was attributed to temporary scouring and filling of the bed, as well as different cross sections being measured. Leopold and Maddock (1953) reported wide variation in slopes of the lines of best fit, with average values of 0.26, 0.40 and 0.34 for width, depth and velocity, respectively. The exponent f was typically greater than b, showing that depth increased faster than width for an increasing discharge. The values also give an indication of channel shape: in a rectangular channel, $b = 0$ and $f + m = 1$, whereas for a triangular channel, $b = f$. With at-a-station relationships, Leopold and Maddock (1953) presented old information in a new way in order to add insight to the nature of rivers.

For downstream hydraulic geometry equations, the difference in regression lines between the trunk of a river and its tributaries was found to be small among various rivers and even canals in India Henderson (1966). Though there was scatter in the data, the general trend was that different river systems have mostly parallel slopes (and different intercepts) across a wide range of physiographic settings. Also, velocity tended to increase in the downstream direction because the increase in depth typically overcomes the decrease in slope. Average values for b, f, and m were 0.5, 0.4, and 0.1 respectively. The averages did not mean to imply homogeneity among the streams, only general trends. The mean annual flow was used as the flow frequency for the downstream HGCs because the data were readily available. Downstream HGCs plotted for other flow frequencies demonstrated that the relationships held for any flow frequency, though the b,f and m values did vary.

Suspended-sediment load was also described as having a log-linear relationship with flow at-a-station in the form:

$$L = pQ^j \tag{2.7}$$

where L is the suspended sediment load, p is an empirical coefficient and j is an empirical exponent. The slope exponent (j) is typically between 2 and 3, meaning that suspended-sediment concentration will increase faster than discharge. This is likely due to suspension being a non-linear function of shear stress or to increased erosion occurring in the watershed during rainfall; the channel must carry away the imposed sediment load. However, flow does not fully account for the suspended sediment load concentration as upstream watershed conditions will also influence sediment load concentration; ie. sediment load depends on erosion and retention happening on the surface, as well the proportion of groundwater entering the stream. By plotting and observing the interactions between b, f, m and L for at-a-station and downstream HGCs, Leopold and Maddock concluded that b, f and m are functions of L. In a specific example, at constant width and discharge, an increase in suspended sediment load is accommodated by an increase in velocity and a decrease in depth (Leopold and Maddock, 1953).

Slope (S) and Manning's bed roughness (n) were also related through Manning's equation. By substituting equations 2.2 and 2.3 for hydraulic radius and velocity, it was shown that $\frac{S^{1/2}}{n}$ increases with discharge at-a-station and decreases downstream for the average stream condition. Wolman (1955) showed that at-a-station slope remained approximately constant across the range of discharges and Leopold and Maddock argued that bed roughness decreases with increasing sediment load, thus proving correct the mathematical interpretation of an increasing at-a-station $\frac{S^{1/2}}{n}$ ratio. Leopold et al. (1964) showed that slope tends to decrease in the downstream direction while roughness remains relatively constant, leading to the explanation of a decreasing $\frac{S^{1/2}}{n}$ ratio in the downstream direction.

The authors include width, depth, velocity, suspended sediment load, slope and roughness parameters in the term "hydraulic geometry". The relationship between the at-a-station and downstream parameters is shown in Figure 2.1 Leopold and Maddock (1953).

Dury (1976) showed that hydraulic geometry equations can be reversed by measuring widths, depths and slopes from existing and former channels to predict bankfull flood, and he confirmed that the relationships were log-linear.

2.3.2 Defining Channel Forming Discharge

Channel forming discharge was first introduced by Leopold et al. (1957). They showed that floodplains are composed mainly of materials carried during lower intensity, higher frequency flows. This led to postulation that a single discharge value can represent the range of flows that form a stream channel. In other words, a discharge can be found that, if allowed to flow continuously in a channel over a long period of time, would shape the channel the same way the actual range of flows have shaped the channel. Wolman and Miller (1960) showed that natural channel shape is determined by flows at or near the bankfull stage rather than infrequent, catastrophic events. They also found these events typically do more work (ie. move more sediment) than any other other flow frequency. Frequencies lower than bankfull can be used in downstream hydraulic geometry curves, however they simply reflect

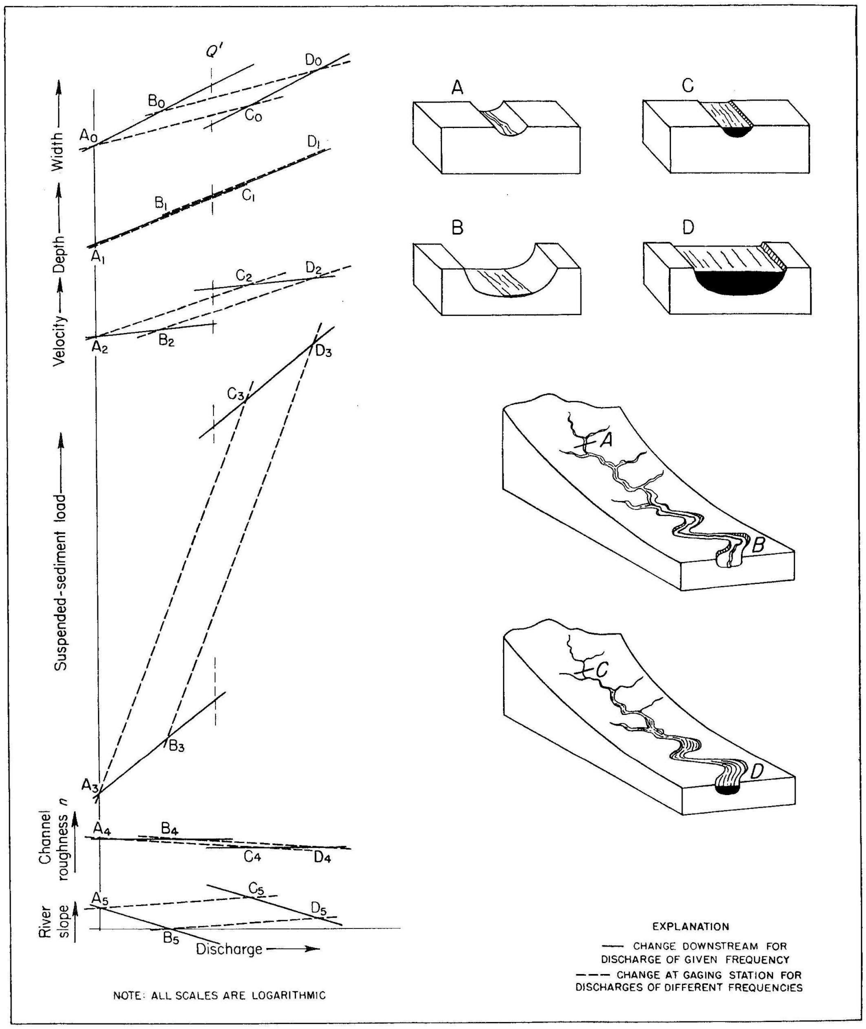

Figure 2.1: Hydraulic Geometry Parameters, (Leopold and Maddock, 1953)

the conditions created by effective flows (defined below as, “that discharge which, over a long period of time, most sediment has been transported” (Benson and Thomas, 1966)) (Leopold et al., 1964). Channel forming discharge, or dominant discharge, is defined in 3 ways: bankfull, effective discharge, and specific recurrence interval (Copeland et al., 2000).

Bankfull Discharge

Bankfull discharge is the maximum flow that a stream can achieve without overtopping its banks and spilling onto the floodplain (Copeland et al., 2000). Wolman (1955) and Pickup and Warner (1976) determined bankfull as the discharge where the width:depth ratio was at a minimum; width would increase dramatically once the flow spilled onto the floodplain. Harvey (1969) found that bankfull discharge was approximately dominant discharge for "normal" streams but found it did not apply to a stream with a large base flow where bankfull was determined by much rarer events. Knighton (1998) concluded that for alluvial or semi-alluvial rivers in humid, temperate regions, bankfull is generally a good approximation of effective discharge.

Though the definition of bankfull being the height of the active floodplain is commonly accepted, physically determining the elevation of the active floodplain is open to interpretation. Williams (1978) compared 16 methods of determining bankfull and provided an empirical equation to calculate based on his results. Radecki-Pawlik (2002) tested the methods Williams outlined on streams in Poland and concluded that bankfull discharges for mountain stream should be reported as a range of values instead of a single value . It is the author's opinion that this conclusion would negate the intent of channel forming discharge, as it is defined as representing the range of flows of a river. Dr. Colin Rennie has found the most useful field determinations to be (1) the elevation of rooted, woody (perennial) vegetation, (2) break in bank slope, and (3) the elevation of the inner bank point bar or mid-channel bars or islands (C. Rennie - personal communication, 2013). Though there are many methods in interpreting bankfull elevation in the field, the definition of bankfull as the elevation of the floodplain is widely accepted.

Effective Discharge

Wolman and Miller (1960) first introduced the concept and method of calculating effective discharge, is defined as "that discharge which, over a long period of time, most sediment has been transported" (Benson and Thomas, 1966). A flow duration curve is multiplied by the sediment load curve. The resulting curve shows the amount of sediment carried at each flow interval, with the maximum being defined as effective flow. Benson and Thomas (1966) applied the definition to nine different streams and found the dominant discharge to be much less than bankfull discharge. Pickup and Warner (1976) also found this to be the case. Emmett and Wolman (2001) calculated effective discharge on five gravel bed rivers in the Colorado and found effective discharge to be between 0.98 and 1.3 of bankfull discharge, concluding that rivers with constraints on bed mobility have effective discharges that are higher than bankfull. Klonsky and Vogel (2011) introduced a consistent method for calculating effective discharge while proposing a new index termed "half-load discharge", the flow at which half the total sediment has been transported over time.

Recurrence Interval

Initially, bankfull discharge was equated to an interval on the frequency duration curve. Leopold and Maddock (1953) also used the flow duration curve to determine flow frequency in their downstream hydraulic relationships (Searcy, 1959a). Nixon and Lacey (1959) measured streams in England and found the average bankfull flow frequency to be 0.6%, or about 2 days per year. Dury (1961) determined that flow duration curve is not a good indicator of bankfull because it varies along the length of a stream. A near-flood event responsible for bankfull flow will have a flashier hydrograph curve in the upstream reaches than in the downstream reaches. This is because the maximum flow attenuates as it moves downstream. Hence, a flood frequency recurrence interval is a more accurate statistical approach for measuring bankfull.

In studies of channel forming discharge as a recurrence interval, many authors have used the term interchangeably with bankfull discharge (Copeland et al., 2000). Leopold et al. (1964) did a survey of 19 rivers and found bankfull recurrence interval varied from 1.07 to 4 years, but with 15 of the stations lying between 1.07 and 1.9 years. *"The effective discharge can often be approximated as bankfull discharge. In many rivers the bankfull discharge is* *one that has a recurrence interval of about 1.5 years"* (Leopold et al., 1964). The outliers are likely attributed to flows whose watersheds cause more flashy hydrology in the channels (Copeland et al., 2000). Andrews (1980) calculated effective discharges for the Yampa River in Colorado using flow and suspended-sediment measurements from 15 gauge stations along with calculated estimates of bed load sediment. He found effective discharge to be nearly identical to measured bankfull discharge. A bankfull recurrence interval of 1-2 years has been found by many more authors (Harman et al., 1999; Dury, 1976; Nixon and Lacey, 1959; Lawlor, 2004; McCandless, 2003), though this value has been questioned over the years (Williams, 1978; Harvey, 1969). One of the reasons for critical concern is that recurrence is typically calculated using an annual maxima series, limiting the minimum recurrence to 1 year (Navratil et al., 2006). Also, as noted above, identification of the bankfull channel geometry can be difficult, leading to uncertainty in the identified relations. Lastly, the recurrence interval of bankfull conditions may depend upon position within the watershed. Adams and Spotila (2005) measured bankfull discharge on nine headwater streams ($<2km^2$ drainage area) in the Appalachian mountains and found bankfull to have a recurrence interval greater than 2 years.

2.3.3 Proposed Hydraulic Geometry Theory

Many theories have been proposed to describe the physics of why streams form themselves in the patterns seen in nature. It is hoped that better understanding of why hydraulic geometry occurs will aid in future design, planning and management. Singh (2003) provides a very strong overview of the theories proposed to date. He divides the theories into six separate categories which will now be discussed. It is by no means meant to be exhaustive, but rather describe the more popular developments in the field.

Regime Theory

As previously mentioned, regime theory was first proposed in the late nineteenth century through observations on irrigation channels in India and was cemented in literature through Leopold and Maddock's famous paper on hydraulic geometry. It says that a river will evolve into an equilibrium or quasi-equilibrium where the flow and sediments introduced to a stream will be carried the length of the reach (Lindley, 1919). Deposition and sedimentation occurs along a reach in equilibrium while the long term average channel geometry (width, depth, and velocity) remains stable (Hey, 1978) (as cited in (Singh, 2003)). Engineers Kennedy and Lindley first proposed an exponent relationship between velocity and depth in 1895(Henderson, 1966). Lacey expanded this relationship by defining hydraulic radius and depth as a function of discharge and introducing a silt factor(Lacey, 1930). Blench (1957) (as cited in (Singh, 2003)) developed relationships similar to Lacey, but differentiated between bed and bank material. These equations did not incorporate any sediment transport theory, but were assumed to have the ability to carry the required sediment load implicitly. There has been large criticism of Lacey's "silt factor" as the units in his equation don't align, and it is contradictory to Strickler equation for roughness(Stevens and Nordin Jr, 1987).

Leopold and Maddock's hydraulic geometry power equations are essentially analogous to regime theory. As described by Singh (2003), "regime theory is a concept and the [hydraulic geometry] regression analysis is a fitting technique to apply the concept to measured data."

Tractive Force Theory

Tractive Force Theory was popularly introduced by Lane (1937) (as cited in (Singh, 2003)). He states the design of a stable channel is based on the premise of ensuring the bed and bank material is secure enough to resist the tractive forces from flow trying to scour or deposit sediment along the reach. In order to calculate the forces, Lane makes five assumptions as to what conditions are holding a particle in place (ex. submerged weight of particle) and the force to cause incipient motion. Channel equilibrium occurs when the tractive forces of the bed equals the lift and drag forces of discharge.

In his thesis defense, Li (1975) (as cited in (Singh, 2003) used the tractive force theory and assumed the threshold discharge to be bankfull discharge. He derived at-a-station downstream hydraulic relationships on this basis and found that they were in good agreement with previously reported findings.

Hydrodynamic Theory

Smith (1974) developed a model for downstream hydraulic geometry based on hydrodynamic theory. He applied three conditions: 1) Sediment is conserved during transport 2) the channel form is just sufficient to transport its water discharge, and 3) the channel form is just sufficient to transport its sediment discharge. He then applied Manning's equation and a sediment transport equation and calculated downstream relationships for non-cohesive sediments. Singh (2003) called this theory "promising and insightful but surprisingly has received little attention". This author surmises that this may be because the calculated b coefficient of 0.6 is higher than the generally accepted average value of 0.5.

Thermodynamic Entropy Theory

Yalin and Da Silva (1999) developed a method for calculating a channel's regime characteristics (width, depth, slope) based on the Second Law of Thermodynamics and the Gibb's equation. The need for a transport rate equation in the formulation is replaced with a regime width relation. A standard method is described and tested using represented discharges on six different rivers.

Minimal Extremal Theories

Many theories have been developed which are based on minimizing some component of a stream's physical interaction. These have been grouped together as minimal extremal theories.

Langbein and Leopold (1964) first proposed a theoretical derivation of hydraulic variables by assuming uniform energy distribution and minimum total energy. A stream having minimum energy would have an elevation profile as concave as physically possible and described by $s = nQ^{-1}$. High slopes occur in the beginning of the reach where flow is low, and low slope in the end of the reach where flow is high. This was described as the lowest possible slope profile, whereas the highest slope profile would be limited by uniform energy distribution: if $w = aQ^{0.5}$, then $s = nQ^{-0.5}$. Actual data showed the coefficient to lie somewhere in between at around -0.75. They explained this in terms of suspended sediment load; high velocity in the upper reaches causes degradation while lower velocity in the lower reaches causes agradation to create the equilibrium profile.

Langbein (1964) also proposed the minimum variance theory in a separate article the same year. It states that a change in hydraulic geometry to accommodate a change in stream power will be shared equally among all the variables. Though this approach has been highly criticized, it does elicit a simple understanding of how constraining one parameter will affect other parameters. Knighton (1977) examined and supported the use of the theory while Williams (1978) identified variables used in the theory: velocity, depth, width, bed shear stress, friction, slope and stream power. Singh et al. (2003), after reviewing all previously proposed theories, used a similar form, the theory of maximum entropy, in deriving theoretical downstream equations.

Also in 1964, Dou proposed that a channel would adjust its dimensions so as to minimize its channel mobility within physical dimensional constraints. He measured channel mobility by the amplitude of the longitudinal and cross-sectional deformation of a channel, and found that it is dependant on velocity and the width:depth ratio. This theory was tested on data from Leopold and Maddock (1953) and found to fit well.

Yang and Song (1979) proposed the theory of minimum energy dissipation, where geometry was constrained by three additional equations: 1) an equation showing total energy dissipation required to transport flow and sediment in the reach, 2) the Manning equation describing flow as a function of wetted perimeter, roughness and slope, and 3) sediment concentration as a function of stream power. Theoretical coefficients of depth and slope were 9/22 and -2/11 respectively, which showed agreement with depth coefficients in the Midwest USA and slope coefficients in Indian canals. It is not clear how the slope coefficient is rectified with the -0.75 value theoretically calculated by Langbein and Leopold (1964).

The theory of minimum stream power was first proposed by Chang (1979). He stated that a channel will adjust its width, depth and slope in order to make unit stream power a minimum. He applied a resistance equation, sediment transport equation and found unique solutions for small discharges. When the minimum calculated slope equaled a valley slope, it was found that the channel was straight. Above a threshold valley slope, the model predicted multiple solutions, indicating sinuous streams eventually leading to braided streams as the valley slope increased. He found with accurate values of grain size and form roughness, his model was able to predict velocity and depth for a specified width. Chang later added the theory of minimum total work to solve for unique solutions of width.

Maximum Extremal Theories

Expectedly, many theories have also been developed based on maximizing a component of a stream's physical interaction. Eaton and Church (2007) state that most of the minimum and maximum theories arrive at the identical solutions, therefore the differences lie not in the predictions, but in the physical implications of the assumptions. Though the theory of maximum entropy is very similar to the theory of minimum variance in practice, it draws from principles in thermodynamics. Leopold and Langbein (1962) drew comparisons to heat transfer when first proposing the theory. Deng and Zhang (1994) used this theory in deriving morphological equations, but assumed width and depth to be independent, and also assigned a hierarchy to the variables. Singh et al. (2003) also used the hierarchy idea in their theory, testing how likely one variable would change over another. This notion explains the greater variation in some exponents than others; for example, the width exponent is consistently around 0.5 while velocity is much less coherent.

After deriving sediment transport equations, White et al. (1981) extended his knowledge to develop morphology equations based on the theory of maximum sediment transport. This assumption was coupled with continuity, sediment transport, and flow resistance equations. Maximum sediment transport occurs when streampower is minimized, making his theory equivalent to the minimum stream power theory in practice. When tested against observed values, depths were found to be in good agreement though slopes showed larger scatter. Caution should be used as only lab data were compared.

Davies and Sutherland (1980) argued that geometry will adjust so that boundary conditions produce a maximum friction factor as defined by the Darcy-Weisbach equation. Their results were consistent for low flow regimes. This author questions this theory as bed form friction is highly dependant on the flow. Bed perturbations will be spaced at different lengths dependant on the speed and turbulence occurring in the water (maximum turbulence without skimming flow). Thus bed form shape and friction will be different at varying lengths and equilibrium will not be achieved on a long term basis.

Huang and Nanson (2000) defined maximum flow efficiency as the maximum sediment transporting flow per unity available stream power. They used this definition, along with an integration of a channel form factor, sediment transport equations and a channel shape factor to calculate equilibrium channel geometry. Results compared well with Lane (1955).

Summary

Singh (2003) concluded in his review that hydraulic geometry theories have not been well compared and contrasted. There is a lot of validity and promise to many of the theories proposed but has been a lack of application on the vast observed data sets available. Though some theories have stood out more than others, Leopold and Maddock (1953)'s paper is still the prominent contribution. As data becomes more readily available and (hopefully) standardized, perhaps more application of the theory will lead to better insight on the complex workings in equilibrium channels.

Over the past 60 years a piecemeal approach has been taken to developing and testing the above mentioned theories. While this is the natural evolution of ideas, no physical theory has come to the forefront and the time seems right for these theories to be tested against one another with standardized methodology. Higher computing power, better sensor technology (ex. remote sensing technologies, improved methods for acoustic measurements, etc.) and increased data sharing provides the opportunity to better test and compare these theories. The collection of an extensive dataset, along with standard implementation of these theories for direct comparison with one another will provide valuable insight into where future research on hydraulic geometry theory should be focused.

2.3.4 Downstream Hydraulic Geometry

Downstream hydraulic geometry can be described as a scaling factor for rivers (Ellis and Church, 2005); as flow increases in the downstream direction, geometry changes in accordance with flow. Leopold and Maddock (1953) first plotted downstream hydraulic geometry for the Powder River in Wyoming and Montana using 23 different stations and mean annual flow (MAF) which is expected to occur $\sim 25\%$ of the time. The difference in lines between the trunk and a river's tributaries was "astonishingly small", leading to the conclusion that for a given river basis, the corresponding geometry values are very consistent for an equal frequency discharge.

Leopold and Maddock (1953) also calculated widths, depths and velocities at various flow frequencies on the Maumee River and the Scioto River which have a common divide and showed "no significant differences in the hydraulic characteristics for the purpose of this paper." These were plotted for the 1% flow and the 50% flow and showed that these relationship could cross drainage divides. Wolman (1955) took measurements on 7 different stations on the Brandywine River, Pennsylvania. He found that downstream exponents were not specifically related to regional differences and concluded that the processes which manifest themselves in these equations are similar in different regions. This study supported the findings of Leopold and Maddock (1953) that the power equations are relatively consistent across a range of physiographic regions. Additional confirmation came when Leopold et al. (1964) plotted hydraulic geometry for many American rivers and also bankfull stage for the Amazon river, one of the largest rivers in the world. The data showed that the hydraulic geometry relationship holds true for many rivers over a wide range of discharges. The width exponent varied from 0.42-0.55 though the Loup river was an outlier which could not be explained.

Leopold et al. (1964) argued that all rivers trend towards a statistically probable equilibrium behavior. There are eight variables involved in downstream changes: width, depth, velocity, slope, sediment size, size of debris, roughness and discharge. Because many of these variables depends so much on landscape factors, it is very difficult to model the processes simultaneously; "there are many hills, many channels, many rills" (Leopold et al., 1964). They concluded that this indeterminacy causes many feedback loops within the system that trends towards mutual adjustment of variables and equilibrium.

Park (1977)'s paper marked a change in thinking about hydraulic geometry. Since Leopold and Maddock's paper, the general consensus and goal had been to find the common average of hydraulic geometry relationships. Park argued that emphasis should be put on explaining the deviations in data rather than simply tying to find average exponent values. By plotting all of the b-f-m relationships to-date on a tri-axial diagram, Park (1977) found the scatter was too substantial to conclude that all streams could be represented by one mean value. He tried grouping the streams by climatic region but found that it did not reduce scatter. He concluded that scatter between the variables likely is due to factors of bank material composition, sediment properties, channel symmetry associated with plan form variations, flow regime and flow variability. In a comment and response to Park's initial paper, both Rhodes (1978) and Park (1978) conclude that there is more to hydraulic geometry than first meets the eye. Park and Rhodes both cited literature (Knighton (1977); Kirkby (1976); Newson and Harrison (1978); Williams (1978); Phillips and Harlin (1984)) that supported the unique nature and characteristics of a broad range of streams that lead to a variation of the coefficients from 0 to 1. They proposed that the research question should be changed from *what is the common factor in rivers?* to *what makes a stream unique and how does that define its geometry?*.

Leopold et al. (1964) attribute the reduced scatter in data for irrigation canals to the uniformity of bed and bank materials. This conclusion aligns with Wolman (1955), who found channel slope and shape to be primarily a function of discharge, incoming sediment load, and bed and bank material. Huang and Nanson (1998) mapped bankfull geometries across the USA and England and delineated them according to bankfull strength as a function of soil type and vegetation. It is clear that bed and bank materials, and vegetation have a critical influence in downstream hydraulic geometry relationships. These influences are further discussed below.

It is important to note sources of error when discussing influences of downstream hydraulic geometry. Harman et al. (2008) quantified uncertainty in the physical measurements at a gauge station. They found for bankfull discharge, width, depth, velocity, and slope, there was a 4%, 3%, 7%, 23% and 42% variation respectively. Osterkamp (1982) and other authors have integrated error analysis into their models through stochastic modelling.

Bed and Bank Material

Schumm (1960) first related sediment type to hydraulic geometry relationships by creating a silt-clay parameter (M) used to described the physical characteristics of bed and bank sediment. For at-a-station geometry, the width:depth ratio could be expressed as a function of this M. He also found M to greatly influence downstream geometry. The width exponent is lower, and depth exponent is greater when M increases in the downstream direction than when compared to M being constant in the downstream direction.

Knighton (1974) measured downstream relationships on a stream of consistent bank erodibility along the length of the reach and found downstream width to primarily be a function of discharge. He found at-a-station geometry to be constrained by cohesive sediment but this was sometimes overcome from sediment carried from upstream being deposited in the form of point bars and islands. Ebisa Fola and Rennie (2010) presented downstream hydraulic equations for regions dominated by cohesive clay beds. Their results suggest that the concept of hydraulic geometry and channel forming discharge can be extended to non-alluvial rivers.

Osterkamp (1982) developed single and multiple variable downstream relationships in Missouri while paying particular attention to sediment sizes. Sediment size was found to play a significant role, though because it does not have uniform effects on geometry, it was treated as a grouping term rather than an independent variable. He calculated larger standard errors in the sand channels and attributed it to the greater erodibility during floods. Moreover, streams with high silt-clay content had lower widths at a given flow and the width exponent was lower than for sand bed and bank streams. The width exponent is said to be influenced by the bed load carried from upstream; because sand size generally accounts for a large part of bed load movement, streams in sandy areas tended to have higher width exponents (ie. width changes faster in the downstream direction). Hey and Thorne (1986) also developed a multi-variate method and found bed load transport to be a significant influence, particularly in regards to slope. Bed material was found to be important in determining downstream depth, velocity and slope.

While it is clear that bed and bank material and bed load significantly influence hydraulic geometry, there is still great uncertainty in their extent. Because comprehensive, consistent, detailed and accurate data is difficult to collect over large areas, most studies have opted to categorize these variables rather than quantify them in the models.

Vegetation

Vegetation has been seen to be an important factor in that it changes bank stability, though its effect and magnitude have been debated. Osterkamp (1982) asserted that riparian vegetation can have a pronounced effect on width-discharge relationships. It should be noted that the vegetation in question typically refers to riparian vegetation, which is the zone whose biome is directly influenced by its proximity to the river. By looking at satellite imagery, Eaton and Church (2007) found that rivers in semiarid environments can have riparian zones almost completely covered by grasses and deciduous trees.

It was generally assumed that forested areas provide greater bank stability because root systems prevent erosion along the banks (Hey and Thorne, 1986). Anderson et al. (2004) analyzed select datasets from 1978 - 2000 with specific regard to vegetation affecting downstream hydraulic geometry. The authors found that in forested areas, widths were smaller than their grassy counterparts for large watersheds. Similarly, Stott (1998) found that forested rivers in Scotland were less susceptible to erosion than their moorland counterparts because tree cover increased the mean temperature and lowered the incidence of frost. Copeland et al. (2005) analyzed 57 bankfull width relationships on sand bed rivers and found streams with greater than 50% tree cover had smaller widths than streams with less than 50% tree cover.

On the other hand, many studies have shown that afforestation is not necessarily correlated with increased bank stability. Hession et al. (2003) measured widths and depths of streams in forested and non-forested environments. They found that vegetation plays a major role in bankfull hydraulic geometry, where forested sections had much larger widths than non-forested sections. Eaton and Church (2007) found that relative bank strength declines in the downstream direction as the channel is incised deeper than the rooting depth of the vegetation. Murgatroyd and Ternan (1983) observed that afforestation practices in England have had large effects on the downstream and at-a-station geometry of rivers in the area. Channel widths within the forest were up to 3 times larger than those predicted from the non-forested catchment upstream. Sinuosity and depth correspondingly increased with the channel markedly showing signs of aggradation to accommodate large widths. Murgatroyd and Ternan (1983) attributed this large impact to the forest suppression of thick turf grass and their roots systems, as well as streams bypassing newly formed log and debris jams. Andrews (1984) calculated hydraulic geometry relationships for 24 rivers in Colorado. He found that rivers with thick bank vegetation had statistically equal exponents to rivers with thin bank vegetation, while the coefficients were slightly different (the coefficient for width was larger in rivers with thin vegetation while the coefficient for depth was larger in rivers with thick vegetation). He compared these results to those already done on British rivers and found the exponents to be the statistically similar.

Montgomery (1997) argued that a site specific approach is required when determining what vegetation is best suited to line the banks of a stream. He cites contradictory studies (as mentioned above) where some forested channels underwent greater erosion and others showed greater stability. Because so many variables are present, a single solution is not possible. On cohesive soil sites, forests may indeed increase bank erosion whereas in alluvial sites the root system would likely increase bank stability. It is clear that the type of vegetation present, and the soil conditions are critical in comparison of bank stability. Afforestion may increase bank stability in barren areas with cohesionless soil, whereas afforestion may decrease bank stability if it is replacing a grassy area composed of cohesive clays.

Attempts to account for vegetation influences have been factored in to theoretical analysis using index methods (Millar, 2005), while other studies have defined the boundary limits by types of sediments to account for soil cohesion. Eaton and Church (2007) accounted for bank strength in their theoretical downstream hydraulic geometry model by defining a coefficient of *influence of vegetation* and *friction strength* of the soil. This study serves to emphasize the importance of soil type and vegetation in bankfull geometry.

2.3.5 Regional Curves

Dunne and Leopold (1978) first introduced regional curves as a way of simplifying bankfull geometry prediction. They found that bankfull discharge and bankfull geometry characteristics are highly correlated to drainage area. They proposed regression equations for bankfull geometry as a power function of drainage area (Chaplin, 2005) and called them regional curves (Keaton et al., 2005). Literature suggests they have almost become interchangeable with hydraulic geometry equations (Doll et al., 2002). The author believes this is an error as hydraulic geometry relationships relate geometry variables to flow rate for all ranges of flow, whereas regional curves relate geometry variables at bankfull to drainage area. Relationships dependant on drainage area are less reliable than discharge because drainage area is a substitute for discharge. Drainage area is used because it is typically easier to determine than discharge (Johnson and Fecko, 2008). The relationship relies on the assumption that all streams within the region exhibit similar morphological characteristics (Johnson and Fecko, 2008). McCandless (2003) found bankfull geometries to be highly correlated with drainage areas in a physiographic province.

Recently there has been a shift in stream design and restoration to utilize natural channel shapes and processes rather than hardening channel surfaces to prevent degradation (Cinotto, 2003). In the 90's, regional curves became a popular tool in natural stream restoration and design (Chaplin, 2005). Regional curves are designed to assist with correct identification of bankfull flow and bankfull flow characteristics, which can often appear vague in the field (Brockman et al., 2012; Mulvihill and Baldigo, 2012). Regional curves have been more widely used than hydraulic geometry relationships for planning because they are regionalized and should have less variability (Mulvihill and Baldigo, 2012). Significant amounts of money are being invested in stream restoration. Between 1990 and 2003, $46 million dollars was invested in stream restoration between 24 states (Johnson and Fecko, 2008). With an increase in funding, there has also been an increase in the amount of literature being produced on the topic.

Dunne and Leopold attempted to build one set of regional geometry curves for the whole eastern USA, but the applicability of one set of curves to span multiple physiographic regions has not been verified (Cinotto, 2003). Study has been devoted into finding the best boundaries for delineating regional curves.

Doll et al. (2002) created regional curves for urban drainage areas in North Carolina and compared them with non-urban drainage areas of similar physiographic characteristics. They found a substantial increase in hydraulic geometry values for urban streams. Chaplin (2005) statistically analyzed physiography, and percentage of rock type in drainage area and concluded that carbonate bedrock has a large influence on regional curves. Regional curves developed from watersheds underlain by 30 percent or less carbonate bedrock generally had steeper slopes than the corresponding relations developed from watersheds underlain by greater than 30 percent carbonate bedrock. Brockman et al. (2012) analyzed 28 streams in two physiographic areas and recorded bankfull dimensions, discharge, slope, median bed material size, sinuosity, Rosgen stream classification and percent impervious area. They found percent impervious area had minimal effect on bankfull dimensions and attributed it to strong bank control due to bedrock, cohesive sediments and vegetation (Brockman et al., 2012).

Regional curves are typically delineated by boundaries based on climatic, physiographic and geologic factors (Keaton et al., 2005). Cinotto (2003) defines a physiographic province as "a region of which all parts are similar in geologic structure and climate and which consequently has a unified geomorphic history". Physiographic provinces are assumed to have homogeneous runoff characteristics (Chaplin, 2005) and have been found to be statistically dissimilar in development of regional curves (McCandless, 2003). Westergard et al. (2005) delineated physiographic regions based on physiography, geology, channel slope, basin storage, precipitation and snow cover. McCandless (2003) analyzed streams spanning several physiographic provinces in the state of Maryland, USA. The results showed that drainage area accounted for 99% of the variablity of bankfull discharge, which had a recurrence interval of 1.48 years.

Johnson and Fecko (2008) explored the limits of physiographic provinces for regional equations. Urban and rural watersheds could result in unique regional geometry relationships (Johnson and Fecko, 2008). Hydraulic geometry could also be affected by precipitation, number of road crossings, imperviousness, vegetation, and bank material (Johnson and Fecko, 2008). "While physiographic provinces provide appropriate boundaries at a large scale, it may be necessary to account for additional sources of equation differences in data collected for equation development"(Johnson and Fecko, 2008). Keaton et al. (2005) found more variability in regional curves for smaller streams for width and area but more variability for depth in larger streams.

The most thorough results in regional curve delineation to date come from the state of New York. Lumia (1991) and Lumia et al. (2006) delineated New York state into seven different hydrologic regions from statistical analysis based on physiography, flood frequency and climate. They found that "generally, the hydrologic-region boundaries coincide with drainage-basin divides". Mulvihill and Baldigo (2012) assessed the variability of regional curves in New York state based on four variables: hydrologic region , mean annual runoff, Rosgen Stream type and water surface slope. Six of the seven regional bankfull discharge curves had lower standard errors of estimation than the regional curve spanning all seven regions. Mean annual runoff proved to be a good delineator for regional curves while the Rosgen classification system and slope accounted for less of the variance. The study concluded that delineation through hydrologic region had the best results and most advantages(Mulvihill and Baldigo, 2012).

Other authors have considered a multivariate approach in regional curves. Though drainage area is considered the primary variable in a drainage basin, other effects such as land cover and bedrock lithology likely accounts for variation in regional curves (Keaton et al., 2005). Inclusion of percentage drainage area forested and percentage drainage area with carbonate bedrock variables resulted in a better coefficient of determination, leading to the conclusion that these two variables are correlated to bankfull variables. Other variables tested showed no correlations Keaton et al. (2005).

Though influences are muddled when hydraulic geometry curves are generalized into regional curves, they offer a simple solution to many practical issues. Research in this field has resulted in more knowledge on the many factors that can possibly affect hydraulic geometry as well as good examples of statistical practices for region delineation.

2.4 Conclusions

The first wide-scale hydrokinetic resource assessment in North America was done by UMA Group (1980) for Canada and was followed a few years later by Miller et al. (1986). Though mandated to a national scale, both studies at the time imposed criteria to limit the scale of the analysis due to the lack of computer automation that was available at the time. In the early 2000's GIS database were employed by the U.S. Department of Energy to conduct hydropower estimations for small hydro. This was preceded and followed by regional studies using similar methods. A national hydrokinetic study for the USA was completed in 2012 by Jacobson (2012). Though the use of GIS and computer has greatly improved the detail and scope of parameters that can be analyzed, many assumptions are still required, such as channel geometry shape, cross sectional velocity field and the calculation of backwater effects.

The great deal of literature on rivers and their hydraulic geometry testifies to the important and complex nature of the topic. Since the late nineteenth century many theoretical and empirical solutions have been proposed describing the mechanics of a river system and the hydraulic geometry conditions they result in. Because there are many variables in a river system (width, depth, velocity, slope, roughness, sediment transport) and fewer equations, many theories have been proposed in order to solve the system. No theory has been able to fully explain or predict all of the hydraulic relations and work remains in testing these theories among a common dataset. Leopold and Maddock's paper 1953 is the most simple and accurate empirical representation of hydraulic geometry and remains the benchmark contribution. Though downstream hydraulic geometry applies to all flow frequencies, the trend has been to use bankfull flow as it is often easiest to determine by identification in the field or equating it to effective flow or a 1.5 year recurrence interval. Many researchers have identified bank stability to be a large influence in hydraulic geometry, which can be mainly attributed to the soil composition of the bank and riparian vegetation along a stream. These variables tend to be categorized (as opposed to quantified) in models because quantifiable field measurement often proves difficult and isolating their influence in a model amongst other variables (such as bed load influence) is extremely challenging. Finally, regional curves, a simplification of hydraulic geometry has been given greater emphasis over the last two decades because of its practical use in stream restoration. Regional curves use drainage area as the independent variable in predicting bankfull geometries. Study in this field has resulted in greater knowledge of the extent of influences in a system, as well as improved methods for regionalization. Much has been learned about hydraulic geometry in the past century. Natural river systems, however, interact in such a complex manner in many varying environments that identifying, let alone quantifying their influence remains a difficult task. The greatest potential for future work currently lies in testing the validity of theories, improved regional identification and increased localization of regional curves.

Chapter 3

Hydrokinetic Estimation

The process used to predict hydrokinetic energy in Canadian rivers is a multi-step process developed by research and development in all three phases of the NRCan study. The first section of this chapter describes the data and techniques employed in the final methodology, including flow prediction techniques (ie. regionalization), the incorporation of measured flow data to improve flow estimates, generation and application of downstream hydraulic geometry, and predicting the uncertainties in both the flow and power estimates and propagating those uncertainties into the final data product. The second section presents the results of these processes.

3.1 Methodology

3.1.1 Data sources

The flow and geometry predictions required to predict hydrokinetic power at ungauged sites required numerous datasets. An ungauged site is defined as a cross section of a stream where flow measurements are not regularly recorded. The large scope of the project required the datasets to be complete and consistent nationwide. A summary of the datasets used is shown in Table 3.1. Detailed descriptions of each the datasets employed in this study follows.

HYDAT

Environment Canada (EC) maintains HYDAT, an archival database that contains all water information collected through the National Hydrometric Program. These data include: daily and monthly means of flow, water levels and sediment concentrations for over 2300 active and 5500 discontinued hydrometric monitoring stations across Canada. The data are available online, and also through the Environment Canada Data Explorer (ECDE) desktop application (Environment Canada, 2010). In this phase of the study HYDAT data were used to calculate historical flow characterizations at gauged sites (see section 3.1.1 for flow characterization details).

Water Survey of Canada - Measurement Database (WSC-MD)

The Water Survey of Canada Measurement Database (WSC-MD) is an MS-Access database developed for internal use by Environment Canada and is a repository for hydrometric field measurement data that were obtained from participating regions of WSC. Data include cross-sectional information on channel geometry and velocity for conducted surveys up to 1997. These datasets were used to fully characterize the hydrokinetic power at each stream gauge location. Data availability for the WSC-MD varied across the country and currently no data for the province of Quebec are found in the WSC-MD. The database contains 1861 of the stations found in the HYDAT database. Only stations with 5 or more flow measurements were employed, which reduced the number of viable stations for hydrokinetic assessment to 1174.

The WSC-MD was employed to create at-a-station hydraulic geometry curves and subsequent regional downstream hydraulic geometry curves (described in Chapter 4). Area and width curves were of primary use as all other critical values could be calculated from these

Table 3.1: Summary of datasets used

Datset	Publisher	Format	Use	Citation
Hydat (Hydrometric Database)	Environment Canada	Microsoft Access database	Flow Measurements	Environment Canada, 2010
WSC-MD (Water Survey of Canada Measurement Database)	Environment Canada	Microsoft Access database	Geometry Measurements	Not published
Digital Stream Network	Natural Resources Canada	Vector Shapefiles	Stream Connectivity	Natural Resources Canada, 2008
Voronoi Catchments	In-house, based on Digital Stream Network	Polygon Shapefiles (0.0004 degree resolution)	Stream Watersheds	Not published
CDED (Canadian Digital Elevation Data)	CTI - Natural Resources Canada	Raster (0.0001 degree resolution)	Watershed Physiographic attributes	CTI, 2000
AAFC SLC (Soil Landscapes Canada)	Agricultre and Agri-Food Canada	Polygon Shapefiles	Watershed Soil, Climatic attributes	AAFC, 2000; AAFC 2008
CTI Landcover (Centre for Topographical Information)	CTI - Natural Resources Canada	Polygon Shapefiles	Watershed Landcover attributes	CTI, 2009
Surficial Geology	Earth Sciences Information Centre	Polygon Shapefiles	Watershed Geology attributes	Turner et al., 2003

relationships.

Digital stream network

The National Atlas of Canada provides a vectorized representation of streams and lakes in Canada in ESRI shapefile format at a 1:1 million scale (Natural Resources Canada, 2008). Approximately 365,000 Canadian river segments are contained in the dataset and have associated drainage areas as small as 0.1 km^2. The data are divided by the major drainage basins in Canada (see Figure 3.1). The stream network has complete connectivity and contain critical attributes used in this study. These attributes are:

- Unique identifiers for each stream segment, called the SkelUID;
- Upstream and downstream connecting nodes for each stream or lake segment; and
- An identifier delineated whether a segment is a river, lake, coastline or inland border.

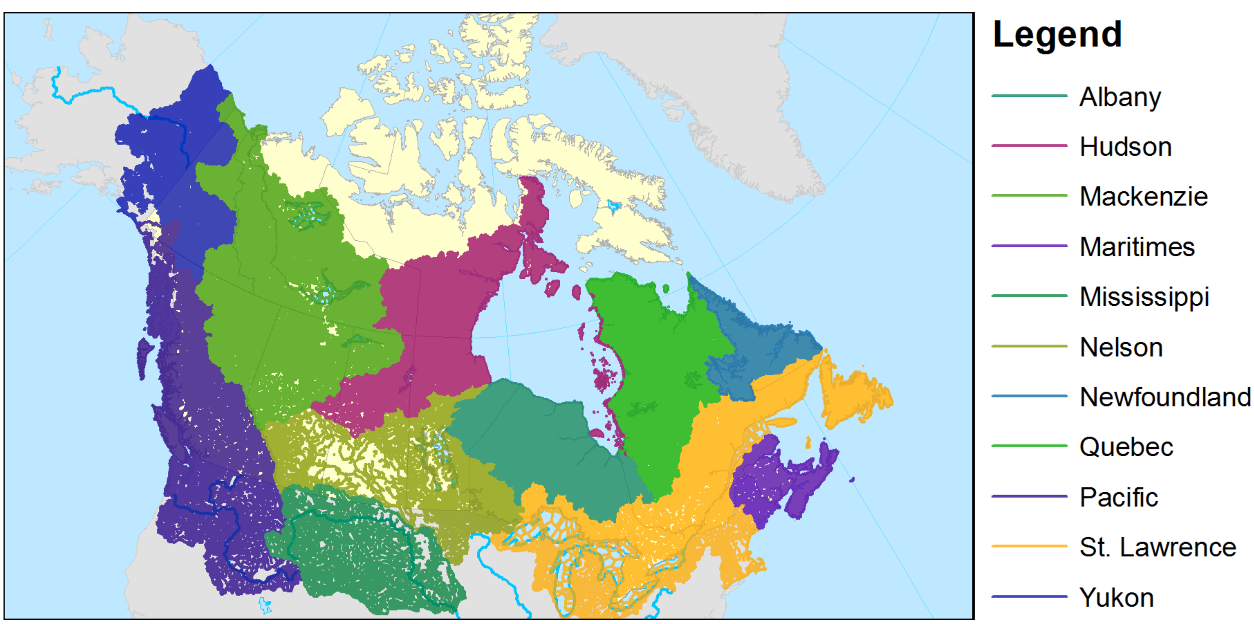

Figure 3.1: National Atlas of Canada - stream segments by major drainage basin.

Voronoi-delineated Catchments

It was identified early in the study that delineating each segment's upstream drainage area, or catchment, from digital elevation models was infeasible due to the high computational expense and data limitations. The best alternative for watershed delineation was the incorporation of a Voronoi Polygon dataset. A Voronoi dataset bases its watershed delineations on distance to nearest stream segments; each stream segment will have one corresponding Voronoi polygon. Though this approach isn't as accurate as DEM delineation because it doesn't account for terrain elevation, it is much less computationally expensive. An example of the differences between the techniques is shown in Figure 3.2.

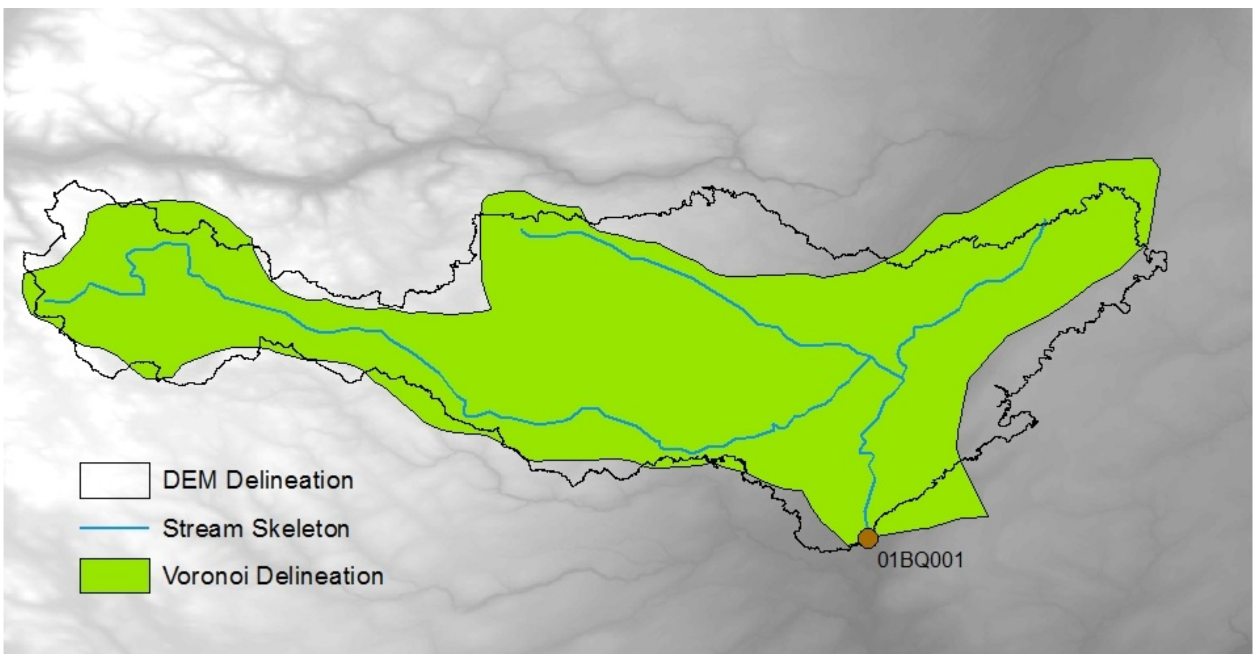

Figure 3.2: Comparison of a DEM and Voronoi Delineated Catchment

NRCan has created a Voronoi dataset based on the National Atlas water network [23]. This dataset provided a set of drainage network segments and associated contributing drainage areas. However, further investigation found that some Voronoi catchments had no corresponding stream while others had multiple corresponding streams. The discrepancy accounted for 4.86% of the total area in Canada and was equally distributed. It was concluded that the dataset could not be used while these discrepancies existed. A new Voronoi polygon dataset was created using the existing stream skeleton network.

A new database of catchments was created based on Voronoi tessellation of the vectorized stream segments using ESRI ArcGIS (ESRI, 2010). This database allowed the authors to associate each stream segment with an upstream drainage area that was calculated with relatively small computational expense. The polygons were created on a regional basis, as defined by major drainage basin (refer to Figure 3.1). Coastal boundaries were included to account for surface areas draining directly to the ocean. Figure 3.3 depicts an example of Voronoi polygons bordering the ocean whose contributing areas feed the ocean instead of a stream. Regional land boundaries were only used to define the region and do not influence the shape of the polygons. Though this methodology works well in the majority of cases, it is limited by variation in the regional boundary. Figure 3.4 shows an example of a Voronoi polygon cut off by a regional boundary. The orange polygons are defined as two distinct polygons with the same corresponding SkelUID. Both polygons will contribute to the watershed area, however, only the larger perimeter will represent the perimeter of the watershed. This influence on flow patterns is minimized because it only occurs at the headwaters of a stream.

In total, approximately 323,000 drainage areas were delineated across 11 regions at a resolution of 0.0004° (between approximately 16-44 m, depending upon geographic location). Each polygon was assigned a SkelUID to reference it to its corresponding stream segment in the National Atlas water network. The delineated drainage areas for the Maritimes region is shown in Figure 3.5.

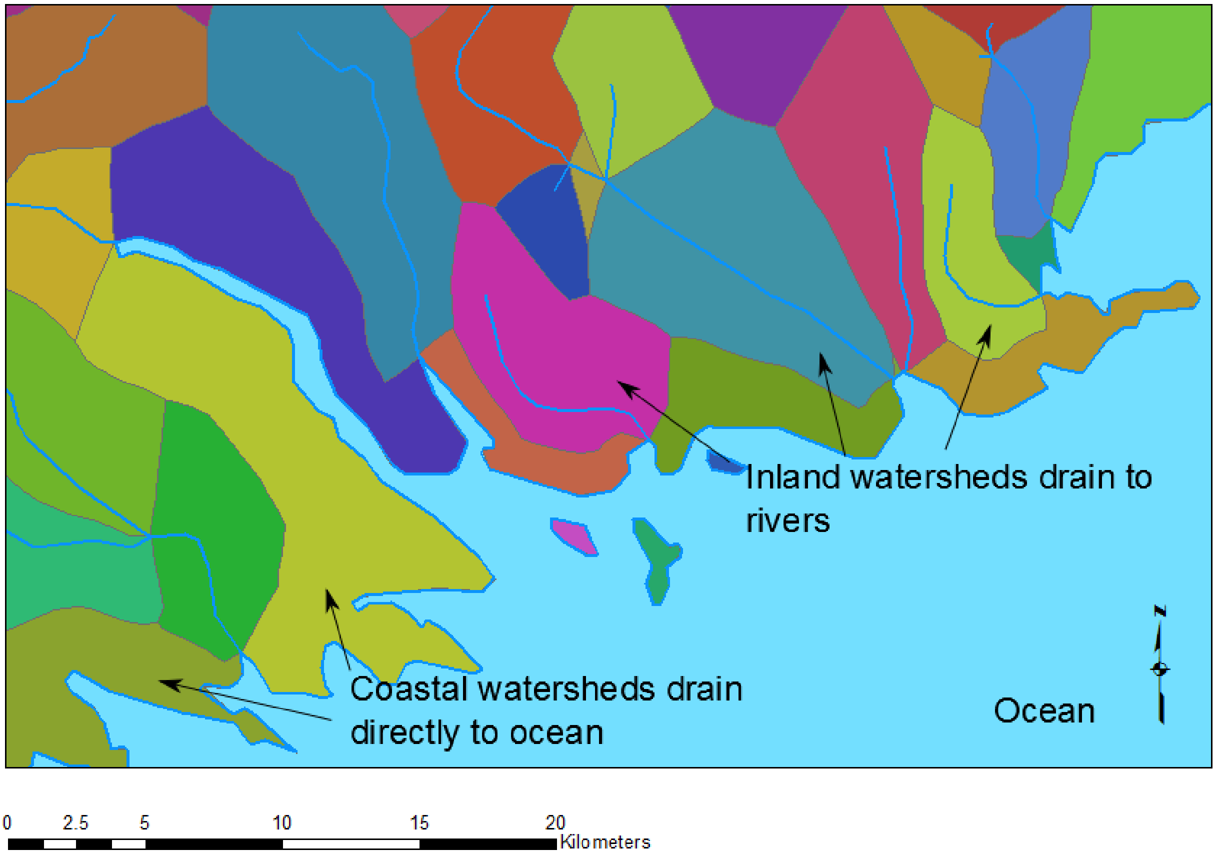

Figure 3.3: Voronoi Delineated Catchments along Coastline

Canadian digital elevation data

The Canadian Digital Elevation Data (CDED) consists of an array of ground elevations or digital elevation model (DEM) extracted from the National Topographic Database and other data sources from the provinces and territories (Centre for Topographic Information (CTI), 2000). The geographic resolution is 0.0001 decimal degrees (maximum 93 m pixel resolution). CDED was used to delineate watersheds for building the calibration dataset and to map elevation attributes of stream catchment databases.

