NRC·CNRC

Assessment of Canada's Hydrokinetic Resources: A Review of Hydrologic Considerations

Report No.: NRC-OCRE-2020-TR-019
Date: March 31, 2020
Authors: M.N. Khaliq, J. Cousineau

Ocean, Coastal and River Engineering Research Center

National Research Council Canada
Conseil national de recherches Canada

Canada

Cat. No. NR16-325/2020E-PDF

ISBN 978-0-660-35532-0

Acknowledgements

The work reported here was undertaken within the framework of an inter-departmental agreement between CanmetENERGY, Natural Resources Canada (NRCan), and the National Research Council Canada (NRC). Project co-ordination support of Ganaysham Panday, from CanmetENERGY, is much appreciated.

Executive Summary

Recently, there has been a growing interest in developing hydrokinetic or in-stream power potential nationally and internationally. This type of resource development using zero-head turbines requires no dams or barrages as in the case of conventional large-scale hydropower production projects. In 2010, Natural Resources Canada (NRCan) commissioned a multiyear project to assess Canada's hydrokinetic potential in an effort to boost renewable energy resources across the country. This project involved development of relevant resource databases that governments and regional entities could use for investment planning and decision-making, as well as for private market development. The National Research Council Canada (NRC) led this effort through an inter-departmental agreement between NRCan and the NRC and completed an assessment of hydrokinetic resources at both regional and national levels. This assessment was completed in three phases as described below:

<u>Phase I Methodology Review and Data Review:</u> In this phase of the study, the NRC undertook a review of various methods, techniques and data sources to identify suitable combinations of techniques and data sources for conducting an assessment of hydrokinetic potential at regional and national levels. Additionally, a number of methods, datasets and locations were also identified to validate selected methodologies. This effort was documented in a comprehensive technical report by the NRC (i.e. Jenkinson 2010), which is available online through NRC's archives.

<u>Phase II Methodology Validation:</u> Following the outcomes from Phase I of the study, a few selected methods and recommendations were evaluated using data from the national hydrometric network, maintained by Water Survey of Canada. A sensitivity/uncertainty analysis was also carried out. The outcomes of this phase and a vision for the national assessment were documented in another technical report (i.e. Jenkinson and Bomhof 2012).

<u>Phase III Assessment Determination:</u> Based on the outcomes and recommendations from both Phase I and Phase II of the study, a nation-wide assessment of the hydrokinetic potential for theoretical energy extraction was carried out. The outcome of this phase along with specific guidelines on the developed databases pertaining to hydrokinetic resources were documented in Jenkinson and Bomhof (2014), which was the third and the final technical report related to the multiyear resource assessment study. This report is available from NRCan on request.

For hydrokinetic resource assessment in a given region of interest, hydrologic investigations and hydraulic modelling are ideally required and both are carried out in tandem. Hydrologic investigations pertain to estimation of various streamflow indices (e.g. monthly and annual mean flows, selected percentiles of flow duration curves, etc.) at all points of interest within a selected region of interest and hydraulic modelling pertains to estimation of river flow velocities at all relevant points. This report specifically reviews most of the methods and data sources pertaining to hydrologic investigations, previously recommended and used for resource assessment in

Jenkinson (2010) and Jenkins and Bomhof (2012, 2014). Where applicable, shortcomings are highlighted and recommendations are made for additional research in order to obtain improved estimates of various streamflow indices, which, in turn, will help improving the quality of resource databases. Hydraulic modelling aspects, which are an integral part of the resource assessment study, are not reviewed in this report. Additional information about the contents and scope of this report is provided below.