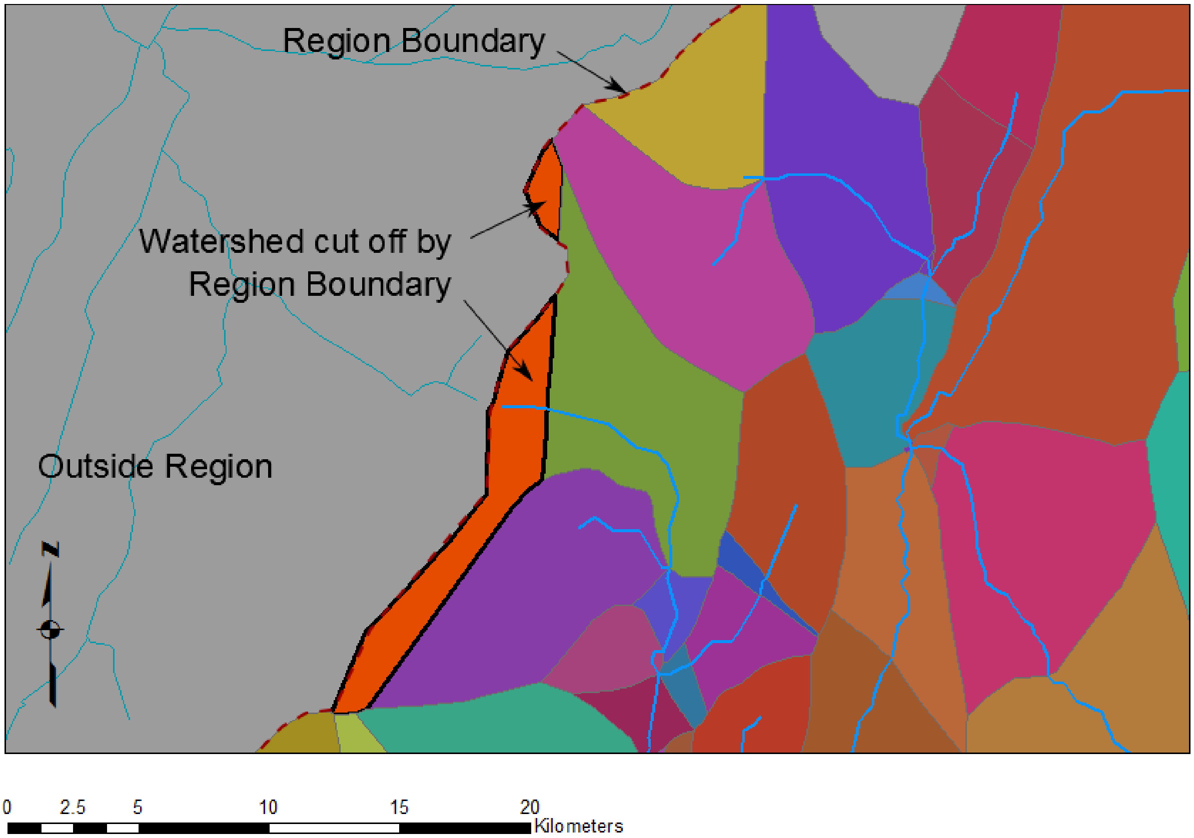

Figure 3.4: Voronoi Delineated Catchment along Regional Boundary

Physiographic and climatic datasets

Soil characteristics were extracted from datasets provided by Agriculture and Agri-Food Canada (AAFC). The AAFC Soil Landscapes of Canada Working Group and the Canadian Soil Information System (CanSIS) created a series of GIS coverage maps that show the major characteristics of soil and land for the entire country. These data were compiled at a scale of 1:1 million [1], and information is organized according to a uniform national set of soil and landscape criteria based on permanent natural attributes (AAFC, 2000) including hydraulic conductivity, drainage and soil water retention. A complete description of the SLC dataset can be found in Appendix B.1.

[1] because the dataset is in polygon format, no raster information is given

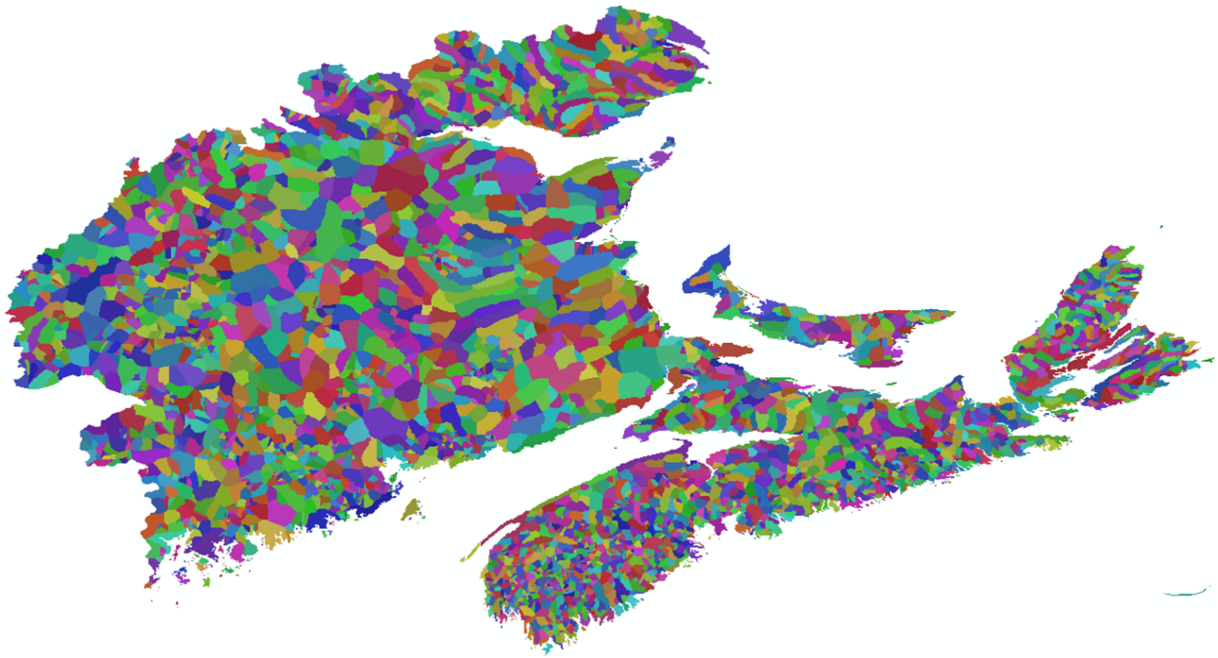

Figure 3.5: Voronoi Delineated Catchments in the Maritimes Region

Land Cover data were obtained from Natural Resources Canada via Geobase (Centre for Topographic Information (CTI), 2009). The vectorized GIS dataset was developed using amalgamated data from classified Landsat ortho-images from the Earth Observation and for Sustainable Development project, the National Land and Water Information Service, and the Canadian Centre of Remote Sensing. Land use variables were classified into 45 different land cover categories. For the purposes of this study, these categories were generalized into 10 unique land classes. See Table 3.3 for the generalized land classes employed in this analysis.

Surficial Geology data was obtained from Earth Sciences Information Centre via GeoGratis and called Geoscape Canada (Turner et al., 2003)[2]. The map contains complete coverage of Canadian surficial geology separated into 26 classes. Each class is also identified

[2]no resolution was provided in the metadata

by a more generic classification based on origin and grain size.

AAFC also provides climate data in GIS format as part of the National Ecological Framework (Agriculture and Agri-Food Canada (AAFC), 2008). Temperature data (annual minimum, annual maximum and annual mean) and precipitation data (annual rainfall, annual snowfall and total precipitation) were interpolated from weather station station from 1961-1990 and attributed to delineated ecodistricts.

The physiographic and climate data described above were stored as attributes. They were for individual stream catchment drainage areas in the same manner that attributes were mapped to the drainage areas in the calibration dataset (described in section 3.1.1. This dataset was later employed in the regression and regionalization analyses for stream flow prediction.

Calibration dataset

A calibration dataset was created to provide a basis for flow regionalization techniques. The dataset consisted of flow characterizations calculated from the HYDAT database and watershed characterizations mapped from the datasets described previously.

Flow was characterized by a 17-point flow duration curve (FDC), mean monthly flows (MMF) and the mean annual flow values (MAF). The FDC is typically employed to describe the flow frequency at a location and provide a graphical representation of stream flow variability (Maidment, 1993; Searcy, 1959b). It is useful in a power prediction study for finding the range of flows and percentage of time they occur where energy can be effectively extracted (ie. it is likely not feasible to extract energy during extreme high or low events). The FDC was divided into 17 percentile flows (shown in Table 3.2) based on studies employed by Shu and Ouarda (2009) and Mohamoud (2008) who found that these points adequately characterize a FDC while making no prior assumptions as to the shape it should take. MMF and MAF were also used to characterize the flows at a station or prediction location. MMF in particular provides a good indication of the seasonal flow patterns which are not preserved in the FDC.

Watershed characteristics were mapped by overlaying delineated boundaries over attribute datasets and extracting the percentage of the watershed area that was covered by each attribute. This was completed using GreenKenue hydrological analysis software (NRC-CHC, 2010). Initially, watershed boundaries were obtained using DEM delineation in GreenKenue on HYDAT stations that were included in the WSC-MD. The dataset was later expanded to include HYDAT stations whose watershed boundaries had been delineated by Environment Canada. These watershed boundary data products are currently being internally reviewed at Environment Canada and are yet to be released publicly (Erika Klyszejko, Environment Canada, personal communication). The mapped attributes are shown in Table 3.3.

FDC Index	**Probability of Exceedence (%)**
1	0.01
2	0.1
3	0.5
4	1
5	5
6	10
7	20
8	30
9	40
10	50
11	60
12	70
13	80
14	90
15	95
16	99
17	99.99

Table 3.2: FDC Exceedence Probability Index

Parameter	**Data Source**
Basin area	Centre for Topographic Information (CTI) (2000)
Basin perimeter	
Basin centroid latitude	
Basin centroid longitude	
Basin mean elevation	
Reach slope [3]	
Reach length	Natural Resources Canada (2008)
Fraction of land cover (10 classes)	Centre for Topographic Information (CTI) (2009)
Fraction of surficial geology (9 classes)	Turner et al. (2003)
Drainage index	AAFC (2000)
Saturated hydraulic conductivity	
Soil water retention	
Annual rainfall	Agriculture and Agri-Food Canada (AAFC) (2008)
Annual snowfall	
Potential evapotranspiration	
Growing season length	
Growing degree day	

Table 3.3: Attributes mapped to watersheds

Station selection was achieved by filtering all possible HYDAT stations to only include:

- Stations having 10+ years data,
- Stations having recent observations (1995 or later), and
- Stations that were unregulated

[3]Reach slope is only available for the calibration dataset, not the Phase III analysis dataset.

Due to the limited watershed delineations available and the time consuming nature of delineating new watersheds, it was not possible to include all of the HYDAT stations that met the aforementioned criteria. A total of 889 stations were collected and analyzed that represent every region and province in Canada, save Quebec. At the time of the study no Quebec data were available in the WSC-MD. Figure 3.6 shows a map of the HYDAT stations included in the calibration dataset.

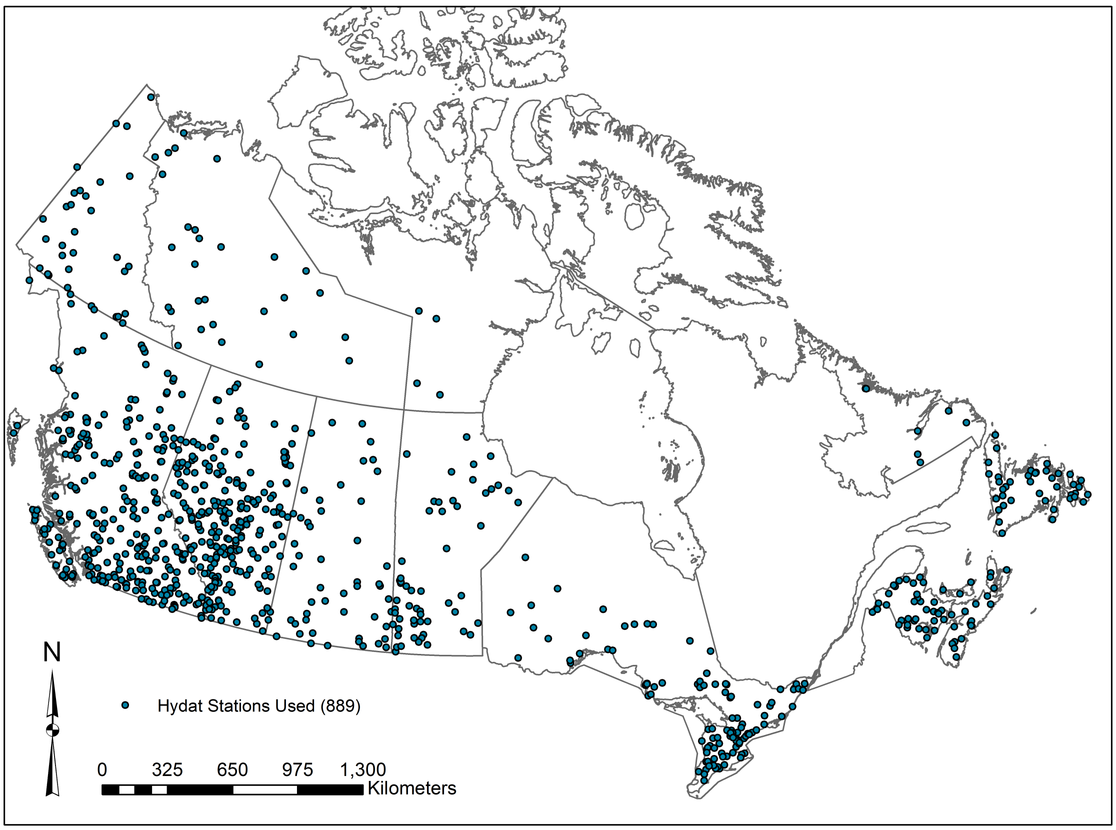

Figure 3.6: HYDAT stations used in calibration dataset

3.1.2 Flow prediction

Flows were predicted for each reach on every river in Canada according to the Digital Stream Network described in Section 3.1.1. The prediction was calculated using multiple linear regression coupled with canonical correlation analysis (MLR-CCA). Though the 17-point FDC and 12 MMFs were predicted using the method explained below, for simplification purposes, only the process to predict the FDC flows is outlined.

All of the calculations and processes described in the following sections were done using R Statistical Programing Language. R is a free software programming environment that lends itself very well to data processing and graphics. The R scripts were written by the author and can be found in the Appendix. The CCA portion of the code was translated from Matlab code originally written by Dr. Ouarda.

Stepwise regression was first used to determine the most statistically significant parametric parameters across the whole range of flow frequencies. Parameters that were significant in 5 or more of the 17 flow frequencies were determined to be key parameters and used in the canonical correlation analysis (CCA) to determine hydrologically similar stations.

CCA is a multivariate statistical technique that permits the establishment of interrelations between two groups of variables by determining linear combinations of one group that are most correlated to linear combinations of the second group (Clark, 1975). The technique has been employed as a regionalization method in flood frequency analysis, where one group of variables represents the flood characteristics and the second represents the physical and climatological characteristics of the basin, the principle being that by knowing the second the first can be predicted (Bobee et al., 1996). Regions can then be defined by examining the proximity of a gauge to other gauges in canonical space. The contributing neighbourhood size is defined using a mahalanobis distance which is set using a chi-squared distribution based on a user defined α. The α value is calibrated to find a good balance between having higher confidence in the prediction and having a large enough neighbourhood size to perform multiple regression. The procedure has been employed for flood frequency analysis in Quebec (Ribeiro-Correa et al., 1995) and Ontario (Ouarda et al., 2001b).

In this study, CCA was used to determine the stations that were the most similar in hydrological response to the target stream reach for which characteristic flows were being predicted. Considering the limited number of predictor variables, only a small number of flow variables could be included to develop the CCA analysis. Of the 17-points included in the FDC the indices 2,3,4,5,8, and 10 for the CCA analysis (representing the probabilities from 0.1% to 50%). This selection of indices was chosen to best match the higher flows in the FDC prediction, accounting for the fact that larger flows have a much greater contribution to the total hydrokinetic power calculations.

The α value was defined on a regional basis (by major drainage basin) and was determined by a balance between the minimization of errors and maximizing the number of stream reaches that could maintain a flow prediction. If α is small, every station in the calibration dataset is used in the MLR analysis to predict flow (ie. there are many neighbouring stations). As α is increased, the number of identified "neighbouring" stations to a river reach is reduced and the subsequent level of confidence of the MLR prediction is generally increased. However, in some cases a high α leads to river reaches not having enough neighbouring stations required to execute an MLR analysis, leaving river reaches without available flow predictions. The CCA-MLR analysis was run for a range of α values (15 values between 0 and 0.75) for each region and the flow predictions were compared with measured flows of nearby HYDAT stations. Average RMSE statistics were computed for each region and compared with the percentage of total reaches which had enough neighbouring stations to perform an MR analysis. The α values for each region were chosen to optimize the two parameters and are shown in Table 3.4.

Once the best neighbouring stations were found, each of the flow estimates (FDC, MMF, MAF) were predicted with multiple linear regression using the best statistical variables for

Region	Calibrated α value
Albany	0.0050
Hudson	0.0001
Mackenzie	0.0001
Maritimes	0.0050
Mississippi	0.0000
Nelson	0.0075
Newfoundland	0.0001
Quebec	0.0010
Pacific	0.0010
St. Lawrence	0.0010
Yukon	0.0001

Table 3.4: Regional calibrated neighbourhood definitions α

each flow frequency, as determined from analysis described above.

In order to preserve the uncertainty associated with the flow estimates the 95% prediction interval was recorded along with the flow estimate. This prediction interval bounds the data containing approximately 95% of the samples. Because a prediction interval accounts for inherent noise in the population, its bounds are always wider than a confidence interval for a regression analysis (Helsel and Hirsch, 1991).

3.1.3 Influence of neighbourhood definition through CCA

The positive influence of the inclusion of CCA as a regionalization technique can be seen through the evaluation of the change in the size of the prediction intervals of the predicted flow values. To compare the prediction intervals among the flow estimates a relative prediction term was employed:

$$C_R = \frac{y_{0.975} - y_{0.025}}{\hat{y}} \tag{3.1}$$

where C_R is the "relative confidence" term which is the 95% prediction interval ($y_{0.975} - y_{0.025}$)normalized to the expected prediction value $\hat{y}$. A small relative confidence interval shows that the regression technique is working correctly (given a large enough sample size). While it is possible that a regression with small relative confidence interval will give a prediction with a large error, this is unlikely. However, it should be noted that a confidence analysis is not a validation. The predictions are validated against observed data in section 5.3. Figures 3.7, 3.8 and 3.9 show maps of prediction intervals for FDC50 flow predictions. To illustrate the influence of regionalization on the analysis one sees a general reduction in the prediction intervals around an estimate with a reduction in the neighbourhood populations. This is illustrated in Figure 3.7, which displays very large prediction intervals, indicating low confidence in results when no regionalization method is used ($\alpha = 0$). By contrast Figures 3.8 and 3.9 shows smaller prediction intervals when using CCA as the α value is increased. However, as the α value is increased larger areas of the stream network emerge with "no estimate", as seen noticeably in areas of southern Manitoba and in the north in Figure 3.9. This highlights the necessary trade-off between improved prediction values, while still developing a comprehensive national dataset with estimates at every stream reach, as required by this study's objectives.

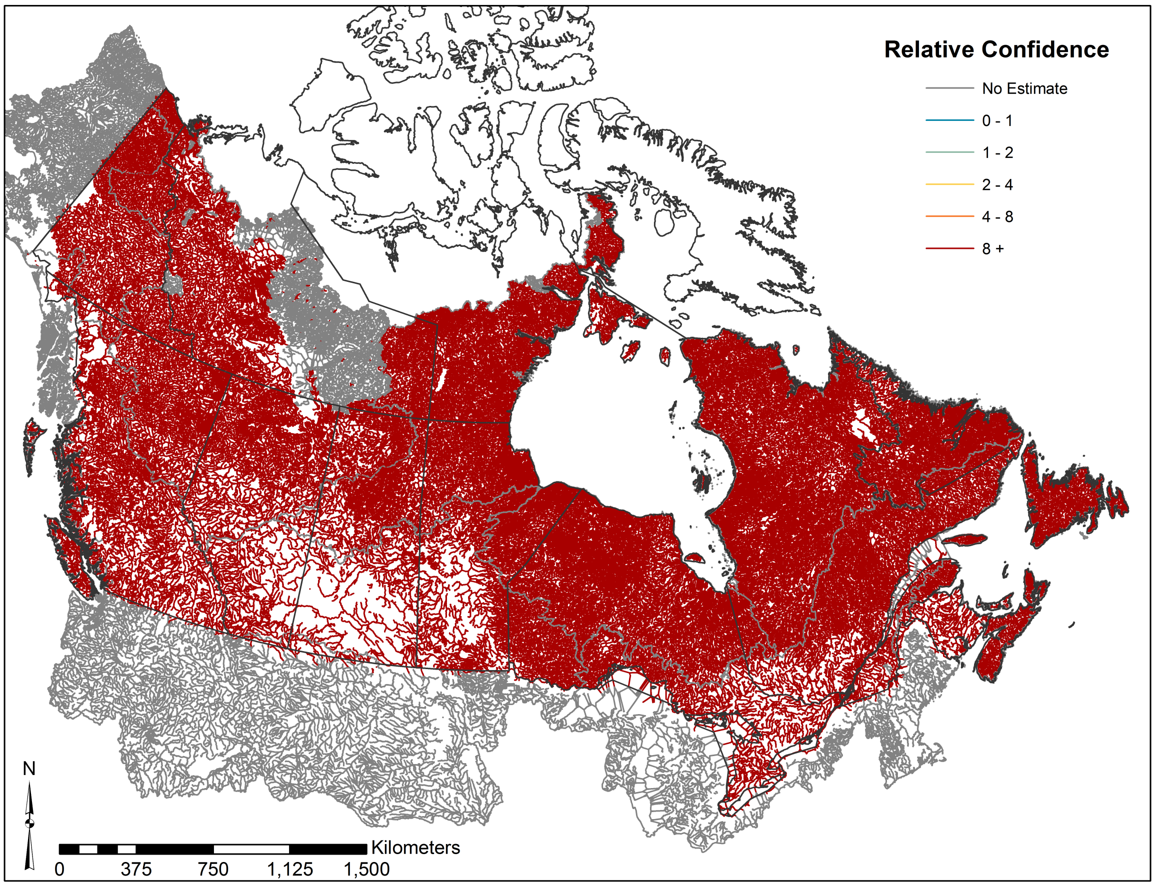

Figure 3.7: Estimate confidence with no regionalization, FDC-50

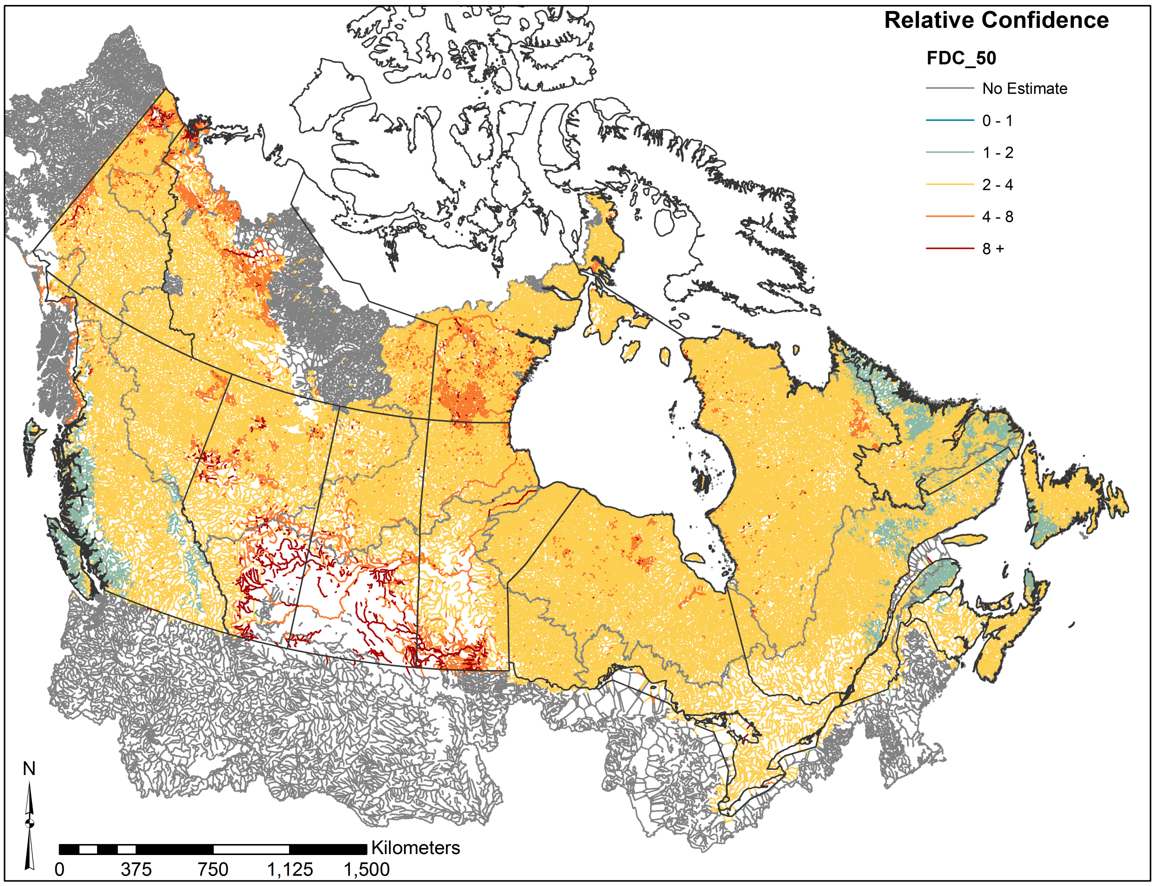

Figure 3.8: Estimate confidence with CCA regionalization ($\alpha = 0.0001$), FDC-50

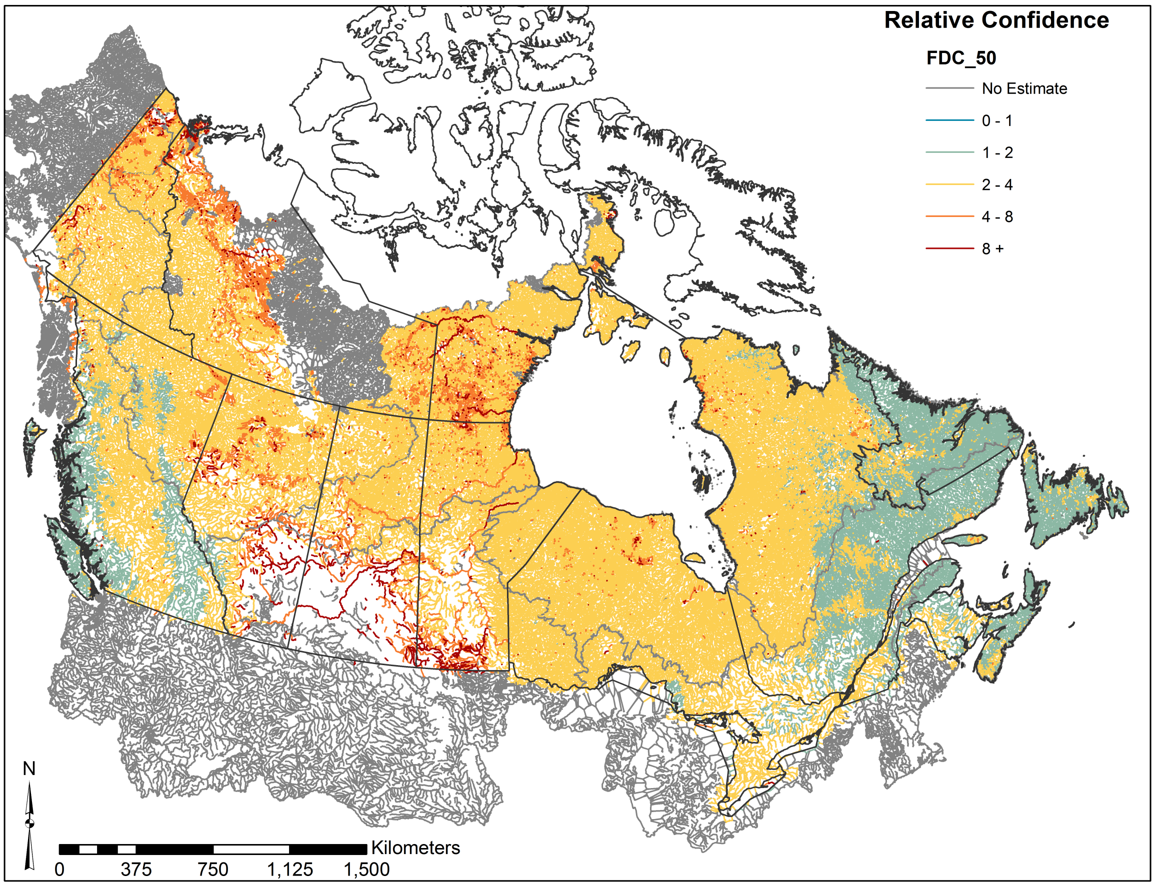

Figure 3.9: Estimate confidence with CCA regionalization ($\alpha = 0.0075$), FDC-50

3.1.4 Augmenting predicted flows with measured flows

The MLR-CCA technique is designed to produce flow estimates across varying landscapes assuming that upstream area of the stream being analyzed is naturally flowing or unregulated, and has adequate physiographic and climatological data available in the upstream watershed to conduct the MLR analysis. In order to estimate flows, and subsequently hydrokinetic power, for a national dataset of river basins as illustrated in Figure 3.1, two challenges emerged:

1. Many of the river basins are international, and the physiographic and climatological basin data outside of Canada were not collected and harmonized with the Canadian data for this study; and

2. Many rivers in Canada are partially or completely regulated by dams or other control structures, invalidating the natural flow requirement for the flow prediction methodology used in validation.

An example of the border issues is illustrated in Figure 3.10, showing a portion of the Columbia River basin. In this figure the shaded areas show watershed that have physiographic and climatological data within Canada so that a prediction can be performed. However, once the river travels into the Montana from British Columbia, it is no longer possible to employ the MLR-CCA methodology alone: the upstream areas are not fully represented from a data perspective and flow cannot be predicted. In this instance the basins with any US headwaters would not be evaluated using the direct MLR-CCA method.

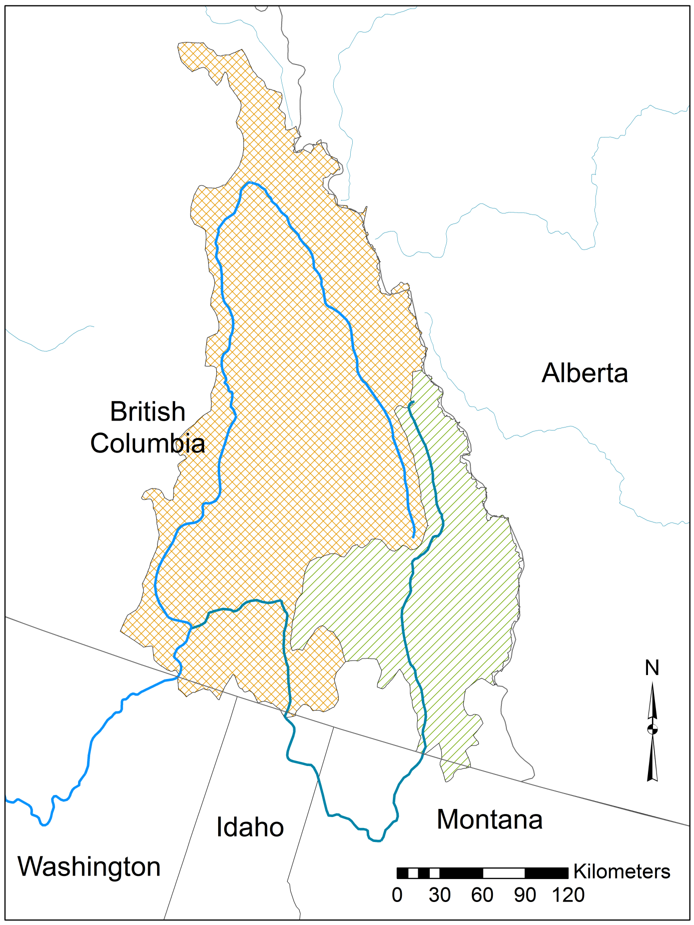

Figure 3.10: Example of watershed data along Canada-USA border

These challenges were overcome by augmenting predicted flows with observed HYDAT data where the data existed; if a HYDAT stream gauge was identified upstream of the stream reach being predicted, the drainage area and corresponding flow contributions were removed from the multiple regression analysis, providing a prediction only on the incremental upstream area not including the drainage area accounted for by the stream gauge. Then the measured flow from the HYDAT stream gauge was then added to the incremental prediction.

This solution addressed both of the aforementioned issues: flow could be predicted on rivers with drainage area partially in the US, at least downstream from the location of the first HYDAT stream gauge, and actual flow data could account for regulation in the rivers as the measured flow data would accurately characterize the regulated flow. An added benefit from this approach is that the relative confidence of the prediction was increased, as the uncertainty in the predictions would be mitigated through the addition of measured flows where available. Note that aggregating FDC curves from multiple watersheds carries the assumption that the watersheds from which the FDC curves were derived are hydrologically similar. While this may not be the case in all circumstances, it is believed that this assumption will have little impact on the results.

HYDAT stations were only used to augment the flow predictions if they contained five or more years of continuous daily data since 1990, and only the flow data from 1990 onward were employed. These criteria ensured that historical flows did not contain both regulated and non-regulated regimes. This requirement allowed for 1400 HYDAT flow stations to augment flows in the analysis (from a total of approximately 6500).

Discrepancies were also found with the connectivity of National Scale Frameworks Hydrology data (NSFH). As a quality control check, the accumulated areas on the network were compared with the published HYDAT drainage areas. In regions with complicated drainage networks or sites where diversions have been constructed, river reaches have been found that do not match the reported HYDAT drainage areas. The connectivity could only be validated where HYDAT drainage areas existed. Where a discrepancy was found, the NSFH network was manually adjusted or the HYDAT flows were not employed to supplement predicted flows.

3.1.5 Propagation of uncertainty in flow estimates

In order to characterize the uncertainty in the estimates of hydrokinetic power in each stream reach as well as nationally, it was necessary to propagate the uncertainty associated with the flow estimates downstream. The approach taken here was to employ a Monte Carlo technique and represent the uncertainty in each estimate by a distribution of flows that match the distribution of the uncertainty in the estimate. For each regression equation developed for the stream reaches, providing an estimate of a predicted flow (MMF, FDC, MAF), 1000 flows were generated that matched the normal distribution defined by the regression prediction interval. As flows were accumulated downstream the measured HYDAT flows were combined with the incremental predicted flows. It was assumed that the HYDAT flows had no associated uncertainty and the 1000 samples were given a delta distribution (standard deviation of zero). To perform the uncertainty analysis, 30 flows (MAF, MMF, FDC) for each of the 365,000 river segments contained 1000 samples in order to capture the flow prediction uncertainty distributions, 10.95 billion data points in total.

The difference between confidence and prediction intervals is explained in Figure 3.11. The confidence interval band is much narrower than the prediction interval band. The confidence interval tells the reader where they can expect the mean of the population to occur 95 % of the time. In other words, it assumes the data in the graph is sampled from a larger distribution, and the mean of that larger distribution is likely to be found in the confidence interval 95 times out of 100. The prediction interval is much wider because it tells the reader where they can expect to see the next sampled point. Again, it assumes that the points are a sample from a larger dataset and that they are distributed around the mean. If a single sample were to be taken from this larger dataset, it would fall within this band 95 times out of 100. In this study the unknown parameters are being predicted based on a sample dataset (the calibration dataset), thus the prediction interval in all the regression equations was used to quantify the uncertainty. The multiple regressions for flow and geometry assumed a log-normal distribution of values, but it could not be assumed that these distributions remained log-normal when they were added together (see the section on augment predicted flows with measured flows, section 3.1.4). It was for this reason that 1000 samples of these distributions were taken at each step, and these sample distributions were carried throughout the process. It is seen as a brute force method because it requires a lot of computational power with few assumptions from the user.

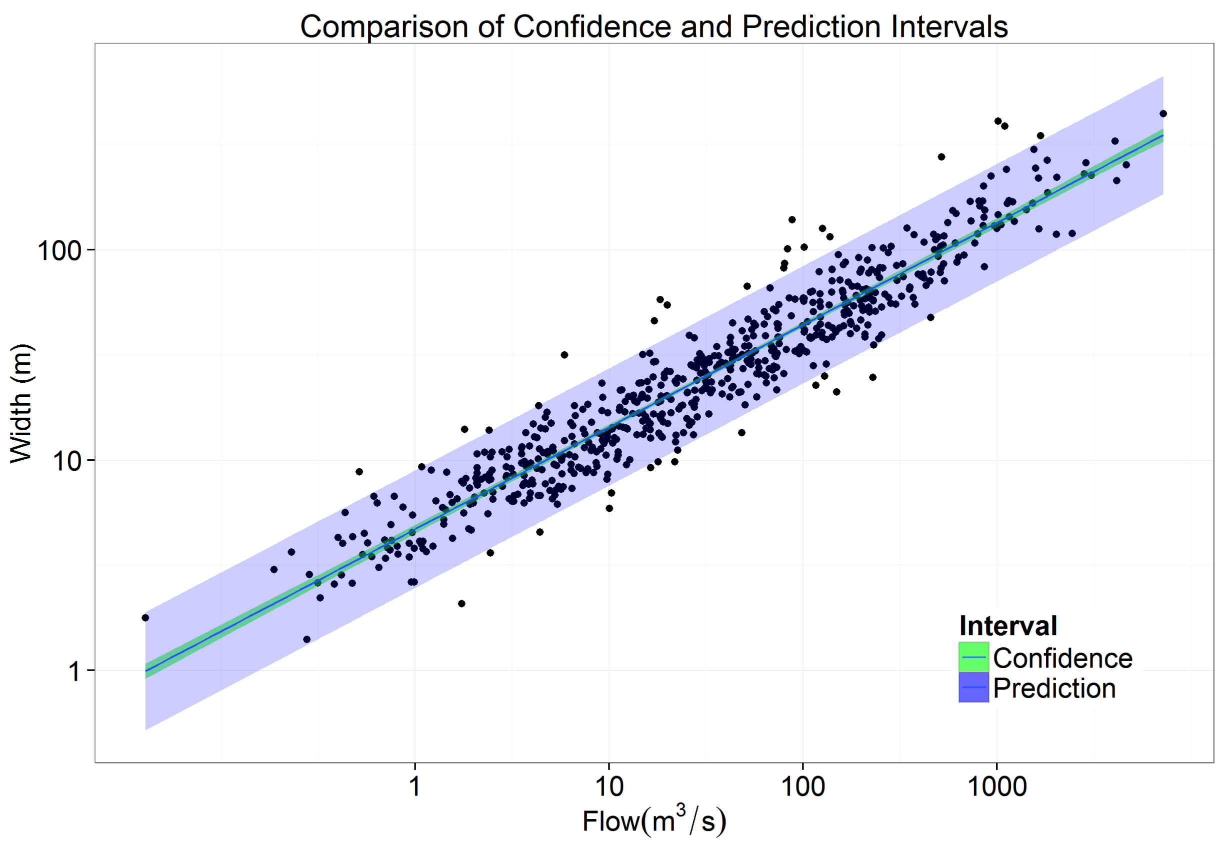

Figure 3.11: Comparison of Confidence and Prediction Intervals

The effect of including HYDAT flows and propagating the uncertainty in the estimates is illustrated in Figure 3.12. When compared to Figure 3.9 it can be seen that the relative prediction intervals are reduced for many of the larger rivers. This is due to the inclusion of the measured flows upstream of these locations which improves the flow estimate and reduces the prediction uncertainty.

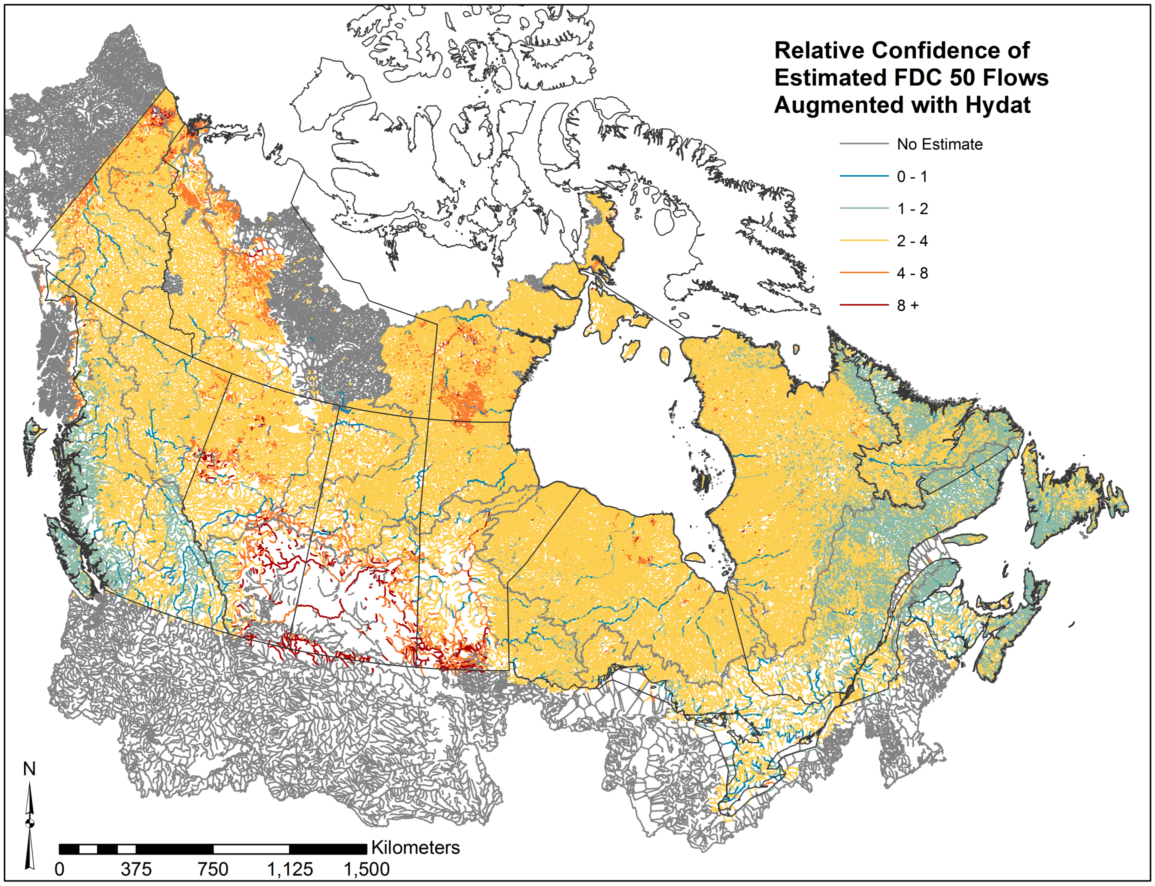

Figure 3.12: Estimate confidence of FDC50 flows including HYDAT flow data

3.1.6 Predicting power at a section from flow estimates

The equation for hydrokinetic power at a given cross section is:

$$P_{KS} = \frac{1}{2}\rho A v^3 \tag{3.2}$$

where P_{KS} is the hydrokinetic power at the cross section, ρ is the density of water, A is the cross sectional area and v is the average stream velocity. For reliable hydrokinetic power, estimates of both the velocity and cross-sectional area of a channel must be known. In some cases, the cross sectional area-averaged power, or cross-sectional specific power (P_{SS}), is useful in evaluating the hydrokinetic potential of a river reach. This is accomplished by dividing equation 3.2 by the cross sectional area:

$$P_{SS} = \frac{P_K S}{A} = \frac{1}{2}\rho v^3 \tag{3.3}$$

Velocity can be computed using either the Manning equation or from continuity relationships. For this study, the dearth of available slope or roughness estimates for the national river network precluded the use of the Manning equation which requires estimates of both (Jenkinson and Bomhof, 2012). In order to develop continuity relationships that could be regionalized, widths and areas were estimated at each flow frequency using downstream hydraulic geometry equations. Downstream hydraulic geometry was proposed by Leopold and Maddock (1953) who demonstrated an empirical log-linear relationship between channel geometric characteristics (width, depth, velocity, slope) and flow for specific flow frequencies. In this study, "at-a-station" geometry and flow relationships at 428 stations were extracted from the WSC-MD described in Section 3.1.1. With the at-a-station relationships defined, the regionalization of the channel geometries were then conducted using traditional downstream geometry relationships as developed by Leopold and Maddock (1953) for each defined flow frequency. The complete procedure for developing the channel geometric relationships is described in Chapter 4.

With geometry estimates available at each stream segment, a continuity relationship could be developed to relate the predicted flow to the predicted velocity using the flow continuity equation:

$$v = \frac{Q}{A} = \frac{Q}{wd} \tag{3.4}$$

where v is the average flow velocity, Q is the volumetric flow, A is the cross sectional area, w is the channel width and d is the average water depth.

As with flow estimates, there is a degree of uncertainty in the width and area estimates produced using the methodology described above. The Monte Carlo simulation approach employed to propagate uncertainty with flow estimates was extended into hydrokinetic power estimates to accurately predict uncertainty associated with the hydrokinetic estimates. For each of the 1000 sample flows, the log-normal distribution of widths, depths, and area estimates were also calculated. Combined with the flow estimate dataset, the final product included dataset of 1000 samples of flows, widths, and areas for each flow frequency for each river reach in Canada. Distributions of velocities, specific power and power at each flow frequency were then computed using equations 3.2 and 3.4. Median and 95% prediction intervals were extracted from these distributions.

The resulting combination of flow and estimates of area at all the exceedence probabilities results in a flow duration curve of hydrokinetic power or Power Duration Curve (PDC). PDC data were summarized by calculating mean annual values through numerical integration and averaged using trapezoid methods.

The cross section velocity, area, power and specific power was computed for each reach in the digital stream network. Examples of velocity maps are shown in the section 3.2. The

GIS files containing all of cross section section will be publicly available through NRCan so that readers will be able to distinguish sites with the highest potential for development.

3.1.7 Total available power

Total power using elevation head

All of the aforementioned methods deal with predicting the theoretical hydrokinetic power at any single section on a river. The total theoretical hydrokinetic power available in Canada is equivalent to the total theoretical hydro power in Canada. The equation for potential hydro power is:

$$P_H = \gamma H Q \tag{3.5}$$

where P_H is the total power, H is the change in elevation head, γ is the specific weight of water, and Q is the volumetric flow rate.

Using flow estimates described earlier in this section, and elevation head as defined by CDED database, total potential hydropower was calculated for each river reach, and summed regionally. Elevation head was calculated by taking the difference between the elevations at the beginning and end of each reach. Reaches were defined by the Digital Stream Network (Section 3.1.1) and have varying lengths, depending on topography and intersections with other streams. Examples of these stream reaches are found in Figures 3.3 and 3.4. A large river is typically composed of thousands of stream segments. The results from this calculation are presented in section 3.2.2. Though the CDED DEM is very accurate on a national scale, its resolution relative to the size of the stream segments is low (elevation is measured to the nearest metre). This means that while hydropower estimates on individual stream segments will tend to be inaccurate, the overall hydropower estimate will be accurate. For example, the true elevation difference of a very short reach will not be captured if it lies completely in a single CDED raster cell (ie. both the beginning and end of the reach will have the same value), but this elevation difference will be made up in the downstream reach that crosses the next raster cell. In essence, a one metre error in elevation has a huge impact on a one kilometre river reach, but can be ignored over the whole river, typically thousands of kilometres in length with hundreds of metres of elevation drop.

Total power using velocity head

A second way of looking at total power is through equation 3.6, which is similar to equation 3.5 but H_V is described by the potential available from velocity whereas H is the potential from gravity. In essence, H_V and H are the same value, but calculated from different means. The main benefit in calculating the total power by this second equation is that it frames total power in terms of velocity of the water, a key parameter we are studying in this hydrokinetic research.

$$P_K = \gamma H_V Q \tag{3.6}$$

Where H_V the head calculated from a modified Manning's equation, derived below in equations 3.7 - 3.6.

$$Q = \frac{AR^{2/3}S^{1/2}}{n} = \frac{AR^{2/3}H_V^{1/2}}{nL^{1/2}} \tag{3.7}$$

$$H_V = \frac{v^2 L n^2}{R^{4/3}} \tag{3.8}$$

where A is area, R is hydraulic radius, S is water surface slope, n is Manning's roughness coefficient, L is length of the reach, H_V is head, γ is specific weight of water, Q is flow and P_K is the total hydro kinetic power in a reach. In its current form, this hydrokinetic power equations assume that no energy is lost due to friction of the moving water because n is defined as the surface roughness in a stream's natural state. In other words, all available potential energy is extracted by the turbines and there is no energy lost due to bed friction. As such, this can be considered an 'inviscid' solution.

Using the above equations, potential hydrokinetic power for each reach has been integrated over its length. In the absence of meaningful nationwide estimates, a Manning's n of 0.035 was used for all reaches. The results have been summed and grouped by province and watershed in similar fashion to the results in section 3.1.7. Technically, equations 3.5 and 3.6 should produce the same total power estimates in Canada, but due to the greater uncertainty in equation 3.6 the total hydropower estimate is approximately twice as large as the hydrokinetic estimate. The uncertainty of equation 3.6 and analysis of the results is discussed in depth in section 5.1. It should also be noted that, as in the hydro power integration, while the majority of river reaches have been calculated, some reaches have not been calculated due to the nature of the CCA-MR analysis.

The benefit of looking at total hydrokinetic power in terms of equation 3.6 becomes apparent when discussing the amount of extractable energy along a given stream reach. Current research on extractable riverine hydrokinetic power is very limited. The best estimate is that 0% to 30% of total power (depending on the flow and slope conditions of the stream reach) can be extracted through hydrokinetic turbines (Jacobson, 2012). Some knowledge can be gleaned from research on wind turbines but there are additional factors that must be accounted for. In actuality, the amount of extractable energy is limited by:

1. Temporal variations in flow;
2. The amount of kinetic energy that can be taken from the water at cross section;
3. Turbine cross sectional area;
4. Turbine efficiency; and
5. Hydraulic influence of other turbines or structures at or downstream the deployment site.

For this study the scope was limited to the actual physical energy available in stream reaches. The factors that were related to turbine design, performance or deployment limitations were not considered as they will vary with turbine design characteristics.

The first assumption made is that all of the water's hydrokinetic energy is extractable all of the time. Currently this is not the case; flows could be too low to actually submerge the turbine, too high and fast to safely operate the turbine or there could be other site conditions such as ice or debris that limit use of the turbine. Because all of these factors are site specific, this study assumes that the hydrokinetic energy is extractable over the whole range of flows. The ability to extract the energy during particular flows is an engineering challenge for the turbine design and not strict natural limitations. Other studies have employed techniques and assumptions to provide limits to the total extractable energy including UMA Group (1980) and Jacobson (2012). The data products that are being released as part of this study will include information on velocity, channel geometry and flow and will allow for any limitations to be applied in post-processing using GIS tools.

The hydraulic influence that a turbine will have on a reach is largely unknown. In terms of the Manning's equation, installing a turbine in a river will increase the roughness coefficient which will in turn increase the water depth and influence the width. Jacobson (2012) built a 1D model to analyze the backwater effects of turbines and developed an empirical equation for equivalent Manning's roughness in terms of flow and slope. Physical studies have been done to find equivalent roughness caused by submerged structures in rivers (Christodoulou, 2013) but no known physical study to date has been undertaken to analyze the effects moving submerged riverine structures such as hydrokinetic turbines. It is also very likely that improved estimates of velocities and hydraulic geometry will be available with further research. As improved estimates of equivalent Manning's roughness, velocities and geometry become available, equation 3.6 may become more practical for estimating extractable hydrokinetic power than equation 3.5.

3.2 Results

The results of applying the flow regionalization, channel geometry estimates and the hydrokinetic calculations outlined in Section 3.1 produced two separate classes of results to meet the study objectives. The first was a set of maps showing areas of highest potential. The national maps show the cross sectional attributes velocity, width, depth, specific power. These were designed for examining targeted locations for high power potential. Secondly, the extractable power product is a theoretical maximum extractable energy and can be integrated over regions and drainage basins to determine the total resource potential by region. It should be noted that this is a national resource assessment and it does not account for hydro energy that is already being extracted by hydro dams and other structures.

3.2.1 Hydrokinetic power and stream velocity estimates

The hydrokinetic power and stream velocity geospatial datasets were produced. Figure 3.13 and 3.14 illustrate two stream flow velocity maps of Canada, the first with a 20% exceedence probability and the second with a 50% exceedence probability. A few characteristics should be noted about the map. First, there are large areas in northern Canada where no power estimates could be computed. This is due to lack of watershed data where resolution of satellite data was too coarse for accurate measurements, or the data required for a MLR-CCA analysis for the region was never obtained. Secondly, some areas in the prairie provinces were not predicted. In general, flow prediction in the prairie provinces was problematic due to the non-linear hydrologic responses due to prairie potholes. In this region it is difficult to predict flow duration curves based only on the physiography and climate. In order to optimize accuracy in the Nelson river basin, which includes the prairies, the alpha value in the CCA was required to be relatively high which in turn increased the number of stream reaches which did not have enough 'neighbouring' stations to conduct a MLR analysis.

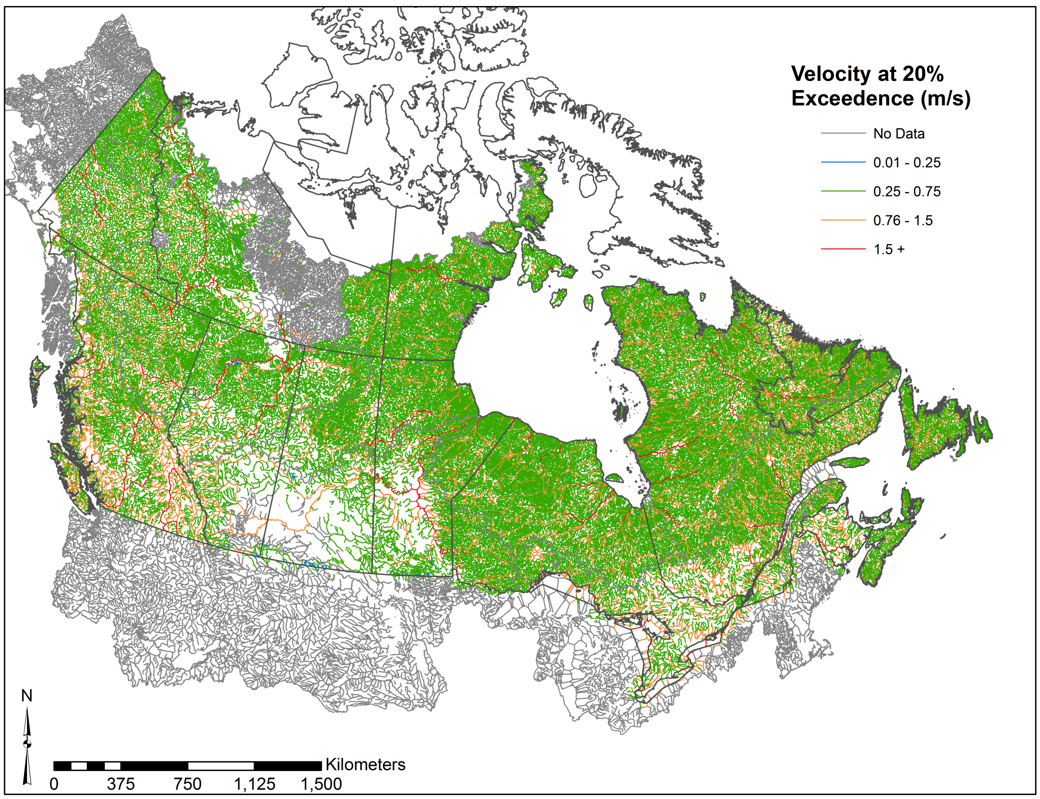

Figure 3.13: National velocity map - 20% exceedence probability

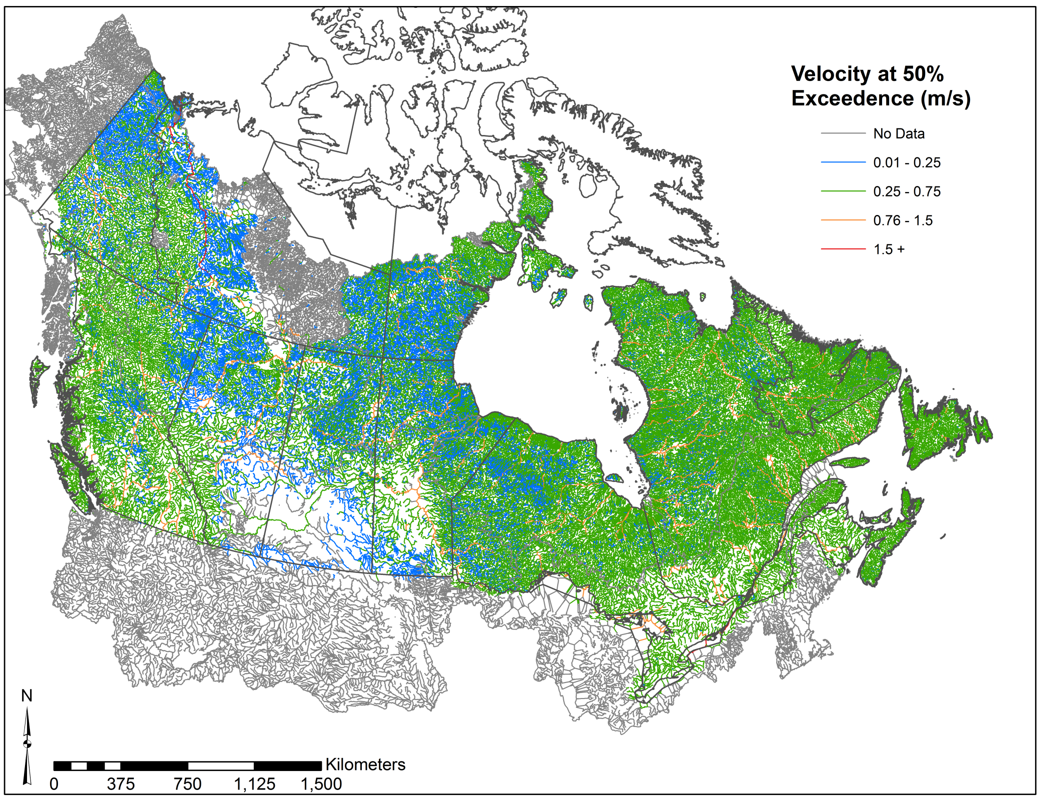

Figure 3.14: National velocity map - 50% exceedence probability

3.2.2 Regional power estimates

The total estimated potential hydropower along with error bars are shown by province and by region in Figures 3.16 and 3.15 respectively. These values were estimated with equation 3.5. The estimated power is shown on a logarithmic scale in order to properly demonstrate the range of power potential across Canada. The bar values represent the expected value for each province or drainage basin, but also plotted are the 95% prediction intervals as whisker plots to illustrate uncertainty in the MLR analysis. Figure 3.15 is of more interest to hydrologists as it separates the potential power by Canada's major watersheds. The author finds the amount of power available in Western and Northern Canada surpising. While the Pacific and Mackenzie basins have more power potential than any other basin, they are some of the least developed basins. The high power potential can be contextualized when one sees a elevation map of Canada. Figure 3.17 shows that the mountains on the Pacific coast are by far the highest in Canada. With high amounts of moisture coming off the neighbouring ocean, the Pacific region shows great hydro and hydrokinetic power potential. But as political lines are rarely drawn according to watershed boundaries, Figure 3.16 shows the same plot but separated by province. There is an estimated 710 GW of total potential power in Canada with the most power concentrated in British Columbia (220 GW) followed by the Quebec with 140 GW. For comparison to installed power, HydroQuebec, BC Hydro and Manitoba Hydro corportations currently have capacities to extract aproximately 36 GW, 12 GW and 6 GW respectively (HydroQuebec, 2012; BC Hydro, 2013; Manitoba Hydro, 2012).

The total estimated hydrokinetic power is shown in Figures 3.18 and 3.19 for regions and provinces respectively. The uncertainty associated with these estimates is much greater than in the hydropower plots, as discussed in section 3.1.7. As mentioned the section, hydropower and hydrokinetic power should theoretically be equal, but there is a discrepancy because of the error associated with velocity estimates. One noteable difference between the hydro and hydrokinetic plots is how much power is estimated in the Pacific region. The Pacific hydropower estimated is much higher than the Pacific hydrokinetic power estimated. Because the hydrokinetic power is estimated from velocity, and velocity is estimated from flow and geometry values (which were predicted with very little influence of elevation differences), the large elevation drops in the mountainous terrains cannot be fully accounted for in the model. Note that because parts of northern Canada did not have sufficient data, Nunavut and the Northwest Territories likely have higher estimated power. These results are discussed in more detail in section 5.

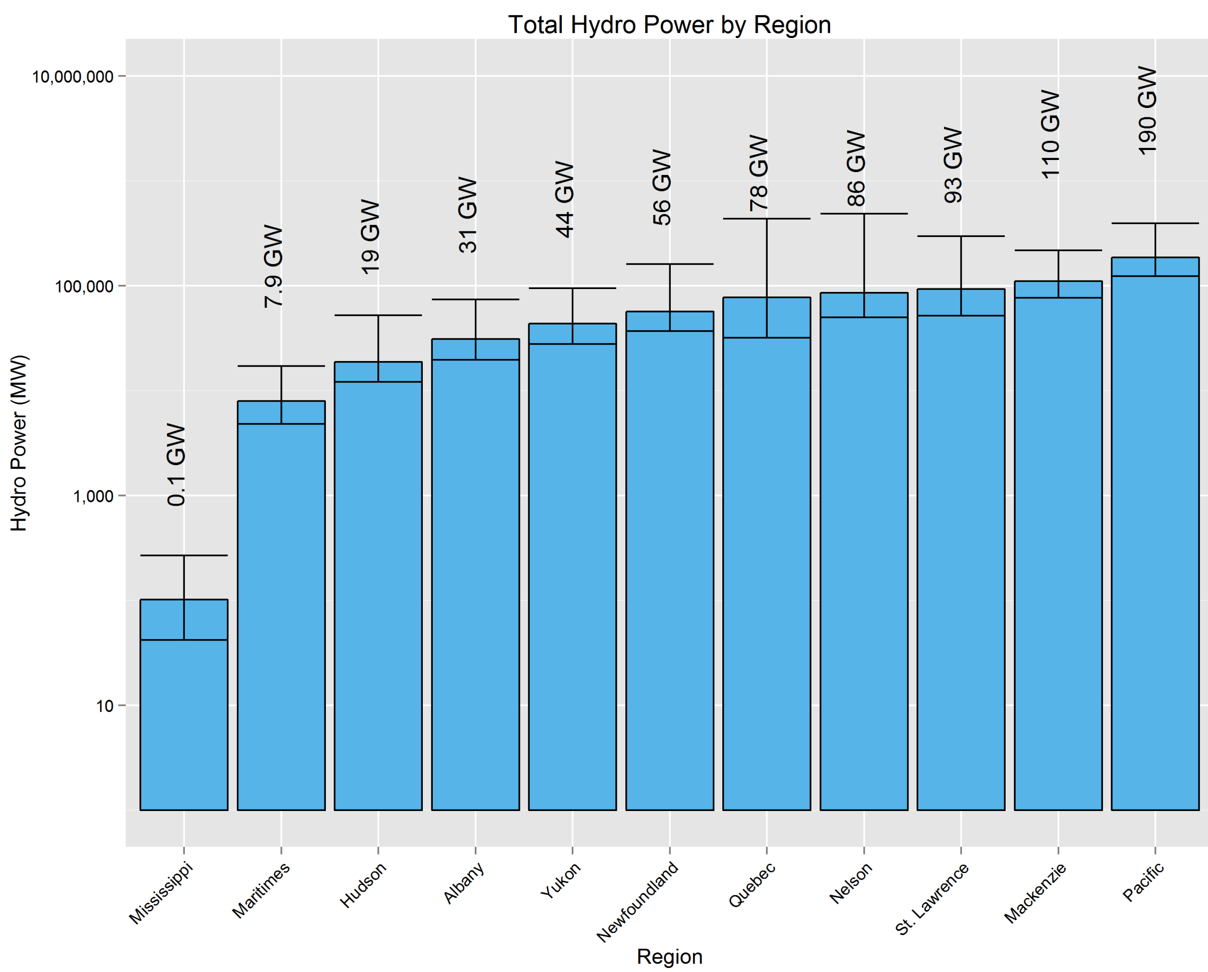

Figure 3.15: Hydropower in Canada by Major Drainage Basin

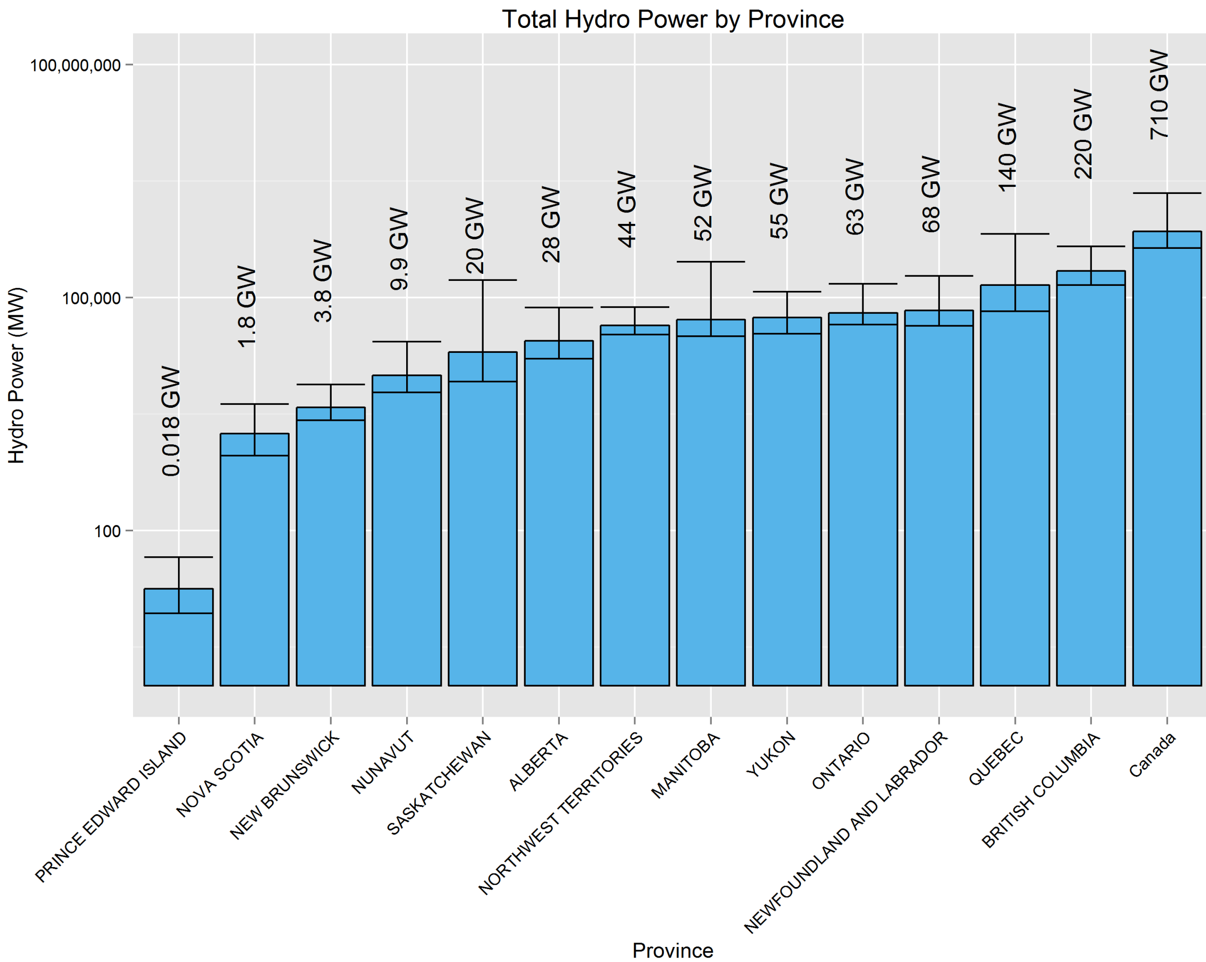

Figure 3.16: Hydropower in Canada by Province

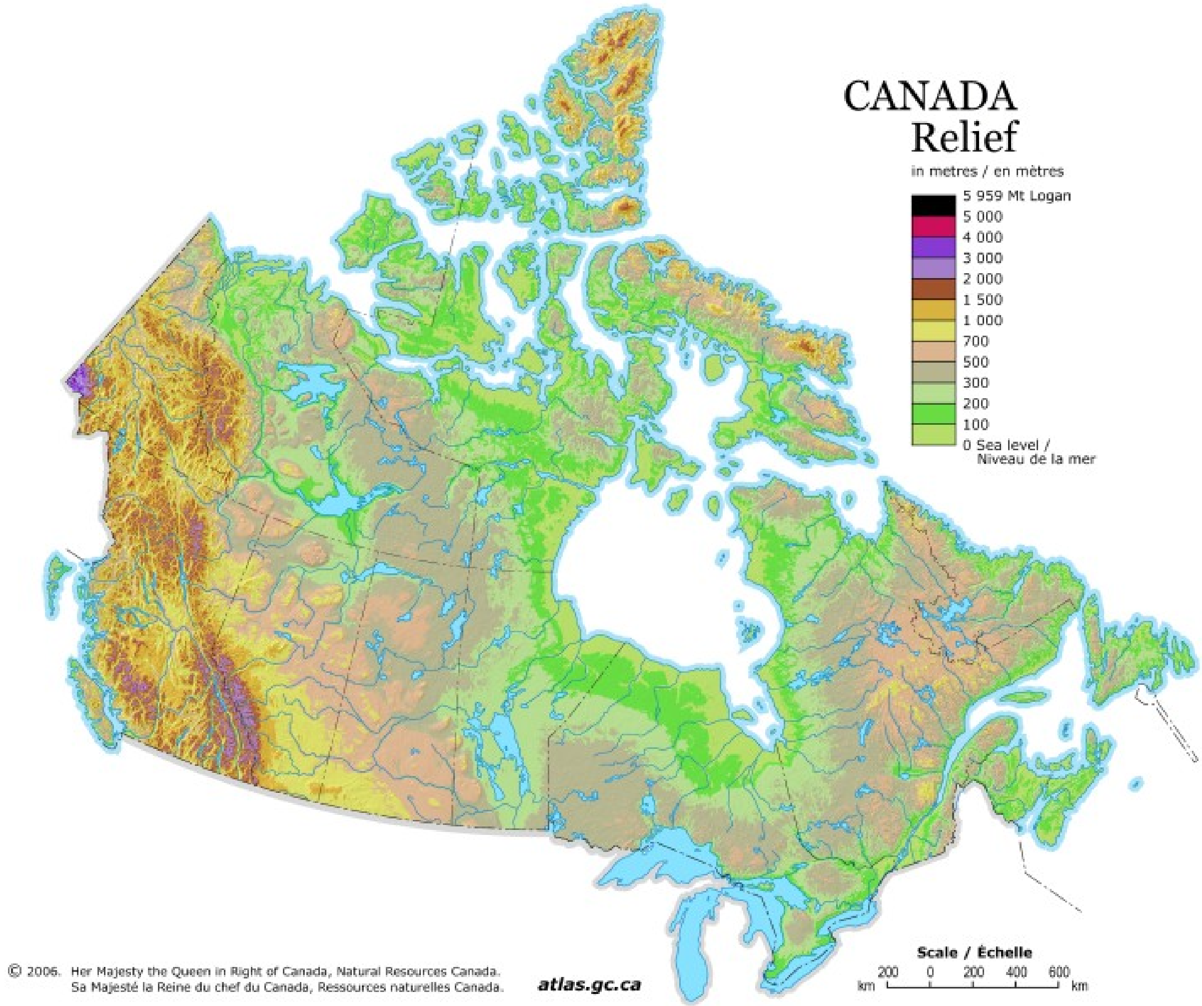

Figure 3.17: Elevations of Canada, Source: NRCan

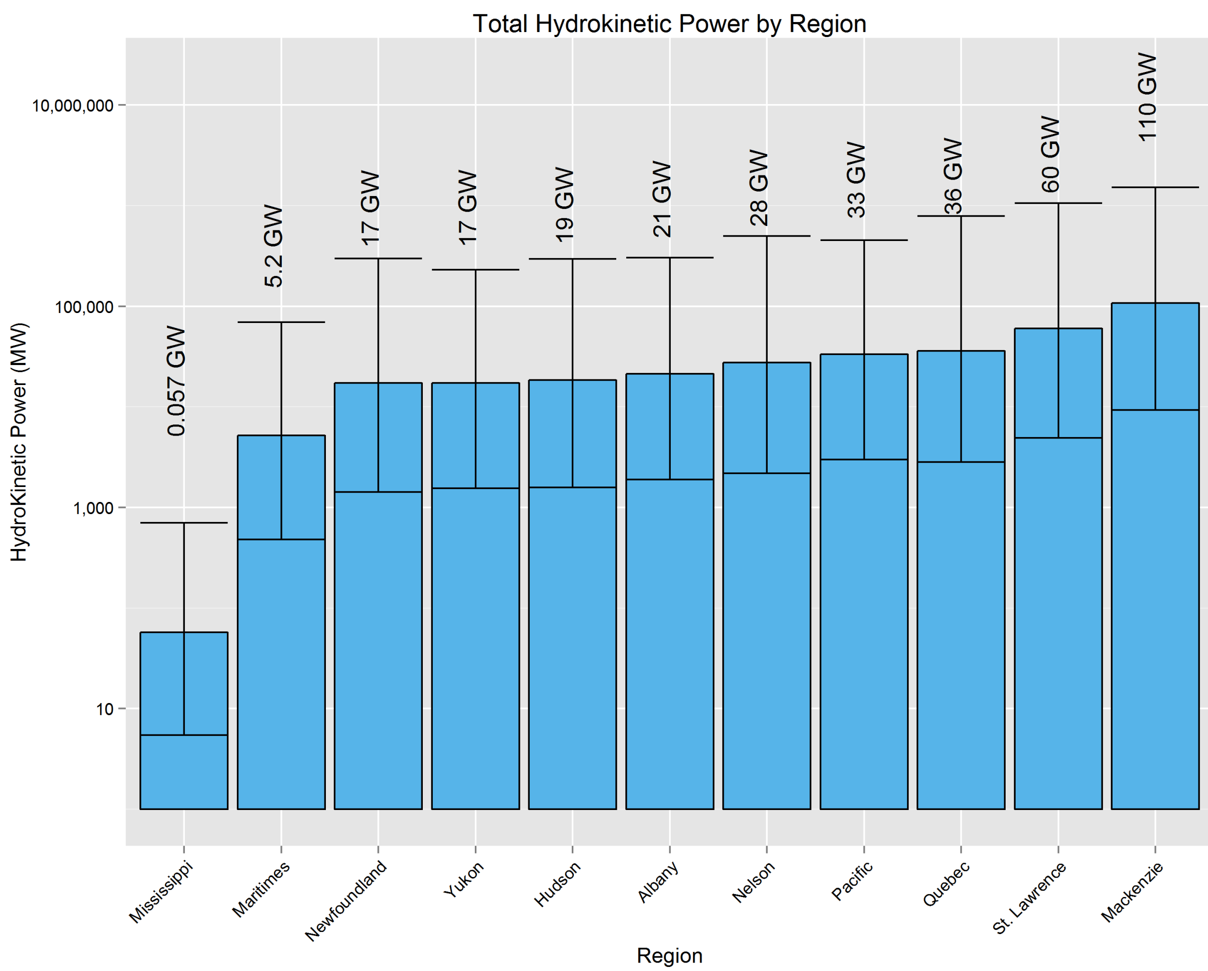

Figure 3.18: Hydrokinetic power in Canada by Major Drainage Basin

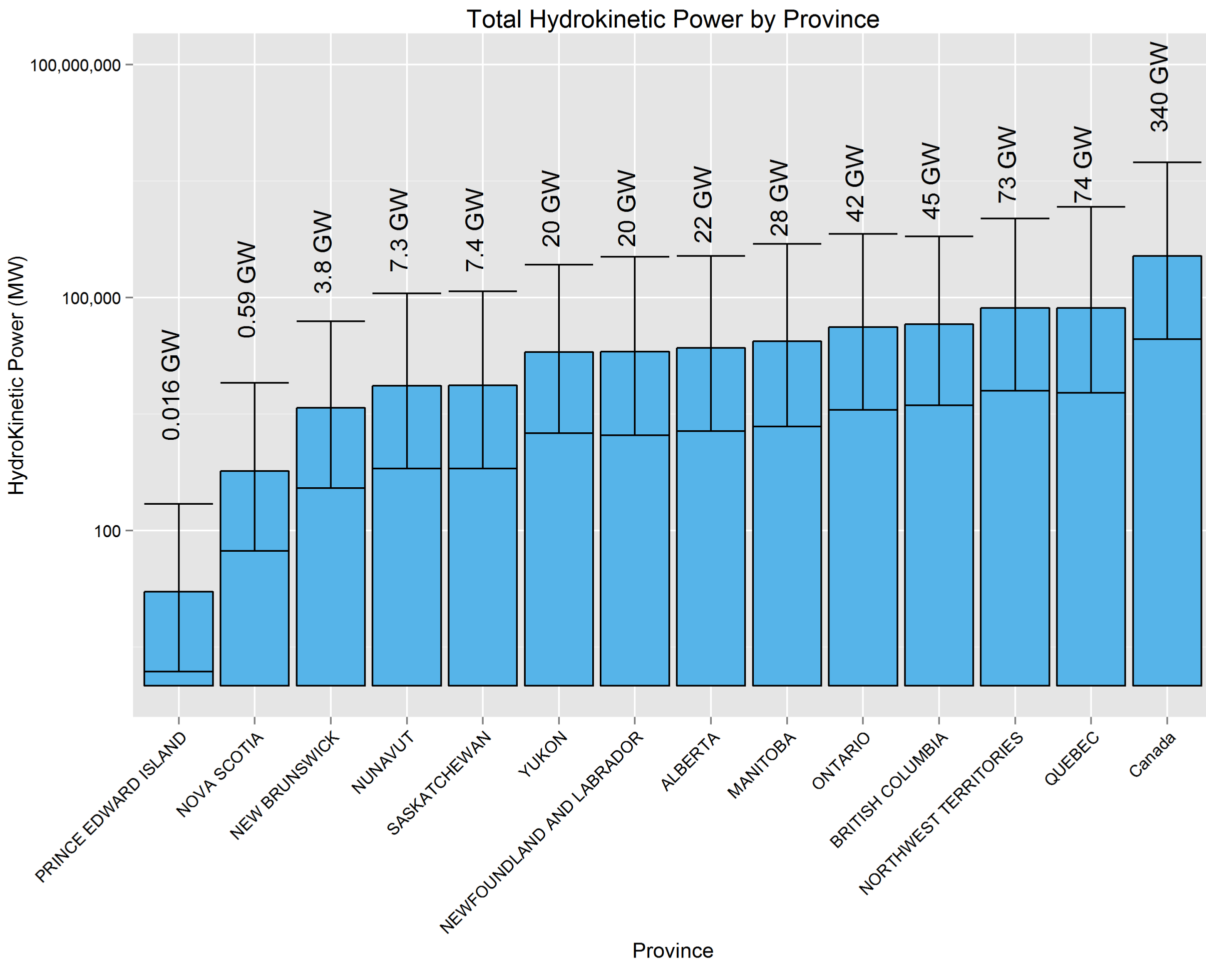

Figure 3.19: Hydrokinetic power in Canada by Province

Chapter 4

Hydraulic Geometry

4.1 Abstract

Regional downstream hydraulic geometry equations are plotted using bankfull flow, as defined by the 1.5-year recurrence interval, and bankfull widths. The study includes 634 gauge stations located across Canada and are grouped by characteristics that affect or are correlated with bank and bed stability. Jack-knife analysis was done to test the predictive capabilities of each grouping. Grouping the data by geology and vegetative characteristics improved predictions in 14 of the 15 categories, with drainage characteristics and mode of deposition groupings showing the best results. These categorizations are investigated further. The study demonstrates that classifying downstream hydraulic geometry equations by bank material and vegetation improves the predictive capabilities of the equations. It also shows that regionalization does not need to occur in a contiguous region. As demonstrated in previous studies, it is also shown that downstream hydraulic equations can be applied to semi-alluvial rivers. Finally, multiple regression was tested with width prediction as a function of flow and up to 3 categorical variables. Results showed no improvement over grouping methods.

4.2 Introduction

Hydraulic geometry equations were first introduced by Leopold and Maddock (1953) and describe hydraulic variables (surface width (w), mean depth (d), mean velocity (v), roughness (n), slope (S) and suspended sediment load (SL)) as log-linear functions of flow (Q) (Knighton, 1998). These relationships are defined independently for at-a-station and for downstream channel geometry relationships, where at-a-station refers to the geometries found at a single cross section on a river with varying flow rates, and downstream refers to geometries found at multiple cross sections on similar rivers with an equal frequency flow. The equations for width, depth, and velocity are shown in 2.1 - 2.3.

$$w = aQ^b \tag{2.1}$$

$$d = cQ^f \tag{2.2}$$

$$v = kQ^m \tag{2.3}$$

Where a, c, k are coefficients and b, f, m are exponents. Though these empirical relationships have typically been applied to alluvial rivers, studies indicate they can also be applied to semi-alluvial and non-alluvial rivers (Montgomery and Gran, 2001; Wolh, 2004; Ebisa Fola and Rennie, 2010; Vachtman and Laronne, 2013). The b exponent for width has been found to be almost universally around 0.50 for alluvial rivers (Knighton, 1998); but this value can vary depending on bank strength. Though there is some doubt as to the role of bank strength in determining channel dimensions (Neill, 1964), the general consent is that while flow is the dominant cause, bank strength plays an important role in influencing channel geometry (Davies and Sutherland, 1980; Osterkamp, 1982; Hey and Thorne, 1986;

Huang and Nanson, 1998; Millar, 2005). Furthermore vegetation and bank sediment play a large role in determining bank strength (Huang and Nanson, 1998). Rivers situated in consolidated clay or rock, which are relatively in-erodible material, have been shown to have higher b values. For example, Ebisa Fola and Rennie (2010) found rivers with clay banks to have a b value of 0.57. A higher b value indicates that river width increases more rapidly with increasing discharge.

Riparian vegetation also has pronounced, but varying, effects on width-discharge relationships (Osterkamp, 1982). Eaton and Church (2007) found that relative bank strength declines in the downstream direction as the channel is incised deeper than the rooting depth of the vegetation. Copeland et al. (2005) analyzed 57 bankfull width relationships on sand bed rivers and found streams with greater than 50 % tree cover had smaller widths than streams with less than 50 % tree cover. Conversely, Murgatroyd and Ternan (1983) found that channel widths within the forest were up to 3 times larger than those predicted from the non-forested catchment upstream, likely due to within-channel woody debris diverting flow toward channel banks and thereby increasing bank erosion.

Bank material composition influences hydraulic geometry (Knighton, 1974). Furthermore, it is believed that bank materials can be assessed and categorized by determining the surface geology characteristics along streams. Previous studies have typically only assessed grain size in areas constrained to contiguous regions of homogeneous surface geology. The aim of this study is twofold. First, to assess the effects of bank stability on downstream hydraulic relations as indicated by surficial geological and vegetative characteristics over a wide range of geological formations. Secondly, as discussed in Ebisa Fola and Rennie (2010), it is believed that streams whose banks are composed of bedrock or cohesive sediments are more likely to show higher rates of width increase with increasing flow (ie. have a higher b value) than sand bank rivers because the shear strength of these banks is more likely to be overcome at greater discharges.

To meet these objectives, flow measurements, hydraulic geometry data, and surface geology and vegetative maps are gathered and analyzed from all regions within Canada. Downstream hydraulic geometry relations are calculated nationally and according to surface geology and vegetative characteristics that potentially influence bank strength. This is the first known attempt to assess the influence of bank strength on downstream hydraulic geometry relations across such a large spatial domain.

4.3 Dataset

This study primarily relies on two databases provided by Environment Canada (EC) as well as multiple land class GIS databases available through various government sources. The EC Hydrometric Database (HYDAT) contains daily average flow data for 6262 active and discontinued gauge stations across Canada dating back to 1860. The database is publicly available online through the EC website. The Water Survey of Canada Measurement Database (WSC-MD) was also utilized. It was developed for internal use by Environment Canada and contains measured widths, depths, and velocities for varying flow rates for 1862 of these stations. These data were originally collected to develop stage-discharge relations at these stations.

Additionally, publicly available GIS datasets were obtained to classify gauge stations by soil and vegetative characteristics. Characteristics were selected on the criteria that nation-wide datasets were available and that the attribute represents bank strength in some form. Soil characteristics were gathered from surficial geology maps published by Natural Resources Canada (Canada, 2010) as well as data published by the Soil Landscapes of Canada (SLC) (AAFC, 1996). The SLC dataset divides Canada into ecodistrics and assigns various attribute values to each district, including surface material, vegetative cover, rooting depth, drainage characteristics, surface form and slope gradient. An exhaustive description of the

SLC classifications can be found in Appendix B.1. Land use data were also found from Natural Resources Canada 2010, which classifies land into nine different types dependant on vegetative characteristics (wetlands, forest, etc.).

Standard flow frequencies were defined in order to accurately assess downstream hydraulic geometry variability in Canada. Bankfull flow is most commonly used in downstream hydraulic geometry equations (Copeland et al., 2000). Bankfull flow can be (1) measured in the field as the flow where the water surface elevation transitions from the channel to the floodplain, (2) equated to effective flow, or (3) assumed to be the 1.5 year flood recurrence interval flow (Copeland et al., 2000). Because it was infeasible to physically measure bankfull flow, and suspended sediment concentration data required for calculating effective flow were limited, bankfull flow was calculated as the 1.5-year annual maximum flood. Flow Duration Curves (FDCs) were also calculated using standard methods (Searcy, 1959a) with HYDAT historical data, and 17 flow percentiles were calculated from this curve ranging from flows that occur less than 0.01 % of the time to 99.99 % of the time.

As the main intent in creating the WSC-MD was to gather accurate flow data, there are biases in the dataset worth discussing. Firstly, the criteria for station locations would have been based on a number of factors: ease of access, stability of reach, nearby control structures, etc. For the purposes of this study, it is assumed that, despite the aforementioned biases, the stations in this dataset represent natural cross sections in Canadian rivers. It is also common practice, during low flows, for flow and geometry measurements to be taken upstream or downstream of the regularly used station cross section in order to achieve the most accurate flow measurement. This problem with the dataset is not uncommon and was pointed out by Park (1977) in United States Geological Survey data.

In order to find w, d, and v at each station for bankfull flow and the 17-point FDC, a robust, log-linear regression was applied to the hydraulic geometry data from the WSC-MD. This regression is equivalent to developing at-a-station hydraulic geometry curves at each station. During the analysis, the trend of large scatter in the data at lower flows became apparent. This is likely caused by the tendency to use the optimal cross sections for flow measurement, as previously discussed. As the data showed more consistency at higher flows, a robust regression was performed on all flows higher than the mean annual flow (MAF). Figure 4.1 shows an example of an at-a-station analysis. Leopold and Maddock showed that in order for continuity to be maintained, the slopes of the width, depth and velocity graphs (b, f and m) must sum to 1. In the present analysis, any station for which exponents did not sum to 1 +/- 0.05 was discarded. The 5 % criterion is commonly used for statistical significance and was also used by Rhodes (1977) as a criterion to include in his tri-axial $b - f - m$ diagrams. The average continuity value for these stations was 1.0023 with a standard deviation of 0.016. Additionally, stations that did not have enough data points (<6) were discarded. A map showing potential stations and filtered stations is shown in Figure 4.2.

4.4 Analysis and Results

The baseline of the study was set by mapping the downstream hydraulic geometry equations with no regionalization. Bankfull widths were plotted against bankfull geometry on a log scale and are shown in Figure 4.3. Even without regionalization a good regression is shown to occur with a R^2 of 0.908.

Downstream hydraulic geometry equations were then regionalized according to different characteristics. These were selected based on nationally available datasets and meant to represent bank strength. A total of 15 different classifications were used. An example of a classification, soil drainage characteristics, is shown in Figure 4.4. Soil Drainage was rated from Very poor to rapid. Very poor drainage is described when the water table is at surface elevation whereas rapid drainage is described for surfaces where water drains away quickly

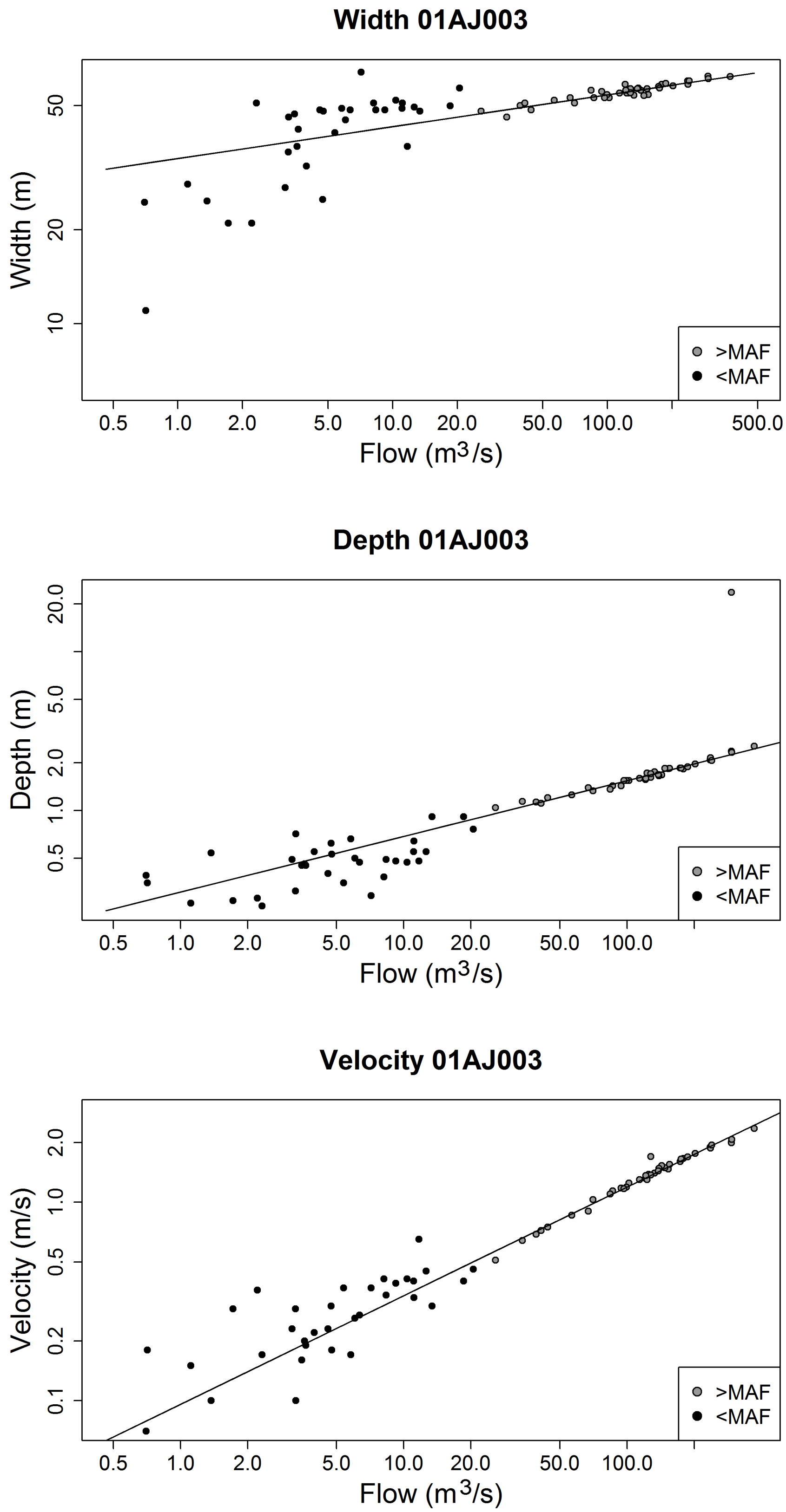

Figure 4.1: Robust regression for at-a-station hydraulic geometry

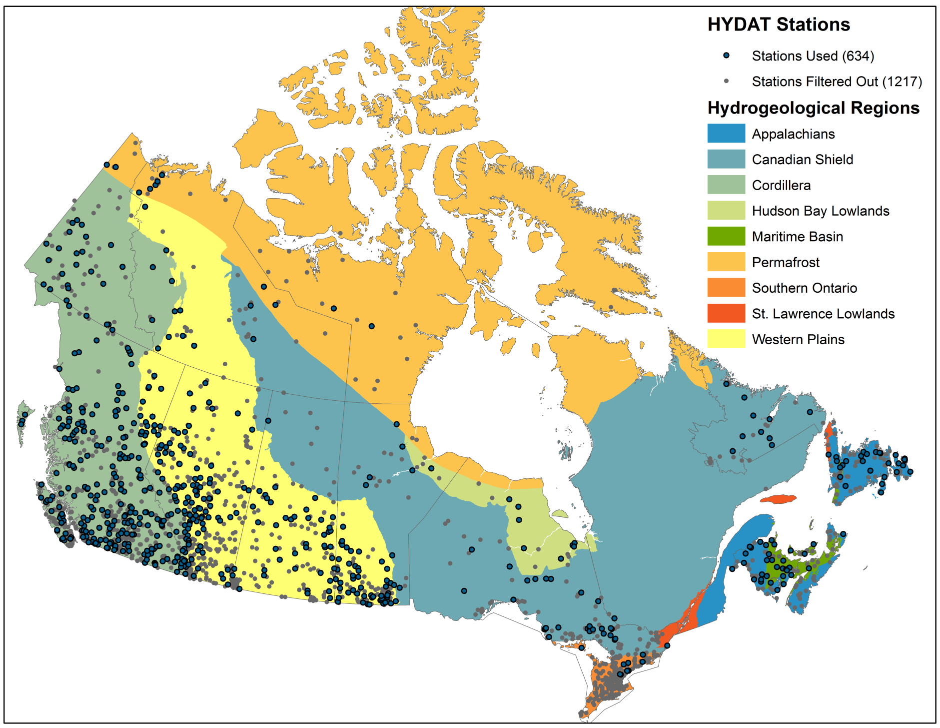

Figure 4.2: Map of HYDAT stations in WSC-MD

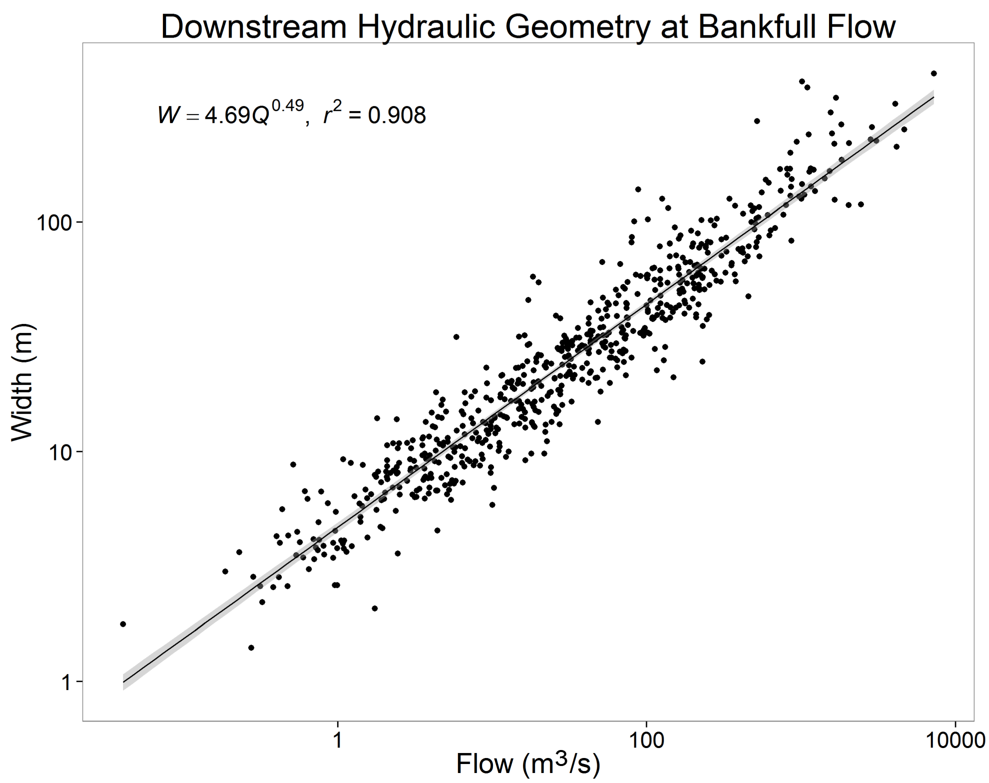

Figure 4.3: Downstream Hydraulic Geometry at Bankfull

through porous soils and the main source of water is from precipition (and not groundwater). A complete description of drainage characteristics, as well as other classifcations, can be found in Appendix B.1. One can see that the R^2 value increases for almost all of the equations. More importantly, the equations adjust themselves to reflect local conditions. For example, higher b values can be seen in soil with poor drainage characteristics.

The model's predictive capability was measured using a jack-knife re-sampling analysis, also know as a leave-one-out analysis. A subset of the data was taken by omitting a single station. The downstream hydraulic geometry equation was calculated on the subset and then used to predict the width of the omitted station. This process was repeated for each station. Observed widths could then be compared with estimated widths. The Nash-Sutcliffe model efficiency coefficient (NASH) was used to assess overall accuracy of the predicted widths and is shown in Equation 4.

$$E = 1 - \frac{\sum (W_a - W_e)^2}{\sum (W_a - \overline{W_a})^2} \tag{4.1}$$

where W_a is the actual width, W_e is the estimated width, and $\overline{W_a}$ is the mean of the all of the actual widths.

The NASH coefficient has a range of $-\infty$ to 1. A value greater than 0 is generally acceptable performance with 1 indicating the model has perfect predictive capabilities. A NASH value less than 0 means that the mean of the observed data is a better predictor than the model (Moriasi et al., 2007).

The NASH results of each regionalization method are shown in Table 4.1. NASH values have a range from 0.765 for random classifications to 0.879 for classification by drainage class. Random classifications were also tested to ensure that improved results were not simply caused by using subsets of data. The baseline results, where no classification was applied, give a NASH value of 0.795, the third least effective category.

Multiple regression was also attempted with up to three categorical variables. All 575

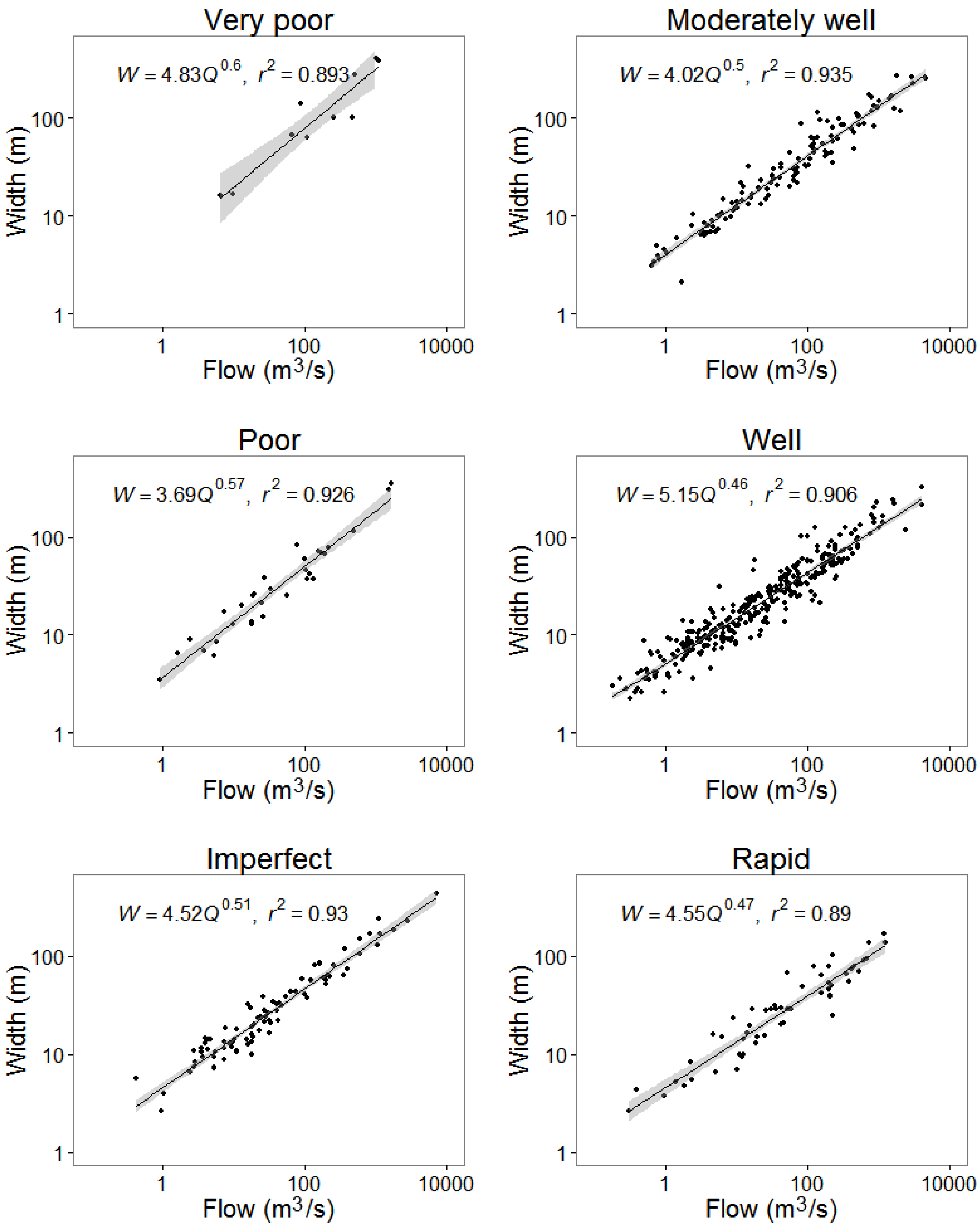

Figure 4.4: Downstream Hydraulic Geometry with Regionalization by Drainage Characterization

Table 4.1: NASH Statistics on Regional Downstream Hydraulic Geometry

Classification	NASH	Number of Classifications	Category
Drainage Class	0.879	6	Geological
Parent Material Mode of Deposition	0.871	10	Geological
Surface Material (Rock - acidic and basic, Soil, etc.)	0.865	5	Geological
Calcerous Class of Parent Material	0.863	5	Geological
Vegetative Landcover	0.861	9	Landcover
Percentage of Rock Fragments greater than 0.2 cm (Volume)	0.859	4	Geological
Geological Regions	0.858	13	Regional
HydroGeological Regions	0.856	10	Regional
Surficial Geology	0.852	18	Geological
Soil Development Classification	0.849	29	Geological
Generic Classes of Surface Material	0.836	7	Geological
LandCover (AAFC)	0.830	9	Landcover
Slope Gradient of Surrounding Region	0.829	6	Landcover
No Classification	0.795		
Rooting Depth	0.793	5	Geological
Random Classifications	0.765	5	

possible combinations of the 15 categorical variables were tested. Widths were predicted by multiple regression using the log form of the downstream hydraulic geometry equation as shown in Equation 4.2.

$$\log w = (b \times \log Q + a) + dC_1 + eC_2 + fC_3 \tag{4.2}$$

where d, e, f are linear coefficients, C_1, C_2, C_3 are categorical variables, and w, b, Q, a were previously defined in Section 4.2. The key results are shown in Table 4.2. The highest NASH value recorded was 0.878 (Table 4.2), slightly lower than when straight categorization was used (Table 4.1). The classifications included in the best multiple regression model were drainage characteristics, slope and hydrogeological region. The chi-squared test was used to see if there was correlation between the classifications (in Table 4.1), however, the tests were inconclusive as the frequency of each classification did not meet the minimum standard, typically considered to be 5 (Camilli and Hopkins, 1978). As in the classification analysis (Table 4.1), drainage characteristics appeared to be the most significant attribute, though when used in the MR equation on its own, it gave a NASH value of only 0.851.