This report is divided into seven chapters and a section on references. The background information on hydrokinetic resources and previous assessment studies is provided in Chapter 1 in order to provide the reader with sufficient background on the topic. Objectives and limitations of the report are also discussed in this chapter. This introductory information was necessary to establish a firm basis for the review and analysis presented in other chapters of the report. An introductory primer on hydrokinetic energy in the form of foundational knowledge is provided in Chapter 2. Chapter 3 of the report provides an overview of the work completed in Phase I of the resource assessment study. Review of existing literature pertaining to transposition of streamflow indices from gauged to ungauged locations within the realm of ungauged hydrology and hydraulic modelling associated with hydrokinetic resource assessment was a significant part of this phase of the study. In the present report, this review is further summarized to highlight important outcomes of that part of the study. Chapter 4 presents a similar overview as given in Chapter 3, but for the work completed within the scope of Phase II of the resource assessment study, wherein the main focus was on validating a suitable methodology, proposed in Phase I, using observational data from hydrometric stations located within selected large geographic regions of Canada. Chapter 5 presents an overview of the work completed in Phase III of the study. In this phase, the focus was on developing estimates of hydrokinetic resources on a national scale, following the methodology which was proposed in Phase I and validated in Phase II. Chapter 6 presents an overall synthesis of the resource assessment study and discusses some of the main findings of the study. Chapter 7 explores avenues of future research on the basis of a set of thoughtful guidelines, and discusses potential recommendations and steps necessary to be followed for updating resource assessment study, by overcoming various shortcomings identified in Chapters 3 to 5 of the report.

The information provided in this report is expected to help pave the way forward for improving estimation of streamflow indices at ungauged locations across Canada, as well as for improving our understanding of geophysical and climatic datasets and their inter-dependencies that form the critical basis for deriving statistical relationships to support assessment of hydrokinetic resources. The contents of this effort will also be useful for developing generalized techniques for estimating hydrologic parameters at ungauged locations through information transposition from gauged to ungauged locations or through direct relationships based on watershed attributes, including topographic, geologic, soil and land use, and climatic attributes. An effort has also been made to reflect on present state of the knowledge in ungauged hydrology with respect to estimation of streamflow indices. However, such reviews should occur on regular basis in order to strengthen and validate existing and emerging approaches based on refined and improved datasets of watershed attributes. These datasets are continuously being refined through dedicated national level initiatives.

Table of Contents

List of Tables

List of Figures

List of Acronyms

Acronym	Description
AAFC	Agriculture and Agri-Food Canada
CanSIS	Canadian Soil Information System
CCA	Canonical Correlation Analysis
CDED	Canadian Digital Elevation Data
DEM	Digital Elevation Model
DFO	Department of Fisheries and Oceans
DHR	Delineation of Homogenous Regions
EC	Environment Canada
ECDE	Environment Canada Data Explorer
EC-MDB	Environment Canada’s Measurement Database
FDC	Flow Duration Curve
GIS	Geographical Information System
GML	Geography Mark-up Language
HYDAT	Hydrometric Database
MAF	Mean Annual Flow
NHN	National Hydro Network
NRC	National Research Council Canada
NRCan	Natural Resources Canada
OCRE	Ocean, Coastal and River Engineering
RBLI	Regression Based Logarithmic Interpolation
REM	Regional Estimation Method
RETScreen	Renewable Energy and Energy Efficient Technologies Screening Tool
RMSE	Root Mean Square Error
ROI	Region of Influence
SLCs	Soil Landscapes of Canada
VDC	Velocity Duration Curve
WSC	Water Survey of Canada

1 Introduction

1.1 Background

Recently, there has been a growing interest in developing hydrokinetic or in-stream power potential nationally and internationally. This type of resource development using zero-head turbines requires no dams or barrages as in the case of conventional large-scale hydropower production projects. Canada has a vast network of rivers and streams (see Figure 1) that have been harnessed for several decades to generate electricity for domestic and industrial consumption. This network of rivers and streams potentially contains a considerable potential for power production from hydrokinetic resources. In the past some efforts have been made to quantify this potential across Canada (e.g. UMA Group 1980, Acres Consulting Services Limited 1984a, Natural Resources Canada 1984a, 1984b, and NRC-CHC 2008b). Though the efforts are continuing, the commercial value of this resource is still largely unknown. In addition, the associated technology for extraction of power is still evolving and is being refined through targeted field testing and controlled laboratory experiments.