Table 4.2: NASH Statistics on Multiple Regression Downstream Hydraulic Geometry

Characteristics	NASH
Drainage Characteristics, Slope, Hydrogeological Region	0.878
Drainage Characteristics, Landcover (AAFC), Surface Material (Generic)	0.878
Drainage Characteristics, Calcerous Class, Surface Material (Specific)	0.878
Drainage Characteristics, % Large Fragments, Hydrogeological Regions	0.877
Drainage Characteristics, Calcerous Class, Hydrogeological Regions	0.877
Drainage Characteristics, Calcerous Class, Slope	0.876
Drainage Characteristics	0.851

4.5 Discussion

The 15 characteristics show in Table 4.1 can be classified into 3 main categories: geological (composition and deposition), landcover, and regional characteristics. Geological characteristics showed the most promising results. This is likely because geological characteristics better correlate with bank strength. The USACE 1994 recommends streams are designed narrower when there is a higher silt/clay content in the soil because cohesive properties make the banks less erodible. Cohesive soils and large materials have high resistance to shear stress while medium sized particles have low resistance to shear stress (Thorne and Tovey, 1981). Vegetation is much harder to correlate. This could be for a variety of reasons. Each type of vegetation has different effects on soil strength; the root systems of large trees may bind a soil together, but they will also cause large erosion if they fall into a stream and impede flow (Montgomery, 1997). Additionally, the vegetative datasets used were national datasets and likely do not have the resolution to differentiate between riparian vegetation and the dominant vegetation in the area (Eaton and Church, 2007).

The categories that showed the best results were drainage characteristics and parent material mode of deposition. The specific results for drainage characteristics are shown in Table 4.3. Soil drainage is controlled by pore space so it can generally be assumed that poor drainage is correlated with smaller particle size. Though the a coefficient shows no discernible trend, soils with poorer drainage exhibit higher rates of width increase (b). A definite pattern emerges when widths are calculated at a flow of 1000 m^3/s (representative higher flow, each classification had a least one flow equal to or greater than 1000 m^3/s); widths are greatest in soil with poor drainage characteristics. These results coincide with other studies that show cohesive soils having higher rates of width increase (Ebisa Fola and Rennie, 2010).

Table 4.3: Downstream width parameters grouped by drainage conditions.

Title	a	b	R2	Stations
Very Poor	4.83	0.60	0.89	10
Poor	3.69	0.57	0.93	31
Imperfect	4.52	0.51	0.93	86
Moderately Well	4.02	0.50	0.94	119
Well	5.15	0.46	0.91	305
Rapid	4.55	0.49	0.89	55

Bankfull widths were also categorized by their mode of deposition, as shown in Table 4.4. The first five parameters (Alluvial to Fluvioglacial) can be summarized as larger grained sediment deposition, likely to have good drainage characteristics. On average the larger grained sediment have lower b values, showing that their widths increase relatively slowly in the downstream direction. Additionally their higher average a coefficient causes them to have larger widths at low bankfull flows. Conversely, the last four parameters are characterized by mud, organic deposition and rock, having properties of poorer drainage and higher cohesiveness. Rock channels were found to be more narrow than alluvial channels for the majority of flows (cf., Wohl and David 2008;Montgomery and Gran 2001). Although lacustrine has more moderate results, this general category has lower initial widths but widths increase rapidly in the downstream direction. On average they showed a 20 % greater width at high flow (1000 m^3/s) than the larger grained sediment deposits.

Table 4.4: Downstream width parameters grouped by mode of deposition.

Title	a	b	R2	Stations
Alluvial	5.15	0.48	0.93	38
Colluvial	4.19	0.47	0.92	75
Eolian	4.39	0.51	0.89	7
Morainal	4.61	0.49	0.91	272
Fluvioglacial	5.35	0.45	0.9	62
Lacustrine	5.42	0.46	0.88	110
Marine	3.77	0.49	0.94	7
Organic	3.71	0.61	0.86	21
Rock	3.59	0.52	0.87	21

4.6 Conclusions

A number of conclusions can be drawn from these results. Firstly, any bank strength classification, save rooting depth, shows better predictive results than no classification. The fact that random classifications show worse predictive results demonstrate that the cause is due to physical reasons and not statistically smaller groupings. Multiple regression using these classifications showed no improvement over the single classification method. While NASH results were comparable to those of the classification analysis, including three characteristics in the equation unnecessarily added complexity with no improvement in predictions. Hey and Thorne (1986) and Davies and Sutherland (1980) also used MR and found discharge to be the dominant variable while Osterkamp (1982) found more success categorizing downstream hydraulic geometry equations based on sediment type than using MR.

Secondly, regional downstream hydraulic equations do not need to be defined by a contiguous region. This is a logical extension of flow regionalization research previously done by Ouarda et al. (2001a) and continued by Jenkinson and Bomhof (2012). It was shown that historical flows could be better predicted using Canonical correlation analysis as a regionalization technique, which does not assume that neighbouring streams need to lie in a contiguous area, only have similar watershed characteristics. The result is a 'regional' hydraulic geometry relation where the 'region' is defined by the local surficial geology and vegetation characteristics, not necessarily geographic location. This is similar in concept to traditional hydraulic geometry relations developed as rivers with specific bed material (eg. sand vs gravel). The advantages of the present approach are that a large variety of local sediment and vegetation characteristics that influence bank strength can be considered, and these characteristics are widely available over large spatial domains. Thus, bank strength specific hydraulic geometry relations can be generated over large domains.

Thirdly, this analysis demonstrates that downstream hydraulic geometry can be extended to semi-alluvial streams. Classically these equations have been applied to streams whose floodplains make up the same material that the river carries. Due to the vast variety of geographic landforms in Canada a large portion of Canadian streams cannot be considered alluvial. Rivers in the Rocky Mountains or in the Canadian Shield are highly likely to be restricted by bedrock while rivers in lacustrine sediment of eastern Ontario are likely be restricted by cohesive clay banks. Though some studies did not find a strong relationship (Ponton, 1972), Montgomery and Gran (2001), Wolh (2004), and Molnar and Ramirez (2002) have shown that downstream hydraulic geometry relationships can be applied to mountain rivers while Ebisa Fola and Rennie (2010) showed that downstream hydraulic geometry equations can be applied for cohesive soils. Reasons for these empirical results could possibly be explained by stream power of these rivers. Wolh (2004) found downstream hydraulic geometry relationships were better developed if the stream power to sediment size ratio dropped below 10,000 kg/s^3. Successful demonstrations of this extension of downstream hydraulic geometry merit further investigation for possible explanations to this phenomenon.

Lastly, surficial geology characteristics were shown to influence regional downstream hydraulic geometry relations. The most successful categorization in this study utilized drainage conditions where presumably low drainage indicated fine soils with minimal pore space. This classification showed a 8.4% (as characterized by the NASH-Sutcliffe coefficient) increase in accuracy over the baseline model. Rivers in locations with poor drainage were found to have higher width exponents, indicating that they become wider more rapidly with increasing bankfull discharge. As discussed in Ebisa Fola and Rennie (2010), rivers where the banks are more resistant to erosion than the bed will be relatively narrower and deeper channels. The results presented herein suggest that rivers incising into poorly drained fine grained soils are relatively narrow and deep at low bankfull discharges (small creeks and rivers), but become increasingly wide at high bankfull discharges (large rivers). In other words, it appears that in large rivers the elevated bank strength associated with cohesion of fine grained soils is insufficient to resist bank erosion and channel widening to accommodate increased discharge. Though the lacustrine classification produced a small b exponent, the mode of deposition classification (Table 4.4) also showed that streams with bedrock and cohesive banks have higher b exponents. One possible reason that the drainage classification could have produced better and more consistent results is because drainage is an easy and standard variable to measure, whereas other classification measurements are more likely to be subjective to a certain degree. Surficial geology classifications have been shown to improve the accuracy of downstream hydraulic geometry relations and merit further investigation into the physical mechanisms at work to achieve this improvement.

Chapter 5

Discussion

5.1 Quantifying Uncertainty

The reader will recall the equation for total hydrokinetic power is (Section 3.1.7):

$$P_K = \gamma H_V Q \tag{3.6}$$

where

$$H_V = \frac{v^2 L n^2}{R^{4/3}} \tag{3.8}$$

and

$$v = \frac{Q}{A} \tag{3.4}$$

$$R = \frac{A}{P} = \frac{A}{W + 2D} = \frac{A}{W + 2\frac{A}{W}}{}^{1} \tag{5.1}$$

[1]assumes rectangular cross section

thus equation 3.6 can be written as:

$$P_K = \gamma(\frac{Q^2 L n_e^2 (W + 2\frac{A}{W})^{\frac{3}{4}}}{A^{\frac{10}{3}}})Q \tag{5.2}$$

The details of the implicit assumptions in this equation were discussed in section 3.1.7. What becomes apparent in equation 5.2 is the large amount of uncertainty in the variables when compared to the classical hydro equation (equation 3.5). In the hydro equation, γ and H carry a high degree of confidence because γ is constant and H is derived from accurate DEM maps.[2] Conversely, the velocity head in the hydrokinetic equation has large uncertainty because it is composed of flow estimates (Q), and geometry estimates (A, w) which have larger uncertainty because they are derived from hydraulic geometry regressions applied to the flow estimates. Additionally, the uncertainty of Manning's n could not be characterized because no methods currently exist to predict this value on a national scale.

A comparison of the national hydropower (P_H) and national hydrokinetic power (P_K) estimates and confidence bounds is shown in figure 3.11 with the corresponding values shown in Table 5.1. It clearly shows that, as expected, the confidence in the hydropower is much higher. While the range of the hydropower estimates is well within an order of magnitude, the hydrokinetic estimates are outside the order of magnitude range. It can also be seen that the hydropower estimates are within the range of the hydrokinetic estimates, if this were not the case there would be serious reason to question the results.

Type	**Estimate**	**Upper (97.5%)**	**Lower (2.5%)**
Hydro (MW)	709,801	2,218,992	433,244
HydroKinetic (MW)	343,143	5,510,234	29,093

Table 5.1: Uncertainty in Hydro and HydroKinetic Estimates

[2]Though the resolution of Canadian DEM is relatively low, they would be considered very accurate for these purposes on a watershed scale.

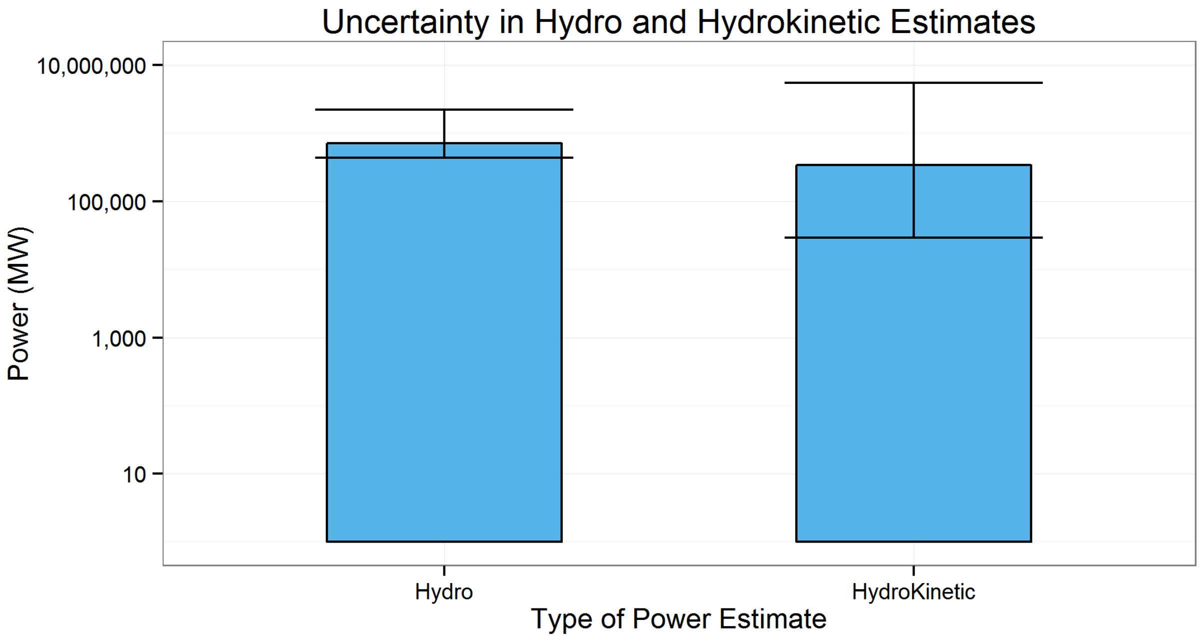

Figure 5.1: Uncertainty in Hydro and HydroKinetic Estimates

5.2 Potential in Reducing Uncertainty

This naturally begs the question of what can be done to improve the hydrokinetic reach estimation. It should be considered at this point whether, for resource development purposes, if the confidence of these estimations are sufficient. The purpose of the study, outlined in Phase I (Jenkinson, 2010) is to provide pre-reconnaissance estimates, showing potential developers how much potential power there is and the sites showing the most potential. The confidence of the estimates obtained from a remote-sensing data approach will never obtain the strict development requirements of investors and regulators; site specific surveys will always be required. Nevertheless, looking at equation 5.2 it can be seen that improved estimates of w, A, and n_e are required to improve the estimates[3], where n_e is introduced as the equivalent roughness of the turbines and natural bed and banks combined.

[3] d could also be substituted for A, it wasn't used in this study because better correlation was found for A than d in the downstream hydraulic geometry equations discussed in Chapter 4

The international scientific community has recognized the need for more and better measurements of the earth's water resources. Remote sensing is a rapidly emerging discipline that may serve to provide a partial solution to this challenge (Mersel et al., 2013). The Surface Water and Ocean Topography (SWOT) is a joint venture between the USA and France that aims to launch a satellite by 2020 to measure ocean and inland water surfaces with a high degree of spatial and temporal resolution (Fu et al., 2009). Among other requirements, the slopes rivers of at least 50 m in the width will measured to within 1 centimetre per kilometre (Fu et al., 2009). Instantaneous measurements of water surface slope, water elevation and inundation area (leading to river widths) will be collected at latitudes as high as 74 degrees (Mersel et al., 2013; Fu et al., 2009). Areas in higher latitudes will be measured up to 4 or 5 times over a 22-day repeat period (Durand et al., 2010a).

With direct measurements of slope and width, depth can be calculated through various methods as outlined by Durand et al. (2010a): 1) Ancillary data and local downstream hydraulic geometry relationships, 2) a joint probability distribution function between width and depth, 3) assuming continuity and that downstream flow changes won't affect upstream flows (ie. backwater) a linear set of equations based on Manning's equation can be solved simultaneously (Durand et al., 2010b) and 4) assimilation of remotely sensed data into a hydrodynamic model (Andreadis et al., 2007). These methods have been derived with the goal of predicting discharge from SWOT data. If we assume that discharge is already predicted from regionalization techniques, depth can be predicted using Manning's equation given width and slope inputs from SWOT and a modeled estimate of Manning's roughness coefficient. It should be re-iterated however, that SWOT data will likely only be available on rivers with at least 50 metres of width. While the stated requirement is to catalog rivers with widths greater than 50 metres, it is unclear at this time whether the SWOT data will produce width estimates of smaller streams. Though the hydraulic measurements that will be available from new remote sensing technologies remain largely untested, there is great potential to substitute these new width and depth predictions in place of the predictions generated from downstream hydraulic geometry equations (Chapter 4), which will produce more accurate instantaneous cross section velocity estimations.

There is also uncertainty associated with finding the amount of power that can be extracted along a stream reach. This can largely be explained by the challenges of calculating the backwater created by the impedance of hydrokinetic turbines in the river. While the hydraulic effects of tidal kinetic turbines have been studied more intensively, very little research has been put into hydraulic effects of riverine kinetic turbines (Jacobson, 2012). In her pHD dissertation, Lalander (2010) found that hydraulic modeling in Norway had not yet accounted changes in velocities and depths due to turbine installations. Lalander and Leijon (2011) used 2D and 3D MIKE models to simulate hydrokinetic turbine installations in downstream of a hydroelectric dam in Dal River, Norway. They found that turbine installation at that specific site would cause a relatively small increase in depth (2 cm). They also concluded that the computer models were not able to simulate energy losses in the wake of the turbine. To calculate recoverable hydrokinetic power in the United States, Jacobson (2012) applied a recovery factor to theoretical hydropower estimates. The recovery factor was calculated as the ratio of calculated hydrokinetic potential in an undeveloped reach to the calculated hydrokinetic potential of a reach with turbines installed, using Manning's equation. Manning's n was modified in the turbine-installed reaches to account for the additional 'roughness' provided from the turbines by an empirical equation based on device efficiency, blockage ratio, number of devices installed, swept area of devices, original water depth, width of the river, segment length and Manning's natural bed roughness. This ratio was calculated at multiple percentile flow rates and varying slopes and is shown in Equation 5.3 (Jacobson, 2012).

$$RF = \frac{0.002647(Q-200)^{0.3426}}{\sqrt{62.4277S^2}} \exp \frac{-(Log(S)-1.498)^2}{19.87} \tag{5.3}$$

where RF is the recovery factor, Q is flow rate in m^3/s and S is slope. The calculated recovery factor is an admirable attempt to simplify a challenging issue. The recovery factor range is within the expected recovery range and is based on rational logic. However, the empirical fit of the equation does make substantive assumptions and has not been physically tested. Additionally, it is sensitive to multiple parameters, among which there are no current means for accurately determining Manning's roughness and slope for in Canada on a national scale. To date, this author is not aware of any experimental lab field studies on the backwater effects of riverine kinetic turbines. It is clear that there is a need for physically based tests to validate theoretical results and further understanding on the hydraulic effects of riverine kinetic turbines. Additionally, improved data sources, such as SWOT, will greatly increase the accuracy estimates on how much energy is actually extractable on a national scale.

5.3 Validation of Results

Along with tracking the uncertainty in the flow, power and geometry predictions, the predictions were also validated against observed stations. Two datasets were used to validate the results. Firstly, the HYDAT database was used to validate the flow regionalization process. Non-regulated HYDAT stations were included if information was available to link their locations to a stream segment where flow was predicted. In total, 1250 stations were utilized and the RRMSE values were computed across the range of FDC flows as well as the MAF.

The calibration dataset was also utilized to validate the estimates. Though it cannot be considered completely independent of the results, the calibration dataset was used because the MLR-CCA regionalization process gives no preference to stations that are located closer to the river reach being predicted. Additionally, it was the only major data source available that contained geometry and power estimates for comparison. A total of 512 stations were used in the calibration dataset for flow validation while 208 stations were used in the calibration dataset for power, area and velocity validation. The discrepancy in the number of stations is due to the fact that while all the stations that were linked to the stream network contained historical flow data, only a portion of those stations contained geometry data (in the WSC-MD database).

In order to contextualize the impact, the results of the flow validation with the HYDAT and calibration datasets is compared in Figures 5.2 for the relative root mean squared error (RRMSE). Generally the flow estimates have a RRMSE of less than 1 for all the flows when compared in both the HYDAT and calibration datasets. The exception to this is the lower magnitude flows (FDC 50% to 99%). These flows contribute less to the overall volume of flow over the long-term. The low RRMSE for the MAF affirms this. While the results show larger errors when validated in the independant HYDAT dataset, the figure shows that calibration dataset has errors of similar magnitude, particularly for the higher magnitude flows and the MAF.

The RRMSE for power, area, and velocity at a cross section is shown in Figure 5.3, 5.4 and 5.5 respectfully. The RRMSEs are much higher in the power estimates than the area and velocity estimates; this is expected. Any errors in the velocity and area estimates are propagated through and multiplied exponentially according to equation 3.3. This uncertainty does not impact the total reported theoretical hydrokinetic power as that is produced from flow estimates using equation 3.5. The figure also confirms that the power estimates are within one order of magnitude for all flows except the extreme magnitude flows, with the predictions for power at MAF showing particularly promising results.

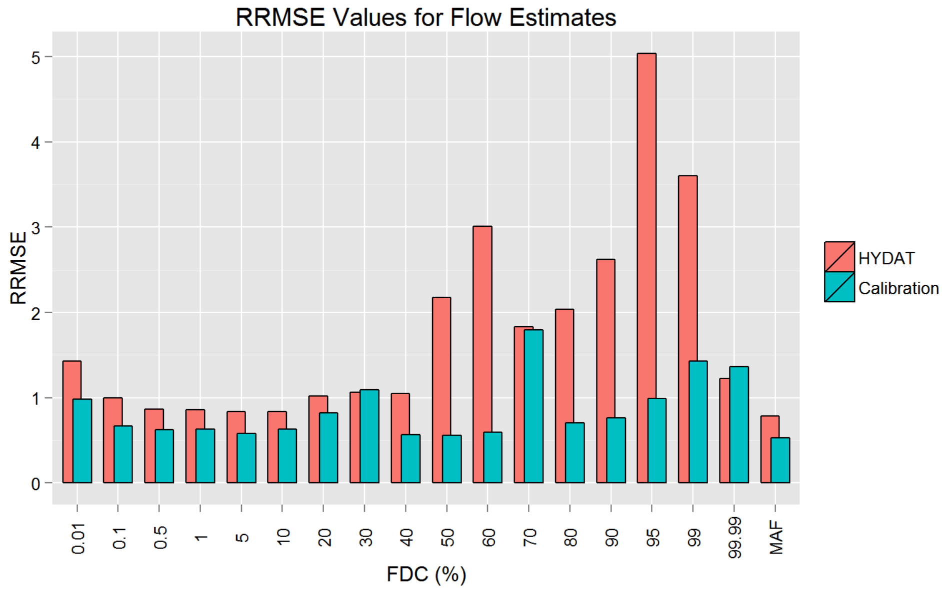

Figure 5.2: Validation of Flow Estimates using the HYDAT and Calibration Dataset: RRMSE

5.4 Impact and Recommendations

This research has successfully predicted hydro and hydrokinetic energy in Canadian Rivers. The predictions confirm that there are massive hydro resources in Canada. The expected hydro power in Canadian rivers is 710 GW with a 97.5% certainty that there is at least 430 GW. In addition, the predictions have been regionalized by region and province, showing the provinces with the greatest power is British Columbia and Quebec. These values do not account for power that has already been developed. Hence it is expected that British Columbia has more potential to develop sites as the region contains greater power potential, and they have only developed 12 GW.

In addition to total power estimates, maps have been created for industry analysts to find sites with the greatest potential. These datasets contain power, flow, velocity and geometry estimates for specific exceedence intervals, monthly values and well as mean annual values for every cross section in Canada. Confidence intervals are provided alongside the estimates.

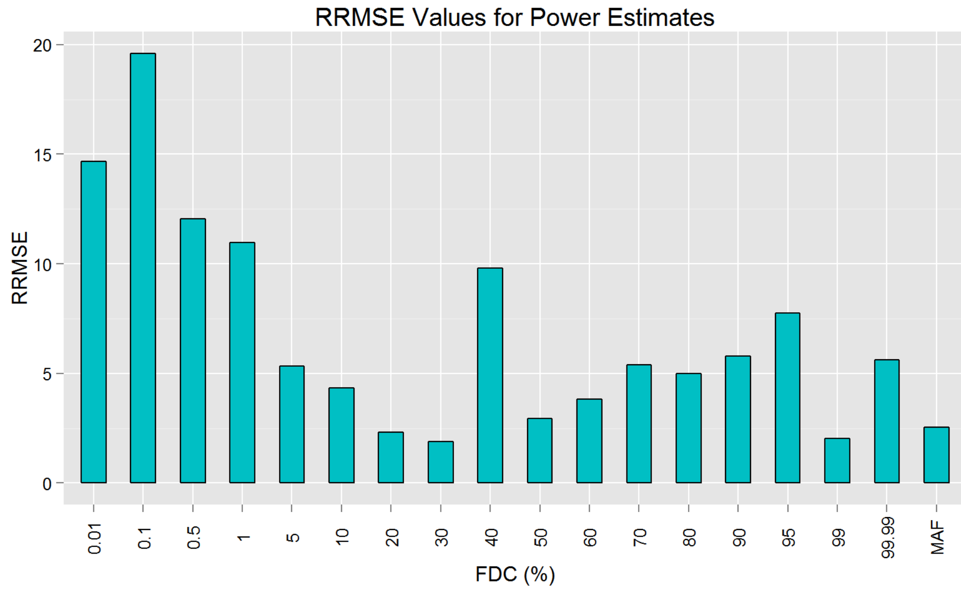

Figure 5.3: Validation of Power Estimates using the Calibration Dataset: RRMSE

These datasets will be available to the public in the near future through NRCan. Users will be able to query the data to suit their needs. For example, a user could filter the data to only include river reaches of pre-specified minimum velocities and depths that are near utilities or communities. When using the dataset, it is recommended that the user consider the confidence intervals of the predictions as the degree of confidence is dependant on how much data was available for a given region. Additionally, a large amount of the prediction uncertainty is due to channel geometry estimates. If a high potential region is identified, this uncertainty can be improved with channel width estimates obtained from areal photography or satellite imagery.

There are also areas of research that show great potential for the improvement of the study results:

1. Due to time constraints, the CCA technique used for regionalization was limited to a small number of FDC exceedence values. The number and selection of these exceedence values could be optimized to find more robust neighbourhood definitions. Additionally,

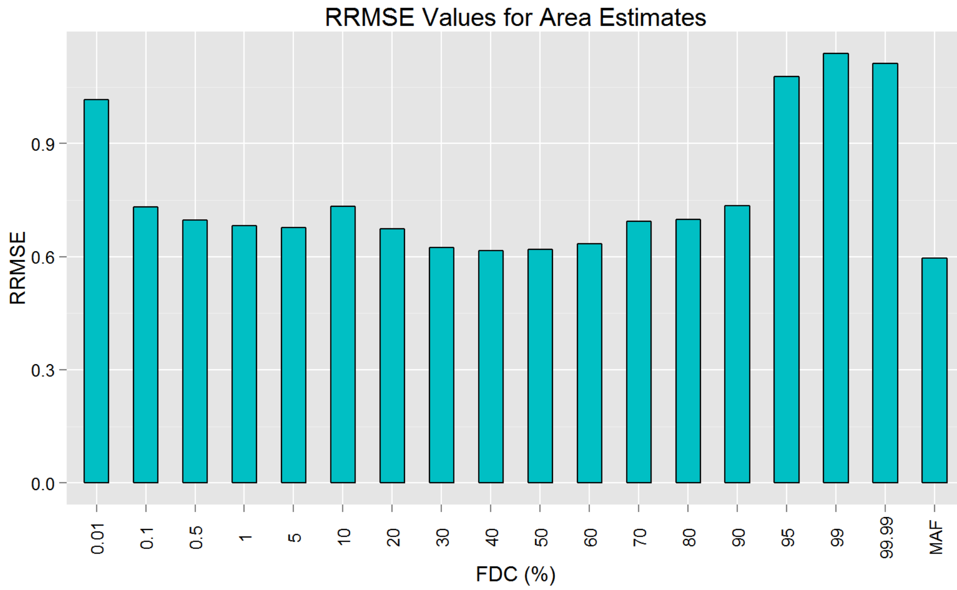

Figure 5.4: Validation of Area Estimates using the Calibration Dataset: RRMSE

other regionalization techniques such as Region of Influence could be implemented and compared.

2. α was optimized on a regional basis (ie. a single optimized α value was used when performing CCA in for all of the stream reaches in a region). Different techniques could be experimented with to vary α with flow regimes or on a reach-by-reach basis

3. An incremental methodology could be developed to more easily re-calculate the estimates on a small scale if more accurate geometry estimates are provided. This would allow practitioners greater confidence of estimates for specific reaches

4. There is limited research on the hydraulic effects of kinetic turbine installation in river reaches. Lab and field tests should be completed to validate and assist in computer modeling efforts.

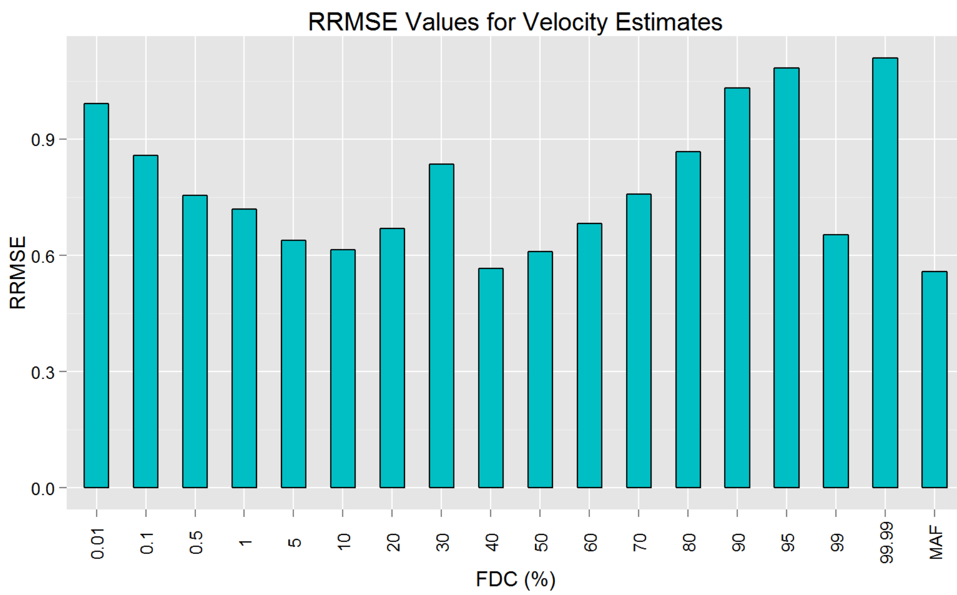

Figure 5.5: Validation of Velocity Estimates using the Calibration Dataset: RRMSE

Chapter 6

Conclusions

6.1 Conclusions and Recommendations

There is renewed interest in extracting electric power from Canada's many rivers. Hydrokinetic turbines, also referred to as zero-head turbines, differ from classic hydro in that they generate electricity from moving water whereas hydro turbines generate electricity from the pressure of the water created in an upstream reservoir. There are some drawbacks to hydrokinetic power; they are less efficient than hydro turbines (Lalander and Leijon, 2011) and produce less consistent power as water cannot be stored upstream.However, there are several advantages to hydrokinetic power over hydro power: hydrokinetic power doesn't inundate upstream areas; it requires less upfront capital; and it can be deployed more easily in remote areas. Canada aims to become a world leader in hydro kinetic technologies (Marine Renewable Energy Technology Roadmap Steering Committee, 2011) and has encouraged private development of hydro kinetic turbines as well as resource assessment studies. To this end, Natural Resources Canada commissioned the National Research Council - Oceans, Coasts and Rivers Engineering division (NRC-OCRE) to undertake a study to assess and quantify riverine hydrokinetic potential. The study was divided into three phases: (I) methodology and data review, (II) methodology validation, and (III) resource assessment determination. Phase II and III were done in collaboration with the University of Ottawa and this thesis discusses the methodologies and results of the resources assessment.

There were two main objectives of the study. The first objective was to provide maps that will aid developers in selecting sites with the highest potential. Hydrokinetic power was calculated at each cross section by predicting flows and geometries at ungauged locations to find velocity and area, which are the requirements to calculate hydrokinetic power. The second objective was to calculate estimates of the total nationwide potential hydrokinetic power available for extraction. Due to the uncertainty of many critical components (slope, bed roughness) the approach taken to solve this objective was to use total theoretical hydro power.

Flow and geometry estimates at ungauged locations are required to assess hydrokinetic potential on a nationwide scale. Flow was characterized by a 17-point Flow Duration Curve (FDC), mean monthly flows and mean annual flow. These flow characterizations were estimated at ungauged locations using regionalization through multiple linear regression coupled with Canonical correlation analysis (MLR-CCA) with flow as the dependant variable and watershed characteristics (drainage area, perimeter, % impermeable area, etc.) as the independant variables. Cross section widths, depths and areas were estimated at ungauged locations by applying estimated flows to downstream hydraulic geometry equations, which empirically relate widths, depths and areas to flows at a specific re-currence interval. Cross section velocity was found using estimated flows and areas in the continuity equation.

Critical datasets were required to estimate cross section power at ungauged locations. The HYDAT database from Environment Canada was used for flow data. The Water Survey of Canada - Measurement Database (WSC-MD) was used to create downstream hydraulic geometry equations. The National Atlas of Canada stream network consists of 365,000 stream segments and was used to identify streams to be analyzed. Voronoi analysis was also applied to the stream network in order to efficiently delineate watersheds. Digital elevation models, and digital physiographic and climate maps were collected. GreenKenue, a GIS program, was used to map watershed attributes by overlaying watershed boundaries on these maps and calculating the percentage of area covered by these attributes.

Flow regionalization is the process of comparing watershed characteristics of gauged and ungauged stations to predict flow at ungauged stations. MLR was applied to each ungauged station using a calibration dataset comprised of gauged stations and their watershed characteristics. The selection of stations to use in each calibration was determined by Canonical correlation analysis, which finds the watersheds in the calibration dataset that are most statistically similar to the ungauged watershed where flow is being estimated. The 'cut-off line' of which stations to include, and which stations to exclude in the MLR is defined by an α value. A very small α considers all of the stations in the calibration dataset to be statistically significant (and thus included in MLR) while a very large α will determine only a few gauged stations to be statistically similar to the ungauged basin. An optimal α value maximized the accuracy of the MLR while not excluding too many stations. Ungauged station flows were also augmented with observed flows when possible. If a HYDAT station was found upstream of the ungauged location, regression was only performed on the watershed area not influenced by the HYDAT station. The regression and HYDAT flows were then summed to get a more accurate flow at the ungauged site.

Flow estimates were applied to downstream hydraulic geometry equations to find cross section velocity. The confidence in the velocity estimates was tracked by applying a Monte Carlo simulation throughout the flow and geometry regressions. 1000 values, following a log-normal distribution, were sampled from the MLR-CCA flow estimates and the augmented HYDAT flows. These flow distributions were applied to the downstream hydraulic geometry equations and re-sampled again, resulting in geometry and velocity estimates in which the 95% confidence intervals were known.

Investigation into downstream hydraulic geometry equations was undertaken in order to achieve more accurate geometry predictions. Bankfull flow and geometry data were gathered from 634 gauged stations across Canada (from the HYDAT and WSC-MD) and categorized according to various surface geology and vegetation classifications. A jack-knife, or leave-one-out, analysis was applied for each classification to determine if any classification was better able to predict bankfull width at an ungauged location. Classifying stations by soil drainage characteristics improved the accuracy of downstream hydraulic geometry equations by 8.4% as measured by the NASH-Sutcliffe coefficient. Almost all of the 15 classifications improved the accuracy of the downstream hydraulic geometry equations. Multiple regression using categorical classifications along flow magnitudes was also tested but found to not provide any significant results.

Total hydrokinetic power was estimated from two methods. The first method equated hydrokinetic power to total hydro power, shown by equation 3.5.

$$P_H = \gamma H Q \tag{3.5}$$

where P_H is the total power, H is the change in elevation head, γ is the specific weight of water, and Q is the volumetric flow rate. Ungauged flow estimates (as previously described) were used for Q and the change in elevation head was taken from the DEM dataset. The low resolution of the DEM meant that while accuracy of power estimates on individual stream reaches were low, the overall regional and power estimates were very accurate. The second method was to calculate the change in hydraulic head in terms of velocity, as shown in equation 3.6.

$$P_K = \gamma H_V Q \tag{3.6}$$

where H_V is the head calculated from a modified Manning's equation. Theoretically, H_V and H are the same value. The benefit in using H_V is that it frames the equation in terms of velocity, the main component being extracted for hydrokinetic power, and allows the effects of hydrokinetic turbine installations on the stream channel to be accounted for. However, due to greater uncertainty in H_V estimates (more accurate slope, Manning's roughness and velocity estimates are required) when compared to the elevation head estimates, the total hydro power equation provides more reliable estimates.

Results show that there is an expected 710 GW of available power in Canada's rivers, with a 97.5% confidence that there is at least 433 GW of available power. When summarized by province, BC and Quebec were shown to have the most power potential. Alongside these estimates, power and velocity maps were produced to show users where the location of stream reaches with the highest power potential.

Calculating the amount of recoverable energy available is limited by the confidence in data sources and the amount of research on hydraulic effects of hydrokinetic turbines on rivers. New remote sensing techniques, such as SWOT (Fu et al., 2009) may hold promise for achieving better slope and geometry data of rivers, which are critical components for finding extractable hydrokinetic energy along a reach. Further research is recommended on understanding the hydraulic effects of hydrokinetic turbine installations. Hydraulic computer models have been built by Jacobson (2012) and Lalander (2010) and are good steps towards understanding the physics, but further modeling is required. There is also a great need for lab and field experiments to be completed in order to test and validate these efforts.

Chapter 7

Bibliography

AAFC (1996). Soil landscapes of canada. `http://sis.agr.gc.ca/cansis/nsdb/slc/index.html`.

AAFC (2000). *Canadian Soil Information System: National Soil Database*. Agriculture and Agri-Food Canada.

Acres Consulting Services Limited (1984). Hydrologic design methodologies for small-scale hydro at ungauged sites – phase i: Applications manual. Technical report, Environment Canada, Inland Waters Directorate.

Adams, R. and Spotila, J. (2005). The form and function of headwater streams based on field and modeling investigations in the southern appalachian mountains. *Earth Surface Processes and Landforms*, 30(12):1521–1546.

Agriculture and Agri-Food Canada (AAFC) (2008). A National Ecological Framework For Canada - Canadian Ecodistrict Climate Normals (1961 - 1990). Online. http://sis.agr.gc.ca/cansis/nsdb/ecostrat/intro.html.

Anderson, R., Bledsoe, B., and Hession, W. (2004). Width of streams and rivers in response to vegetation, bank material, and other factors. *Journal of the American Water Resources Association*, 40(5):1159–1172.

Andreadis, K. M., Clark, E. A., Lettenmaier, D. P., and Alsdorf, D. E. (2007). Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. *Geophysical Research Letters*, 34(10):L10403.

Andrews, E. (1980). Effective and bankfull discharges of streams in the yampa river basin, colorado and wyoming. *Journal of Hydrology*, 46(3-4):311–330.

Andrews, E. (1984). Bed-material entrainment and hydraulic geometry of gravel-bed rivers in colorado. *Geological Society of America Bulletin*, 95(3):371–378.

BC Hydro (2013). Annual report. `http://www.bchydro.com/content/dam/BCHydro/customer-portal/documents/corporate/accountability-reports/financial-reports/annual-reports/bc-hydro-annual-report-2013.pdf`.

Benson, M. and Thomas, D. (1966). Definition of dominant discharge. *International Association of Scientific Hydrology – Bulletin*, 11(2):76 – 82.

Blench, T. (1957). *Regime behaviour of canals and rivers*. Butterworths Scientific Publications.

Bobee, B., Mathier, L., Perron, H., Trudel, P., Rasmussen, P., Cavadias, G., Bernier, J., Nguyen, V.-T.-V., Pandey, G., Ashkar, F., Ouarda, T., Adamowski, K., Alila, Y., Daviau, J., Gingras, D., Liang, G., Rousselle, J., Birikundavyi, S., Ribeiro-Corra, J., Roy, R., and Pilon, P. (1996). Inter-comparison of regional flood frequency procedures for canadian rivers. *Journal of Hydrology*, 186(1-4):85–103.

Brockman, R. R., Agouridis, C. T., Workman, S. R., Ormsbee, L. E., and Fogle, A. W.

(2012). Bankfull regional curves for the inner and outer bluegrass regions of kentucky. *Journal of the American Water Resources Association*, 48(2):391 – 406.

Camilli, G. and Hopkins, K. D. (1978). Applicability of chi-square to 2× 2 contingency tables with small expected cell frequencies. *Psychological Bulletin*, 85(1):163.

Canada, N. R. (2010). Geoscape canada: A map of canadas earth materials. Online.

Cavadias, G. (1990). The canonical correlation approach to regional flood estimation. *Regionalization in Hydrology*, 191:171–178.

Centre for Topographic Information (CTI) (2000). Canadian Digital Elevation Data (CDED). Internet. http://www.geobase.ca/geobase/en/data/cded/index.html.

Centre for Topographic Information (CTI) (2009). Land cover, circa 2000-vector data product, edition 1.0. Technical report, Natural Resources Canada.

Chang, H. (1979). Minimum stream power and river channel patterns. *Journal of Hydrology*, 41(3):303–327.

Chaplin, J. (2005). *Development of regional curves relating bankfull-channel geometry and discharge to drainage area for streams in Pennsylvania and selected areas of Maryland*. US Geological Survey.

Christodoulou, G. C. (2013). Equivalent roughness of submerged obstacles in open channel flows. *Journal of Hydraulic Engineering*.

Cinotto, P. (2003). *Development of regional curves of bankfull-channel geometry and discharge for streams in the non-urban, Piedmont Physiographic Province, Pennsylvania and Maryland*. US Department of the Interior, US Geological Survey.

Clark, D. (1975). Understanding canonical correlation analysis: Concepts and techniques in modern geography no. 3. Online.

Copeland, R., Soar, P., and Thorne, C. (2005). Channel-forming discharge and hydraulic geometry width predictors in meandering sand-bed rivers. In *Proc., 2005 World Water and Environmental Resources Congress: Impacts of Global Change*. ASCE Reston, Va.

Copeland, R. R., Biedenharn, D. S., and Fischenich, J. C. (2000). *Channel-forming discharge*. US Army Corps of Engineers.

Davies, T. and Sutherland, A. (1980). Resistance to flow past deformable boundaries. *Earth Surface Processes*, 5(2):175–179.

Deng, Z. and Zhang, K. (1994). Morphologic equations based on the principle of maximum entropy. *International Journal of Sediment Research*, 1.

Doll, B. A., Wise-Frederick, D. E., Buckner, C. M., Wilkerson, S. D., Harman, W. A., Smith, R. E., and Spooner, J. (2002). Hydraulic geometry relationships for urban streams throughout the piedmont of north carolina. *Journal of the American Water Resources Association*, 38(3):641 – 651.

Dou, G. (1964). Hydraulic geometry of plain alluvial rivers and tidal river mouth. *J Hydraul Eng*, 2:1–13.

Dunne, T. and Leopold, L. (1978). *Water in environmental planning*. WH Freeman.

Durand, M., Fu, L.-L., Lettenmaier, D. P., Alsdorf, D. E., Rodriguez, E., and Esteban-Fernandez, D. (2010a). The surface water and ocean topography mission: Observing terrestrial surface water and oceanic submesoscale eddies. *Proceedings of the IEEE*, 98(5):766–779.

Durand, M., Rodriguez, E., Alsdorf, D. E., and Trigg, M. (2010b). Estimating river depth from remote sensing swath interferometry measurements of river height, slope, and width. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 3(1):20–31.

Dury, G. (1961). Bankfull discharge: an example of its statistical relationships. *Hydrological Sciences Journal*, 6(3):48–55.

Dury, G. (1976). Discharge prediction, present and former, from channel dimensions. *Journal of Hydrology*, 30(3):219 – 245.

Eaton, B. and Church, M. (2007). Predicting downstream hydraulic geometry: a test of rational regime theory. *Journal of Geophysical Research*, 112(F3):03025 – 1.

Ebisa Fola, M. and Rennie, C. D. (2010). Downstream hydraulic geometry of clay-dominated cohesive bed rivers. *Journal of Hydraulic Engineering*, 136(8):524 – 527.

Ellis, E. R. and Church, M. (2005). Hydraulic geometry of secondary channels of lower fraser river, british columbia, from acoustic doppler profiling. *Water Resources Research*, 41(8):1 – 15.

Emmett, W. W. and Wolman, M. G. (2001). Effective discharge and gravel-bed rivers. *Earth Surface Processes and Landforms*, 26(13):1369 – 1380.

Environment Canada (2010). *EC Data Explorer Reference Manual - Version 1.2.* http://www.ec.gc.ca/rhc-wsc/default.asp?lang=En&n=0A47D72F-1.

ESRI (2010). *ArcGIS Help Library - ArcGIS 10.* ESRI.

Fu, L.-L., Alsdorf, D., Rodriguez, E., Morrow, R., Mognard, N., Lambin, J., Vaze, P., and Lafon, T. (2009). The swot (surface water and ocean topography) mission: spaceborne radar interferometry for oceanographic and hydrological applications. *Space Agency, Venice, Italy*, pages 21–25.

Harman, C., Stewardson, M., and DeRose, R. (2008). Variability and uncertainty in reach bankfull hydraulic geometry. *Journal of Hydrology*, 351:13 – 25.

Harman, W., Jennings, G., Patterson, J., Clinton, D., Slate, L., Jessup, A., Everhart, J., and Smith, R. (1999). Bankfull hydraulic geometry relationships for north carolina streams. *Wildland Hydrology*, pages 401–408.

Harvey, A. M. (1969). Channel capacity and adjustment of streams to hydrologic regime. *Journal of Hydrology*, 8(1):82 – 98.

Helsel, D. and Hirsch, R. (1991). Statistical methods in water resources. In *Techniques in Water Resources Investigations of the United States Geological Survey*, volume 4, chapter A3 Hydrologic Analysis and Interpretation. USGS.

Henderson, F. (1966). *Open channel flow*. Macmillan series in civil engineering. Macmillan.

Hession, W., Pizzuto, J., Johnson, T., and Horwitz, R. (2003). Influence of bank vegetation on channel morphology in rural and urban watersheds. *Geology*, 31(2):147–150.

Hey, R. (1978). Determinate hydraulic geometry of river channels. *Journal of the Hydraulics Division*, 104(6):869–885.

Hey, R. and Thorne, C. (1986). Stable channels with mobile gravel beds. *Journal of Hydraulic Engineering*, 112(8):671–689.

Huang, H. and Nanson, G. (1998). The influence of bank strength on channel geometry: an integrated analysis of some observations. *Earth Surface Processes and Landforms*, 23(10):865–876.

Huang, H. and Nanson, G. (2000). Hydraulic geometry and maximum flow efficiency as products of the principle of least action. *Earth Surface Processes and Landforms*, 25(1):1–16.

HydroQuebec (2012). Annual report. `http://www.hydroquebec.com/publications/en/annual_report/pdf/annual-report-2012.pdf`.

Jacobson, P. (2012). Assessment and mapping of the riverine hydrokinetic energy resource in the continental united states. Technical report, Electric Power Research Institute.

Jenkinson, R. W. (2010). Assessment of Canadas Hydrokinetic Power Potential – Phase I Report: Methodology and Data Review. Technical report, Natural Resources Canada and National Research Council - Canadian Hydraulics Centre (NRC-CHC).

Jenkinson, R. W. and Bomhof, J. (2012). Assessment of Canadas Hydrokinetic Power Potential – Phase II Methodology Validation. Technical report, Natural Resources Canada and National Research Council - Ocean Coastal and River Engineering (NRC-OCRE).

Johnson, P. A. and Fecko, B. J. (2008). Regional channel geometry equations; a statistical comparison for physiographic provinces in the eastern us. *River Research and Applications*, 24(6):823 – 834.

Keaton, J., Messinger, T., and Doheny, E. (2005). *Development and analysis of regional curves for streams in the non-urban valley and ridge physiographic province, Maryland, Virginia, and West Virginia*. US Department of the Interior, US Geological Survey.

Kerr Wood Leidal Associates (2008). Computer Modelling for Water - Finding Run-of-River Hydroelectricty Potential in B.C. *Canadian Consulting Engineer*, pages 16–17.

Kirkby, M. (1976). *Maximum sediment efficiency as a criterion for alluvial channels*. University of Leeds, School of Geography.

Klonsky, L. and Vogel, R. M. (2011). Effective measures of" effective" discharge. *Journal of Geology*, 119(1):1–14.

Knighton, A. (1977). Short-term changes in hydraulic geometry. *John Wiley & Sons: New York*, pages 101–119.

Knighton, D. (1974). Variation in width-discharge relation and some implications for hydraulic geometry. *Geological Society of America Bulletin*, 85(7):1069–1076.

Knighton, D. (1998). *Fluvial Forms and Processes: A New Perspective*. Arnold Publishers.

Lacey, G. (1930). Stable channels in alluvium (includes appendices). In *Minutes of the Proceedings*, volume 229, pages 259–292. Thomas Telford.

Lalander, E. (2010). *Modelling the hydrokinetic energy resource for in-stream energy converters*. PhD thesis, Uppsala University.

Lalander, E. and Leijon, M. (2011). In-stream energy converters in a river–effects on upstream hydropower station. *Renewable energy*, 36(1):399–404.

Lane, E. (1937). Stable channels in erodible materials. *Transactions of the American Society of Civil Engineers*, 102(1):123–142.

Lane, E. (1955). Design of stable channels. *Transactions of the American Society of Civil Engineers*, 120(1):1234–1260.

Langbein, W. (1964). Geometry of river channels. *Journal of the Hydraulics Division, ASCE*, 90(HY2).

Langbein, W. and Leopold, L. (1964). Quasi-equilibrium states in channel morphology. *American Journal of Science*, 262(6):782–&.

Lawlor, S. M. (2004). Determination of channel-morphology characteristics, bankfull discharge, and various design-peak discharges in western montana. Technical report, US Geological Survey.

Leopold, L. and Langbein, W. (1962). *The concept of entropy in landscape evolution*. US Government Printing Office Washington, DC.

Leopold, L., Wolman, M., and Miller, J. (1964). *Fluvial Processes in Geomorphology*. W.H. Freeman and Company.

Leopold, L. B. and Maddock, T. (1953). The hydraulic geometry of stream channels and some physiographic implications.

Leopold, L. B., Wolman, M. G., Wolman, M. G., and Wolman, M. G. (1957). *River channel patterns: braided, meandering, and straight*. US Government Printing Office Washington, DC.

Li, R. (1975). Mathematical modeling of response from small watershed. *Unpublished Ph.D. Dissertation*, page 212.

Lindley, E. (1919). Regime channels. In *proc. Punjab Engineering Congress*, volume 7.

Lumia, R. (1991). *Regionalization of flood discharges for rural, unregulated streams in New York, excluding Long Island*. US Geological Survey.

Lumia, R., Freehafer, D., and Smith, M. (2006). Magnitude and frequency of floods in new york. Technical report, U. S. Geological Survey.

Maidment, D. (1993). *Handbook of hydrology*. McGraw-Hill New York.

Manitoba Hydro (2012). Annual report. `http://www.hydro.mb.ca/corporate/ar/index.shtml`.

Marine Renewable Energy Technology Roadmap Steering Committee (2011). Charting the course : Canada's marine renewable energy technology roadmap. Technical report, Natural Resources Canada.

McCandless, T. (2003). Maryland stream survey: Bankfull discharge and channel characteristics of streams in the allegheny plateau and the valley and ridge hydrologic regions. *US Fish and Wildlife Service CBFO-S03-01, Annapolis, Maryland*.

Mersel, M. K., Smith, L. C., Andreadis, K. M., and Durand, M. T. (2013). Estimation of river depth from remotely sensed hydraulic relationships. *Water Resources Research*.

Millar, R. (2005). Theoretical regime equations for mobile gravel-bed rivers with stable banks. *Geomorphology*, 64(3):207–220.

Miller, G., Franceschi, J., Lese, W., and Rico, J. (1986). The allocation of kinetic hydro energy conversion systems (khecs) in usa drainage basins: Regional resource and potential power. *Final report, NYUDAS*, pages 86–151.

Mohamoud, Y. (2008). Prediction of daily flow duration curves and streamflow for ungauged catchments using regional flow duration curves. *Hydrological Sciences Journal*, 53(4):706–724.

Molnar, P. and Ramirez, J. A. (2002). On downstream hydraulic geometry and optimal energy expenditure: case study of the ashley and taieri rivers. *Journal of Hydrology*, 259(1):105–115.

Monk, R., Joyce, S., and Homenuke, M. (2009). Rapid hydropower assessment model: Identify hydroelectric sites using geographic information systems. In *Small Hydro 2009 Conference, Vancouver, B.C., April 28-29, 2009*. Kerr Wood Leidal Associates Ltd.

Montgomery, D. (1997). River management: What's best on the banks? *Nature*, 388(6640):328–329.

Montgomery, D. and Gran, K. (2001). Downstream variations in the width of bedrock channels. *Water Resources Research*, 37(6):1841–1846.

Moriasi, D., Arnold, J., Van Liew, M., Bingner, R., Harmel, R., and Veith, T. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3):885–900.

Mulvihill, C. I. and Baldigo, B. P. (2012). Optimizing bankfull discharge and hydraulic geometry relations for streams in new york state. *Journal of the American Water Resources Association*, 48(3):449 – 463.

Murgatroyd, A. and Ternan, J. (1983). The impact of afforestation on stream bank erosion and channel form. *Earth Surface Processes and Landforms*, 8(4):357–369.

Natural Resources Canada (2008). Atlas of canada 1,000,000 national frameworks data, hydrology - drainage network. Online.

Navratil, O., Albert, M.-B., Herouin, E., and Gresillon, J.-M. (2006). Determination of bankfull discharge magnitude and frequency: Comparison of methods on 16 gravel-bed river reaches. *Earth Surface Processes and Landforms*, 31(11):1345 – 1363.

Neill, C. (1964). Alluvial processes and river channel regime. *Trans. Eng. Inst. Can*, 7:3–8.

Newson, M. and Harrison, J. (1978). Channel studies in the plynlimon experimental catchments. *Institute of Hydrology*.

Nixon and Lacey, G. (1959). The study of bankfull discharges of rivers in england and wales. In *ICE Proceedings*, volume 14, pages 395–425. Ice Virtual Library.

NRC-CHC (2010). *Green Kenue Reference Manual.* National Research Council Canada - Canadian Hydraulics Centre, Ottawa, Ontario.

Osterkamp, W. (1982). Perennial-steamflow characteristics related to channel geometry and sediment in missouri river basin. *US Geological Survey professional paper.*

Ouarda, T. B., Girard, C., Cavadias, G. S., and Bobe, B. (2001a). Regional flood frequency estimation with canonical correlation analysis. *Journal of Hydrology*, 254:157 – 173.

Ouarda, T. B. M. J., Girard, C., Cavadias, G. S., and Bobe, B. (2001b). Regional flood frequency estimation with canonical correlation analysis. *Journal of Hydrology*, 254(1-4):157–173.

Park, C. C. (1977). World-wide variations in hydraulic geometry exponents of stream channels: An analysis and some observations. *Journal of Hydrology*, 33:133 – 146.

Park, C. C. (1978). World-wide variations in hydraulic geometry exponents of stream channels: An analysis and some observations reply. *Journal of Hydrology*, 39:199 – 202.

Phillips, P. J. and Harlin, J. M. (1984). Spatial dependency of hydraulic geometry exponents in a subalpine stream. *Journal of Hydrology*, 71:277 – 283.

Pickup, G. and Warner, R. (1976). Effects of hydrologic regime on magnitude and frequency of dominant discharge. *Journal of Hydrology*, 29(1-2):51 – 75.

Ponton, J. (1972). Hydraulic geometry in the green and birkenhead basins, british columbia. *Mountain Geomorphology: Geomorphological Processes in the Canadian Cordillera*, pages 151–160.

Radecki-Pawlik, A. (2002). Bankfull discharge in mountain streams: Theory and practice. *Earth Surface Processes and Landforms*, 27(2):115 – 123.

Rhodes, D. (1977). The bfm diagram: graphical representation and interpretation of at-a-station hydraulic geometry. *American Journal of Science*, 277:73 – 95.

Rhodes, D. D. (1978). World-wide variations in hydraulic geometry exponents of stream channels: An analysis and some observations: Comments. *Journal of Hydrology*, 39:193 – 197.

Ribeiro-Correa, J., Cavadias, G., Clement, B., and Rousselle, J. (1995). Identification of hydrological neighborhoods using canonical correlation analysis. *Journal of Hydrology*, 173(1-4):71–89.

Schumm, S. (1960). *The shape of alluvial channels in relation to sediment type*. US Government Printing Office.

Searcy, J. (1959a). *Flow-duration curves*. US Government Printing Office.

Searcy, J. K. (1959b). Flow-duration curves. Water Supply Paper 1542-A, U. S. Geological Survey.

Shu, C. and Ouarda, T. (2009). Improved methods for daily streamflow estimates at ungauged sites. *Water Resources Research*. In Press.

Singh, V. (2003). On the theories of hydraulic geometry. *International journal of sediment Research*, 18(3):196–218.

Singh, V., Yang, C., and Deng, Z. (2003). Downstream hydraulic geometry relations: 1. Theoretical development. *Water Resources Research*, 39(12).

Smith, T. R. (1974). A derivation of the hydraulic geometry of steady-state channels from conservation principles and sediment transport laws. *Journal of Geology*, 82(1):98 – 104.

Stevens, M. and Nordin Jr, C. (1987). Critique of the regime theory for alluvial channels. *Journal of Hydraulic Engineering*, 113(11):1359–1380.

Stott, T. (1998). A comparison of stream bank erosion processes on forested and moorland streams in the balquhidder catchments, central scotland. *Earth Surface Processes and Landforms*, 22(4):383–399.

Thorne, C. R. and Tovey, N. K. (1981). Stability of composite river banks. *Earth Surface Processes and Landforms*, 6(5):469–484.

Turner, R., Clague, J., and Hastings, N. (2003). Geoscape Canada : a map of Canada's earth materials. Miscellaneous Report 81, Geological Survey of Canada.

UMA Group (1980). An Evaluation of The Kinetic Energy of Canadian Rivers & Estuaries. Technical report, Canadian National Research Council - Canadian Hydraulics Centre.

U.S. Army Corps of Engineers (1994). *CHANNEL STABILITY ASSESSMENT FOR FLOOD CONTROL PROJECTS*. Washington, DC 20314-1000, em 1110-2-1418 edition.

USDOE (2002a). Low Head/Low Power Hydropower Resources Assessment of the Arkansas White Red Hydrologic Region. Technical Report DOE/ID-11019, U. S. Department of Energy.

USDOE (2002b). Low Head/Low Power Hydropower Resources Assessment of the Pacific Northwest Hydrologic Region. Technical report, U. S. Department of Energy.

USDOE (2003). Low Head/Low Power Hydropower Resources Assessment of the North Atlantic and Mid-Atlantic Hydrologic Regions. Technical Report DOE/ID-11077, U. S. Department of Energy.

USDOE (2004). Water Energy Resources of the United States with Emphasis on Low Head/Low Power Resources. Technical Report DOE/ID-11111, U.S. Department of Energy Idaho Operations Office.

USDOE (2006). Feasibility Assessment of the Water Energy Resources of the United States for New Low Power and Small Hydro Classes of Hydroelectric Plants. Technical Report DOE-ID-11263, US Department of Energy, Energy Efficiency and Renewable Energy; Wind and Hydropower Technologies.

USGS and USEPA (2006). *NHDPlus User Guide*. US Geological Survey and US Environmental Protection Agency.

Vachtman, D. and Laronne, J. B. (2013). Hydraulic geometry of cohesive channels undergoing base level drop. *Geomorphology*.

Vermaak, H. J., Kusakana, K., and Koko, S. P. (2013). Status of micro-hydrokinetic river technology in rural applications: A review of literature. *Renewable and Sustainable Energy Reviews*, 29:625–633.

Watson, C., Biedenharn, D., and Scott, S. (1999). Channel rehabilitation: Processes, design, and implementation. Technical report, DTIC Document.

Westergard, B., Mulvihill, C., Ernst, A., and Baldigo, B. (2005). Regionalized equations for bankfull-discharge and channel characteristics of streams in new york state: Hydrologic region 5 in central new york. Technical report, U. S. Geological Survey.

White, D., Bettess, R., and Paris, E. (1981). *River regime based on sediment transport concepts*. Hydraulics Research Station.

Williams, G. P. (1978). Bank-full discharge of rivers. *Water Resources Research*, 14(6):1141 – 1154.

Wohl, E. and David, G. (2008). Consistency of scaling relations among bedrock and alluvial channels. *Journal of Geophysical Research*, 113(F4):F04013.

Wolh, E. (2004). Limits of downstream hydraulic geometry. *Geology*.

Wolman, M. G. (1955). *The natural channel of Brandywine creek, Pennsylvania*. US Government Printing Office.

Wolman, M. G. and Miller, J. P. (1960). Magnitude and frequency of forces in geomorphic processes. *Journal of Geology*, 68(1):54 – 74.

Yalin, M. and Da Silva, A. (1999). Regime channels in cohesionless alluvium. *Journal of Hydraulic Research*, 37(6):725–742.

Yang, C. and Song, C. (1979). Theory of minimum rate of energy dissipation. *Journal of the Hydraulics Division*, 105(7):769–784.

Appendices

Appendix A

R-Scripts

A.1 Script to predict aggregate flows using MLR-CCA

```
#********************************************
#Name: Phase_3.R
#Written By: James Bomhof and Wayne Jenkinson
#Date Written: January 8, 2013
#Description: Performs multiple regression for
#flow and power prediction using Canonical Correlation Analysis
#
#Edite by: Ivana Vouk, July 29, 2013
#********************************************
print("Application of CCA to Phase III Datasets")
Sys.time()
options(scipen=5)

#Get libraries and set options
rm(list=ls())
#Libraries
library(getopt)
source("rlib/libCCA_MR.R")
source("rlib/libENSIM_IO.r")


#Get the details about region, output, etc. in n set of arguments (for batching)
#validation database
spec <- c("hkdb.file", "h", 1, "character")
#region code (to filter validation database)
spec <- rbind(spec, c("region.code", "c", 2, "integer"))
#region file - Phase III dataset
spec <- rbind(spec, c("region.file", "f", 3, "character"))
```

```
#diretory to place output files
spec <- rbind(spec, c("out.dir", "o", 4, "character"))
#comparative analysis - FDC (1), MMF (4), Specific Power(3), Total Power(2)
spec <- rbind(spec, c("y.type", "y", 5, "integer"))
#what attributes get used in regression analysis
spec <- rbind(spec, c("attrib.used","u",6,"character"))
#optimal alpha index
spec <- rbind(spec, c("alpha.index", "a", 7, "integer"))
#5 letter name of region
spec <- rbind(spec, c("name", "n", 8, "character"))
#toggles test or analysis dataset
spec <- rbind(spec, c("toggle", "t", 9, "character"))
opt = getopt(spec)


#specify the default values if the arguments are not supplied.

if ( is.null(opt$hkdb.file ) ) {
  opt$hkdb.file =
    "../../SourceFiles/ValidationStations/HKDB_FDC_MMF_PVA_ALLPHYSIO_rev4.csv"
}
if ( is.null(opt$region.code ) ) {
  opt$region.code = NA
} #applies only when toggle=1
if ( is.null(opt$region.file ) ) {
  opt$region.file =
    "../../SourceFiles/CHCVoronoiAggregationCatchmentsV4/missi/AggregateCatchmentTable.tb0"
}
if ( is.null(opt$out.dir ) ) {
  opt$out.dir = "../Aggregate/AggEstimates/FDCAGGmissi"
}
if ( is.null(opt$y.type ) ) {
  opt$y.type = 1
  } #type of comparative analysis - FDC (1), Total Power(2),Specific Power(3),
    #MMF (4), MMP(5),MMSP(6)
if ( is.null(opt$attrib.used ) ) {
  opt$attrib.used = "../../SourceFiles/Attributes/Attributes_All.csv"
  } #attributes (watershed characteristics) to use in regression analysis
if ( is.null(opt$alpha.index ) ) {
  opt$alpha.index = 1
}
if ( is.null(opt$name)) {
  opt$name = "missi"
}
if ( is.null(opt$toggle)) {
  opt$toggle = 0
} #see below for explanation
# when toggle = 0, the skeleton network (Voronoi polygons) is used for Phase III application
# when toggle = 1, the validation dataset (delineated using GK) is used for Phase II validation
```

```
#assign arguments to local variables.
hkdb.file.name <- opt$hkdb.file
region.code <- opt$region.code
region.file.name <- opt$region.file
out.path <- opt$out.dir
out.name <-opt$name
y.type <- opt$y.type
attrib.incl<-as.character(read.csv(opt$attrib.used,header=F)$V1)
alpha.optim <- opt$alpha.index
toggle<-opt$toggle
if(y.type==4||y.type==5||y.type==6){FlowRange<-12}else{FlowRange<-17}



# Define Constants
#Define Exceedence Ranges
if(FlowRange==12){
  exceed <- seq(1,12,1)
}else{
  exceed <- c(0.0001,.001,.005,.01,.05,.1,.2,.3,.4,.5,.6,.7,.8,.9,.95,.99,.9999)
}
exceed.len <- length(exceed)
#Define Alpha Ranges
alpha <- c(0,0.0001,0.001,0.0025,0.005,0.0075,0.01,0.03,0.05,0.08,0.1,0.12,0.15,
           0.2,0.25,0.3,0.35,0.4,0.45,0.5,0.55,0.6,0.65,0.7,0.75)
#shorten alpha temporarily for debugging
alpha.len <- length(alpha)

#create the output path (if it doesn't already exist)
dir.create(file.path(out.path), showWarnings = F)


#******************************************************************************
#Load and Sort Data
#load data from spreadsheet
rawdata<- read.csv(hkdb.file.name)

#filter by region if a region code was specified
if(is.finite(region.code)){rawdata<-subset(rawdata,Region_Code==region.code)}

#parse out flow and physiographic data
MMF <-rawdata[,c(8:19)]
QFDC<-rawdata[,c(20:36)]
PDC<-rawdata[,c(37:53)]
SPDC <-rawdata[,c(54:70)]
ADC<-rawdata[,71:87]
MMP<-rawdata[,134:145]
MMSP<-rawdata[,146:157]
MMA<-rawdata[,158:169]
```

```
#set predictive variables (could be flow, power, specific power) and specify
# the neighbourhood indices to use.
if(y.type==1){
  FDC <- QFDC
  col.list<-c(2,4,5,6,8,10)
}else if(y.type==2){
  FDC <- PDC
  col.list<-c(2,4,5,6,8,10)
}else if (y.type==3){
  FDC <- SPDC
  col.list<-c(2,4,5,6,8,10)
}else if (y.type==4){
  FDC<-MMF
  col.list<-c(2,4,5,6,8,10)
}else if (y.type==5){
  FDC<-MMP
  col.list<-c(2,4,5,6,8,10)
}else if (y.type==6){
  FDC<-MMSP
  col.list<-c(2,4,5,6,8,10)
}

Index <- c(1:nrow(FDC))

#Filter for any flows that shouldn't be included
row.has.na <- apply(FDC, 1, function(x){any(is.na(x))}) #find flows with NAs
FDC<-FDC[!row.has.na,]
Index <- c(1:nrow(FDC))


#Get Physio Data
Physio<-rawdata[,c(88:133)]
Physio[Physio==0]<-0.0001
Physio<-Physio[!row.has.na,attrib.incl]#subset(Physio,)#select=attrib.incl)



#Normalize flows
FDC<-abs(FDC)
FDC[FDC==0]<-0.0001
logFDC <- log(FDC)
FDCs<-FDC/Physio[,1]


#load an aggregate catchment table to get the fields that are in the Phase III
# analysis and combine geoclasses
x <- ReadTB0(region.file.name)
{
  x$data.table$GeoGlaciers<-x$data.table$GeoClass3
  x$data.table$GeoLakeMud<-x$data.table$GeoClass25+x$data.table$GeoClass15
```

```
  x$data.table$GeoLakeSand<-x$data.table$GeoClass19+x$data.table$GeoClass21+
    x$data.table$GeoClass9
  x$data.table$GeoMud<-x$data.table$GeoClass6+x$data.table$GeoClass7
  x$data.table$GeoPeat<-x$data.table$GeoClass22
  x$data.table$GeoRock<-x$data.table$GeoClass2+x$data.table$GeoClass4+
    x$data.table$GeoClass8+x$data.table$GeoClass13+x$data.table$GeoClass14+
    x$data.table$GeoClass18+x$data.table$GeoClass20+x$data.table$GeoClass23
  x$data.table$GeoSandGravel<-x$data.table$GeoClass5+x$data.table$GeoClass11+
    x$data.table$GeoClass12+x$data.table$GeoClass16+x$data.table$GeoClass17+
    x$data.table$GeoClass24
  x$data.table$GeoTill<-x$data.table$GeoClass10
  x$data.table$Geowater<-x$data.table$GeoClass1+x$data.table$GeoClass0

  x$data.table <- x$data.table[,c(1:28,55:81)]
}

#determine the common columns (intersection) for analysis.
col.names.hkdb <- colnames(Physio)
col.names.region <- colnames(x$data.table)
col.names <- intersect(col.names.hkdb, col.names.region)

#Strip datasets to common columns and store names for each observation
Physio <- Physio[,col.names]
if(toggle==0){Physio.region <- x$data.table[,c("SkelUID", col.names)]
              Stationnames<-Physio.region[,1]} else{Physio.region<-Physio
             Stationnames<-rawdata[,2]}


#Clean up physio datasets
row.is.bad <- which(Physio.region$DrainageArea>=0&Physio.region$AnnPrecRain
                    >=0&Physio.region$MeanElev>=0)
Physio.region <- Physio.region[row.is.bad,] # remove zero area elements (errors)
Stationnames<-Stationnames[row.is.bad]


# for debugging - produce a smaller subset for quick analysis
#Physio.region<- Physio.region[1:50,] ;if(toggle==0){Stationnames<-Physio.region[,1]}
#Physio.region<- Physio.region[Physio.region$SkelUID==c(427105),] ;if(toggle==0){Stationnames<-Physio.r


#Log transformed variables
log.vars <- c("DrainageArea")
Physio[,log.vars]<-log(Physio[,log.vars])
Physio.region[,log.vars]<-log(Physio.region[,log.vars])



#******************************************************************************
#find the key variables that contribute to each of the flow predictions
keyvars<-vector("list",ncol(FDC)) #initialize variable to store key variables
allvars<-vector(mode="character")#initialize variable to store key variables
```

```
print("Determining Model Variables...")
pb <- txtProgressBar(min = 0, max = ncol(FDC), style = 3)


for(i in 1:ncol(FDC))
{
  #only concerned about 1 flow frequency
  Y<-logFDC[,i]
  #all the physiographic data
  X<-Physio
  #run regression model
  Model<-lm(formula=Y~.,data=X)
  #trim regression model to most important variables
  Model.min<-step(Model,direction="backward",trace=0)
  #get the trimmed model stats
  Dev<-as.data.frame(summary(Model.min)$coefficients[-1,c(2,4)])
  #record only variables that meet standards
  keyvars[[i]]<-rownames(subset(Dev,Dev[,1]<5&Dev[,2]<0.05))

  if(length(keyvars[[i]])==1){keyvars[[i]]<-c("DrainageArea","Perimeter")}
  allvars <- c(allvars,rownames(subset(Dev,Dev[,1]<5&Dev[,2]<0.05)))

# short suspension to update progress bar
  Sys.sleep(0.1)
  # update progress bar
  setTxtProgressBar(pb, i)
}
close(pb)

#limit both datasets to the set of common predictor variables.
tmp<-sort(table(allvars),decreasing=T)
allkeyvars <-names(tmp[tmp>=6])
cullPhysio<- Physio[,allkeyvars]
cullPhysio.region <- Physio.region[,allkeyvars]




#****************************************************************************
#Find Nearest Neighbours

#initialize variables
neigh.count <- rep(NA, length=nrow(cullPhysio.region))
NeighBours<-list()

ptm<-proc.time()
print(paste("Finding Neighbours - Alpha:",alpha.optim))
pb <- txtProgressBar(min = 0, max =nrow(cullPhysio.region), style = 3)
for(i in 1:nrow(cullPhysio.region)) #for element in the target dataset
  {
```

```
  Xtargneigh <- cullPhysio.region[i,]
  X <- cullPhysio
  Y <- log(FDCs)
  tmpIndex <-Index

  #if using validation dataset, remove target from validation dataset (jack-knife)
  if(toggle==1){X<-X[-i,];Y<-Y[-i,];tmpIndex <-Index[-i]}

  #CCA nearest neighbour function
  neigh <- neighbour(X,Y,Xtargneigh,alpha[alpha.optim],tmpIndex)
  #relates the temp index to the main index
  Control <- neigh$index[neigh$NeighYN==1]
  #record how many neighbouring stations each station has
  neigh.count[i] <- sum(neigh$NeighYN)
  NeighBours[[i]]<-Control

  # short suspension to update progress bar
  Sys.sleep(0.1)
  # update progress bar
  setTxtProgressBar(pb, i)
  }
      close(pb)
  proc.time()-ptm


#********************************************************************************
#Now do Multiple Regression using the neighbours found
  #Initialize Variables
  FDCest.P3 <- array(NA, c(nrow(cullPhysio.region), FlowRange))
  pval.P3 <- array(NA, c(nrow(cullPhysio.region), FlowRange))
  FDCest.P3.lower <- array(NA, c(nrow(cullPhysio.region), FlowRange))
  FDCest.P3.upper <- array(NA, c(nrow(cullPhysio.region), FlowRange))

options(warn=0)
print(paste("Performing Multiple Regression:",alpha.optim))
pb <- txtProgressBar(min = 0, max =nrow(cullPhysio.region), style = 3)
for(i in 1:nrow(cullPhysio.region))
  {
  if(neigh.count[i]>=5) #if there are fewer than five (arbitrarily picked)
                        #neighbouring stations, can't do linear regression

    for(j in 1:ncol(FDC)) #for each flow
    {

      sameFlow<-subset(logFDC[NeighBours[[i]],j],logFDC[NeighBours[[i]],j]
                       ==log(0.0001))
      if(length(sameFlow)!=length(logFDC[NeighBours[[i]],j]))
        #if statement: ensure enough flow data is available, require (>5) to do
        # regression, also ensure that all the flows <> 0, which is sometimes
        # possible w/FDC 17)
        {
```

```
      #j=1
       Xtarg <- Physio.region[i,keyvars[[j]]]
       #create dataset used for lm(), 1st station is target and the rest are
       # nearest neighbours
       XYSUB<-cbind(Flow=logFDC[NeighBours[[i]],j],
                    Physio[NeighBours[[i]],keyvars[[j]]])

       #linear regression, subset of all data except target station
       Model<-lm(formula=Flow~.,data=XYSUB)
       #exp(f.est <- predict(Model.min,Xtarg,interval="predict"))

       #trim regression model to most important variables
       #Model<-step(Model,direction="backward",trace=0)
       fstat<-summary(Model)$fstatistic #get f-statistic
       #calculate pvalue of regression
       pval<-pf(fstat[1], fstat[2], fstat[3], lower.tail=FALSE)
       pval.P3[i,j] <- pval
       if(pval<0.05&is.finite(pval)){
         f.est <- predict(Model,Xtarg,interval="predict")
         FDCest.P3[i,j]<-exp(f.est[1])
         FDCest.P3.lower[i,j]<-exp(f.est[2])
         FDCest.P3.upper[i,j]<-exp(f.est[3])
         } #predict flow if pvalue<0.05
      }
   }


  # short suspension to update progress bar
  Sys.sleep(0.1)
  # update progress bar
  setTxtProgressBar(pb, i)
}
close(pb)




#********************************************************************************
#Calculate Statistics (if Toggle = 1)
stat.out<-NA

if(toggle==1){

  FDC[FDCest.P3==NA]<-NA

  meanflow<-apply(FDC,2,mean,na.rm = TRUE)
  meanflowmatrix <- t(matrix(meanflow,ncol(FDC),nrow(FDC)))
  Dev <- FDC - meanflowmatrix
  SST <- apply(Dev^2,2,sum,na.rm=T)

  #initialize variables before loop
```

```
  NASH<-matrix(NA,length(alpha),ncol(FDC))
  RMSE<-matrix(NA,length(alpha),ncol(FDC))
  RRMSE<-matrix(NA,length(alpha),ncol(FDC))
  Bias<-matrix(NA,length(alpha),ncol(FDC))

  #Calculate Stats
  Err <- FDC - FDCest.P3
  SSE <- apply(Err^2,2,sum,na.rm=T)
  NASH <- 1-SSE/SST
  RMSE <- (apply(Err^2,2,mean,na.rm=T))^.5
  RRMSE <- (apply((Err/FDC)^2,2,mean,na.rm=T))^.5
  Bias <- apply(Err,2,mean,na.rm=T)

  #Plot Errors
  png(paste(out.path, "/", 'Stats',alpha.optim,'.png',sep=""),
      width=1500,height=750)
  par(mfrow=c(1,2))

  plot(NASH ~ exceed, main="NASH for MR Regional Flows",
       xlab = "FDC Flows", type="l",xlim=c(0,max(exceed)),ylim=c(-.25,1),
       xaxt="n")
  points(NASH ~ exceed)
  axis(1,at=exceed)
  grid()

  y.lim<- c(1,100000)
  options(scipen=3)

  plot(RRMSE ~ exceed, log="y",main="RRMSE for MR Regional Flows",
       xlab = "FDC Flows", type="l",xlim=c(0,max(exceed)),ylim=y.lim,xaxt="n",
       yaxt="n")
  points(RRMSE ~ exceed)
  axis(1,at=exceed)
  axis(2,at=c(1,10,100,1000,10000,100000,1000000,10000000))
  grid()
  dev.off()

  stat.out<-data.frame(rbind(NASH,RMSE,RRMSE))
  write.csv(stat.out,paste(out.path,"/stats", alpha.optim,".csv",sep=""))

  }




#******************************************************************************
#Output Results

#Calculate Daily Average (calculation depends on FDC or MMF, functions are in library)
if(y.type!=4||y.type!=5||y.type!=6){
  DailyAverage<-apply(FDCest.P3,1,DailyMeanFDC)
```

```
}else{
  DailyAverage<-apply(FDCest.P3,1,DailyMeanMMF)
}
if(y.type!=4||y.type!=5||y.type!=6){
  DailyAveragelower<-apply(FDCest.P3.lower,1,DailyMeanFDC)
}else{
  DailyAveragelower<-apply(FDCest.P3.lower,1,DailyMeanMMF)
}
if(y.type!=4||y.type!=5||y.type!=6){
  DailyAverageupper<-apply(FDCest.P3.upper,1,DailyMeanFDC)
}else{
  DailyAverageupper<-apply(FDCest.P3.upper,1,DailyMeanMMF)
}

#Create data frames
out.neigh.count <-as.data.frame(cbind(Stationnames,(neigh.count)))
names(out.neigh.count)<-c("UID","Num_Neigh")
out.FDCest <- as.data.frame(cbind(Stationnames, FDCest.P3,DailyAverage))
names(out.FDCest) <- c("SkelUID",names(FDC),"DailyAverage")
out.pval <- as.data.frame(cbind(Stationnames, (pval.P3)))
names(out.pval) <- c("SkelUID", names(FDC))
out.FDCest.lower <- as.data.frame(cbind(Stationnames, FDCest.P3.lower,
                                        DailyAveragelower))
names(out.FDCest.lower) <- c("SkelUID",names(FDC),"DailyAverage")
out.FDCest.upper <- as.data.frame(cbind(Stationnames, FDCest.P3.upper,
                                        DailyAverageupper))
names(out.FDCest.upper) <- c("SkelUID",names(FDC),"DailyAverage")

#******************************************************************************
#Remove all estimates that have really big confidence intervals
#Make copy of originals
out.FDCest.orig<-out.FDCest
out.FDCest.lower.orig<-out.FDCest.lower
out.FDCest.upper.orig<-out.FDCest.upper

#Calculate confidence interval
Conf<-(out.FDCest.upper-out.FDCest.lower)/out.FDCest

#Set Limit
MaxConf<-50

#Remove bad estimates
out.FDCest[Conf>MaxConf]<-(NA)
out.FDCest.lower[Conf>MaxConf]<-(NA)
out.FDCest.upper[Conf>MaxConf]<-(NA)
#******************************************************************************



if(y.type==1){Flow.out<-"FDC"}
```

```
if(y.type==2){Flow.out<-"PDC"}
if(y.type==3){Flow.out<-"SPDC"}
if(y.type==4){Flow.out<-"MMF"}
if(y.type==5){Flow.out<-"MMP"}
if(y.type==6){Flow.out<-"MMSP"}

#Save data
save(Conf,out.FDCest,out.FDCest.lower,out.FDCest.upper,out.FDCest.orig,
     out.FDCest.lower.orig,out.FDCest.upper.orig,out.pval,neigh.count,
     stat.out,file=paste(out.path, "/",Flow.out,"AGG",out.name,
                         alpha.optim,".Rdata", sep=""))

#Write csvs (optional)
# write.csv(out.FDCest, file=paste(out.path, "/", "out_FDCest.csv", sep=""))
# write.csv(out.FDCest.lower, file=paste(out.path, "/", "out_FDCestlower.csv", sep=""))
# write.csv(out.FDCest.upper, file=paste(out.path, "/", "out_FDCestupper.csv", sep=""))
# write.csv(out.pval, file=paste(out.path, "/", "out_pval.csv", sep=""))
# write.csv(out.neigh.count, file=paste(out.path, "/", "neigh_count.csv", sep=""))
# write.csv(allkeyvars,file=paste(out.path,"/","allkeyvars.csv",sep=""))
Sys.time()




#
# #*****************************************************************************
# #Map the Results (Optional)
#
# source("rlib/libMaps.R") #get library package
# library(maptools)
# library(RColorBrewer)
# library(classInt)
# library(plyr)
# can.map <- CanadaMap()
# usa.map <- USAMap()
#
# #out.neigh.count<-read.csv("Output/output_pacif_q_a=7/neigh_count.csv")
#
#
# rivers <- readShapeLines(paste("Maps/Skeletons/",out.name,"_skel.shp",sep=""), repair=TRUE)
# projection(rivers) <-proj.latlong
#
# rivers <- spTransform(rivers, proj.albers)
# jam<-rivers
# can.map <- spTransform(can.map, proj.albers)
# usa.map <- spTransform(usa.map, proj.albers)
#
# rivers@data <- join(rivers@data,out.neigh.count,by=intersect("UID","UID"))
# rivers$Num_Neigh[is.na(rivers$Num_Neigh)]<- (0.1)
# #colours <-brewer.pal(7,"Dark2")
# colours<- c("black", "cornflowerblue", "aquamarine4", "chartreuse4", "darkgoldenrod", "firebrick", "li
```

```
#
# brks<-c(0,5,50,100,150,200,250)
#
# png(file=paste(out.path,"/","RiverNeigh", alpha.optim,".png",sep=""),pointsize=5,res=2000,width=5,heig
# plot(can.map, lwd=.3, border="white", col="beige", axes=F,xlim=rivers@bbox[1,],ylim=rivers@bbox[2,])
# plot(usa.map, lwd=.3, border="white", col="beige", add=T)
# plot(rivers, lwd=.1, col=colours[findInterval(rivers$Num_Neigh,brks)],add=T,axes=F)
# legend(rivers@bbox[1,1]+diff(rivers@bbox[1,])*.15,rivers@bbox[2,1]+diff(rivers@bbox[2,])*.15,box.lwd =
# dev.off()
```

A.2 Script to predict increment flows using MLR-CCA

```
#********************************************
#Name: Phase_3.R
#Written By: James Bomhof and Wayne Jenkinson
#Date Written: January 8, 2013
#Description: Performs multiple regression for
#flow and power prediction using Canonical Correlation Analysis
#
#Edited by: Ivana Vouk, July 29, 2013
#********************************************
#setwd("Y:\\NRCan-Hydrokinetic\\Phase3Proc\\R")
print("Application of CCA to Phase III Datasets")
Sys.time()
options(scipen=5)

#Get libraries and set options
rm(list=ls())
#Libraries
library(getopt)
source("rlib/libCCA_MR.R")
source("rlib/libENSIM_IO.R")



#Get the details about region, output, etc. in an set of arguments.
#validation database
spec <- c("hkdb.file", "h", 1, "character")
#region code (to filter validation database)
spec <- rbind(spec, c("region.code", "c", 2, "integer"))
#region file - Phase III dataset
spec <- rbind(spec, c("region.file", "f", 3, "character"))
#diretory to place output files
spec <- rbind(spec, c("out.dir", "o", 4, "character"))
#comparative analysis - FDC (1), MMF (4), Specific Power(3), Total Power(2)
spec <- rbind(spec, c("y.type", "y", 5, "integer"))
#what attributes get used in regression analysis
```

```
spec <- rbind(spec, c("attrib.used","u",6,"character"))
#the optimal alpha index.
spec <- rbind(spec, c("alpha.index", "a", 7, "integer"))
#5 letter name of regions
spec <- rbind(spec, c("name", "n", 8, "character"))
#toggles test or analysis dataset
spec <- rbind(spec, c("toggle", "t", 9, "character"))
opt = getopt(spec)


#specify the default values if the arguments are not supplied.
if ( is.null(opt$hkdb.file ) ) {
  opt$hkdb.file = "../../SourceFiles/ValidationStations/HKDB_FDC_MMF_PVA_ALLPHYSIO_rev4.csv"
}
if ( is.null(opt$region.code ) ) {
  opt$region.code = NA
}
if ( is.null(opt$region.file ) ) {
  opt$region.file = "../../SourceFiles/CHCVoronoiIncrementalAggregationV1/missi/IncrementalAggregateCat
}
if ( is.null(opt$out.dir ) ) {
  opt$out.dir = "FDCINCMissitest"
}
if ( is.null(opt$y.type ) ) {
  opt$y.type = 4
}
if ( is.null(opt$attrib.used ) ) {
  opt$attrib.used = "../../SourceFiles/Attributes/Attributes_All.csv"
}
if ( is.null(opt$alpha.index ) ) {
  opt$alpha.index = 1
}
if ( is.null(opt$name)) {
  opt$name = "missi"
}
#0 means use skeleton network, 1 means use validation dataset
if ( is.null(opt$toggle)) {
  opt$toggle = 0
}



#assign argumentns to local variables.
hkdb.file.name <- opt$hkdb.file
region.code <- opt$region.code
region.file.name <- opt$region.file
out.path <- opt$out.dir
out.name <-opt$name
y.type <- opt$y.type
attrib.incl<-as.character(read.csv(opt$attrib.used,header=F)$V1)
alpha.optim <- opt$alpha.index
```

```
toggle<-opt$toggle
if(y.type==4||y.type==5||y.type==6){FlowRange<-12}else{FlowRange<-17}



# Define Constants
#Define Exceedence Ranges
if(FlowRange==12){
  exceed <- seq(1,12,1)
}else{
  exceed <- c(0.0001,.001,.005,.01,.05,.1,.2,.3,.4,.5,.6,.7,.8,.9,.95,.99,.9999)
}
exceed.len <- length(exceed)
#Define Alpha Ranges
alpha <- c(0,0.0001,0.001,0.0025,0.005,0.0075,0.01,0.03,0.05,0.08,0.1,0.12,0.15,
           0.2,0.25,0.3,0.35,0.4,0.45,0.5,0.55,0.6,0.65,0.7,0.75)
#shorten alpha temporarily for debugging
alpha.len <- length(alpha)

#create the output path (if it doesn't already exist)
dir.create(file.path(out.path), showWarnings = FALSE)



#********************************************************************************
#Load and Sort Data
#load data from spreadsheet
rawdata<- read.csv(hkdb.file.name)

#filter by region if a region code was specified
if(is.finite(region.code)){rawdata<-subset(rawdata,Region_Code==region.code)}

#parse out flow and physiographic data
QFDC<-rawdata[,c(20:36)]
PDC<-rawdata[,c(37:53)]
SPDC <-rawdata[,c(54:70)]
MMF <-rawdata[,c(8:19)]


#set predictive variables (could be flow, power, specific power) and specify
# the neighbourhood indices to use.
if(y.type==1){
  FDC <- QFDC
  col.list<-c(2,4,5,6,8,10)
}else if(y.type==2){
  FDC <- PDC
  col.list<-c(2,4,5,6,8,10)
}else if (y.type==3){
  FDC <- SPDC
  col.list<-c(2,4,5,6,8,10)
}else if (y.type==4){
```

```
  FDC<-MMF
  col.list<-c(2,4,5,6,8,10)
}

Index <- c(1:nrow(FDC))

#Filter for any flows that shouldn't be included
row.has.na <- apply(FDC, 1, function(x){any(is.na(x))}) #find flows with NAs
FDC<-FDC[!row.has.na,]
Index <- c(1:nrow(FDC))


#Get Physio Data
Physio<-rawdata[,c(88:133)]
Physio[Physio==0]<-0.0001
Physio<-Physio[!row.has.na,attrib.incl]#subset(Physio,)#select=attrib.incl)



#Normalize flows
FDC<-abs(FDC)
FDC[FDC==0]<-0.0001
logFDC <- log(FDC)
FDCs<-FDC/Physio[,1]


#load an aggregate catchment table to get the fields that are in the Phase III
# analysis and combine geoclasses
x <- ReadTB0(region.file.name)
{
  x$data.table$GeoGlaciers<-x$data.table$GeoClass3
  x$data.table$GeoLakeMud<-x$data.table$GeoClass25+x$data.table$GeoClass15
  x$data.table$GeoLakeSand<-x$data.table$GeoClass19+x$data.table$GeoClass21+
    x$data.table$GeoClass9
  x$data.table$GeoMud<-x$data.table$GeoClass6+x$data.table$GeoClass7
  x$data.table$GeoPeat<-x$data.table$GeoClass22
  x$data.table$GeoRock<-x$data.table$GeoClass2+x$data.table$GeoClass4+
    x$data.table$GeoClass8+x$data.table$GeoClass13+x$data.table$GeoClass14+
    x$data.table$GeoClass18+x$data.table$GeoClass20+x$data.table$GeoClass23
  x$data.table$GeoSandGravel<-x$data.table$GeoClass5+x$data.table$GeoClass11+
    x$data.table$GeoClass12+x$data.table$GeoClass16+x$data.table$GeoClass17+
    x$data.table$GeoClass24
  x$data.table$GeoTill<-x$data.table$GeoClass10
  x$data.table$Geowater<-x$data.table$GeoClass1+x$data.table$GeoClass0

  x$data.table <- x$data.table[,c(1:30,57:83)]
}

#determine the common columns (intersection) for analysis.
col.names.hkdb <- colnames(Physio)
col.names.region <- colnames(x$data.table)
```

```
col.names <- intersect(col.names.hkdb, col.names.region)

#Strip datasets to common columns and store names for each observation
Physio <- Physio[,col.names]
if(toggle==0){Physio.region <- x$data.table[,c("SkelUID", col.names)]
              Stationnames<-Physio.region[,1]} else{Physio.region<-Physio
             Stationnames<-rawdata[,2]}


#Clean up physio datasets
row.is.bad <- which(Physio.region$DrainageArea>=0&Physio.region$AnnPrecRain>=
                      0&Physio.region$MeanElev>=0)
Physio.region <- Physio.region[row.is.bad,] # remove zero area elements (errors)
Stationnames<-Stationnames[row.is.bad]


# for debugging - produce a smaller subset for quick analysis
#Physio.region<- Physio.region[1:50,] ;if(toggle==0){Stationnames<-Physio.region[,1]}

#Log transformed variables
log.vars <- c("DrainageArea")
Physio[,log.vars]<-log(Physio[,log.vars])
Physio.region[,log.vars]<-log(Physio.region[,log.vars])



#******************************************************************************
#find the key variables that contribute to each of the flow predictions
keyvars<-vector("list",ncol(FDC)) #initialize variable to store key variables
allvars<-vector(mode="character")#initialize variable to store key variables

print("Determining Model Variables...")
pb <- txtProgressBar(min = 0, max = ncol(FDC), style = 3)


for(i in 1:ncol(FDC))
{
  #only concerned about 1 flow frequency
  Y<-logFDC[,i]
  #all the physiographic data
  X<-Physio
  #run regression model
  Model<-lm(formula=Y~.,data=X)
  #trim regression model to most important variables
  Model.min<-step(Model,direction="backward",trace=0)
  #get the trimmed model stats
  Dev<-as.data.frame(summary(Model.min)$coefficients[-1,c(2,4)])
  #record only variables that meet standards
  keyvars[[i]]<-rownames(subset(Dev,Dev[,1]<5&Dev[,2]<0.05))

  if(length(keyvars[[i]])==1){keyvars[[i]]<-c("DrainageArea","Perimeter")}
```

```
  allvars <- c(allvars,rownames(subset(Dev,Dev[,1]<5&Dev[,2]<0.05)))

# short suspension to update progress bar
  Sys.sleep(0.1)
  # update progress bar
  setTxtProgressBar(pb, i)
}
close(pb)

#limit both datasets to the set of common predictor variables.
tmp<-sort(table(allvars),decreasing=T)
allkeyvars <-names(tmp[tmp>=6])
cullPhysio<- Physio[,allkeyvars]
cullPhysio.region <- Physio.region[,allkeyvars]




#****************************************************************************
#Find Nearest Neighbours

#initialize variables
neigh.count <- rep(NA, length=nrow(cullPhysio.region))
NeighBours<-list()

ptm<-proc.time()
print(paste("Finding Neighbours - Alpha:",alpha.optim))
pb <- txtProgressBar(min = 0, max =nrow(cullPhysio.region), style = 3)
for(i in 1:nrow(cullPhysio.region)) #for element in the target dataset
  {
  Xtargneigh <- cullPhysio.region[i,]
  X <- cullPhysio
  Y <- log(FDCs)
  tmpIndex <-Index

  #if using validation dataset, remove target from validation dataset (jack-knife)
  if(toggle==1){X<-X[-i,];Y<-Y[-i,];tmpIndex <-Index[-i]}

  #CCA nearest neighbour function
  neigh <- neighbour(X,Y,Xtargneigh,alpha[alpha.optim],tmpIndex)
  #relates the temp index to the main index
  Control <- neigh$index[neigh$NeighYN==1]
  #record how many neighbouring stations each station has
  neigh.count[i] <- sum(neigh$NeighYN)
  NeighBours[[i]]<-Control

  # short suspension to update progress bar
  Sys.sleep(0.1)
  # update progress bar
  setTxtProgressBar(pb, i)
  }
```

```
    close(pb)
  proc.time()-ptm


#******************************************************************************
#Now do Multiple Regression using the neighbours found
  #Initialize Variables
  FDCest.P3 <- array(NA, c(nrow(cullPhysio.region), FlowRange))
  pval.P3 <- array(NA, c(nrow(cullPhysio.region), FlowRange))
  FDCest.P3.lower <- array(NA, c(nrow(cullPhysio.region), FlowRange))
  FDCest.P3.upper <- array(NA, c(nrow(cullPhysio.region), FlowRange))

options(warn=0)
print(paste("Performing Multiple Regression:",alpha.optim))
pb <- txtProgressBar(min = 0, max =nrow(cullPhysio.region), style = 3)
for(i in 1:nrow(cullPhysio.region))
  {
  if(neigh.count[i]>=5) #if there are fewer than five (arbitrarily picked)
    # neighbouring stations, can't do linear regression

    for(j in 1:ncol(FDC)) #for each flow
    {

      sameFlow<-subset(logFDC[NeighBours[[i]],j],logFDC[NeighBours[[i]],j]==
                         log(0.0001))
      if(length(sameFlow)!=length(logFDC[NeighBours[[i]],j]))
        #if statement: ensure enough flow data is available, require (>5) to do
        # regression, also ensure that all the flows <> 0, which is sometimes
        # possible w/FDC 17)
        {
        Xtarg <- Physio.region[i,keyvars[[j]]]
        #create dataset used for lm(), 1st station is target and the rest are
        # the nearest neighbours
        XYSUB<-cbind(Flow=logFDC[NeighBours[[i]],j],
                     Physio[NeighBours[[i]],keyvars[[j]]])
        #linear regression, subset of all data except target station
        Model<-lm(formula=Flow~.,data=XYSUB)
        #trim regression model to most important variables
        #Model<-step(Model,direction="backward",trace=0)
        fstat<-summary(Model)$fstatistic #get f-statistic
        #calculate pvalue of regression
        pval<-pf(fstat[1], fstat[2], fstat[3], lower.tail=FALSE)
        pval.P3[i,j] <- pval
        if(pval<0.05&is.finite(pval)){
          f.est <- predict(Model,Xtarg,interval="predict")
          FDCest.P3[i,j]<-exp(f.est[1])
          FDCest.P3.lower[i,j]<-exp(f.est[2])
          FDCest.P3.upper[i,j]<-exp(f.est[3])
          } #predict flow if pvalue<0.05
       }
    }
```

```
  # short suspension to update progress bar
  Sys.sleep(0.1)
  # update progress bar
  setTxtProgressBar(pb, i)
}
close(pb)



#*****************************************************************************
#Calculate Statistics (if Toggle = 1)
stat.out<-NA

if(toggle==1){

  FDC[FDCest.P3==NA]<-NA

  meanflow<-apply(FDC,2,mean,na.rm = TRUE)
  meanflowmatrix <- t(matrix(meanflow,ncol(FDC),nrow(FDC)))
  Dev <- FDC - meanflowmatrix
  SST <- apply(Dev^2,2,sum,na.rm=T)

  #initialize variables before loop
  NASH<-matrix(NA,length(alpha),ncol(FDC))
  RMSE<-matrix(NA,length(alpha),ncol(FDC))
  RRMSE<-matrix(NA,length(alpha),ncol(FDC))
  Bias<-matrix(NA,length(alpha),ncol(FDC))

  #Calculate Stats
  Err <- FDC - FDCest.P3
  SSE <- apply(Err^2,2,sum,na.rm=T)
  NASH <- 1-SSE/SST
  RMSE <- (apply(Err^2,2,mean,na.rm = T))^.5
  RRMSE <- (apply((Err/FDC)^2,2,mean,na.rm = T))^.5
  Bias <- apply(Err,2,mean,na.rm=T)

  #Plot Errors
  png(paste(out.path, "/", 'Stats',alpha.optim,'.png',sep=""),width=1500,
      height=750)
  par(mfrow=c(1,2))

  plot(NASH ~ exceed, main="NASH for MR Regional Flows", xlab = "FDC Flows",
       type="l",xlim=c(0,max(exceed)),ylim=c(-.25,1),xaxt="n")
  points(NASH ~ exceed)
  axis(1,at=exceed)
  grid()

  y.lim<- c(1,100000)
```

```
  options(scipen=3)

  plot(RRMSE ~ exceed, log="y",main="RRMSE for MR Regional Flows",
       xlab = "FDC Flows", type="l",xlim=c(0,max(exceed)),ylim=y.lim,xaxt="n",
       yaxt="n")
  points(RRMSE ~ exceed)
  axis(1,at=exceed)
  axis(2,at=c(1,10,100,1000,10000,100000,1000000,10000000))
  grid()
  dev.off()

  stat.out<-data.frame(rbind(NASH,RMSE,RRMSE))
  write.csv(stat.out,paste(out.path,"/stats", alpha.optim,".csv",sep=""))

  }




#****************************************************************************
#Output Results

#Calculate Daily Average (calculation depends on FDC or MMF, functions are in library)
if(y.type!=4){
  DailyAverage<-apply(FDCest.P3,1,DailyMeanFDC)
}else{
  DailyAverage<-apply(FDCest.P3,1,DailyMeanMMF)
}
if(y.type!=4){
  DailyAveragelower<-apply(FDCest.P3.lower,1,DailyMeanFDC)
}else{
  DailyAveragelower<-apply(FDCest.P3.lower,1,DailyMeanMMF)
}
if(y.type!=4){
  DailyAverageupper<-apply(FDCest.P3.upper,1,DailyMeanFDC)
}else{
  DailyAverageupper<-apply(FDCest.P3.upper,1,DailyMeanMMF)
}

#Create data frames
out.neigh.count <-as.data.frame(cbind(Stationnames,(neigh.count)))
names(out.neigh.count)<-c("UID","Num_Neigh")
out.FDCest <- as.data.frame(cbind(Stationnames, FDCest.P3,DailyAverage))
names(out.FDCest) <- c("SkelUID",names(FDC),"DailyAverage")
out.pval <- as.data.frame(cbind(Stationnames, (pval.P3)))
names(out.pval) <- c("SkelUID", names(FDC))
out.FDCest.lower <- as.data.frame(cbind(Stationnames, FDCest.P3.lower,
                                        DailyAveragelower))
names(out.FDCest.lower) <- c("SkelUID",names(FDC),"DailyAverage")
out.FDCest.upper <- as.data.frame(cbind(Stationnames, FDCest.P3.upper,
                                        DailyAverageupper))
```

```
names(out.FDCest.upper) <- c("SkelUID",names(FDC),"DailyAverage")


#******************************************************************************
#Remove all estimates that have really big confidence intervals
#Make copy of originals
out.FDCest.orig<-out.FDCest
out.FDCest.lower.orig<-out.FDCest.lower
out.FDCest.upper.orig<-out.FDCest.upper

#Calculate confidence interval
Conf<-(out.FDCest.upper-out.FDCest.lower)/out.FDCest

#Set Limit
MaxConf<-50

#Remove bad estimates
out.FDCest[Conf>MaxConf]<-(NA)
out.FDCest.lower[Conf>MaxConf]<-(NA)
out.FDCest.upper[Conf>MaxConf]<-(NA)
#******************************************************************************


if(y.type==1){Flow.out<-"FDC"}
if(y.type==4){Flow.out<-"MMF"}

#Save data
save(Conf,out.FDCest,out.FDCest.lower,out.FDCest.upper,out.FDCest.orig,
     out.FDCest.lower.orig,out.FDCest.upper.orig, out.pval,
     neigh.count, stat.out,file=paste(out.path, "/",Flow.out,"INC",
                                  out.name,alpha.optim,".Rdata", sep=""))

#Write csvs (optional)
# write.csv(out.FDCest, file=paste(out.path, "/", "out_FDCest.csv", sep=""))
# write.csv(out.FDCest.lower, file=paste(out.path, "/", "out_FDCestlower.csv", sep=""))
# write.csv(out.FDCest.upper, file=paste(out.path, "/", "out_FDCestupper.csv", sep=""))
# write.csv(out.pval, file=paste(out.path, "/", "out_pval.csv", sep=""))
# write.csv(out.neigh.count, file=paste(out.path, "/", "neigh_count.csv", sep=""))
# write.csv(allkeyvars,file=paste(out.path,"/","allkeyvars.csv",sep=""))
Sys.time()




#
# #****************************************************************************
# #Map the Results (Optional)
#
# source("rlib/libMaps.R") #get library package
# library(maptools)
# library(RColorBrewer)
```

```
# library(classInt)
# library(plyr)
# can.map <- CanadaMap()
# usa.map <- USAMap()
#
# #out.neigh.count<-read.csv("Output/output_pacif_q_a=7/neigh_count.csv")
#
#
# rivers <- readShapeLines(paste("Maps/Skeletons/",out.name,"_skel.shp",sep=""), repair=TRUE)
# projection(rivers) <-proj.latlong
#
# rivers <- spTransform(rivers, proj.albers)
# jam<-rivers
# can.map <- spTransform(can.map, proj.albers)
# usa.map <- spTransform(usa.map, proj.albers)
#
# rivers@data <- join(rivers@data,out.neigh.count,by=intersect("UID","UID"))
# rivers$Num_Neigh[is.na(rivers$Num_Neigh)]<- (0.1)
# #colours <-brewer.pal(7,"Dark2")
# colours<- c("black", "cornflowerblue", "aquamarine4", "chartreuse4", "darkgoldenrod", "firebrick", "li
#
# brks<-c(0,5,50,100,150,200,250)
#
# png(file=paste(out.path,"/","RiverNeigh", alpha.optim,".png",sep=""),pointsize=5,res=2000,width=5,heig
# plot(can.map, lwd=.3, border="white", col="beige", axes=F,xlim=rivers@bbox[1,],ylim=rivers@bbox[2,])
# plot(usa.map, lwd=.3, border="white", col="beige", add=T)
# plot(rivers, lwd=.1, col=colours[findInterval(rivers$Num_Neigh,brks)],add=T,axes=F)
# legend(rivers@bbox[1,1]+diff(rivers@bbox[1,])*.15,rivers@bbox[2,1]+diff(rivers@bbox[2,])*.15,box.lwd =
# dev.off()
```

A.3 Script to combine the aggregrate and increment flows

```
#***********************************************************
#Name:IncrementStation.R
#Author: James Bomhof
#Date: January 22, 2013
#Description: Imports estimated flows from Phase_3_Increm.R. Checks to
#see if river segments have any HYDAT stations upstream
#and adds the HYDAT flows to the estimated flow. If there is a HYDAT station
#upstream that doesn't contain good data, the estimated flows are summed. If it turns out that no HYDAT
#the computer pulls in the original aggregate prediction and uses that. The estimates being summed are 
#distribution. The samples are stored in a list that are used in later scripts to calculate powers whil
#intervals when the form of the distribution is unknown.
#***********************************************************
rm(list=ls())
```

```
#Get Libraries
source("rlib/libENSIM_IO.r")
source("rlib/libCCA_MR.r")
library(gdata)
library(getopt)


#Get the details about region, output, etc. in an set of arguments.
spec <- c("name", "n", 1, "character") #5 letter name of regions
spec <- rbind(spec, c("y.type", "y", 1, "integer"))  #what type of comparative analysis to do - FDC (1)
spec <- rbind(spec, c("alpha.index", "a", 1, "integer"))  #the optimal alpha index.

opt = getopt(spec)

#specify the default values if the arguments are not supplied.
if ( is.null(opt$name)) { opt$name = "missi"}
if ( is.null(opt$y.type ) ) { opt$y.type = 1} #what type of comparative analysis to do - FDC (1), MMF (4
if ( is.null(opt$alpha.index ) ) { opt$alpha.index = 2}


#Define variables
y.type<-opt$y.type
Region<-opt$name
if(y.type==4){Flow.type="MMF"}else{Flow.type="FDC"}
AlphaLoop=opt$alpha.index


a=2
#Loop through each alpha and add calc'd to estimated flows
for(a in AlphaLoop){
Alpha=a

#Import Incremental Aggregated Estimated Flows
load(paste("Inc_AggEstimates/",Flow.type,"INC",Region,"/",Flow.type,"INC",Region,Alpha,".Rdata",sep=""))


rep.row<-function(x,n){matrix(rep(x,each=n),nrow=n)}



#Import Incremental Aggregate .tb0
IncAgg <- ReadTB0(paste("../../SourceFiles/CHCVoronoiIncrementalAggregationV1/",Region,"/IncrementalAgg


#Import FDC and MMF Flows from HYDAT Stations
if(y.type==4)
{
  MMF <- read.csv("../../SourceFiles/HydatFlows/MMF1990.csv")
  MMF <- na.omit(MMF) #reduce record to only stations that have flows for each month of the year
  #put in list form for distribution sampling
  MMFList<-list()
```

```
  MMFList[as.character(MMF$Station)]<-c(1:nrow(MMF))
  MMFRep<-function(i){rep.row(x=as.numeric(MMF[i,c(-1,-2)]),n=1000)}
  FlowList<-lapply(MMFList,MMFRep)
}

if(y.type==1)
{
  FDC <- read.csv("../../SourceFiles/HydatFlows/FDC1990.csv")
  FDC <- na.omit(FDC[,1:19])
  #put in list form for distribution sampling
  FDCList<-list()
  FDCList[as.character(FDC$Station)]<-c(1:nrow(FDC))
  FDCRep<-function(i){rep.row(x=as.numeric(FDC[i,c(-1,-2)]),n=1000)}
  FlowList<-lapply(FDCList,FDCRep)
}


#merge datasets to find out which skeluids are linked with what hydat stations
VorFlow<-merge(out.FDCest,IncAgg,by.x="SkelUID",by.y="SkelUID")
VorFlow.lower<-merge(out.FDCest.lower,IncAgg,by.x="SkelUID",by.y="SkelUID")
VorFlow.upper<-merge(out.FDCest.upper,IncAgg,by.x="SkelUID",by.y="SkelUID")


Ustations<-(VorFlow$UpstreamHydatStnNums)
Hstations<-(VorFlow$HydatStnNum)
levels(Ustations)[1]<-NA

#parse data depending on MMF or FDC
if(y.type==4)
{VorFlow<-VorFlow[,1:13];HydatFlow<-MMF;VorFlow.lower<-VorFlow.lower[,1:13];VorFlow.upper<-VorFlow.uppe
}else{VorFlow<-VorFlow[,1:18];HydatFlow<-FDC;VorFlow.lower<-VorFlow.lower[,1:18];VorFlow.upper<-VorFlow

#Standard Error
SE<-(log(VorFlow)-log(VorFlow.lower))/qnorm(0.975)

Sample<-function(Mean,SD){exp((rnorm(1000,mean=Mean,SD)))} #function to sample distribution

SampleStore<-function(i)
{
  Mean<-log(VorFlow[i,c(-1)])
  SD<-SE[i,c(-1)]
  Output<-mapply(Sample,Mean,SD) #get and store distribution in matrix for one station
  return(Output)
}



#Define new output
VorFlowNewList<-list()
VorFlowNewList[as.character(VorFlow$SkelUID)]<-NA
```

```
print(paste("Calculating Incremented Flows",Region,"Alpha",Alpha,"Flow",y.type))
pb <- txtProgressBar(min = 0, max = nrow(VorFlow), style = 3)

#Loop to check if flows are augmented by a hydat station and add hydat flows if it is
for(i in 1:nrow(VorFlow))
{
  #check to see if flow estimate is augmented by hydat station

  if(!is.na(Ustations[i]))
  {
    #Ustations[i]<-trim(Ustations[i])
    #set initial indexes for while loop
    HydatStnsget<-vector()
    VorStnsget<-i
    tmp<-0


    #get list of stations to include
    getstations<-unlist(strsplit(as.character(trim(Ustations[i])),split=" "))
    #getstations<-c(getstations,"01DB002") #for testing


    while(length(getstations)!=0)
    {
      #Find out which Stations have Hydat Flows
      FlowInd<-match(as.factor(getstations),HydatFlow$Station)

      #Record index to get flows from hydat db (later)
      HydatStnsget<-c(HydatStnsget,FlowInd[!is.na(FlowInd)])

      #get names of stns still needed
      stnsneeded<-getstations[is.na(FlowInd)]

      #Record index to get these flows from incremental aggregate database
      VorStnsget<-c(VorStnsget,match(stnsneeded,Hstations))

      #check to see if more stations are nested in this aggregate database
      morestations<-unlist(strsplit(as.character(trim(Ustations[match(stnsneeded,Hstations)])),split="

      #refresh getstations
      if(is.null(morestations)){getstations<-morestations}else{getstations<-morestations[!is.na(moresta

    }

    if(length(HydatStnsget!=0))
    {
      #gather matrices to add
      #first get Voronoi matrices (create if needed and add to IncAggList)
```

```
      IncAggList<-lapply(as.list(VorStnsget),SampleStore)

      #get Hydat List
      HydatList<-FlowList[HydatStnsget]

      #Sum the matrices together
      VorFlowNewList[[i]]<-Reduce('+',c(HydatList,IncAggList))

      #if there are any NaNs, reset matrix back to NA
      if(any(is.nan(VorFlowNewList[[i]][1,1:13]))){VorFlowNewList[i]<-NA}
    }
  }

  #else do nothing
  else{}

  # short suspension to update progress bar
  Sys.sleep(0.1)
  # update progress bar
  setTxtProgressBar(pb, i)
}

print(paste(length(VorFlowNewList[!is.na(VorFlowNewList)]),"of",length(VorFlowNewList), "SkelUIDs were 

#**********************************************************************
#Check if the Incremental Aggregated flows output NA values, if so then substitute original aggregated 

#Import Aggregated Estimated Flows
load(paste("../Aggregate/AggEstimates/",Flow.type,"AGG",Region,"/",Flow.type,"AGG",Region,Alpha,".Rdata


#get list of SkelUIDs with NA incremental aggregate flows
BadSkels<-names(VorFlowNewList[is.na(VorFlowNewList)])


SE<-(log(out.FDCest)-log(out.FDCest.lower))/qnorm(0.975)

Sample<-function(Mean,SD){exp((rnorm(1000,mean=Mean,SD)))} #function to sample distribution

SampleStore<-function(i)
{
  Mean<-log(out.FDCest[i,c(-1)])
  SD<-SE[i,c(-1)]
  Output<-mapply(Sample,Mean,SD) #get and store distribution in matrix for one station
  return(Output)
}


tmp<-match(BadSkels,out.FDCest$SkelUID)

VorFlowNewList[BadSkels]<-lapply(as.list(tmp),SampleStore)
```

```
#Save output file
#split into 2 files
VorFlowOrig<-VorFlowNewList

VorFlowNewList<-VorFlowOrig[1:trunc(length(VorFlowOrig)/2)]
file.name1<-paste("Inc_SumEstimates/",Flow.type,"INC_SUM_",Region,Alpha,"_1.Rdata",sep="")
save(VorFlowNewList, out.pval,file=file.name1)

VorFlowNewList<-VorFlowOrig[trunc((length(VorFlowOrig)/2)+1):length(VorFlowOrig)]
file.name2<-paste("Inc_SumEstimates/",Flow.type,"INC_SUM_",Region,Alpha,"_2.Rdata",sep="")
save(VorFlowNewList, out.pval,file=file.name2)
}
```

A.4 Script to apply hydraulic geometry to flow results and estimate hydrokinetic power

```
#*****************************************************************
#Code: PowerConversion_E.R
#Written on: February 22/13
#Written By: James Bomhof
#Description: Script to import samples of flows and perform a monte carlo simulation
#while regressing them against downstream hydraulic areas and geometry.
#Sp Pwr and Pwr distributions are then created with the flow and area distributions
#Finally, stats are calculated from these distributions. This is similar to PowerConversion.R,
#except it uses the manning's equation to account for backwater.
#**************************************************

#Set defaults and load libraries (if needed)
rm(list=ls())
options(scipen=10)
library(getopt)
source("rlib/libCCA_MR.R")
source("rlib/libENSIM_IO.R")

spec <- c("name", "n", 1, "character") #5 letter name of regions
spec <- rbind(spec, c("alpha.index", "a", 2, "integer"))  #the optimal alpha index.
```

```
spec <- rbind(spec, c("type", "t", 3, "character"))
spec <- rbind(spec, c("split", "s", 4, "integer"))
opt = getopt(spec)

if ( is.null(opt$name ) ) { opt$name = "marit"}
if ( is.null(opt$alpha.index ) ) { opt$alpha.index = 4}
if ( is.null(opt$type)) { opt$type = "ADC"}
if ( is.null(opt$split)){opt$split = 2}


#get best region estimates
Density<-1000
Region<-opt$name
Alpha<-opt$alpha.index
Type<-opt$type #"MMA" #Can be ADC,PDC,SPDC,MMA,MMP,MMSP
Split<-opt$split
n<-0.035

if(Type=="ADC"||Type=="PDC"||Type=="SPDC"){Flow="FDC";Length=17}else{Flow="MMF";Length=12}
if(Type=="ADC"||Type=="MMA"){TypeRegress="Area"}


#Define Monte Carlo Function
MonteCarlo<-function(Flow,Model){
  df<-data.frame(Flows=Flow)
  Distribution<-predict(Model,df,se.fit=T)
  Estimate<-rnorm(Distribution$fit,n=rep(1,length(Distribution$fit)),sd=rep(Distribution$residual.scale
  return(Estimate)
}



#get Incremental Aggregate estimates*******************************************
load(paste("../../Inc_SumEstimates/",Flow,"INC_SUM_",Region,Alpha,"_",Split,".Rdata",sep=""))
VorFlowOrigList<-VorFlowNewList
#load(paste("../../Inc_SumEstimates/",Flow,"INC_SUM_",Region,Alpha,"_",2,".Rdata",sep=""))
#VorFlowNewList<-c(VorFlowOrigList,VorFlowNewList)
rm(VorFlowOrigList)
for(i in 1:length(VorFlowNewList)){VorFlowNewList[[i]]<-VorFlowNewList[[i]][,-(Length+1)]}



#Calculate Area Distribution****************************************
#Get Downstream Area Regression
load(paste("../../../DownstreamRegression/",TypeRegress,"Model.Rdata",sep=""))

#Set output list
blankmatrix<-matrix(NA,1000,Length)
VorAreaList<-list()
VorAreaList[names(VorFlowNewList)]<-NA
for(i in 1:length(VorAreaList)){VorAreaList[[i]]<-blankmatrix}
```

```
if(Flow=="FDC"){AModel<-AModelFDC}else{AModel<-AModelMMF}

#Loop through each skelUID and calculate area distribution for each flow frequency
print("Estimating Regressed Areas")
pb <- txtProgressBar(min = 0, max = length(VorAreaList), style = 3)
system.time( #time code(optional)
for(i in 1:length(VorAreaList))
  {
  for(j in 1:Length)
    {VorAreaList[[i]][,j]<-exp(MonteCarlo(VorFlowNewList[[i]][,j],Model=AModel[[j]]))}
  # update progress bar
  setTxtProgressBar(pb, i)
  }
)


#Calculate Width Distribution**************************************
#Get Downstream Width Regression

load(paste("../../../DownstreamRegression/WidthModel.Rdata",sep=""))

#Set output list
blankmatrix<-matrix(NA,1000,Length)
VorWidthList<-list()
VorWidthList[names(VorFlowNewList)]<-NA
for(i in 1:length(VorWidthList)){VorWidthList[[i]]<-blankmatrix}
if(Flow=="FDC"){AModel<-AModelFDC}else{AModel<-AModelMMF}

#Loop through each skelUID and calculate area distribution for each flow frequency
print("Estimating Regressed Widths")
pb <- txtProgressBar(min = 0, max = length(VorWidthList), style = 3)
system.time( #time code(optional)
  for(i in 1:length(VorWidthList))
  {
    for(j in 1:Length){VorWidthList[[i]][,j]<-exp(MonteCarlo(VorFlowNewList[[i]][,j],Model=AModel[[j]])
    # update progress bar
    setTxtProgressBar(pb, i)
  }
)




#**************************************************************************
#Calculate Power Distribution
print("Calc Power")
VorPwrList<-list()
VorPwrList[names(VorFlowNewList)]<-NA
for(i in 1:length(VorPwrList)){VorPwrList[[i]]<-Density/2*((VorFlowNewList[[i]]/VorAreaList[[i]])^3)*Vo
#save(VorPwrList,file=paste("FinalROutput/EstimatePwrs_",Region,Alpha,Flow,"_",Split,".Rdata",sep=""))
```

```
#Calculate Velocity Distribution
print("Calc Velocity")
VorVelList<-list()
VorVelList[names(VorFlowNewList)]<-NA
for(i in 1:length(VorVelList)){VorVelList[[i]]<-(VorFlowNewList[[i]]/VorAreaList[[i]])}

#Calculate IntPower Distribution
#get length of SkelUID
Lengths<-ReadTB0(paste("../../../../SourceFiles/CHCVoronoiAggregationCatchmentsV4/",Region,"/AggregateCa
OrderLengths<-merge(Lengths,as.data.frame(names(VorFlowNewList)),by.x="SkelUID",by.y="names(VorFlowNewLi

print("Calc PK")
VorPKList<-list()
VorPKList[names(VorFlowNewList)]<-NA
for(i in 1:length(VorPKList)){VorPKList[[i]]<-10^(-6)*9806*VorFlowNewList[[i]]*(VorVelList[[i]]^2*Order

#Calculate MAF for all variables
ListMAF<-function(VorList){
  output<-list()
  for(i in 1:length(VorList)){VorList[[i]]<-cbind(VorList[[i]][,1:17],apply(VorList[[i]][,1:17],1,Daily
  return(VorList)}
print("Calculate MAF Values")
VorFlowNewList<-ListMAF(VorFlowNewList)
VorPwrList<-ListMAF(VorPwrList)
VorVelList<-ListMAF(VorVelList)
VorAreaList<-ListMAF(VorAreaList)
VorWidthList<-ListMAF(VorWidthList)
VorPKList<-ListMAF(VorPKList)


#*****************************************************************
#Calculate magnitude of Flow confidence
print(paste("Calc Flow Stats"))
FlowStats<-MCStats(VorFlowNewList)

#Calculate magnitude of Power confidence
print(paste("Calc Pwr Stats"))
PwrStats<-MCStats(VorPwrList)

#Calculate magnitude of Vel confidence
print(paste("Calc Vel Stats"))
VelStats<-MCStats(VorVelList)

#Calculate magnitude of Area confidence
print(paste("Calc Area Stats"))
AreaStats<-MCStats(VorAreaList)

#Calculate magnitude of Width confidence
print(paste("Calc Width Stats"))
WidthStats<-MCStats(VorWidthList)
```

```
#Calculate magnitude of PK confidence
print(paste("Calc PK Stats"))
PKStats<-MCStats(VorPKList)



#Save output
save(FlowStats,PwrStats,VelStats,AreaStats,WidthStats,PKStats,file=paste("Estimates_",Region,Alpha,Flow

#save(VorFlowNewList,VorAreaList,VorPwrList,VorVelList,VorAreaList,VorPKList,file=paste("EstimateFlows_
```

Appendix B

Attribute Descriptions

B.1 Soil Landscapes of Canada

ISO 19131 Soil Landscapes of Canada – Data Product Specification

Revision: A

Data specification: Soil Landscapes of Canada

- Table of Contents -

1. OVERVIEW 9
1.1. Informal description 9
1.2. Data product specification metadata 9
1.3. Terms and definitions 9
1.4. Abbreviations 10
2. SPECIFICATION SCOPE 10
3. DATA PRODUCT IDENTIFICATION 11
3.1. Data Series Identification 11
3.1.1. Soil Landscapes of Canada Version 2.2 11
3.1.2. Soil Landscapes of Canada Version 3.2 11
3.2. Data product identification 13
3.2.1. Soil Landscapes of Canada Version 2.2 - Boundaries 13
3.2.2. Soil Landscapes of Canada Version 2.2 – Component Table (CMP) 14
3.2.3. Soil Landscapes of Canada Version 2.2 – Landscape Table (LAT) 14
3.2.4. Soil Landscapes of Canada Version 2.2 – EcoZone Table 15
3.2.5. Soil Landscapes of Canada Version 2.2 – EcoRegion Table 15
3.2.6. Soil Landscapes of Canada Version 2.2 – EcoDistrict Table 16
3.2.7. Soil Landscapes of Canada Version 3.2 - Boundaries 17
3.2.8. Soil Landscapes of Canada Version 3.2 – Component Table (CMP) 17
3.2.9. Soil Landscapes of Canada Version 3.2 – Ecological Framework Table (EFT) 18
3.2.10. Soil Landscapes of Canada Version 3.2 – Landform Definition Table (LDT) 19
3.2.11. Soil Landscapes of Canada Version 3.2 – Landform Extent Table (LET) 19
3.2.12. Soil Landscapes of Canada Version 3.2 – Landscape Area Table (LAT) 20
3.2.13. Soil Landscapes of Canada Version 3.2 – Landscape Segmentation Table (LST) 21
3.2.14. Soil Landscapes of Canada Version 3.2 – Soil Layer Table (SLT) 21
3.2.15. Soil Landscapes of Canada Version 3.2 – Soil Name Table (SNT) 22
3.2.16. Soil Landscapes of Canada Version 3.2 – Component Rating Table (CRT) 23
4. CONTENT AND STRUCTURE 24
4.1. Feature-based application schema 24
4.2. Feature catalogue – Soil Landscapes of Canada Version 2.2 – Feature Catalogue .. 25
4.2.1. Feature attributes 25
4.2.1.1. Soil Landscapes Identifier 25

4.2.1.2. Ecodistrict Identifier 25
4.2.1.3. Component Number 26
Component of a Polygon 26
4.2.1.4. Percent of Polygon 26
Percent of polygon occupied by the component 26
4.2.1.5. Slope Gradient 26
4.2.1.6. Surface Stoniness 27
4.2.1.7. Local Surface Form 27
4.2.1.8. Province 34
4.2.1.9. Soil Code 34
4.2.1.10. Soil Code Modifier 35
4.2.1.11. Kind of Surface Material 35
4.2.1.12. Soil Name Identifier 36
4.2.1.13. Vegetation cover 36
4.2.1.14. Parent Material Mode of Deposition 39
4.2.1.15. Coarse Fragment Percentage 47
Coarse fragment content of the control section 47
4.2.1.16. Rooting Depth 48
Rooting depth, unrestricted 48
4.2.1.17. Drainage 48
4.2.1.18. Soil development 50
4.2.1.19. Calcareous class of parent material 53
Coarse fragment content of the control section 53
4.2.1.20. Land Area 54
4.2.1.21. Area of ocean and fresh water (ha) 54
4.2.1.22. Area of Fresh Water (ha) 54
4.2.1.23. Area Covered by Ocean (ha) 54
4.2.1.24. Area of the entire SLC polygon (ha) 55
4.2.1.25. Area of fresh water (ha) 55
4.2.1.26. Area of small water bodies (ha) 55
4.2.1.27. Area of small islands (ha) 56
4.2.1.28. Available water capacity (dominant soil landscape) 56
4.2.1.29. Available water capacity (subdominant soil landscape) 56
4.2.1.30. Reliability class of polygon 57
4.2.1.31. Complexity class of polygon 57

4.2.1.32. Percentage of the polygon occupied by the dominant soil landscape 59
4.2.1.33. Percentage of the polygon occupied by the subdominant soil landscape...... 59
4.2.1.34. Landform Segment Percent 59
4.2.1.35. Landform Segment Slope Percent 59
4.2.1.36. Landform Segment Slope Length 60
4.2.1.37. Landform Segment Name 60
4.2.1.38. Polygon Landform Identifier 60
4.2.1.39. Polygon Landform Percent 61
4.2.1.40. Agricultural Crop Use 61
4.2.1.41. Layer Number 61
4.2.1.42. Upper Depth 62
4.2.1.43. Lower Depth 62
4.2.1.44. Polygon Component Identifier 62
4.2.1.45. Ecoregion Identifier 62
4.2.1.46. Ecoprovince Identifier 63
4.2.1.47. Ecozone Identifier 63
4.2.1.48. Horizon Lithological Discontinuity 63
4.2.1.49. Horizon Master 64
4.2.1.50. Horizon Suffix 64
4.2.1.51. Horizon Modifier 64
4.2.1.52. Coarse Fragment Percentage 65
4.2.1.53. Dominant Sand Fraction 65
4.2.1.54. Very Fine Sand Percentage 65
4.2.1.55. Total Percentage Sand 66
4.2.1.56. Total Percentage Silt 66
4.2.1.57. Total Clay Percentage 66
4.2.1.58. Organic Carbon Percentage 66
4.2.1.59. pH in Calcium Chloride 67
4.2.1.60. Project Report pH 67
4.2.1.61. Base Saturation Percentage 67
4.2.1.62. Cation Exchange Capacity 68
4.2.1.63. Saturated Hydraulic Conductivity 68
4.2.1.64. Water Retention at 0 Kilopascals 68
4.2.1.65. Water Retention at 10 Kilopascals 68
4.2.1.66. Water Retention at 33 Kilopascals 69

4.2.1.67. Water Retention at 1500 Kilopascals ... 69
4.2.1.68. Bulk Density ... 69
4.2.1.69. Electrical Conductivity ... 70
4.2.1.70. Calcium Carbonate Percentage ... 70
4.2.1.71. Von Post Decomposition ... 70
4.2.1.72. Woody Material Percentage ... 70
4.2.1.73. Soil Name ... 71
4.2.1.74. Kind of Material ... 71
4.2.1.75. Water Table Type ... 71
4.2.1.76. Restrictive Layer ... 72
4.2.1.77. Soil Layer Restrictive Soil Type ... 72
4.2.1.78. Drainage Type ... 73
4.2.1.79. Parent Material 1 Texture ... 73
4.2.1.80. Parent Material 2 Texture ... 74
4.2.1.81. Parent Material 3 Texture ... 74
4.2.1.82. Parent Material 1 Chemical Composition ... 75
4.2.1.83. Parent Material 2 Chemical Composition ... 76
4.2.1.84. Parent Material 3 Chemical Composition ... 76
4.2.1.85. Parent Material 1 Mode of Deposition ... 77
4.2.1.86. Parent Material 2 Mode of Deposition ... 78
4.2.1.87. Parent Material 3 Mode of Deposition ... 79
4.2.1.88. CSSC 2nd Edition Soil Order ... 80
4.2.1.89. CSSC 2nd Edition Soil Great Group ... 80
4.2.1.90. CSSC 2nd Edition Soil Subgroup ... 80
4.2.1.91. CSSC 3rd Edition Soil Order ... 81
4.2.1.92. CSSC 3rd Edition Soil Great Group ... 81
4.2.1.93. CSSC 3rd Edition Soil Subgroup ... 81
4.2.1.94. 0 to 50 cm Coarse Fragment Content ... 81
4.2.1.95. 50 to 100 cm Coarse Fragment Content ... 82
4.2.1.96. Depth to Bedrock ... 82
4.2.1.97. Component Restrictive Soil Type ... 83
4.2.1.98. Water Capacity ... 83
4.3. Feature catalogue – Soil Landscapes of Canada Version 3.2 – Feature Catalogue ... 84
4.3.1. Feature attributes ... 84
4.3.1.1. Soil Landscapes Identifier ... 84

4.3.1.2. Ecodistrict Identifier 84
4.3.1.3. Component Number 85
4.3.1.4. Percent of Polygon 85
4.3.1.5. Slope Gradient 85
4.3.1.6. Surface Stoniness 86
4.3.1.7. Local Surface Form 86
4.3.1.8. Province 93
4.3.1.9. Soil Code 93
4.3.1.10. Soil Code Modifier 94
4.3.1.11. Soil Profile Type 94
4.3.1.12. Soil Name Identifier 94
4.3.1.13. Polygon Component Identifier 94
4.3.1.14. Ecoregion Identifier 95
4.3.1.15. Ecoprovince Identifier 95
4.3.1.16. Ecozone Identifier 95
4.3.1.17. Landform Segment Identifier 96
4.3.1.18. Landform Type 96
4.3.1.19. Landform Slope Type 96
4.3.1.20. Landform Segment Position 97
4.3.1.21. Landform Segment Percent 97
4.3.1.22. Landform Segment Slope Percent 97
4.3.1.23. Landform Segment Slope Length 98
4.3.1.24. Landform Segment Name 98
4.3.1.25. Polygon Landform Identifier 98
4.3.1.26. Polygon Landform Percent 99
4.3.1.27. Land Area 99
4.3.1.28. Total Water Area 99
4.3.1.29. Fresh Water Area 99
4.3.1.30. Ocean Water Area 100
4.3.1.31. Total Area 100
4.3.1.32. Agricultural Crop Use 100
4.3.1.33. Layer Number 101
4.3.1.34. Upper Depth 101
4.3.1.35. Lower Depth 101
4.3.1.36. Horizon Lithological Discontinuity 101

4.3.1.37. Horizon Master ... 102
4.3.1.38. Horizon Suffix ... 102
4.3.1.39. Horizon Modifier ... 103
4.3.1.40. Coarse Fragment Percentage ... 103
4.3.1.41. Dominant Sand Fraction ... 103
4.3.1.42. Very Fine Sand Percentage ... 104
4.3.1.43. Total Percentage Sand ... 104
4.3.1.44. Total Percentage Silt ... 104
4.3.1.45. Total Clay Percentage ... 104
4.3.1.46. Organic Carbon Percentage ... 105
4.3.1.47. pH in Calcium Chloride ... 105
4.3.1.48. Project Report pH ... 105
4.3.1.49. Base Saturation Percentage ... 105
4.3.1.50. Cation Exchange Capacity ... 106
4.3.1.51. Saturated Hydraulic Conductivity ... 106
4.3.1.52. Water Retention at 0 Kilopascals ... 106
4.3.1.53. Water Retention at 10 Kilopascals ... 107
4.3.1.54. Water Retention at 33 Kilopascals ... 107
4.3.1.55. Water Retention at 1500 Kilopascals ... 107
4.3.1.56. Bulk Density ... 107
4.3.1.57. Electrical Conductivity ... 108
4.3.1.58. Calcium Carbonate Percentage ... 108
4.3.1.59. Von Post Decomposition ... 108
4.3.1.60. Woody Material Percentage ... 109
4.3.1.61. Soil Name ... 109
4.3.1.62. Kind of Material ... 109
4.3.1.63. Water Table Type ... 110
4.3.1.64. Restrictive Layer ... 110
4.3.1.65. Soil Layer Restrictive Soil Type ... 110
4.3.1.66. Drainage Type ... 111
4.3.1.67. Parent Material 1 Texture ... 111
4.3.1.68. Parent Material 2 Texture ... 112
4.3.1.69. Parent Material 3 Texture ... 113
4.3.1.70. Parent Material 1 Chemical Composition ... 113
4.3.1.71. Parent Material 2 Chemical Composition ... 114

4.3.1.72. Parent Material 3 Chemical Composition 114
4.3.1.73. Parent Material 1 Mode of Deposition 115
4.3.1.74. Parent Material 2 Mode of Deposition 116
4.3.1.75. Parent Material 3 Mode of Deposition 117
4.3.1.76. CSSC 2nd Edition Soil Order 118
4.3.1.77. CSSC 2nd Edition Soil Great Group 118
4.3.1.78. CSSC 2nd Edition Soil Subgroup 118
4.3.1.79. CSSC 3rd Edition Soil Order 119
4.3.1.80. CSSC 3rd Edition Soil Great Group 119
4.3.1.81. CSSC 3rd Edition Soil Subgroup 119
4.3.1.82. 0 to 50 cm Coarse Fragment Content 120
4.3.1.83. 50 to 100 cm Coarse Fragment Content 120
4.3.1.84. Depth to Bedrock 120
4.3.1.85. Component Restrictive Soil Type 121
4.3.1.86. Water Capacity 121
5. REFERENCE SYSTEM 122
5.1. Spatial reference system 122
5.2. Temporal reference system 122
6. DATA QUALITY 122
6.1. Completeness 122
6.2. Logical consistency 122
6.3. Positional accuracy 122
6.4. Temporal accuracy 122
6.5. Thematic accuracy 122
6.6. Lineage statement 122
7. DATA CAPTURE 125
8. DATA MAINTENANCE 125
9. PORTRAYAL 125
10. DATA PRODUCT DELIVERY 125
11. METADATA 126

Data specification: Soil Landscapes of Canada

1. OVERVIEW

1.1. Informal description

The Soil Landscapes of Canada (SLC) are a series of datasets that show the major characteristics of soil and land for the whole country.

These datasets are based on existing soil survey maps that have been recompiled at a scale of 1:1 million. Each map area (a polygon) is described by a standard set of attributes. The full array of attributes that describe a distinct type of soil and its associated landscape (for example, surface form, slope, water table depth, permafrost and lakes) is called a soil landscape. SLC polygons may contain one or more distinct soil landscape components and may also contain small but highly contrasting components. However, the location of these components within the polygon is not defined.

SLC information was originally conceived as a standardized database consisting of major attributes important to plant growth, land management and soil degradation. This information has since turned out to be a useful framework to support other databases, including Environment Canada's Ecological Land Classification System.

SLC version 3.2 is the latest revision of the Soil Landscapes of Canada. It was developed by Agriculture and Agri-Food Canada to provide information about the country's agricultural soils at the provincial and national levels. SLC version 3.2 replaces SLC version 3.1.1.

1.2. Data product specification metadata

This section provides metadata about the creation of this data product specification.

Data product specification title:	Soil Landscapes of Canada
Data product specification reference date:	February 19, 2013
Data product specification responsible party:	Canadian Soil Information System (CanSIS)
Data product specification language:	English, French
Data product specification topic category:	Geoscientific Information

1.3. Terms and definitions

Additional technical and processing information can be found on the Canadian Soil Information System (CanSIS) web pages:

http://sis2.agr.gc.ca/cansis/nsdb/slc/index.html

- Feature attribute
  characteristic of a feature
- Class
  description of a set of objects that share the same attributes, operations, methods,

relationships, and semantics [UML Semantics]
NOTE: A class does not always have an associated geometry (e.g. the metadata class).

- Feature
  abstraction of real world phenomena

- Object
  entity with a well-defined boundary and identity that encapsulates state and behaviour [UML Semantics]
  NOTE: An object is an instance of a class.

- Package
  grouping of a set of classes, relationships, and even other packages with a view to organizing the model into more abstract structures

1.4. Abbreviations

AAFC	Agriculture and Agri-Food Canada
CanSIS	Canadian Soil Information System
CSSC	Canadian System of Soil Classification
SLC	Soil Landscapes of Canada
GIS	Geographic Information System
EPSG	European Petroleum Survey Group
NAD 83	North American Datum of 1983
HTTP	Hypertext Transfer Protocol
CMP	Component Table
CRT	Component Rating Table
SNT	Soil Names Table
SLT	Soil Layer Table
LST	Landscape Segmentation Table
LAT	Landscape Area Table
LDT	Landform Definition Table
LET	Landform Extent Table
EFT	Ecological Framework Table
NAHARP	National Agri-Environmental Health Analysis and Reporting Program
NCGAVS	National Carbon and Greenhouse Gas Emission Accounting and Verification System

2. SPECIFICATION SCOPE

This data specification has only one scope, the general scope.

NOTE: The term ‘specification scope’ originates from the International Standard ISO19131. ‘Specification scope’ does not express the purpose for the creation of a data specification or the potential use of data, but identifies partitions of the data specification where specific requirements apply.

3. DATA PRODUCT IDENTIFICATION

3.1. Data Series Identification

3.1.1. Soil Landscapes of Canada Version 2.2

Title	Soil Landscapes of Canada Version 2.2
Alternate Title	
Abstract	The “Soil Landscapes of Canada (SLC) Version 2.2” dataset series provides a set of geo-referenced soil areas (polygons) that are linked to attribute data found in the associated Component Table (CMP), Landscape Table (LAT), Carbon Layer Table (CLYR), and Dom/Sub File (DOM_SUB). Together, these datasets describe the spatial distribution of soils and associated landscapes for Canada.
Purpose	This version of SLC was developed by Agriculture and Agri-Food Canada to provide information about the country's soils at the provincial and national levels.
Topic Category	Geoscientific Information
Spatial Reference Type	vector
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	This version includes the following significant changes to the quality of the data and the structure of the dataset. 1. the contents of the Component Table have undergone an extensive manual review by regional CLRN pedologists. 2. The dataset has now gone through a formal set of validation procedures. 3. Water bodies are now found only in the HYDRO coverage, and the new Landscape Table must be used for all area calculations. Although SLC polygons and numbers are effectively unchanged by this release, they can be mapped only by displaying the SLC coverage in conjunction with the HYDRO coverage(see the rationale, and the new data model).
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope identification	series

3.1.2. Soil Landscapes of Canada Version 3.2

Title	Soil Landscapes of Canada Version 3.2
Alternate Title	
Abstract	The “Soil Landscapes of Canada (SLC) Version 3.2” dataset series provides a set of geo-referenced soil areas (polygons) that are linked to attribute data found in the associated Component Table (CMP), Component Rating Table (CRT), Soil Names Table (SNT), Soil Layer Table (SLT), Landscape Segmentation Table (LST), Landform Extent Table (LET), Landform Definition Table and Ecological Framework Table (EFT). Together, these datasets describe the spatial distribution of soils and associated landscapes for the agricultural areas of Canada. However, some provinces (Alberta, Nova Scotia, and Prince Edward Island) contain CMP, SNT and SLT data for the entire province (that is, beyond the agricultural areas). This version is complemented by the previous SLC version 2.2, which covers the entire country.
Purpose	This version of SLC was developed by Agriculture and Agri-Food Canada to provide information about the country's agricultural soils at the provincial and national levels.
Topic Category	Geoscientific Information
Spatial Reference Type	vector
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	The Soil Landscapes of Canada (SLC) is a series of datasets that show the major characteristics of soil and land for the whole country. They are based on existing soil survey maps which have been recompiled at a scale of 1:1,000,000. SLC polygons are the most detailed spatial entities within the ecological hierarchy employed in Canada: Ecozones --> Ecoprovinces --> Ecoregions--> Ecodistricts --> Soil Landscapes of Canada. The information is organized according to a uniform national set of soil and landscape criteria based on permanent natural attributes. Each area (or polygon) on the map is described by a standard set of attributes. The full array of attributes that describe a distinct type of soil and its associated landscape, such as surface form, slope, water table depth, permafrost and lakes, is called a soil landscape. SLC polygons may contain one or more distinct soil landscape components and may also contain small but highly contrasting inclusion components. The location of these components within the polygon is not defined. The Soil Landscapes of Canada was originally conceived as a standardized database consisting of major attributes important to plant growth, land management, and soil degradation. They provide an excellent ecological framework for integrating and characterizing the biological, physical, climatic and demographic dimensions of land-related information on a regional basis. This common national framework for base maps, soil landscape maps and ecological stratification maps allows for better integration of environmental modeling and assessment activities at national and regional scales. SLCs were compiled for the National Agri-Environmental Health Analysis and Reporting Program (NAHARP), National Carbon and Greenhouse Gas Emission Accounting and Verification System (NCGAVS) and other related national programs. These data have since turned out to be a useful framework to support other databases, including Environment Canada's National Ecological Framework for Canada. The creation of the SLC database has taken a number of years. Updates and corrections to boundaries, attributes and file structures have taken place over the years. New versions are released as major structural or attribute changes are implemented. SLC version 3.2 was released in February, 2011. No changes were made to the polygon line work from the earlier SLC 3.x versions. Version 3.2 retains essentially the data base tables found in the previous SLCv3.1.1 release, with some specific updates and changes to the data model as described below. It also includes a new Landform Definition Table (LDT), a Landform Extent Table (LET) and a Landscape Segmentation Table (LST). These tables indicate the major

	landform types within each SLC polygon, as well as the predicted location of the soil components within each landform. The landform data is provided primarily for agricultural polygons within the national SLCv3.2 coverage. Information in these tables is essentially equivalent to that found in the Component Table, but with greater spatial resolution. See the Cross-Reference section for more details on the association between the attribute and polygon data.
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope identification	series

3.2. Data product identification

3.2.1. Soil Landscapes of Canada Version 2.2 - Boundaries

Title	Soil Landscapes of Canada Version 2.2 – Boundaries (PED_CA_SLC_1M_V2_2)
Alternate Title	
Abstract	The “Soil Landscapes of Canada (SLC) Version 2.2 – Boundaries” dataset contains boundary information for the soil landscape polygons.
Purpose	The Boundaries dataset provides location information for the soil landscape polygons and links them to the associated tables of attributes.
Topic Category	Geoscientific Information
Spatial Reference Type	vector
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Soil Landscapes Identifier, Ecodistrict Identifier

3.2.2. Soil Landscapes of Canada Version 2.2 – Component Table (CMP)

Title	Soil Landscapes of Canada Version 2.2 - Component Table (CMP) (PED_CA_SLC_CMP_V2_2)
Alternate Title	
Abstract	The "Soil Landscapes of Canada (SLC) Version 2.2 – Component Table (CMP)" is a table that contains attributes that describe the individual soil and landscape features (components) of a soil landscape polygon. The attributes include the percentage of the polygon occupied by the component and slope, stoniness and local surface form information. The table also contains the Soil Landscapes Identifier attribute, which is used to join the CMP table to the Soil Names Table (SNT).
Purpose	The Component Table provides soil and landscape feature information for soil landscape polygons and a link to the Soil Names Table (SNT).
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Soil Landscapes Identifier, Component Number, Percent of Polygon, Kind of Surface Material, Vegetation Cover, Parent Material Mode of Deposition, Coarse Fragment Content, Rooting Depth, Drainage Class, Soil Development, Parent Material Calcareous Class, Local Surface Form, Slope Gradient, Soil Name Identifier, Soil Code, Soil Code Modifier

3.2.3. Soil Landscapes of Canada Version 2.2 – Landscape Table (LAT)

Title	Soil Landscapes of Canada Version 2.2 - Landscape Table (EFT) (PED_CA_SLC_LAT_V2_2)
Alternate Title	
Abstract	The "Soil Landscapes of Canada (SLC) version 2.2 – Landscape Table (LAT)" dataset is a table that provides the attributes of the polygon, and attributes which apply to the entire landscape represented by that polygon. All areas are in hectares.
Purpose	The Landscape Attribute Table was developed to simplify the calculation of land and water areas for the SLC polygons, and remove ambiguities in the way water is represented in previous versions. This table was derived from the SLC and HYDRO coverages.
Topic Category	Geoscientific Information
Spatial Reference Type	textTable

Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Soil Landscape Identifier, Total Area of Fresh Water (HA), Area Covered by Ocean (HA), Land Area (HA), Total Area (HA), Area of Small Islands (HA), Area of Small Water Bodies (HA), Fresh Water (HA), Area of Ocean and Fresh Water

3.2.4. Soil Landscapes of Canada Version 2.2 – EcoZone Table

Title	Soil Landscapes of Canada Version 2.2 - EcoZone Table (PED_CA_SLC_ECOZONE_V2_2)
Alternate Title	
Abstract	The "Soil Landscapes of Canada (SLC) version 2.2 – EcoZone Table" dataset is a table that provides the linking attributes from the SLC to the Ecozone.
Purpose	To link the SLC polygons to the EcoZone polygons.
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	EcoZone, EcoZone Name, EcoZone Nom

3.2.5. Soil Landscapes of Canada Version 2.2 – EcoRegion Table

Title	Soil Landscapes of Canada Version 2.2 - EcoRegion Table

	(PED_CA_SLC_ECOREGION_V2_2)
Alternate Title	
Abstract	The "Soil Landscapes of Canada (SLC) version 2.2 – EcoRegion Table" dataset is a table that provides the linking attributes from the SLC to the EcoRegion.
Purpose	To link the SLC polygons to the EcoRegion polygons.
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	EcoRegion, EcoRegion Name, EcoRegion Nom

3.2.6. Soil Landscapes of Canada Version 2.2 – EcoDistrict Table

Title	Soil Landscapes of Canada Version 2.2 - EcoDistrict Table (PED_CA_SLC_ECODISTRICT_V2_2)
Alternate Title	
Abstract	The "Soil Landscapes of Canada (SLC) version 2.2 – EcoDistrict Table" dataset is a table that provides the linking attributes from the SLC to the EcoDistrict.
Purpose	To link the SLC polygons to the EcoDistrict polygons.
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial

	ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	EcoDistrict, EcoDistrict Name, EcoDistrict Nom

3.2.7. Soil Landscapes of Canada Version 3.2 - Boundaries

Title	Soil Landscapes of Canada Version 2.2 - EcoZone Table (PED_CA_SLC_ECOZONE_V2_2)
Alternate Title	
Abstract	The "Soil Landscapes of Canada (SLC) version 2.2 – EcoZone Table" dataset is a table that provides the linking attributes from the SLC to the Ecozone.
Purpose	
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Ecozone, Ecozone Name, Ecozone Nom

3.2.8. Soil Landscapes of Canada Version 3.2 – Component Table (CMP)

Title	Soil Landscapes of Canada Version 3.2 - Component Table (CMP) (PED_CA_SLC_CMP_V3_2)
Alternate Title	
Abstract	The "Soil Landscapes of Canada (SLC) Version 3.2 – Component Table (CMP)" is a table that contains attributes that describe the individual soil and landscape features (components) of a soil landscape polygon. The attributes include the percentage of the polygon occupied by the component and slope, stoniness and local surface form information. The table also contains the Soil Landscapes Identifier attribute, which is used to join the CMP table to the Soil Names Table (SNT).
Purpose	The Component Table provides soil and landscape feature information for soil landscape polygons and a link to the Soil Names Table (SNT).
Topic Category	Geoscientific Information
Spatial Reference	textTable

Type	
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Soil Landscapes Identifier, Component Number, Percent of Polygon, Slope Gradient, Surface Stoniness, Local Surface Form, Province, Soil Code, Soil Code Modifier, Soil Profile Type, Soil Name Identifier, Polygon Component Identifier

3.2.9. Soil Landscapes of Canada Version 3.2 – Ecological Framework Table (EFT)

Title	Soil Landscapes of Canada Version 3.2 - Ecological Framework Table (EFT) (PED_CA_SLC_EFT_V3_2)
Alternate Title	
Abstract	The “Soil Landscapes of Canada (SLC) version 3.2 – Ecological Framework Table (EFT)” dataset is a table that provides the attributes of ecodistricts. Ecodistricts are made up of groups of soil landscape polygons and can, in turn, be grouped into larger ecoregions, ecoprovinces and ecoregions. An ecodistrict is defined by characteristics such as relief, landforms, geology, soil, vegetation, water bodies and fauna.
Purpose	The Ecological Framework Table (EFT) contains the attributes of ecodistricts.
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Ecodistrict Identifier, Ecoregion Identifier, Ecoprovince Identifier, Ecozone Identifier

3.2.10. Soil Landscapes of Canada Version 3.2 – Landform Definition Table (LDT)

Title	Soil Landscapes of Canada Version 3.2 - Landform Definition Table (LDT) (PED_CA_SLC_LDT_V3_2)
Alternate Title	
Abstract	The “Soil Landscapes of Canada (SLC) Version 3.2 - Landform Definition Table (LDT)” dataset defines a set of 84 possible landscape segments. Each of 21 landform types is defined using four segments that represent the upper slope, mid slope, lower slope and depression points.
Purpose	The Landform Definition Table (LDT) provides descriptions of landform segments found in soil landscape polygons.
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Landform Segment Identifier, Landform Type, Landform Slope Type, Landform Segment Position, Landform Segment Percent, Landform Segment Slope Percent, Landform Segment Slope Length, Landform Segment Name

3.2.11. Soil Landscapes of Canada Version 3.2 – Landform Extent Table (LET)

Title	Soil Landscapes of Canada Version 3.2 - Landform Extent Table (LET) (PED_CA_SLC_LET_V3_2)
Alternate Title	
Abstract	The “Soil Landscapes of Canada (SLC) Version 3.2 - Landform Extent Table (LET)” dataset is a table that contains attributes that describe the predominant landforms found within each soil landscape polygon. For example, the landform type can be level, rolling, ridged or steep and its slope can be gentle, moderate or steep. LET also contains information about the percentage of the polygon that each landform occupies.
Purpose	The Landform Extent Table provides information about the predominant landforms found within each soil landscape polygon and the percentage of the polygon each landform occupies.
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic	This specification is applicable within the extent of Canada.

Description	
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Polygon Landform Identifier, Soil Landscapes Identifier, Landform Type, Landform Slope Type, Polygon Landform Percent

3.2.12. Soil Landscapes of Canada Version 3.2 – Landscape Area Table (LAT)

Title	Soil Landscapes of Canada Version 3.2 - Landscape Area Table (LAT) (PED_CA_SLC_LAT_V3_2)
Alternate Title	
Abstract	The “Soil Landscapes of Canada (SLC) Version 3.2 - Landscape Area Table” dataset is a table that contains information about the total area, land area and area covered by water in a soil landscape polygon. The area attribute values for each soil landscape polygon were derived from the HYDRO coverage, which represents land, fresh water and marine water within Canada.
Purpose	This dataset provides information about the total area of a soil landscape polygon and its fresh, ocean and total water area.
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Soil Landscapes Identifier, Land Area, Total Water Area, Fresh Water Area, Ocean Water Area, Total Area

3.2.13. Soil Landscapes of Canada Version 3.2 – Landscape Segmentation Table (LST)

Title	Soil Landscapes of Canada Version 3.2 - Landscape Segmentation Table (LST) (PED_CA_SLC_LST_V3_2)
Alternate Title	
Abstract	The "Soil Landscapes of Canada (SLC) Version 3.2 - Landscape Segmentation Table (LST)" dataset is a table that describes the main landform and slope combinations in each soil landscape polygon, and the representative soils found in each of the associated landscape segments. Each soil landscape polygon with soil data has a dominant landform, and some complex polygons may also have a second and third significant landform. The landscape segment is identified by the Landform Segment Identifier attribute, while the estimated extent of the landscape segment within the soil landscape polygon is specified by the Percent of Polygon attribute. For polygons with a single landscape segment, the value of Percent of Polygon is 100.
Purpose	The Landscape Segmentation Table provides information about the main landform and slope combinations located in a soil landscape polygon.
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Soil Landscapes Identifier, Landform Type, Landform Slope Type, Landform Segment Position, Percent of Polygon, Agricultural Crop Use, Province, Soil Code, Soil Code Modifier, Soil Profile Type, Polygon Landform Identifier, Landform Segment Identifier, Soil Name Identifier

3.2.14. Soil Landscapes of Canada Version 3.2 – Soil Layer Table (SLT)

Title	Soil Landscapes of Canada Version 3.2 - Soil Layer Table (SLT) (PED_CA_SLC_SLT_V3_2)
Alternate Title	
Abstract	The "Soil Landscapes of Canada (SLC) Version 3.2 - Soil Layer Table (SLT)" dataset is a table that describes the chemical and physical properties of layers (horizons) associated with each soil name in the Soil Name Table (SNT).
Purpose	The Soil Layer Table provides attributes that describe the chemical and physical properties of layers (horizons) associated with soil names in the Soil Names Table.
Topic Category	Geoscientific Information
Spatial	textTable

Reference Type	
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Soil Name Identifier, Province, Soil Code, Soil Code Modifier Soil Profile Type, Layer Number, Upper Depth, Lower Depth, Horizon Lithological Discontinuity, Horizon Master, Horizon Suffix, Horizon Modifier, Coarse Fragment Percentage, Dominant Sand Fraction, Very Fine Sand Percentage, Total Percentage Sand, Total Percentage Silt, Total Clay Percentage, Organic Carbon Percentage, pH in Calcium Chloride, Project Report pH, Base Saturation Percentage, Cation Exchange Capacity, Saturated Hydraulic Conductivity, Water Retention at 0 Kilopascals, Water Retention at 10 Kilopascals, Water Retention at 33 Kilopascals, Water Retention at 1500 Kilopascals, Bulk Density, Electrical Conductivity, Calcium Carbonate Percentage, Von Post Decomposition, Woody Material Percentage

3.2.15. Soil Landscapes of Canada Version 3.2 – Soil Name Table (SNT)

Title	Soil Landscapes of Canada Version 3.2 - Soil Name Table (SNT) (PED_CA_SLC_SNT_V3_2)
Alternate Title	
Abstract	The “Soil Landscapes of Canada (SLC) Version 3.2 – Soil Name Table (SNT)” is a table that provides general soil information.
Purpose	The Soil Names Table provides general soil information.
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset

Feature Attribute Names	Soil Name Identifier, Province, Soil Code, Soil Code Modifier, Soil Profile Type, Soil Name, Kind of Material, Water Table Type, Restrictive Layer, Soil Layer Restrictive Soil Type, Drainage Type, Parent Material 1 Texture, Parent Material 2 Texture, Parent Material 3 Texture, Parent Material 1 Chemical Composition, Parent Material 2 Chemical Composition, Parent Material 3 Chemical Composition, Parent Material 1 Mode of Deposition, Parent Material 2 Mode of Deposition, Parent Material 3 Mode of Deposition, CSSC 2nd Edition Soil Order, CSSC 2nd Edition Soil Great Group, CSSC 2nd Edition Soil Subgroup, CSSC 3rd Edition Soil Order, CSSC 3rd Edition Soil Great Group, CSSC 3rd Edition Soil Subgroup

3.2.16. Soil Landscapes of Canada Version 3.2 – Component Rating Table (CRT)

Title	Soil Landscapes of Canada Version 3.2- Component Rating Table (CRT) (PED_CA_SLC_CRT_V3_2)
Alternate Title	
Abstract	The “Soil Landscapes of Canada (SLC) Version 3.2 - Component Rating Table (CRT)” is a table that contains soil and landform suitability ratings that describe individual components within soil landscape polygon. These ratings are derived from the contents of the Soil Names Table (SNT) and Soil Layers Table (SLT).
Purpose	The Component Rating Table (CRT) provides soil and landform suitability ratings for soil landscape polygons.
Topic Category	Geoscientific Information
Spatial Reference Type	textTable
Spatial Resolution	1:1,000,000
Geographic Description	This specification is applicable within the extent of Canada.
Supplemental Information	
Constraints	Data are subject to the Government of Canada Open Data License Agreement: http://www.data.gc.ca Her Majesty the Queen in right of Canada (2011), as represented by the Minister of Agriculture and Agri-Food Canada. All rights reserved.
Keywords	Thesaurus: Government of Canada Core Subject Thesaurus (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html) Date: February 1, 2000 Keywords: Earth sciences, Agriculture, Soil Taxonomy, Soil, Terrestrial ecosystems, Wetlands
Scope Identification	dataset
Feature Attribute Names	Polygon Component Identifier, 0 to 50 Coarse Fragment Content, 50 to 100 Coarse Fragment Content, Depth to Bedrock, Component Restrictive Soil Type, Water Capacity

4. CONTENT AND STRUCTURE

4.1. Feature-based application schema

4.2. Feature catalogue – Soil Landscapes of Canada Version 2.2 – Feature Catalogue

Title	Soil Landscapes of Canada Version 2.2 - Feature Catalogue
Scope	
Version Number	1.0
Version Date	2013-01-23
Producer	Agri-Geomatics

System-generated attributes (for example, OBJECTID, Shape, Shape Length and Area) are not defined in the feature catalog.

4.2.1. Feature attributes

4.2.1.1. Soil Landscapes Identifier

Name	Soil Landscapes Polygon Identifier (SL)		
Definition	This polygon identifier is used to link the digital map data to attribute data in soil landscape files and other associated files. This number is not necessarily unique within the country. This identifier contains two parts: • the first two digits identify the map sheet, and • the remaining four digits identify the number assigned to each soil landscape polygon of the map sheet.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.2. Ecodistrict Identifier

Name	Ecodistrict Identifier (ECODISTRICT)
Definition	This four digit integer field indicates which • EcoDistrict, and thus • EcoRegion and • EcoZone the polygon lies within.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Integer
Value Domain Type	0 (not enumerated)
Value Domain	

	Feature Attribute Value		
	Label	Code	Definition

4.2.1.3. Component Number

Name	Component Number (CMP)		
Definition	An arbitrarily assigned number that uniquely identifies the component within the polygon.		
Aliases	Component of a Polygon		
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.4. Percent of Polygon

Name	Percent of Polygon (PERCENT)		
Definition	This figure identifies the percentage of the polygon that is occupied by the component. In the original version 1.0 Component Table, components occupying less than 16% of the polygon were coded only if they had a significant impact on the spatial estimation of soil carbon (see list below). In these cases, a maximum of two components less than 16% were coded.		
Aliases	Percent of polygon occupied by the component		
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.5. Slope Gradient

Name	Slope Gradient (SLOPE)
Definition	SLOPE contains categorized slope gradient (in percent) of polygon components for the low end of range (Shields, 1982).
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Character
Value Domain Type	1 (enumerated)
Value Domain	

	Feature Attribute Value		
	Label	Code	Definition
	Little or none	*A*	< 4%
	Gentle	*B*	4 - 9 %
	Moderate	*C*	10 - 15 %
	Steep	*D*	16 - 30 %
	Extremely steep	*E*	31 - 60%
	Excessively steep	*F*	> 60%
		n/a	Not applicable (water)

4.2.1.6. Surface Stoniness

Name	Surface Stoniness (STONE_TYPE_CODE)		
Definition	Describes the stoniness of the soil surface, which has an impact on a soil's agricultural capabilities.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	None	*N*	
	Slightly Stony	*S*	
	Very Stony	*V*	
	Unknown	*U*	

4.2.1.7. Local Surface Form

Name	Local Surface Form (LOCSF)		
Definition	Descriptions define classes of local physical surface forms (assemblage of slopes) or recurring patterns of forms which occur at the earth's surface. When applied to consolidated materials, form refers to the product of their modification by geological processes. Select only one code per soil landscape, either from mineral surface forms or from wetland surface forms.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Dissected	*D*	A dissected (or gullied) pattern providing external drainage for an area.
	Hummocky (or irregular)	*H*	A very complex sequence of slopes extending from

			somewhat rounded concavities (or swales) of various sizes to irregular conical knolls (or knobs) and short discontinuous ridges; there is a general lack of concordance between knolls and swales. Slopes are generally 4-70%. Examples are hummocky moraines and hummocky fluvioglacial landforms.
	Inclined	*I*	A sloping, unidirectional surface with a generally constant slope unbroken by marked irregularity or gullies; a weakly developed dissected pattern provides external drainage for the local area. Slopes are 2-70%; the form of inclined slopes is not related to the initial mode of origin of the underlying material.
	Knoll and kettle	*K*	A chaotic sequence of knolls and kettles (or sloughs), which occupies 15-20% of an area and has no external drainage. Slopes are generally >3%. Examples are morainal plains and hill lands.
	Level	*L*	A flat or very gently sloping, unidirectional surface with a generally constant slope unbroken by marked elevations and depressions. Slopes are generally <2%. Examples are floodplains and lake plains.
	Rolling	*M*	A very regular sequence of moderate slopes

			extending from rounded and, in some places, confined concave depressions to broad, rounded convexities producing a wavelike pattern of moderate relief. Slope gradients are generally >5% but may be less. This surface form is usually controlled by the underlying bedrock.
	Ridged	*R*	A long, narrow elevation of the surface, usually sharp crested with steep sides; ridges may be parallel, subparallel, or intersecting. Examples are eskers, crevasse fillings, washboard moraines and some drumlins.
	Steep	*S*	Erosional slopes of >70%, present on both consolidated and unconsolidated materials. An example is an escarpment.
	Terraced	*T*	Scarp face and the horizontal or gently inclined surface (or tread) above it. An example is an alluvial terrace.
	Undulating	*U*	A very regular sequence of gentle slopes that extends from rounded and, in some places, confined concavities to broad, rounded convexities producing a wavelike pattern of low local relief. Slope length is generally <0.8 km and the dominant gradient of slopes is usually 2-5%. The terrain lacks an external drainage

			pattern. Examples are some ground moraines and lacustrine material of varying textures.
	Domed bog	*B04*	A large bog (diameter usually >500 m) with a convex surface rising several metres above the surrounding terrain. The centre usually drains in all directions; small crescentic pools commonly form around the highest point; a concentric pattern is formed if the highest point is in the centre, while an eccentric pattern is formed if the highest point is off-centre. Peat development is usually >3 m.
	Polygonal peat bog	*B05*	A perennially frozen bog rising approximately 1 m above the surrounding fen. The surface is relatively flat, scored by a polygonal pattern of trenches that developed over ice wedges. The permafrost and ice wedges developed in peat originally deposited in a nonpermafrost environment.
	Peat plateau bog	*B07*	A bog composed of perennially frozen peat rising abruptly about 1 m from the surrounding unfrozen fen. The surface is relatively flat and even, and commonly covers large areas. The peat was originally deposited in a nonpermafrost environment and is associated in many

			places with collapse scar bogs or fens.
	Atlantic plateau bog	*B09*	A bog with a flat to undulating surface raised above the surrounding terrain. The bog edges commonly slope steeply downwards to the mineral soil terrain. Large pools scattered on the bog reach depths of 2-4 m.
	Basin bog	*B13*	A bog situated in a basin with essentially closed drainage which receives water from precipitation and runoff from the immediate surroundings. The surface of the bog is flat with peat generally deepest at the centre.
	Flat bog	*B14*	A bog having a flat, featureless surface and occurring in broad, poorly defined depressions. The depth of peat is generally uniform.
	String bog	*B15*	A pattern of narrow (2-3 m wide), low (<1 m high) ridges oriented at right angles to the direction of drainage; wet depressions or pools occur between the ridges. The water and peat are very low in nutrients because the water has been derived from other ombrotrophic wetlands. The peat thickness is >1 m.
	Blanket bog	*B16*	A bog consisting of extensive peat deposits that occur more or less uniformly over gently sloping hills and valleys. The peat thickness is usually

			<2 m.
	Slope bog	*B18*	A bog occurring in areas of high rainfall on appreciably sloping land surfaces. The bog is fed by rainwater and by water draining from other nutrient-poor peatlands. The peat may exceed 1 m in thickness.
	Veneer bog	*B19*	A bog occurring on gently sloping terrain underlain by generally discontinuous permafrost. Although drainage is predominantly below the surface, overland flow occurs in poorly defined drainways during peak runoff. Peat thickness is usually <1.5 m.
	Lowland polygon bog	*B20*	A bog with flat-topped or convex peat surfaces (often referred to as "high-centre polygons") separated by trenches over ice wedges that form a polygonal pattern when viewed from above. The peat was deposited in a permafrost environment as shown by internal structures.
	Northern ribbed fen	*F01*	A fen with parallel, low peat ridges ("strings") alternating with wet hollows or shallow pools, oriented across the major slope at right angles to water movement. The depth of peat is >1 m.
	Shore fen	*F07*	A fen with an anchored surface mat that forms the shore of a pond or lake. The rooting zone is affected by the water of the lake

			at both normal and flood levels.
	Slope fen	*F11*	A fen occurring mainly on slow-draining, nutrient-enriched seepage slopes. Pools are usually absent, but wet seepage tracks may occur. Peat thickness is usually <2 m.
	Horizontal fen	*F13*	A fen with a gently sloping, featureless surface. This fen occupies broad, often ill-defined depressions and may interconnect with other fens. Peat accumulation is generally uniform.
	Stream swamp	*S01*	A swamp occurring along the banks of permanent or semipermanent streams. The high water table is maintained by the level of water in the stream. The swamp is seasonally inundated with subsequent sediment deposition.
	Basin swamp	*S04*	A swamp developed in a topographically defined basin where water derived locally may be augmented by drainage from other parts of the watershed. Accumulation of well-decomposed peat is shallow (<0.5 m) at the edge but may reach 2 m at the centre.
	Stream marsh	*M06*	A marsh occupying shorelines, bars, stream beds, or islands in continuously flowing watercourses. The marsh is subject to prolonged annual flooding and is commonly covered by thick layers of

			sediments.
	Shallow basin marsh	*M11*	A marsh occurring in a uniformly shallow marsh depression or swale, having a gradual gradient from the edge to the deepest portion; the marsh edge may be poorly defined. Water levels fluctuate rapidly.
	Shore marsh	*M14*	A marsh occupying the contact zone between high and low water marks bordering semipermanent or permanent lakes. The marsh,usually found along protected shorelines, behind barrier beaches in lagoons, on islands, or in embayments, is subject to flooding by a rise in lake levels, wind waves, or surface runoff.

4.2.1.8. Province

Name	Province (PROVINCE_CODE)		
Definition	Specifies the province.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.9. Soil Code

Name	Soil Code (SOILCODE)
Definition	A representative soil code (and modifier when present) is included in this detailed legend to provide a link to the national CanSIS soil name and soil layer files. Data in the soil layer file is required for agricultural and nonagricultural use interpretations, simulation modeling and general land evaluation studies. For areas where no soil name is established, # is used for non-applicable.
Aliases	
Producer	Canadian Soil Information System (CanSIS)

Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.10. Soil Code Modifier

Name	Soil Code Modifier (MODIFIER)		
Definition	Soil code modifiers are used with the Soil Code to define a named soil with more detailed information. As with theSoil Code, modifiers provide a link to the CanSIS soil name and soil layer files.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.11. Kind of Surface Material

Name	Kind of Surface Material (KINDMAT)		
Definition	Kind of soil, rock outcrop or other material at the surface.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Ice and snow	*IC*	Glacial ice and permanent snow.
	Organic soil	*OR*	Contains >30% organic matter as measured by weight.
	Soft rock	*R1*	Rock that can be dug with a shovel (i.e. undifferentiated shales, upper Cretaceous and Tertiary materials).
	Hard rock, acidic	*R2*	Granite.
	Hard rock, basic	*R3*	Limestone.
	Hard rock	*R4*	Hard rock of unspecified origin and undifferentiated properties.
	Mineral soil	*SO*	Predominantly

			mineral particles; contains <30% organic matter as measured by weight.
	Water	*WA*	Water.
	Urban	*UR*	Urban areas. Note that only a few major urban area polygons are included on maps, therefore, do not use for tabulating total urban coverage.
	Rock field	*F*	Frost-heaved bedrock, rock slides and talus material.

4.2.1.12. Soil Name Identifier

Name	Soil Name Identifier (SNF)		
Definition	This code identifies the provincial Soil Name File to which the soil code and modifier refer.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.13. Vegetation cover

Name	Vegetation cover (VEGET)		
Definition	Vegetation cover or land use, or both.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Agricultural crops	*A*	Cultivated field crops.
	Bog	*B*	Bogs may be treed or treeless and are usually covered with Sphagnum spp. and ericaceous shrubs.
	Coniferous forest	*C*	Dominated by needle-leaved, cone-bearing species. In the

			Subarctic Ecoclimatic Regions this refers to an open lichen coniferous forest; in the Boreal Ecoclimatic Regions this refers to a closed canopy coniferous forest.
	Deciduous forest	*D*	Dominated by broadleaf species.
	Fen	*F*	Dominated by sedges, grasses, reeds, and brown mosses with some shrubs and, at times, a sparse tree layer.
	Grassland	*G*	Perennial native grassland or improved pasture.
	Arctic desert	*H*	Unvegetated areas in the High Arctic; may be caused by either climatic (too cold or too dry) or edaphic (low soil nutrients or toxic substrates such as salt) factors, or a combination of both.
	Lichen	*L*	Dominated by lichens with significant amounts of mosses and usually including low to medium ericaceous shrubs.
	Mixed forest	*M*	Composed of both coniferous and deciduous tree species (refer to codes C and D).
	Parkland	*P*	A forest - grassland transition consisting of a mosaic of trembling aspen stands interspersed with patches of cropland, grassland, and meadow.
	Marshland	*R*	A mosaic surface pattern composed of pools or channels interspersed with clumps of emergent sedges, grasses, rushes, and reeds, and bordered by

			grassy meadows and peripheral bands of shrubs or trees; submerged and floating aquatics flourish in open water areas.
	Shrubland	*S*	Dominated by shrub species.
	Sedge peat	*SP*	Dominated by Carex spp. and generally moderately decomposed and matted; the sedge leaves are readily identifiable to the naked eye.
	Tundra, alpine	*TA*	Treeless terrain occurring at high altitudes, immediately above the forest zone and the upper altitudinal timberline; vegetation consists of lichens, mosses, sedges, grasses, forbs, and low shrubs (<20 cm) such as heath, dwarf willows, and birches.
	Tundra, high shrub	*TH*	Dominated by 20-60 cm high shrubs occurring in the Low Arctic Ecoclimatic Region.
	Tundra, medium shrub	*TM*	Dominated by 10 - 20 cm high shrubs occurring in the Mid-Arctic Ecoclimatic Region.
	Tundra, low shrub	*TL*	Dominated by <10 cm high shrubs occurring in the High Arctic Ecoclimatic Region.
	Tundra, broken herb - low shrub	*TB*	Dwarf shrubs, Dryas and willows are important components of the vegetation. Other vascular plants are present but have low total coverage. Crustose lichens are common on the ground surface. Ground cover ranges from 10-20%.

	Unvegetated surface	*U*	Unvegetated surface
	Meadow, wet	*W*	Dominated by sedges and cotton grass, with "wet" mosses (e.g., Mnium spp., Sphagnum spp.) and occasional herbs.
		#	Non-applicable

4.2.1.14. Parent Material Mode of Deposition

Name	Parent Material Mode of Deposition (PMDEP)		
Definition	The mode of deposition of mineral materials and undifferentiated organic materials is shown by a single alpha code whereas the origin of specified organic material is given by a numeric, two-digit code.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Alluvial	*A*	Sediment, generally consisting of gravel and sand with a minor fraction of silt and clay. The gravels are typically rounded and contain interstitial sand; alluvial sediments are moderately to well sorted and display stratification. Examples are channel deposits, overbank deposits, terraces, alluvial fans, and deltas.
	Bog	*B*	Bogs consist of unspecified organic materials formed in an ombrotrophic (nutrient-poor) environment caused by the slightly elevated nature of the bog which dissociates it from nutrient-rich ground water or surrounding mineral soils. Near the surface, materials are usually

			undecomposed (fibric), yellowish to pale brown, loose and spongy in consistence, and entire sphagnum plants are readily identified. These materials are extremely acid, with low bulk density and high fibre content; at lower depths they become darker, compacted, and somewhat layered. Bogs are associated with slopes or depressions on topography with a water table at or near the surface in the spring and slightly below it during the rest of the year. They are usually covered with sphagnum mosses, but sedges may also grow on them. Bogs may be treed or treeless and many are characterized by a layer of ericaceous shrubs.
	Colluvial	*C*	Massive to moderately-well stratified, nonsorted to poorly sorted sediments with any range of particle sizes from clay to boulders that have reached their present position only by direct, gravity-induced movement (except snow avalanches). Processes include slow displacements such as creep and solifluction and rapid movements such as earth flows.
	Residual	*D*	Unconsolidated, weathered, or partly weathered mineral soil material that accumulates by in situ disintegration of

			bedrock.
	Eolian	*E*	Sediment, generally consisting of medium to fine sand and coarse silt particles, which is well sorted and poorly compacted. It may be massive, or show internal structures such as cross bedding or ripple laminae. Individual grains may be rounded and show signs of frosting. These materials have been transported and deposited by wind action. Examples are dunes, shallow deposits of sand and coarse silt, and loess.
	Fluvioglacial	*F*	Material moved by glaciers and subsequently sorted and deposited by streams flowing from the melting ice. Deposits are stratified and may occur in the form of outwash plains, deltas, kames, eskers, and kame terraces.
	Marsh	*H*	Mineral wetland or peatland that is periodically inundated by standing or slow-moving water. Surface water levels may fluctuate seasonally, with declining levels exposing drawdown zones of matted vegetation or mudflats. Waters are rich in nutrients, varying from fresh to highly saline. Substratum usually consists of mineral material, although in some places it consists of well-

			decomposed peat. Soils are predominantly Gleysols, with some Humisols and Mesisols. Marshes characteristically show zonal or mosaic surface patterns composed of pools or channels interspersed with clumps of emergent sedges, grasses, rushes, and reeds, and are bordered by grassy meadows and peripheral bands of shrubs or trees. Submerged and floating aquatics flourish where open-water areas occur.
	Folic	*I*	Upland organic material.
	Lacustrine	*L*	Sediment, generally consisting of either stratified fine sand, silt, and clay deposited on the lake bed, or moderately-well sorted, stratified sand and coarse materials that consist of near lake shore or beach deposits. These materials have either settled from suspension in bodies of standing fresh water or accumulated at their margins through wave action.
	Morainal	*M*	Sediment, generally consisting of well-compacted material that is nonstratified and contains a heterogeneous mixture of sand, silt, and clay particles and coarse fragments in a mixture that has been transported beneath, beside, on, within, or in front of a

			glacier but not modified by any intermediate agent. Examples are basal till (ground moraine), lateral and terminal moraines, rubbly moraines of cirque glaciers, hummocky ice-disintegration moraines, and pre-existing, unconsolidated sediments reworked by a glacier so that their original character is largely or completely destroyed.
	Fen	*N*	Fens consist of unspecified organic materials formed in a minerotrophic environment due to the close association of the material with mineral-rich waters. The material is usually moderately-well to well decomposed, dark brown to black, with fine- to medium-sized fibres; decomposition often becomes greater at lower depths. The materials are covered with a dominant component of sedges, but grasses and reeds may be associated in local pools.
	Organic	*O*	A layered sequence of more than three undifferentiated types of organic material (>30% organic matter by weight).
	Rock	*R*	A consolidated bedrock layer that is too hard to break with the hands (>3 on Mohs' scale) or to dig with a spade when moist.
	Swamp	*S*	Minerotrophic wetlands with the

			water table at or above the peat surface. The dominant unspecified organic materials are forest and fen peat formed in a eutrophic (nutrient-rich) environment due to strong water movement from the margins or other mineral sources. The material is usually moderately well to well decomposed and has a dark brown to reddish brown matrix; the more decomposed materials are black. It has an amorphous or very fine-fibred structure containing a random distribution of woody fragments and trunks of coniferous tree species. The vegetation cover may consist of coniferous or deciduous trees, tall shrubs, herbs, and mosses; in some regions sphagnum mosses are abundant.
	Anthropogenic	*T*	Materials modified by people,including those associated with mineral exploitation and waste disposal. They include materials deposited as a result of human activities or geological materials modified artificially so that their physical properties (structure, cohesion, compaction) have been drastically altered. Examples are areas of landfill, spoil heaps, open-pit mines and levelled irrigated areas.

	Undifferentiated	*U*	A sequence of more than three types of genetic mineral materials outcropping on a steep erosional escarpment. This complex class is to be used where units relating to individual genetic materials cannot be delimited separately at the scale of mapping. It may include colluvium derived from the various genetic materials and resting upon the scarp slope.
	Volcanic	*V*	Volcanic pumice and ash.
	Marine	*W*	Unconsolidated deposits of clay, silt, sand, or gravel that are well to moderately well sorted and well to moderately well stratified (in some places containing shells). The deposits have settled from suspension in salt or brackish water bodies or have accumulated at their margins through shoreline processes such as wave action and longshore drift. Nonfossiliferous deposits may be judged marine if they are located in an area that may reasonably be considered to have contained salt water at the time the deposits were formed.
	Fibric sphagnum	*11*	Sphagnum organic material having a fibric (slight) degree of decomposition in which the fibric materials are readily identifiable as to botanical origin. The

			peat is usually undecomposed (or fibric), light yellowish brown to pale brown, and loose and spongy in consistency with the entire sphagnum plant being readily identifiable.
	Fibric sedge and moss	*20*	Sedge and brown moss organic material having a fibric (slight) degree of decomposition. The peat is composed dominantly of sedge (Carex spp.) and brown mosses (Depranocladus spp.). It is usually slightly decomposed and brown to dark brown; consistency is variable and ranges from loose to matted. The material commonly contains large amounts of very fine sedge roots.
	Mesic sedge	*21*	Sedge organic material having a mesic (intermediate) degree of decomposition. The peat is composed dominantly of sedge (Carex spp.) and is moderately decomposed and matted. The sedge leaves are readily identifiable by the naked eye. This material commonly contains large amounts of very fine roots of the above species.
	Mesic woody sedge	*22*	Woody sedge organic material having a mesic(intermediate) degree of decomposition. The peat is composed dominantly of sedge peat (see code 21)

			with subdominant amounts of woody materials.
	Mesic woody forest	*23*	Woody forest organic material having a mesic(intermediate) degree of decomposition. The peat contains significant amounts of woody materials derived from both coniferous and deciduous tree species. In general, wood fragments are easily identifiable in this peat.
	Marine	*23*	Sphagnum organic material having a mesic (intermediate) degree of decomposition.
	Humic sedge	*31*	Sedge organic material having a humic (most advanced) degree of decomposition in which most of the material is humified and there are few recognizable fibres.
	Non applicable	*#*	Non applicable

4.2.1.15. Coarse Fragment Percentage

Name	Coarse Fragment Content (CFRAG)		
Definition	Rounded, subrounded, flat, angular or irregular rock fragments from 0.2 to 60 cm or more in size expressed in percent by volume.		
Aliases	Coarse fragment content of the control section		
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	<10%	*A*	
	10-30%	*B*	
	31-65%	*C*	
	>65%	*D*	
	Non applicable	*#*	

4.2.1.16. Rooting Depth

Name	Rooting Depth (ROOTDP)
Definition	The unrestricted rooting depth in centimetres.
Aliases	Rooting depth, unrestricted
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Character
Value Domain Type	1 (enumerated)
Value Domain	

Feature Attribute Value

Label	Code	Definition
<20 cm	*A*	
20-75 cm	*B*	
>75-150 cm	*C*	
>150 cm	*D*	
Non-applicable (e.g. rock, ice)	*#*	

4.2.1.17. Drainage

Name	Drainage (DRAIN)
Definition	How well water moves through the soil.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Character
Value Domain Type	1 (enumerated)
Value Domain	

Feature Attribute Value

Label	Code	Definition
Excessive	*E*	Water is removed from the soil very rapidly in relation to supply; excess water flows downward very rapidly if underlying material is pervious; subsurface flow may be very rapid during heavy rainfall provided the gradient is steep; source of water is precipitation.
Rapid	*R*	Water is removed from the soil rapidly in relation to supply; excess water flows downward if underlying material is pervious; subsurface flow may occur on steep gradients during

			heavy rainfall; source of water is precipitation.
	Well	*W*	Water is removed from the soil readily but not rapidly; excess water flows downward into underlying pervious material or laterally as subsurface flow. These soils commonly retain optimum amounts of moisture for plant growth after rains or addition of irrigation water.
	Moderately well	*M*	Water is removed from the soil somewhat slowly in relation to supply due to low perviousness, a shallow water table, lack of gradient, or a combination of these factors; precipitation is the dominant source of water in medium to fine textured soils; precipitation and significant additions by subsurface flow are necessary in coarse-textured soils.
	Imperfect	*I*	Water is removed from the soil sufficiently slowly in relation to supply leaving the soil wet for a significant part of the growing season; excess water moves slowly downward if precipitation is the major supply; if subsurface water, groundwater, or both are the main source the flow rate may vary.
	Poor	*P*	Water is removed so slowly in relation to supply that the soil remains wet for a comparatively large

			part of the time that the soil is not frozen; excess water is evident in the soil for much of the time; subsurface flow, groundwater flow, or both, in addition to precipitation, are the main sources of water; a perched water table may also be present.
	Very poor	*V*	Water is removed from the soil so slowly that the water table remains at or on the surface for a majority of the time the soil is not frozen; groundwater flow and subsurface flow are the major sources of water; precipitation is less important except where there is a perched water table.
	Non applicable	*#*	Non-applicable

4.2.1.18. Soil development

Name	Soil development (DEVEL)		
Definition	Soil development.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Brown Chernozemic	*A*	Dominantly Orthic Brown subgroup with inclusions of other subgroups within the Brown great group.
	Dark Brown Chernozemic	*B*	Dominantly Orthic Dark Brown subgroup with inclusions of other subgroups within the Dark Brown great group.
	Black Chernozemic	*C*	Dominantly Orthic Black subgroup with inclusions of other subgroups within the

			Black great group.
	Dark Gray Chernozemic or Dark Gray Luvisolic	*D*	Dominantly Orthic Dark Gray Chernozemic subgroup or Dark Gray Luvisolic subgroup with inclusions of other subgroups within the Dark Gray great group or of the gleyed Dark Gray Luvisolic subgroup.
	Gray Brown Luvisolic	*E*	Dominantly Orthic Gray Brown Luvisolic subgroup group with inclusions of other subgroups within the Gray Brown Luvisolic great group.
	Gray Luvisolic	*F*	Dominantly Orthic Gray Luvisolic subgroup with inclusions of other Gray Luvisolic subgroups.
	Brown Solonetzic	*G*	May be dominantly Brown Solonetzic, Brown Solodized Solonetzic, or Brown Solod subgroups with inclusions of these subgroups (e.g. dominantly Brown Solodized Solonetzic with inclusions of Brown Solod).
	Dark Brown Solonetzic	*H*	May be dominantly Dark Brown Solonetzic, Dark Brown Solodized Solonetzic, or Dark Brown Solod subgroups with inclusions of these subgroups.
	Brunisolic Gray Luvisolic	*I*	Dominantly Brunisolic Gray Luvisolic subgroup with inclusions of its gleyed subgroup.
	Black Solonetzic	*J*	May be dominantly Black Solonetzic, Black Solodized Solonetzic, or Black Solod subgroups with inclusions of

			these subgroups and their gleyed subgroups.
	Gray Solonetzic	*K*	Dominantly Gray Solodized Solonetzic or Gray Solod subgroups with inclusions of their gleyed subgroups.
	Melanic Brunisolic	*L*	Dominantly Melanic Brunisolic great group.
	Eutric Brunisolic	*M*	Dominantly Eutric Brunisolic great group.
	Sombric Brunisolic	*N*	Dominantly Sombric Brunisolic great group.
	Organic Cryosolic	*O*	Dominantly Organic Cryosolic great group.
	Dystric Brunisolic	*P*	Dominantly Dystric Brunisolic great group.
	Humic Podzolic	*Q*	Dominantly Humic Podzolic great group.
	Regosolic	*R*	Dominantly Regosolic order.
	Static Cryosolic	*S*	Dominantly Static Cryosolic great group.
	Turbic Cryosolic	*T*	Dominantly Turbic Cryosolic great group.
	Gleysolic	*U*	Dominantly Gleysolic order.
	Ferro-Humic Podzolic	*V*	Dominantly Ferro-humic Podzolic great group.
	Humo-Ferric Podzolic	*W*	Dominantly Humo-Ferric Podzolic great group.
	Fibrisol	*X*	Dominantly Fibrisol great group.
	Mesisol	*Y*	Dominantly Mesisol great group.
	Humisol	*Z*	Dominantly Humisol great group.
	Folisol	*2*	Dominantly Folisol great group.
	Podzolic Gray Luvisolic	*3*	Podzolic Gray Luvisolic subgroup (only occurs as subdominant soil).

	Orthic Turbic Cryosolic	*4*	Dominantly Orthic Turbic Cryosolic subgroup.
	Brunisolic Turbic Cryosolic	*5*	Dominantly Brunisolic Turbic Cryosolic subgroup.
	Regosolic Turbic Cryosolic	*6*	Dominantly Regosolic Turbic Cryosolic subgroup.
	Gleysolic Turbic Cryosolic	*	Dominantly Gleysolic Turbic Cryosolic subgroup.
	Orthic Static Cryosolic	*7*	Dominantly Orthic Static Cryosolic subgroup.
	Brunisolic Static Cryosolic	*8*	Dominantly Brunisolic Static Cryosolic subgroup.
	Regosolic Static Cryosolic	*9*	Dominantly Regosolic Static Cryosolic subgroup.
	Gleysolic Static Cryosolic	*$*	Dominantly Gleysolic Static Cryosolic subgroup.
	Non-applicable	*#*	Non-applicable

4.2.1.19. Calcareous class of parent material

Name	Calcareous class of parent material (CALC)		
Definition	Rounded, subrounded, flat, angular or irregular rock fragments from 0.2 to 60 cm or more in size expressed in percent by volume.		
Aliases	Coarse fragment content of the control section		
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Noncalcareous	*0*	No $CaCO_3$ detectable with dilute HCl.
	Weakly	*1*	1-5% $CaCO_3$ equivalents (weak effervescence with dilute HCl).
	Strongly	*2*	6-40% $CaCO_3$ equivalents (moderate to strong effervescence with dilute HCl).
	Extremely	*3*	>40% $CaCO_3$ equivalents

			(very strong effervescence with dilute HCl).
	Non applicable	#	(water, rock, ice)

4.2.1.20. Land Area

Name	Land Area (LAND_AREA)		
Definition	This is the total area of land in the SLC polygon.		
Aliases	Area applicable to the Component Table (ha)		
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.21. Area of ocean and fresh water (ha)

Name	Area of ocean and fresh water (ha) (WATER_TOTAL)		
Definition	This is the total area of ocean and fresh water in the entire SLC polygon.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.22. Area of Fresh Water (ha)

Name	Area of Fresh Water (ha) (HYDRO_FRESH)		
Definition	This is the total area of freshwater lakes and rivers in hectares.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.23. Area Covered by Ocean (ha)

Name	Area Covered by Ocean (ha) (HYDRO_OCEAN)		
Definition	This is the total area of the SLC polygon that forms part of an ocean.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.24. Area of the entire SLC polygon (ha)

Name	Area of the entire SLC polygon (ha) (POLYGON_AREA)		
Definition	This is the total area, including both land, fresh water, and ocean, of the entire SLC polygon.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.25. Area of fresh water (ha)

Name	Area of fresh water (ha) (WATER_FRESH)		
Definition	This is the total area of fresh water in the entire SLC polygon.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.26. Area of small water bodies (ha)

Name	Area of small water bodies (ha) (SMALL_WATER)
Definition	This is the total area, in kilohectares, of small freshwater lakes in the the entire SLC polygon. Small lakes are perrenial water bodies which are too small to be shown on the hydro coverage.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Integer
Value Domain Type	0 (not enumerated)
Value Domain	
	Feature Attribute Value

	Label	Code	Definition

4.2.1.27. Area of small islands (ha)

Name	Area of small islands (ha) (SMALL_LAND)		
Definition	This is the total area, in kilohectares, of small islands in the the entire SLC polygon. These islands are too small to depict at the 1:1,000,000 scale of the HYDRO coverage, and are classified as water in that coverage.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.28. Available water capacity (dominant soil landscape)

Name	Available water capacity (dominant soil landscape) (AVWATDOM)		
Definition	That portion of water in a soil that can be readily absorbed by plant roots; most workers consider it to be the water held in the soil between field capacity and a pressure of up to about 15 bars.		
Aliases	Available water capacity in upper 120 cm (dominant soil landscape)		
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	50 mm	*1*	
	100 mm	*2*	
	150 mm	*3*	
	200 mm	*4*	
	250 mm	*5*	
	Not applicable	*6*	Solonetzic or saline soils
	Not applicable	*7*	High water table
	Not applicable	*8*	Perennially frozen subsoils
	Not applicable	*#*	Water, ice, rock
	Not identified	-	

4.2.1.29. Available water capacity (subdominant soil landscape)

Name	Available water capacity (subdominant soil landscape) (AVWATSUB)
Definition	That portion of water in a soil that can be readily absorbed by plant roots; most workers consider it to be the water held in the soil between

	field capacity and a pressure of up to about 15 bars.		
Aliases	Available water capacity in upper 120 cm (subdominant soil landscape)		
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	50 mm	*1*	
	100 mm	*2*	
	150 mm	*3*	
	200 mm	*4*	
	250 mm	*5*	
	Not applicable	*6*	Solonetzic or saline soils
	Not applicable	*7*	High water table
	Not applicable	*8*	Perennially frozen subsoils
	Not applicable	*#*	Water, ice, rock
	Not identified	-	

4.2.1.30. Reliability class of polygon

Name	Reliability class of polygon (RELIA)		
Definition	Reliability of generalized maps is required to provide an indication of the confidence that can be placed in derivative interpretations and to assist in assigning priorities for future field-mapping projects. Four levels of map reliability have been established in relation to inspection intensity and publication scale of source documents and whether aerial photographs are used during field mapping.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Very low	*V*	
	Low	*L*	
	Medium	*M*	
	High	*H*	
	not identified	-	

4.2.1.31. Complexity class of polygon

Name	Complexity class of polygon (COMPLEXITY)
Definition	Complexity of soil landscape attribute classes is determined from

	information provided on source maps and the accompanying soil reports. The concept of complexity provides an indication of attribute variability within a polygon, particularly with respect to the classes of parent material deposition modes and soil development. Three levels of complexity have been established. They provide a warning of variability to anyone interpreting the information.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Low	*L*	Soil and landscape attributes within the polygon are uniform for most interpretations; in most cases the polygon has only a dominant component.
	Medium	*M*	Soil and landscape attributes are moderately variable but predictable; there are generally dominant and subdominant components, each of which usually have been generalized from no more than two classes of parent material or soil development, or both; there may also be an inclusion in the polygon.
	High	*H*	Soil and landscape attributes are highly variable and unpredictable; dominant, subdominant , and inclusion components are present, each of which has been generalized from more than two classes of parent material or soil development, or both; this class indicates extreme oversimplification in any interpretations

			from the attribute tables.
	not identified	-	Complexity was not identified.

4.2.1.32. Percentage of the polygon occupied by the dominant soil landscape

Name	Percentage of the polygon occupied by the dominant soil landscape (DISTRDOM)		
Definition	Percentage of the polygon occupied by the dominant soil landscape.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.33. Percentage of the polygon occupied by the subdominant soil landscape

Name	Percentage of the polygon occupied by the subdominant soil landscape (DISTRSUB)		
Definition	Percentage of the polygon occupied by the subdominant soil landscape.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.34. Landform Segment Percent

Name	Landform Segment Percent (LFS_EXTENT)		
Definition	Specifies the extent of this landform segment within the landform, in percent.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.35. Landform Segment Slope Percent

Name	Landform Segment Slope Percent (LFS_SLOPE)

Definition	Specifies the typical slope for this landform segment, in percent.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.36. Landform Segment Slope Length

Name	Landform Segment Slope Length (LFS_LENGTH)		
Definition	Specifies the typical length in meters of a landform segment.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.37. Landform Segment Name

Name	Landform Segment Name (LFS_NAME)		
Definition	Specifies the name of the landform segment (landform type, slope class and toposequence position/segment). (???Not sure what the information in brackets is saying – it's not identified as a concatenation.)		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.38. Polygon Landform Identifier

Name	Polygon Landform Identifier (LANDFORM_ID)
Definition	Identifies each landform within the soil landscapes polygon coverage. It is created by concatenating the Soil Landscapes Identifier, Landform Type and Landform Slope Type attributes.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Character
Value Domain Type	0 (not enumerated)

Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.39. Polygon Landform Percent

Name	Polygon Landform Percent (LANDFRM_EXT_PERCENT_NUM)		
Definition	Specifies the estimated extent that a landform occupies within a soil landscape polygon, in percent. For soil landscape polygons with a single landform, the value for Polygon Landform Percent is 100.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.40. Agricultural Crop Use

Name	Agricultural Crop Use (CROP_PRD_TYPE_CODE)		
Definition	Specifies how likely it is that the landscape segment is used for agricultural crop production.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Low	*L*	
	Moderate	*M*	
	High	*H*	

4.2.1.41. Layer Number

Name	Layer Number (HRZN_NUM)		
Definition	Specifies the horizon number.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.42. Upper Depth

Name	Upper Depth (HRZN_UDEPTH_CM_NUM)		
Definition	Specifies the depth of the top of the horizon.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.43. Lower Depth

Name	Lower Depth (HRZN_LDEPTH_CM_NUM)		
Definition	Specifies the depth of the bottom of the horizon.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.44. Polygon Component Identifier

Name	Polygon Component Identifier (CMP_ID)		
Definition	This field uniquely identifies an instance of a component within a polygon. It is a concatenation of the Soil Landscapes Identifier (SLC_ID) and the Component Number (CMP_NUM).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.45. Ecoregion Identifier

Name	Ecoregion Identifier (ECOREGION_ID)
Definition	Specifies an ecoregion. An ecoregion is the third level of generalization used in the ecological framework of Canada. It is a collection of ecodistricts and a subdivision of an ecoprovince.
Aliases	
Producer	Canadian Soil Information System (CanSIS)

Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.46. Ecoprovince Identifier

Name	Ecoprovince Identifier (ECOPROVINCE_ID)		
Definition	Specifies an ecoprovince. An ecoprovince is the second level of generalization used in the ecological framework of Canada. It is a collection of ecoregions and a subdivision of an ecozone.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.47. Ecozone Identifier

Name	Ecozone Identifier (ECOZONE_ID)		
Definition	Specifies an ecozone. An ecozone is the highest level of generalization used in the ecological framework of Canada. It is a collection of ecoprovinces.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.48. Horizon Lithological Discontinuity

Name	Horizon Lithological Discontinuity (HRZN_LIT_NUM)
Definition	Identifies the presence of lithological discontinuity horizons. When present, this attribute uses a value between 2 and 5 to identify the number of the lithological discontinuity.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Character
Value Domain Type	0 (not enumerated)
Value Domain	
	Feature Attribute Value

	Label	Code	Definition

4.2.1.49. Horizon Master

Name	Horizon Master (HRZN_MASTER_TYPE_CODE)		
Definition	Identifies the master horizon. For more information, see Chapter 2 of The Canadian System of Soil Classification, Third edition.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Surface mineral horizon	*A*	
	Subsurface mineral horizon	*B*	
	Parent material horizon	*C*	
	Organic horizon partially decomposed	*F*	
	Organic horizon of indiscernible origin	*H*	
	Organic horizon poorly decomposed	*L*	
	Organic horizon from mosses, rushes and woody materials	*O*	
	Rock	*R*	
	Water	*W*	
	n/a	-	

4.2.1.50. Horizon Suffix

Name	Horizon Suffix (HRZN_SUFFIX_TEXT)		
Definition	Specifies the horizon suffix. For more information, see Chapter 2 of The Canadian System of Soil Classification, Third edition.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.51. Horizon Modifier

Name	Horizon Modifier (HZN_MODIFIER_TEXT)		
Definition	Specifies a soil horizon modifier.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.52. Coarse Fragment Percentage

Name	Coarse Fragment Percentage (COFRAG_PERCENT_NUM)		
Definition	Specifies the amount of coarse fragments in a layer as a percentage of volume.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.53. Dominant Sand Fraction

Name	Dominant Sand Fraction (DOMSAND_TYPE_CODE)		
Definition	Identifies the dominant sand fraction of a layer.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Very Coarse	*VC*	
	Coarse	*C*	
	Medium	*M*	
	Fine	*F*	
	Very Fine	*VF*	
	n/a	-	

4.2.1.54. Very Fine Sand Percentage

Name	Very Fine Sand Percentage (VFSAND_PERCENT_NUM)
Definition	Indicates the percentage of very fine sand in a layer, by weight.
Aliases	

Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.55. Total Percentage Sand

Name	Total Percentage Sand (TSAND_PERCENT_NUM)		
Definition	Indicates the percentage of total sand in a layer, by weight.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.56. Total Percentage Silt

Name	Total Percentage Silt (TSILT_PERCENT_NUM)		
Definition	Indicates the percentage of total silt in a layer, by weight.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.57. Total Clay Percentage

Name	Total Clay Percentage (TCLAY_PERCENT_NUM)		
Definition	Indicates the percentage of total clay in a layer, by weight.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.58. Organic Carbon Percentage

Name	Organic Carbon Percentage (ORGCARBON_PERCENT_NUM)		
Definition	Indicates the percentage of organic carbon in a layer, by weight.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.59. pH in Calcium Chloride

Name	pH in Calcium Chloride (PH_CACL2_NUM)		
Definition	Specifies the soil pH, measured using pH in calcium chloride.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.60. Project Report pH

Name	Project Report pH (PH_PRJRPT_NUM)		
Definition	Indicates the pH as specified in the project report.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.61. Base Saturation Percentage

Name	Base Saturation Percentage (BASE_STRT_PERCENT_NUM)		
Definition	Indicates the base saturation in percent.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.62. Cation Exchange Capacity

Name	Cation Exchange Capacity (CEC_NUM)		
Definition	Identifies the cation exchange capacity.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.63. Saturated Hydraulic Conductivity

Name	Saturated Hydraulic Conductivity (KSAT_NUM)		
Definition	Indicates the saturated hydraulic conductivity, in centimeters per hour (cm/h).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.64. Water Retention at 0 Kilopascals

Name	Water Retention at 0 Kilopascals (KP0_NUM)		
Definition	Identifies water retention at 0 kilopascals, in percent of total soil volume.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.65. Water Retention at 10 Kilopascals

Name	Water Retention at 10 Kilopascals (KP10_NUM)
Definition	Identifies water retention at 10 kilopascals, in percent of total soil volume.
Aliases	

Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.66. Water Retention at 33 Kilopascals

Name	Water Retention at 33 Kilopascals (KP33_NUM)		
Definition	Identifies water retention at 33 kilopascals, in percent of total soil volume.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.67. Water Retention at 1500 Kilopascals

Name	Water Retention at 1500 Kilopascals (KP1500_NUM)		
Definition	Identifies water retention at 1500 kilopascals, in percent of total soil volume.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.68. Bulk Density

Name	Bulk Density (BULK_DENSITY_NUM)		
Definition	Identifies the bulk density of the fine earth fraction of a layer, excluding coarse fragment content.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.69. Electrical Conductivity

Name	Electrical Conductivity (ELEC_CONDUCT_NUM)		
Definition	Specifies the electrical conductivity of a layer in decisiemens per meter (dS/m).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.70. Calcium Carbonate Percentage

Name	Calcium Carbonate Percentage (CACO3_PERCENT_NUM)		
Definition	Specifies a value that is equivalent to a percentage of pure calcium carbonate.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.71. Von Post Decomposition

Name	Von Post Decomposition (DECOMP_VONPOST_CODE)		
Definition	Specifies peat decomposition using the Von Post method of asessment.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.72. Woody Material Percentage

Name	Woody Material Percentage (WOODY_MAT_PERCENT_NUM)
Definition	Specifies the volume of woody material in a layer as a percentage.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Integer

Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.73. Soil Name

Name	Soil Name (SOIL_NAME)		
Definition	Specifies the complete soil name, without abbreviations. Each Soil Name attribute can be associated with up to two entries in the Soil Names Table (SNT).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.74. Kind of Material

Name	Kind of Material (KIND_TYPE_CODE)		
Definition	Specifies the kind of surface soil, rock outcrop or other material or feature.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Mineral Soil	*M*	Predominantly mineral particles. Contains less than 30% organic matter as measured by weight.
	Organic Soil	*O*	Contains more than 30% organic matter as measured by weight.
	Non-soil	*N*	Rockland, quarry, unclassified, not surveyed and so on.
	Urban	*U*	Urban areas, airports and other artificial features.

4.2.1.75. Water Table Type

Name	Water Table Type (WTRTBL_TYPE_CODE)		
Definition	Indicates the presence of a water table and identifies its characteristics.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Always	*YB*	Present during both seasons
	Growing Season	*YG*	Present during the growing season
	Non growing season	*YN*	Present outside the growing season
	Unspecified time	*YU*	Present during unspecified time
	Never	*NO*	Not present anytime
	Not applicable	-	Not applicable

4.2.1.76. Restrictive Layer

Name	Restrictive Layer (RESTRICT_LAYER_CODE)		
Definition	Indicates the soil layer that restricts root growth. A value between 1 and 9 indicates that the layer is restrictive and can be linked to the Layer Number attribute in the Soil Layer Table. A value of 0 indicates that no restricting layer is present. N/A indicates "not applicable".		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.77. Soil Layer Restrictive Soil Type

Name	Soil Layer Restrictive Soil Type (RESTRICT_TYPE_CODE)		
Definition	Indicates the type of soil in the root restricting layer.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Not applicable	-	

	Solonetzic B horizon	*BN*	
	Cryic (frozen) horizon	*CR*	
	Compact (basal) till	*CT*	
	Duric horizon	*DU*	
	Fragipan	*FP*	
	Lithic	*LI*	
	Ortstein horizon	*OR*	
	Placic horizon	*PL*	
	Salinity	*SA*	
	Undifferentiated	*UN*	

4.2.1.78. Drainage Type

Name	Drainage Type (DRAINAGE_TYPE_CODE)		
Definition	Indicates the type of drainage found within a soil profile.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Very rapidly drained	*VR*	
	Rapidly drained	*R*	
	Well drained	*W*	
	Moderately well drained	*MW*	
	Imperfectly drained	*I*	
	Poorly drained	*P*	
	Very poorly drained	*VP*	
	Not applicable	-	

4.2.1.79. Parent Material 1 Texture

Name	Parent Material 1 Texture (PMTEXTURE1_TYPE_CODE)		
Definition	Specifies the texture of the underlying parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Very Coarse	*VC*	
	Coarse	*C*	
	Moderately Coarse	*MC*	

	Medium	*M*	
	Moderately Fine	*MF*	
	Fine	*F*	
	Very Fine	*VF*	
	Coarse Skeletal	*CS*	
	Medium Skeletal	*MS*	
	Fine Skeletal	*FS*	
	Fragmental	*FR*	
	Stratified (Mineral)	*SM*	
	Fibric	*FI*	
	Mesic	*ME*	
	Humic	*HU*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.2.1.80. Parent Material 2 Texture

Name	Parent Material 2 Texture (PMTEXTURE2_TYPE_CODE)		
Definition	Specifies the texture of the second underlying parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Very Coarse	*VC*	
	Coarse	*C*	
	Moderately Coarse	*MC*	
	Medium	*M*	
	Moderately Fine	*MF*	
	Fine	*F*	
	Very Fine	*VF*	
	Coarse Skeletal	*CS*	
	Medium Skeletal	*MS*	
	Fine Skeletal	*FS*	
	Fragmental	*FR*	
	Stratified (Mineral)	*SM*	
	Fibric	*FI*	
	Mesic	*ME*	
	Humic	*HU*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.2.1.81. Parent Material 3 Texture

Name	Parent Material 3 Texture (PMTEXTURE3_TYPE_CODE)

Definition	Specifies the texture of the third underlying parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Very Coarse	*VC*	
	Coarse	*C*	
	Moderately Coarse	*MC*	
	Medium	*M*	
	Moderately Fine	*MF*	
	Fine	*F*	
	Very Fine	*VF*	
	Coarse Skeletal	*CS*	
	Medium Skeletal	*MS*	
	Stratified (Mineral)	*SM*	
	Fibric	*FI*	
	Mesic	*ME*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.2.1.82. Parent Material 1 Chemical Composition

Name	Parent Material 1 Chemical Composition (PMCHEM1_TYPE_CODE)		
Definition	Specifies the chemical composition of the upper layer parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Extremely/Strongly acidic	*EA*	
	Medium acid to neutral	*AN*	
	Weakly calcareous	*WC*	
	Moderately/Very strongly calcareous	*VC*	
	Extremely calcareous	*EC*	
	Calcareous and saline	*SA*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.2.1.83. Parent Material 2 Chemical Composition

Name	Parent Material 2 Chemical Composition (PMCHEM2_TYPE_CODE)		
Definition	Specifies the chemical composition of the second underlying parent material (mineral material that is transformed into soil), if present.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Extremely/Strongly acidic	*EA*	
	Medium acid to neutral	*AN*	
	Weakly calcareous	*WC*	
	Moderately/Very strongly calcareous	*VC*	
	Extremely calcareous	*EC*	
	Calcareous and saline	*SA*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.2.1.84. Parent Material 3 Chemical Composition

Name	Parent Material 3 Chemical Composition (PMCHEM3_TYPE_CODE)		
Definition	Specifies the chemical composition of the third underlying parent material (mineral material that is transformed into soil), if present.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Extremely/Strongly acidic	*EA*	
	Medium acid to neutral	*AN*	
	Weakly calcareous	*WC*	
	Moderately/Very strongly calcareous	*VC*	
	Extremely calcareous	*EC*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.2.1.85. Parent Material 1 Mode of Deposition

Name	Parent Material 1 Mode of Deposition (MDEPOSIT1_TYPE_CODE)		
Definition	Specifies the mode of deposition of the uppermost underlying parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Not applicable	-	
	Colluvial	*COLL*	
	Fluvioeolian	*FLEO*	
	Undifferentiated Organic	*UNDO*	
	Bedrock, undifferentiated	*RKUD*	
	Igneous, acidic bedrock (dominantly granite)	*RKIA*	
	Igneous, basic bedrock (dominantly basalt)	*RKIB*	
	Calcareous bedrock (dominantly sandstone and shale)	*RKCA*	
	Limestone bedrock (limestone and dolomite)	*RKLS*	
	Shale bedrock	*RKSH*	
	Sandstone bedrock	*RKSS*	
	Slate bedrock	*RKSA*	
	Anthropogenic	*ANTH*	
	Eolian	*EOLI*	
	Fluviolacustrine	*FLLC*	
	Fluvial	*FLUV*	
	Fen Peat	*FNPT*	
	Forest Peat	*FOPT*	
	Glaciofluvial	*GLFL*	
	Glaciolacustrine	*GLLC*	
	Glaciomarine	*GLMA*	
	Lacustrine	*LACU*	
	Lacustro-Till	*LATL*	
	Marine	*MARI*	
	Residual	*RESD*	
	Saprolite	*SAPR*	
	Sphagnum Peat	*SPPT*	

	Till (morainal)	*TILL*	
	Undifferentiated Mineral	*UNDM*	
	Volcanic	*VOLC*	

4.2.1.86. Parent Material 2 Mode of Deposition

Name	Parent Material 2 Mode of Deposition (MDEPOSIT2_TYPE_CODE)		
Definition	Specifies the mode of deposition of the second underlying parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Not applicable	-	
	Colluvial	*COLL*	
	Undifferentiated Organic	*UNDO*	
	Bedrock, undifferentiated	*RKUD*	
	Igneous, acidic bedrock (dominantly granite)	*RKIA*	
	Igneous, basic bedrock (dominantly basalt)	*RKIB*	
	Igneous, basic coarse bedrock (coarse grained basalts)	*RKIC*	
	Igneous, basic fine grained bedrock (dominantly andesites)	*RKIF*	
	Calcareous bedrock (dominantly sandstone and shale)	*RKCA*	
	Limestone bedrock (limestone and dolomite)	*RKLS*	
	Shale bedrock	*RKSH*	
	Sandstone bedrock	*RKSS*	
	Siltstone and mudstone bedrock	*RKSM*	
	Schist and phyllite bedrock	*RKSP*	
	Gneiss bedrock	*RKGN*	
	Eolian	*EOLI*	

	Fluviolacustrine	*FLLC*	
	Fluvial	*FLUV*	
	Fen Peat	*FNPT*	
	Forest Peat	*FOPT*	
	Glaciofluvial	*GLFL*	
	Glaciolacustrine	*GLLC*	
	Glaciomarine	*GLMA*	
	Lacustrine	*LACU*	
	Lacustro-Till	*LATL*	
	Marine	*MARI*	
	Residual	*RESD*	
	Saprolite	*SAPR*	
	Sphagnum Peat	*SPPT*	
	Till (morainal)	*TILL*	
	Undifferentiated Mineral	*UNDM*	

4.2.1.87. Parent Material 3 Mode of Deposition

Name	Parent Material 3 Mode of Deposition (MDEPOSIT3_TYPE_CODE)		
Definition	Specifies the mode of deposition of the third underlying parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Not applicable	-	
	Undifferentiated Organic	*UNDO*	
	Bedrock, undifferentiated	*RKUD*	
	Igneous, acidic bedrock (dominantly granite)	*RKIA*	
	Limestone bedrock (limestone and dolomite)	*RKLS*	
	Shale bedrock	*RKSH*	
	Slate bedrock	*RKSA*	
	Fluviolacustrine	*FLLC*	
	Fluvial	*FLUV*	
	Fen Peat	*FNPT*	
	Glaciofluvial	*GLFL*	
	Glaciolacustrine	*GLLC*	
	Glaciomarine	*GLMA*	
	Lacustrine	*LACU*	

	Marine	*MARI*	
	Residual	*RESD*	
	Till (morainal)	*TILL*	
	Undifferentiated Mineral	*UNDM*	

4.2.1.88. CSSC 2nd Edition Soil Order

Name	CSSC 2nd Edition Soil Order (SLC_V2_ORDER_CODE)		
Definition	Specifies the soil order according to The Canadian System of Soil Classification (CSSC), Second edition.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.89. CSSC 2nd Edition Soil Great Group

Name	CSSC 2nd Edition Soil Great Group (SLC_V2_GR_GROUP_CODE)		
Definition	Specifies the soil great group according to The Canadian System of Soil Classification (CSSC), Second edition.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.90. CSSC 2nd Edition Soil Subgroup

Name	CSSC 2nd Edition Soil Subgroup (SLC_V2_SUB_GROUP_CODE)		
Definition	Specifies the soil subgroup according to The Canadian System of Soil Classification (CSSC), Second edition. A soil subgroup is a subdivision of a great group.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.91. CSSC 3rd Edition Soil Order

Name	CSSC 3rd Edition Soil Order (SLC_V3_ORDER_CODE)		
Definition	Specifies the soil order according to the Canadian System of Soil Classification, Third edition.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.92. CSSC 3rd Edition Soil Great Group

Name	CSSC 3rd Edition Soil Great Group (SLC_V3_GR_GROUP_CODE)		
Definition	Specifies the soil great group according to The Canadian System of Soil Classification (CSSC), Third edition.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.93. CSSC 3rd Edition Soil Subgroup

Name	CSSC 3rd Edition Soil Subgroup (SLC_V3_SUB_GROUP_CODE)		
Definition	Specifies the soil subgroup according to The Canadian System of Soil Classification (CSSC), Third edition. A subgroup is a subdivision of a great group.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.2.1.94. 0 to 50 cm Coarse Fragment Content

Name	0 to 50 cm Coarse Fragment Content (CFRAG1_TYPE_CODE)
Definition	Specifies the average coarse fragment content of the soil profile, from 0 to 50 centimeters of depth. For root-restricting layers, at less than 50 centimeters of depth.

Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	More than 65%	*D*	
	31-65%	*C*	
	10-30%	*B*	
	Less than 10%	*A*	
	Not Applicable	-	

4.2.1.95. 50 to 100 cm Coarse Fragment Content

Name	50 to 100 cm Coarse Fragment Content (CFRAG2_TYPE_CODE)		
Definition	Specifies the average coarse fragment content of the soil profile, from 50 to 100 centimeters of depth. For a root-restricting layer, at greater than 50 centimeters of depth.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	More than 65%	*D*	
	31-65%	*C*	
	10-30%	*B*	
	Less than 10%	*A*	
	Not Applicable	-	

4.2.1.96. Depth to Bedrock

Name	Depth to Bedrock (DEPTH_TYPE_CODE)		
Definition	Specifies the depth to bedrock or a root-restricting layer.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Less than 25 cm	*1*	
	25-49 cm	*2*	
	50-74 cm	*3*	
	75-99 cm	*4*	

	100 cm or more	*5*	
	n/a (for example, rock or ice)	-	

4.2.1.97. Component Restrictive Soil Type

Name	Component Restrictive Soil Type (RESTRICTION_TYPE_CODE)		
Definition	Specifies the type of soil in a root-restricting layer.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Not applicable	-	
	Solonetzic B horizon	*BN*	
	Cryic (frozen) horizon	*CR*	
	Compact (Basal) till	*CT*	
	Duric horizon	*DU*	
	Fragipan	*FP*	
	Lithic	*LI*	
	Ortstein horizon	*OR*	
	Placic horizon	*PL*	
	Salinity	*SA*	
	Undifferenciated	*UN*	

4.2.1.98. Water Capacity

Name	Water Capacity (WTR_CPCTY_TYPE_CODE)		
Definition	Specifies a soil's capacity for holding water.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	<50 mm	*1*	
	50-99 mm	*2*	
	100-149 mm	*3*	
	150-199 mm	*4*	
	200-250 mm	*5*	
	Water, ice or rock	-	

4.3. Feature catalogue – Soil Landscapes of Canada Version 3.2 – Feature Catalogue

Title	Soil Landscapes of Canada Version 3.2 - Feature Catalogue
Scope	
Version Number	1.0
Version Date	2013-01-23
Producer	Agri-Geomatics

System-generated attributes (for example, OBJECTID, Shape, Shape Length and Area) are not defined in the feature catalog.

4.3.1. Feature attributes

4.3.1.1. Soil Landscapes Identifier

Name	Soil Landscapes Identifier (SLC_ID, SL)		
Definition	Joins attribute datasets (LAT, CMP and LST) to the Boundaries dataset. More than one polygon in the spatial dataset can share the same Soil Landscapes Identifier. The value of SLC_ID is the Ecodistrict Identifier (one to four digits) followed by a three-digit SLC polygon number. The polygon numbers for each Ecodistrict start at 001.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.2. Ecodistrict Identifier

Name	Ecodistrict Identifier (ECODISTRICT_ID, ECODISTRICT)		
Definition	Specifies the ecodistrict for a soil landscape polygon. An ecodistrict is the fourth level of generalization used in the ecological framework of Canada. Ecodistricts are groups of soil landscape polygons and can in turn be grouped into ecoregions.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.3. Component Number

Name	Component Number (CMP_NUM, CMP)		
Definition	A number that uniquely identifies a component of a polygon.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.4. Percent of Polygon

Name	Percent of Polygon (CMP_PERCENT_NUM, PERCENT)		
Definition	Specifies the percentage of the land area of the polygon that the component occupies.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.5. Slope Gradient

Name	Slope Gradient (SLOPE_TYPE_CODE, SLOPE)		
Definition	SLOPE contains categorized slope gradient (in percent) of polygon components for the low end of range (Shields, 1982). First used in SLC 3.0. Identical class definitions to version 1.0 but "Not Applicable" code differs.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Little or none	*A*	Little or no slope: 0 - 3 % gradient.
	Gentle	*B*	Gentle slopes: 4 - 9 % gradient.
	Moderate	*C*	Moderate slopes: 10 - 15 % gradient.
	Steep	*D*	Steep slopes: 16 - 30 % gradient.

	Extremely steep	*E*	Extremely steep slopes: 31 - 60% gradient.
	Excessively steep	*F*	Excessively steep slopes: > 60% gradient.
		n/a	Not applicable (water)

4.3.1.6. Surface Stoniness

Name	Surface Stoniness (STONE_TYPE_CODE)		
Definition	Describes the stoniness of the soil surface, which has an impact on a soil's agricultural capabilities.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	None	*N*	
	Slightly Stony	*S*	
	Very Stony	*V*	
	Unknown	*U*	

4.3.1.7. Local Surface Form

Name	Local Surface Form (LOCSF_TYPE_CODE, LOCSF)		
Definition	Describes assemblages of slopes or recurring patterns of forms that occur at the earth's surface. When applied to consolidated materials (material that has been transformed to hard rock), it refers to the form produced after modification by geological processes.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Dissected	*D*	A dissected (or gullied) pattern providing external drainage for an area.
	Hummocky (or irregular)	*H*	A very complex sequence of slopes extending from somewhat rounded concavities (or swales) of various sizes to irregular conical knolls (or

			knobs) and short discontinuous ridges; there is a general lack of concordance between knolls and swales. Slopes are generally 4-70%. Examples are hummocky moraines and hummocky fluvioglacial landforms.
	Inclined	*I*	A sloping, unidirectional surface with a generally constant slope unbroken by marked irregularity or gullies; a weakly developed dissected pattern provides external drainage for the local area. Slopes are 2-70%; the form of inclined slopes is not related to the initial mode of origin of the underlying material.
	Knoll and kettle	*K*	A chaotic sequence of knolls and kettles (or sloughs), which occupies 15-20% of an area and has no external drainage. Slopes are generally >3%. Examples are morainal plains and hill lands.
	Level	*L*	A flat or very gently sloping, unidirectional surface with a generally constant slope unbroken by marked elevations and depressions. Slopes are generally <2%. Examples are floodplains and lake plains.
	Rolling	*M*	A very regular sequence of moderate slopes extending from rounded and, in some places, confined concave depressions to

			broad, rounded convexities producing a wavelike pattern of moderate relief. Slope gradients are generally >5% but may be less. This surface form is usually controlled by the underlying bedrock.
	Ridged	*R*	A long, narrow elevation of the surface, usually sharp crested with steep sides; ridges may be parallel, subparallel, or intersecting. Examples are eskers, crevasse fillings, washboard moraines and some drumlins.
	Steep	*S*	Erosional slopes of >70%, present on both consolidated and unconsolidated materials. An example is an escarpment.
	Terraced	*T*	Scarp face and the horizontal or gently inclined surface (or tread) above it. An example is an alluvial terrace.
	Undulating	*U*	A very regular sequence of gentle slopes that extends from rounded and, in some places, confined concavities to broad, rounded convexities producing a wavelike pattern of low local relief. Slope length is generally <0.8 km and the dominant gradient of slopes is usually 2-5%. The terrain lacks an external drainage pattern. Examples are some ground moraines and lacustrine material of varying textures.

	Domed bog	*B04*	A large bog (diameter usually >500 m) with a convex surface rising several metres above the surrounding terrain. The centre usually drains in all directions; small crescentic pools commonly form around the highest point; a concentric pattern is formed if the highest point is in the centre, while an eccentric pattern is formed if the highest point is off-centre. Peat development is usually >3 m.
	Polygonal peat bog	*B05*	A perennially frozen bog rising approximately 1 m above the surrounding fen. The surface is relatively flat, scored by a polygonal pattern of trenches that developed over ice wedges. The permafrost and ice wedges developed in peat originally deposited in a nonpermafrost environment.
	Peat plateau bog	*B07*	A bog composed of perennially frozen peat rising abruptly about 1 m from the surrounding unfrozen fen. The surface is relatively flat and even, and commonly covers large areas. The peat was originally deposited in a nonpermafrost environment and is associated in many places with collapse scar bogs or fens.
	Atlantic plateau bog	*B09*	A bog with a flat to undulating surface raised above the

			surrounding terrain. The bog edges commonly slope steeply downwards to the mineral soil terrain. Large pools scattered on the bog reach depths of 2-4 m.
	Basin bog	*B13*	A bog situated in a basin with essentially closed drainage which receives water from precipitation and runoff from the immediate surroundings. The surface of the bog is flat with peat generally deepest at the centre.
	Flat bog	*B14*	A bog having a flat, featureless surface and occurring in broad, poorly defined depressions. The depth of peat is generally uniform.
	String bog	*B15*	A pattern of narrow (2-3 m wide), low (<1 m high) ridges oriented at right angles to the direction of drainage; wet depressions or pools occur between the ridges. The water and peat are very low in nutrients because the water has been derived from other ombrotrophic wetlands. The peat thickness is >1 m.
	Blanket bog	*B16*	A bog consisting of extensive peat deposits that occur more or less uniformly over gently sloping hills and valleys. The peat thickness is usually <2 m.
	Slope bog	*B18*	A bog occurring in areas of high rainfall on appreciably sloping land surfaces. The bog is

			fed by rainwater and by water draining from other nutrient-poor peatlands. The peat may exceed 1 m in thickness.
	Veneer bog	*B19*	A bog occurring on gently sloping terrain underlain by generally discontinuous permafrost. Although drainage is predominantly below the surface, overland flow occurs in poorly defined drainways during peak runoff. Peat thickness is usually <1.5 m.
	Lowland polygon bog	*B20*	A bog with flat-topped or convex peat surfaces (often referred to as "high-centre polygons") separated by trenches over ice wedges that form a polygonal pattern when viewed from above. The peat was deposited in a permafrost environment as shown by internal structures.
	Northern ribbed fen	*F01*	A fen with parallel, low peat ridges ("strings") alternating with wet hollows or shallow pools, oriented across the major slope at right angles to water movement. The depth of peat is >1 m.
	Shore fen	*F07*	A fen with an anchored surface mat that forms the shore of a pond or lake. The rooting zone is affected by the water of the lake at both normal and flood levels.
	Slope fen	*F11*	A fen occurring mainly on slow-draining, nutrient-enriched seepage

			slopes. Pools are usually absent, but wet seepage tracks may occur. Peat thickness is usually <2 m.
	Horizontal fen	*F13*	A fen with a gently sloping, featureless surface. This fen occupies broad, often ill-defined depressions and may interconnect with other fens. Peat accumulation is generally uniform.
	Stream swamp	*S01*	A swamp occurring along the banks of permanent or semipermanent streams. The high water table is maintained by the level of water in the stream. The swamp is seasonally inundated with subsequent sediment deposition.
	Basin swamp	*S04*	A swamp developed in a topographically defined basin where water derived locally may be augmented by drainage from other parts of the watershed. Accumulation of well-decomposed peat is shallow (<0.5 m) at the edge but may reach 2 m at the centre.
	Stream marsh	*M06*	A marsh occupying shorelines, bars, stream beds, or islands in continuously flowing watercourses. The marsh is subject to prolonged annual flooding and is commonly covered by thick layers of sediments.
	Shallow basin marsh	*M11*	A marsh occurring in a uniformly shallow marsh depression or swale, having a gradual gradient

			from the edge to the deepest portion; the marsh edge may be poorly defined. Water levels fluctuate rapidly.
	Shore marsh	*M14*	A marsh occupying the contact zone between high and low water marks bordering semipermanent or permanent lakes. The marsh,usually found along protected shorelines, behind barrier beaches in lagoons, on islands, or in embayments, is subject to flooding by a rise in lake levels, wind waves, or surface runoff.

4.3.1.8. Province

Name	Province (PROVINCE_CODE)		
Definition	Specifies the province.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.9. Soil Code

Name	Soil Code (SOIL_CODE)		
Definition	A value used as part of the link to a representative soil in the Soil Names Table (SNT) and Soil Layer Table (SLT). For areas that are not associated with a soil name, the value is '#' (number sign; that is, not applicable).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.10. Soil Code Modifier

Name	Soil Code Modifier (MODIFIER_CODE)		
Definition	Describes a particular set of characteristics related to a soil code. Used as part of the link to a representative soil in the Soil Names Table (SNT) and Soil Layer Table (SLT).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.11. Soil Profile Type

Name	Soil Profile Type (PROFILE_TYPE_CODE)		
Definition	Indicates whether the soil profile type is agricultural or native. Used as part of the link to a representative soil in the Soil Names Table (SNT) and Soil Layer Table (SLT).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Agricultural soil profile	*A*	
	Native soil profile	*N*	

4.3.1.12. Soil Name Identifier

Name	Soil Name Identifier (SOIL_ID)		
Definition	Joins the Soil Names Table (SNT) to other tables such as the Component Table (CMP). There is only one record in the SNT for each Soil Name Identifier attribute. It is created by concatenating the Province, Soil Code, Soil Code Modifier and Soil Profile Type attributes.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.13. Polygon Component Identifier

Name	Polygon Component Identifier (CMP_ID)		
Definition	This field uniquely identifies an instance of a component within a polygon. It is a concatenation of the Soil Landscapes Identifier (SLC_ID) and the Component Number (CMP_NUM).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.14. Ecoregion Identifier

Name	Ecoregion Identifier (ECOREGION_ID)		
Definition	Specifies an ecoregion. An ecoregion is the third level of generalization used in the ecological framework of Canada. It is a collection of ecodistricts and a subdivision of an ecoprovince.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.15. Ecoprovince Identifier

Name	Ecoprovince Identifier (ECOPROVINCE_ID)		
Definition	Specifies an ecoprovince. An ecoprovince is the second level of generalization used in the ecological framework of Canada. It is a collection of ecoregions and a subdivision of an ecozone.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.16. Ecozone Identifier

Name	Ecozone Identifier (ECOZONE_ID)
Definition	Specifies an ecozone. An ecozone is the highest level of generalization used in the ecological framework of Canada. It is a collection of ecoprovinces.
Aliases	

Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.17. Landform Segment Identifier

Name	Landform Segment Identifier (LANDFORM_SEG_ID)		
Definition	Identifies landform segments and their positions. It is created by concatenating the Landform Type, Landform Slope Type and Landform Segment Position attributes, and is used to link to the Landscape Segmentation Table (LST).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.18. Landform Type

Name	Landform Type (LANDFRM_TYPE_CODE)		
Definition	A one-character code that identifies the type of landform.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Hummocky(or irregular)	*H*	
	Inclined or dissected	*I*	
	Level	*L*	
	Rolling	*M*	
	Ridged	*R*	
	Steep	*S*	
	Terraced	*T*	
	Undulating	*U*	

4.3.1.19. Landform Slope Type

Name	Landform Slope Type (LANDFRM_SLOPE_TYPE_CODE)
Definition	Specifies the type of slope for a landform.

Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Little or none	*A*	
	Gentle	*B*	
	Moderate	*C*	
	Steep	*D*	
	Extremely steep	*E*	
	Excessively steep	*F*	

4.3.1.20. Landform Segment Position

Name	Landform Segment Position (LANDFRM_SEG_TYPE_CODE)		
Definition	Identifies one of four possible positions for a landform segment: upper slope, mid slope, lower slope or depression.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Upper	*U*	
	Middle	*M*	
	Lower	*L*	
	Depression	*D*	

4.3.1.21. Landform Segment Percent

Name	Landform Segment Percent (LFS_EXTENT)		
Definition	Specifies the extent of this landform segment within the landform, in percent.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.22. Landform Segment Slope Percent

Name	Landform Segment Slope Percent (LFS_SLOPE)

Definition	Specifies the typical slope for this landform segment, in percent.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.23. Landform Segment Slope Length

Name	Landform Segment Slope Length (LFS_LENGTH)		
Definition	Specifies the typical length in meters of a landform segment.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.24. Landform Segment Name

Name	Landform Segment Name (LFS_NAME)		
Definition	Specifies the name of the landform segment (landform type, slope class and toposequence position/segment).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.25. Polygon Landform Identifier

Name	Polygon Landform Identifier (LANDFORM_ID)
Definition	Identifies each landform within the soil landscapes polygon coverage. It is created by concatenating the Soil Landscapes Identifier, Landform Type and Landform Slope Type attributes.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Character
Value Domain Type	0 (not enumerated)
Value Domain	
	Feature Attribute Value

	Label	Code	Definition

4.3.1.26. Polygon Landform Percent

Name	Polygon Landform Percent (LANDFRM_EXT_PERCENT_NUM)		
Definition	Specifies the estimated extent that a landform occupies within a soil landscape polygon, in percent. For soil landscape polygons with a single landform, the value for Polygon Landform Percent is 100.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.27. Land Area

Name	Land Area (LAND_HA_AREA)		
Definition	Specifies the total area of land in the soil landscapes polygon.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.28. Total Water Area

Name	Total Water Area (WATER_HA_AREA)		
Definition	Specifies the total area in soil landscapes polygon occupied by ocean and fresh water.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.29. Fresh Water Area

Name	Fresh Water Area (FRESH_HA_AREA)
Definition	Specifies the total area of a soil landscapes polygon occupied by fresh

	water.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.30. Ocean Water Area

Name	Ocean Water Area (OCEAN_HA_AREA)		
Definition	Specifies the total area of a soil landscapes polygon that forms part of an ocean.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.31. Total Area

Name	Total Area (TOTAL_HA_AREA)		
Definition	Specifies the total area of a soil landscapes polygon, including land and fresh and ocean water.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.32. Agricultural Crop Use

Name	Agricultural Crop Use (CROP_PRD_TYPE_CODE)		
Definition	Specifies how likely it is that the landscape segment is used for agricultural crop production.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

	Low	*L*	
	Moderate	*M*	
	High	*H*	

4.3.1.33. Layer Number

Name	Layer Number (HRZN_NUM)		
Definition	Specifies the horizon number.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.34. Upper Depth

Name	Upper Depth (HRZN_UDEPTH_CM_NUM)		
Definition	Specifies the depth of the top of the horizon.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.35. Lower Depth

Name	Lower Depth (HRZN_LDEPTH_CM_NUM)		
Definition	Specifies the depth of the bottom of the horizon.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.36. Horizon Lithological Discontinuity

Name	Horizon Lithological Discontinuity (HRZN_LIT_NUM)
Definition	Identifies the presence of lithological discontinuity horizons. When present, this attribute uses a value between 2 and 5 to identify the number of the lithological discontinuity.

Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.37. Horizon Master

Name	Horizon Master (HRZN_MASTER_TYPE_CODE)		
Definition	Identifies the master horizon. For more information, see Chapter 2 of The Canadian System of Soil Classification, Third edition.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Surface mineral horizon	*A*	
	Subsurface mineral horizon	*B*	
	Parent material horizon	*C*	
	Organic horizon partially decomposed	*F*	
	Organic horizon of indiscernible origin	*H*	
	Organic horizon poorly decomposed	*L*	
	Organic horizon from mosses, rushes and woody materials	*O*	
	Rock	*R*	
	Water	*W*	
	n/a	-	

4.3.1.38. Horizon Suffix

Name	Horizon Suffix (HRZN_SUFFIX_TEXT)
Definition	Specifies the horizon suffix. For more information, see Chapter 2 of The Canadian System of Soil Classification, Third edition.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Integer
Value Domain Type	0 (not enumerated)

Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.39. Horizon Modifier

Name	Horizon Modifier (HZN_MODIFIER_TEXT)		
Definition	Specifies a soil horizon modifier.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.40. Coarse Fragment Percentage

Name	Coarse Fragment Percentage (COFRAG_PERCENT_NUM)		
Definition	Specifies the amount of coarse fragments in a layer as a percentage of volume.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.41. Dominant Sand Fraction

Name	Dominant Sand Fraction (DOMSAND_TYPE_CODE)		
Definition	Identifies the dominant sand fraction of a layer.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Very Coarse	*VC*	
	Coarse	*C*	
	Medium	*M*	
	Fine	*F*	
	Very Fine	*VF*	

	n/a	-	

4.3.1.42. Very Fine Sand Percentage

Name	Very Fine Sand Percentage (VFSAND_PERCENT_NUM)		
Definition	Indicates the percentage of very fine sand in a layer, by weight.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.43. Total Percentage Sand

Name	Total Percentage Sand (TSAND_PERCENT_NUM)		
Definition	Indicates the percentage of total sand in a layer, by weight.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.44. Total Percentage Silt

Name	Total Percentage Silt (TSILT_PERCENT_NUM)		
Definition	Indicates the percentage of total silt in a layer, by weight.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.45. Total Clay Percentage

Name	Total Clay Percentage (TCLAY_PERCENT_NUM)
Definition	Indicates the percentage of total clay in a layer, by weight.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Integer

Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.46. Organic Carbon Percentage

Name	Organic Carbon Percentage (ORGCARBON_PERCENT_NUM)		
Definition	Indicates the percentage of organic carbon in a layer, by weight.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.47. pH in Calcium Chloride

Name	pH in Calcium Chloride (PH_CACL2_NUM)		
Definition	Specifies the soil pH, measured using pH in calcium chloride.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.48. Project Report pH

Name	Project Report pH (PH_PRJRPT_NUM)		
Definition	Indicates the pH as specified in the project report.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.49. Base Saturation Percentage

Name	Base Saturation Percentage (BASE_STRT_PERCENT_NUM)

Definition	Indicates the base saturation in percent.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.50. Cation Exchange Capacity

Name	Cation Exchange Capacity (CEC_NUM)		
Definition	Identifies the cation exchange capacity.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.51. Saturated Hydraulic Conductivity

Name	Saturated Hydraulic Conductivity (KSAT_NUM)		
Definition	Indicates the saturated hydraulic conductivity, in centimeters per hour (cm/h).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.52. Water Retention at 0 Kilopascals

Name	Water Retention at 0 Kilopascals (KP0_NUM)
Definition	Identifies water retention at 0 kilopascals, in percent of total soil volume.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Integer
Value Domain Type	0 (not enumerated)
Value Domain	
	Feature Attribute Value

	Label	Code	Definition

4.3.1.53. Water Retention at 10 Kilopascals

Name	Water Retention at 10 Kilopascals (KP10_NUM)		
Definition	Identifies water retention at 10 kilopascals, in percent of total soil volume.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.54. Water Retention at 33 Kilopascals

Name	Water Retention at 33 Kilopascals (KP33_NUM)		
Definition	Identifies water retention at 33 kilopascals, in percent of total soil volume.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.55. Water Retention at 1500 Kilopascals

Name	Water Retention at 1500 Kilopascals (KP1500_NUM)		
Definition	Identifies water retention at 1500 kilopascals, in percent of total soil volume.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.56. Bulk Density

Name	Bulk Density (BULK_DENSITY_NUM)
Definition	Identifies the bulk density of the fine earth fraction of a layer, excluding

	coarse fragment content.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.57. Electrical Conductivity

Name	Electrical Conductivity (ELEC_CONDUCT_NUM)		
Definition	Specifies the electrical conductivity of a layer in decisiemens per meter (dS/m).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.58. Calcium Carbonate Percentage

Name	Calcium Carbonate Percentage (CACO3_PERCENT_NUM)		
Definition	Specifies a value that is equivalent to a percentage of pure calcium carbonate.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.59. Von Post Decomposition

Name	Von Post Decomposition (DECOMP_VONPOST_CODE)
Definition	Specifies peat decomposition using the Von Post method of asessment.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Integer
Value Domain Type	0 (not enumerated)
Value Domain	
	Feature Attribute Value

	Label	Code	Definition

4.3.1.60. Woody Material Percentage

Name	Woody Material Percentage (WOODY_MAT_PERCENT_NUM)		
Definition	Specifies the volume of woody material in a layer as a percentage.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Integer		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.61. Soil Name

Name	Soil Name (SOIL_NAME)		
Definition	Specifies the complete soil name, without abbreviations. Each Soil Name attribute can be associated with up to two entries in the Soil Names Table (SNT).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.62. Kind of Material

Name	Kind of Material (KIND_TYPE_CODE)		
Definition	Specifies the kind of surface soil, rock outcrop or other material or feature.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Mineral Soil	*M*	Predominantly mineral particles. Contains less than 30% organic matter as measured by weight.
	Organic Soil	*O*	Contains more than

			30% organic matter as measured by weight.
	Non-soil	*N*	Rockland, quarry, unclassified, not surveyed and so on.
	Urban	*U*	Urban areas, airports and other artificial features.

4.3.1.63. Water Table Type

Name	Water Table Type (WTRTBL_TYPE_CODE)		
Definition	Indicates the presence of a water table and identifies its characteristics.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Always	*YB*	Present during both seasons
	Growing Season	*YG*	Present during the growing season
	Non growing season	*YN*	Present outside the growing season
	Unspecified time	*YU*	Present during unspecified time
	Never	*NO*	Not present anytime
	Not applicable	-	Not applicable

4.3.1.64. Restrictive Layer

Name	Restrictive Layer (RESTRICT_LAYER_CODE)		
Definition	Indicates the soil layer that restricts root growth. A value between 1 and 9 indicates that the layer is restrictive and can be linked to the Layer Number attribute in the Soil Layer Table. A value of 0 indicates that no restricting layer is present. N/A indicates "not applicable".		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.65. Soil Layer Restrictive Soil Type

Name	Soil Layer Restrictive Soil Type (RESTRICT_TYPE_CODE)		
Definition	Indicates the type of soil in the root restricting layer.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Not applicable	-	
	Solonetzic B horizon	*BN*	
	Cryic (frozen) horizon	*CR*	
	Compact (basal) till	*CT*	
	Duric horizon	*DU*	
	Fragipan	*FP*	
	Lithic	*LI*	
	Ortstein horizon	*OR*	
	Placic horizon	*PL*	
	Salinity	*SA*	
	Undifferentiated	*UN*	

4.3.1.66. Drainage Type

Name	Drainage Type (DRAINAGE_TYPE_CODE)		
Definition	Indicates the type of drainage found within a soil profile.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Very rapidly drained	*VR*	
	Rapidly drained	*R*	
	Well drained	*W*	
	Moderately well drained	*MW*	
	Imperfectly drained	*I*	
	Poorly drained	*P*	
	Very poorly drained	*VP*	
	Not applicable	-	

4.3.1.67. Parent Material 1 Texture

Name	Parent Material 1 Texture (PMTEXTURE1_TYPE_CODE)
Definition	Specifies the texture of the underlying parent material (mineral material that is transformed into soil).

Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Very Coarse	*VC*	
	Coarse	*C*	
	Moderately Coarse	*MC*	
	Medium	*M*	
	Moderately Fine	*MF*	
	Fine	*F*	
	Very Fine	*VF*	
	Coarse Skeletal	*CS*	
	Medium Skeletal	*MS*	
	Fine Skeletal	*FS*	
	Fragmental	*FR*	
	Stratified (Mineral)	*SM*	
	Fibric	*FI*	
	Mesic	*ME*	
	Humic	*HU*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.3.1.68. Parent Material 2 Texture

Name	Parent Material 2 Texture (PMTEXTURE2_TYPE_CODE)		
Definition	Specifies the texture of the second underlying parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Very Coarse	*VC*	
	Coarse	*C*	
	Moderately Coarse	*MC*	
	Medium	*M*	
	Moderately Fine	*MF*	
	Fine	*F*	
	Very Fine	*VF*	
	Coarse Skeletal	*CS*	
	Medium Skeletal	*MS*	
	Fine Skeletal	*FS*	

	Fragmental	*FR*	
	Stratified (Mineral)	*SM*	
	Fibric	*FI*	
	Mesic	*ME*	
	Humic	*HU*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.3.1.69. Parent Material 3 Texture

Name	Parent Material 3 Texture (PMTEXTURE3_TYPE_CODE)		
Definition	Specifies the texture of the third underlying parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Very Coarse	*VC*	
	Coarse	*C*	
	Moderately Coarse	*MC*	
	Medium	*M*	
	Moderately Fine	*MF*	
	Fine	*F*	
	Very Fine	*VF*	
	Coarse Skeletal	*CS*	
	Medium Skeletal	*MS*	
	Stratified (Mineral)	*SM*	
	Fibric	*FI*	
	Mesic	*ME*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.3.1.70. Parent Material 1 Chemical Composition

Name	Parent Material 1 Chemical Composition (PMCHEM1_TYPE_CODE)		
Definition	Specifies the chemical composition of the upper layer parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Extremely/Strongly	*EA*	

	acidic		
	Medium acid to neutral	*AN*	
	Weakly calcareous	*WC*	
	Moderately/Very strongly calcareous	*VC*	
	Extremely calcareous	*EC*	
	Calcareous and saline	*SA*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.3.1.71. Parent Material 2 Chemical Composition

Name	Parent Material 2 Chemical Composition (PMCHEM2_TYPE_CODE)		
Definition	Specifies the chemical composition of the second underlying parent material (mineral material that is transformed into soil), if present.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Extremely/Strongly acidic	*EA*	
	Medium acid to neutral	*AN*	
	Weakly calcareous	*WC*	
	Moderately/Very strongly calcareous	*VC*	
	Extremely calcareous	*EC*	
	Calcareous and saline	*SA*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.3.1.72. Parent Material 3 Chemical Composition

Name	Parent Material 3 Chemical Composition (PMCHEM3_TYPE_CODE)
Definition	Specifies the chemical composition of the third underlying parent material (mineral material that is transformed into soil), if present.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Character
Value Domain Type	1 (enumerated)
Value Domain	
	Feature Attribute Value

	Label	Code	Definition
	Extremely/Strongly acidic	*EA*	
	Medium acid to neutral	*AN*	
	Weakly calcareous	*WC*	
	Moderately/Very strongly calcareous	*VC*	
	Extremely calcareous	*EC*	
	Undifferentiated	*UD*	
	Not applicable	-	

4.3.1.73. Parent Material 1 Mode of Deposition

Name	Parent Material 1 Mode of Deposition (MDEPOSIT1_TYPE_CODE)		
Definition	Specifies the mode of deposition of the uppermost underlying parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Not applicable	-	
	Colluvial	*COLL*	
	Fluvioeolian	*FLEO*	
	Undifferentiated Organic	*UNDO*	
	Bedrock, undifferentiated	*RKUD*	
	Igneous, acidic bedrock (dominantly granite)	*RKIA*	
	Igneous, basic bedrock (dominantly basalt)	*RKIB*	
	Calcareous bedrock (dominantly sandstone and shale)	*RKCA*	
	Limestone bedrock (limestone and dolomite)	*RKLS*	
	Shale bedrock	*RKSH*	
	Sandstone bedrock	*RKSS*	
	Slate bedrock	*RKSA*	
	Anthropogenic	*ANTH*	
	Eolian	*EOLI*	
	Fluviolacustrine	*FLLC*	

	Fluvial	*FLUV*	
	Fen Peat	*FNPT*	
	Forest Peat	*FOPT*	
	Glaciofluvial	*GLFL*	
	Glaciolacustrine	*GLLC*	
	Glaciomarine	*GLMA*	
	Lacustrine	*LACU*	
	Lacustro-Till	*LATL*	
	Marine	*MARI*	
	Residual	*RESD*	
	Saprolite	*SAPR*	
	Sphagnum Peat	*SPPT*	
	Till (morainal)	*TILL*	
	Undifferentiated Mineral	*UNDM*	
	Volcanic	*VOLC*	

4.3.1.74. Parent Material 2 Mode of Deposition

Name	Parent Material 2 Mode of Deposition (MDEPOSIT2_TYPE_CODE)		
Definition	Specifies the mode of deposition of the second underlying parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Not applicable	-	
	Colluvial	*COLL*	
	Undifferentiated Organic	*UNDO*	
	Bedrock, undifferentiated	*RKUD*	
	Igneous, acidic bedrock (dominantly granite)	*RKIA*	
	Igneous, basic bedrock (dominantly basalt)	*RKIB*	
	Igneous, basic coarse bedrock (coarse grained basalts)	*RKIC*	
	Igneous, basic fine grained bedrock (dominantly andesites)	*RKIF*	
	Calcareous bedrock (dominantly	*RKCA*	

	sandstone and shale)		
	Limestone bedrock (limestone and dolomite)	*RKLS*	
	Shale bedrock	*RKSH*	
	Sandstone bedrock	*RKSS*	
	Siltstone and mudstone bedrock	*RKSM*	
	Schist and phyllite bedrock	*RKSP*	
	Gneiss bedrock	*RKGN*	
	Eolian	*EOLI*	
	Fluviolacustrine	*FLLC*	
	Fluvial	*FLUV*	
	Fen Peat	*FNPT*	
	Forest Peat	*FOPT*	
	Glaciofluvial	*GLFL*	
	Glaciolacustrine	*GLLC*	
	Glaciomarine	*GLMA*	
	Lacustrine	*LACU*	
	Lacustro-Till	*LATL*	
	Marine	*MARI*	
	Residual	*RESD*	
	Saprolite	*SAPR*	
	Sphagnum Peat	*SPPT*	
	Till (morainal)	*TILL*	
	Undifferentiated Mineral	*UNDM*	

4.3.1.75. Parent Material 3 Mode of Deposition

Name	Parent Material 3 Mode of Deposition (MDEPOSIT3_TYPE_CODE)		
Definition	Specifies the mode of deposition of the third underlying parent material (mineral material that is transformed into soil).		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Not applicable	-	
	Undifferentiated Organic	*UNDO*	
	Bedrock, undifferentiated	*RKUD*	
	Igneous, acidic bedrock (dominantly granite)	*RKIA*	

	Limestone bedrock (limestone and dolomite)	*RKLS*	
	Shale bedrock	*RKSH*	
	Slate bedrock	*RKSA*	
	Fluviolacustrine	*FLLC*	
	Fluvial	*FLUV*	
	Fen Peat	*FNPT*	
	Glaciofluvial	*GLFL*	
	Glaciolacustrine	*GLLC*	
	Glaciomarine	*GLMA*	
	Lacustrine	*LACU*	
	Marine	*MARI*	
	Residual	*RESD*	
	Till (morainal)	*TILL*	
	Undifferentiated Mineral	*UNDM*	

4.3.1.76. CSSC 2nd Edition Soil Order

Name	CSSC 2nd Edition Soil Order (SLC_V2_ORDER_CODE)		
Definition	Specifies the soil order according to The Canadian System of Soil Classification (CSSC), Second edition.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.77. CSSC 2nd Edition Soil Great Group

Name	CSSC 2nd Edition Soil Great Group (SLC_V2_GR_GROUP_CODE)		
Definition	Specifies the soil great group according to The Canadian System of Soil Classification (CSSC), Second edition.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.78. CSSC 2nd Edition Soil Subgroup

Name	CSSC 2nd Edition Soil Subgroup (SLC_V2_SUB_GROUP_CODE)

Definition	Specifies the soil subgroup according to The Canadian System of Soil Classification (CSSC), Second edition. A soil subgroup is a subdivision of a great group.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.79. CSSC 3rd Edition Soil Order

Name	CSSC 3rd Edition Soil Order (SLC_V3_ORDER_CODE)		
Definition	Specifies the soil order according to the Canadian System of Soil Classification, Third edition.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.80. CSSC 3rd Edition Soil Great Group

Name	CSSC 3rd Edition Soil Great Group (SLC_V3_GR_GROUP_CODE)		
Definition	Specifies the soil great group according to The Canadian System of Soil Classification (CSSC), Third edition.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.81. CSSC 3rd Edition Soil Subgroup

Name	CSSC 3rd Edition Soil Subgroup (SLC_V3_SUB_GROUP_CODE)
Definition	Specifies the soil subgroup according to The Canadian System of Soil Classification (CSSC), Third edition. A subgroup is a subdivision of a great group.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Character

Value Domain Type	0 (not enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition

4.3.1.82. 0 to 50 cm Coarse Fragment Content

Name	0 to 50 cm Coarse Fragment Content (CFRAG1_TYPE_CODE)		
Definition	Specifies the average coarse fragment content of the soil profile, from 0 to 50 centimeters of depth. For root-restricting layers, at less than 50 centimeters of depth.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	More than 65%	*D*	
	31-65%	*C*	
	10-30%	*B*	
	Less than 10%	*A*	
	Not Applicable	-	

4.3.1.83. 50 to 100 cm Coarse Fragment Content

Name	50 to 100 cm Coarse Fragment Content (CFRAG2_TYPE_CODE)		
Definition	Specifies the average coarse fragment content of the soil profile, from 50 to 100 centimeters of depth. For a root-restricting layer, at greater than 50 centimeters of depth.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	More than 65%	*D*	
	31-65%	*C*	
	10-30%	*B*	
	Less than 10%	*A*	
	Not Applicable	-	

4.3.1.84. Depth to Bedrock

Name	Depth to Bedrock (DEPTH_TYPE_CODE)
Definition	Specifies the depth to bedrock or a root-restricting layer.

Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Less than 25 cm	*1*	
	25-49 cm	*2*	
	50-74 cm	*3*	
	75-99 cm	*4*	
	100 cm or more	*5*	
	n/a (for example, rock or ice)	-	

4.3.1.85. Component Restrictive Soil Type

Name	Component Restrictive Soil Type (RESTRICTION_TYPE_CODE)		
Definition	Specifies the type of soil in a root-restricting layer.		
Aliases			
Producer	Canadian Soil Information System (CanSIS)		
Value Data Type	Character		
Value Domain Type	1 (enumerated)		
Value Domain			
	Feature Attribute Value		
	Label	Code	Definition
	Not applicable	-	
	Solonetzic B horizon	*BN*	
	Cryic (frozen) horizon	*CR*	
	Compact (Basal) till	*CT*	
	Duric horizon	*DU*	
	Fragipan	*FP*	
	Lithic	*LI*	
	Ortstein horizon	*OR*	
	Placic horizon	*PL*	
	Salinity	*SA*	
	Undifferenciated	*UN*	

4.3.1.86. Water Capacity

Name	Water Capacity (WTR_CPCTY_TYPE_CODE)
Definition	Specifies a soil's capacity for holding water.
Aliases	
Producer	Canadian Soil Information System (CanSIS)
Value Data Type	Character
Value Domain Type	1 (enumerated)
Value Domain	

	Feature Attribute Value		
	Label	Code	Definition
	<50 mm	*1*	
	50-99 mm	*2*	
	100-149 mm	*3*	
	150-199 mm	*4*	
	200-250 mm	*5*	
	Water, ice or rock	-	

5. REFERENCE SYSTEM

5.1. Spatial reference system

Horizontal coordinate reference system: WGS 84
Map projection: Geographic; EPSG: 3857 Version 8.1.4

5.2. Temporal reference system

Gregorian calendar

6. DATA QUALITY

6.1. Completeness

6.2. Logical consistency

6.3. Positional accuracy

6.4. Temporal accuracy

6.5. Thematic accuracy

6.6. Lineage statement

6.6.1. Soil Landscapes of Canada Version 3.2 – Boundaries

Lineage Statement	SLC version 3.2 is the latest revision of the Soil Landscapes of Canada, which was developed by Agriculture and Agri-Food Canada to provide information about the country's agricultural soils at the provincial and national levels. SLC v3.2 replaces SLC v3.1.1 and retains the linkage to the national Ecological Stratification System for Canada that was established in SLC v2.2.
Scope	Dataset (Soil Landscapes of Canada Version 3.2 – Boundaries

6.6.2. Soil Landscapes of Canada Version 3.2 – Component Table (CMP)

Lineage Statement	The Soil Landscape Component Table was updated for version 3.2 to meet the requirements of a number of AAFC initiatives including the National Greenhouse Gas and Carbon Accounting Validation System (NGCAVS), and the National Agri-Environmental Health Assessment and Reporting Program (NAHARP). It is based on the SLC version 2.2 component table, but has been updated to reflect new knowledge in the agricultural areas of the country. It has been compiled for all map coverages in Canada. It is a generalized version of the information that can be found in the detailed soil surveys.
Scope	Dataset (Soil Landscapes of Canada Version 3.2 – Component Table (CMP))

6.6.3. Soil Landscapes of Canada Version 3.2 – Ecological Framework Table (EFT)

Lineage Statement	The Ecological Framework Table was new for SLC version 3.2 and contains information formerly found in the Polygon Attribute Table (PAT).
Scope	Dataset (Soil Landscapes of Canada Version 3.2 – Ecological Framework Table (EFT))

6.6.4. Soil Landscapes of Canada Version 3.2 – Landform Definition Table (LDT), Soil Landscapes of Canada Version 3.2 - Landform Extent Table (LET), Soil Landscapes of Canada Version 3.2 - Landscape Segmentation Table (LST)

Lineage Statement	This is a new set of tables introduced with SLCv3.2. A similar Landform Table was provided for specific clients with SLCv3.1 and v3.1.1. The SLCv3.2 Landform Table is the first version provided to all clients with the official CanSIS download files. Significant changes from the previous SLCv3.1.1 Landform table include: 1. The tables have been restructured into a more normalized form, to make it easier to understand the data and how it was developed. 2. The soil MODIFIER codes were changed from 3 to 5 characters, to match the new 5 character MODIFIER codes used in the SLCv3.2 CMP, SNT and SLT. 3. A new, 11 character SOIL_ID was added, to provide a convenient linkage to the other CanSIS files. 4. New Landform and Landform segment identifier codes were provided, which can be used to link to the specific attributes of each landform segment as defined in the Landform Segmentation Table (LST). 5. Many specific bugs in the assignment of soils to specific SLCv3.1.1 landscape segments were identified and fixed in the SLCv3.2 Landform table. Many of these were in "cross border" SLC polygons, which had landforms and soil components listed for two provinces.
Scope	Dataset (Soil Landscapes of Canada Version 3.2 – Landform Definition Table (LDT), Soil Landscapes of Canada Version 3.2 - Landform Extent Table (LET), Soil Landscapes of Canada Version 3.2 - Landscape Segmentation Table (LST))

6.6.5. Soil Landscapes of Canada Version 3.2 – Landscape Area Table (LAT)

Lineage Statement	The Landscape Area Table was developed to simplify the calculation of land and water areas for the SLC polygons, and remove ambiguities in the way water was represented in previous versions. This table was

	derived from the SLC and HYDRO coverages. In earlier SLC versions this table was called the Land Area Table. For version 3.2, field names in this table were renamed in order to improve clarity.
Scope	Dataset (Soil Landscapes of Canada Version 3.2 – Landscape Area Table (LAT))

6.6.6. Soil Landscapes of Canada Version 3.2 – Soil Layer Table (SLT)

Lineage Statement	Updates to this table in SLC version 3.2 include: 1. Duplicate soil component records were eliminated. 2. Records for each soil were reordered so that they appear in sequence. 3. Soils now have complete soil horizon data to 100 cm depth, or to the depth of a significant root restricting layer, such as bedrock. 4. Additional checks were added to ensure that the sand, silt and clay percentages for each soil horizon were valid, and to ensure that the percentage of very fine sand was less than or equal to the total sand percentage. 5. The Soil Layer Table also contains additional soil records for each province that do not occur in the Component Table. Some of these "extra" soils may occur with the Landscape Segmentation Table, while others are minor provincial soils that only occur in provincial detailed soil map polygons. A copy of the most recent provincial Soil Layer file is used as the source for the soil attribute data in the Soil Layer Table. The extra provincial soil records in the Soil Layer Table have also been updated from the SLCv3.1.1 Soil Layer Table.
Scope	Dataset (Soil Landscapes of Canada Version 3.2 – Soil Layer Table (SLT))

6.6.7. Soil Landscapes of Canada Version 3.2 – Soil Name Table (SNT)

Lineage Statement	Updates to this table in SLC version 3.2 include: 1.The Soil Order, Great Group and Subgroup codes in the Soil Name Table for all provinces but Saskatchewan have been rechecked and updated to conform to the Canadian System of Soil Classification Third Edition. 2.An additional check was made to ensure that data fields that identify a lithic, cryic or saline root restricting layer and layer number (RESTRICT_TYPE and ROOTRESTRI) in the Soil Name Table are linked to a valid root restricting layer number for the same soil in the Soil Layer table, with the appropriate horizon designation and attribute values. 3.The chemical property fields for the first, second and third parent materials (PMCALC1, PMCALC2, and PMCALC3) were renamed PMCHEM1, PMCHEM2, and PMCHEM3. The new names better reflect the range of values, which includes acidic classes as well as calcareous classes. The data field code sets are otherwise unchanged. The revised PMCHEM1 and PMCHEM2 data field names are also provided in the SLCv3.2 Component table. 4.Additional checks were made for more complete and consistent identification of soil parent materials (MDEP1,2,3, PMTEX1,2,3, and PMCHEM 1,2,3 fields). A soil with one, two or three parent materials should have valid values in all three of these parent material data field types. 5.The Soil Name Table also contains additional soil records for each province that do not occur in the Component Table. Some of these "extra" soils may occur with the Landscape Segmentation Table, while others are minor provincial soils that only occur in provincial detailed soil map polygons. A copy of the most recent provincial Soil Name file is

	used as the source for the soil attribute data in the Soil Name Table. The extra provincial soils in the SLCv3.2 Soil Name Table have also been updated from the SLCv3.1.1 Soil Name Table.
Scope	Dataset (Soil Landscapes of Canada Version 3.2 – Soil Name Table (SNT))

6.6.8. Soil Landscapes of Canada Version 3.2 – Component Rating Table (CRT)

Lineage Statement	This is a new table introduced with SLCv3.2. It contains basic interpretations for each component, generated from the contents of the Soil Name and Soil Layer tables. Fields in this table were formerly found in the CMP.
Scope	Dataset (Soil Landscapes of Canada Version 3.2 – Component Rating Table (CRT))

7. DATA CAPTURE

8. DATA MAINTENANCE

AsNeeded

9. PORTRAYAL

Not applicable.

10. DATA PRODUCT DELIVERY

Delivery medium information:

units of delivery: package
medium name: online via HTTP, online via direct access

Delivery format information:

File Geodatabase
format name: Esri Geodatabase (File-based)
format version: 10.1
specification: A collection of various types of GIS datasets held in a file system folder. (http://arcgis.com)
languages: eng
character set: utf8

Delivery format information:

GML
format name: Geography Markup Language
format version: 2.0
specification: Open Geospatial Consortium Inc., OpenGIS®Geography Markup Language (GML) Implementation Specification, Version 3.1.1, 2004-02-07, Reference number of this OGC® project document: 03-105r1 (http://portal.opengeospatial.org/files/?artifact_id=4700)

languages: eng
character set: utf8

Delivery format information:

csv

format name: Comma Delimited
format version: 1.0
specification: A delimited data format that has fields/columns separated by the comma character
languages: eng
character set: utf8

11. METADATA

The metadata requirements follow the Government of Canada's Treasury Board Standard on Geospatial Data (ISO 19115).