In 2010, Natural Resources Canada (NRCan) commissioned a multiyear project to assess Canada’s hydrokinetic potential in an effort to boost renewable energy resources across the country. To support this initiative, a database of hydrokinetic resources was required to be developed to help government and regional entities for policy-making and investment planning. For industry, knowledge of the resource potential, and where it is located, are key pieces of information for technology and market development. The knowledgebase and the outputs of such an effort are also useful for northern regions of Canada where decentralized power production from renewable energy sources is an economically viable option for offsetting the high costs of diesel power production. However, development and identification of an accurate method for assessing hydrokinetic potential of rivers and streams using hydrometric, physiographic and climatic datasets remains a significant challenge. In collaboration with NRCan, the National Research Council Canada (NRC) led this effort and completed a nation-wide assessment of hydrokinetic resources. This assessment was completed in three phases as described below:

<u>Phase I Methodology Review and Data Review:</u> In this phase of the study, the NRC undertook a review of various methods, techniques and data sources for conducting a regional assessment of hydrokinetic potential. Additionally, a number of approaches, datasets and locations were identified to validate a few preferred methods. This effort was documented in the form of a technical report in 2010 (i.e. Jenkinson 2010).

<u>Phase II Methodology Validation:</u> Following the outcomes from Phase I of the study, a few selected methods were evaluated using data from a number of hydrometric stations. A sensitivity/uncertainty analysis was also carried out. The outcomes of this phase were documented in another technical report (i.e. Jenkinson and Bomhof 2012).

Phase III Assessment Determination: Based on the results and recommendations that came out from Phase I and Phase II of the hydrokinetic resource assessment study, a nation-wide assessment of the hydrokinetic potential for theoretical energy extraction was conducted in Phase III. The outcome of this phase along with guidelines on the datasets pertaining to hydrokinetic resources were documented in Jenkinson and Bomhof (2014), the third of a series of three technical reports.

The above studies were financially supported by NRCan through inter-departmental agreements between NRCan and the NRC. Some of the related information is taken directly from these studies to fulfill objectives of this report, which is prepared within the framework of another inter-departmental agreement between NRCan and the NRC. The focus of this effort is to improve assessment of hydrokinetic resources at the regional and national levels based on new technological developments, new analysis tools and improved scientific understanding of hydrologic investigations and hydraulic modelling, as well as by developing refined geophysical and geospatial datasets, where applicable.

For assessing hydrokinetic resources of a given region of interest, information on both hydrological and hydraulic aspects for the entire regional river network is generally required. This information is derived mostly through hydrologic investigations combined with hydraulic modelling. For hydrologic investigations, long-term streamflow observations play a critical role. However, most of the Canadian river network is ungauged and recorded observations are available frequently in southern parts of the country and much less so for northern areas, north of the 60 degrees parallel. In the absence of recorded observations, information on relevant streamflow indices at ungauged locations is obtained through indirect means, e.g. by transposing known or processed information from gauged to ungauged locations, following a set of established scientific and hydrologic principles. The streamflow indices often considered for hydrokinetic resource assessment are mean monthly or mean annual flows and flow duration curves (FDCs). The FDC establishes a time-independent relationship between various streamflow magnitudes and their frequencies of occurrence. Hydraulic modelling component pertains to obtaining information on channel velocities or velocity duration curves, which are analogous to FDCs, at various points of interest within a river reach. It is important to note that hydrokinetic power at a given point in a river reach is proportional to cube of the velocity of moving water.

Conventionally, all of the above analyses and investigations are performed using observed or transposed data assuming a stationary climate. Due to climate change as projected by Global Climate Models (GCMs) and documented in various reports of the Intergovernmental Panel on Climate Change (IPCC) (IPCC 2007, 2013), the assumption of a stationary climate has become questionable and therefore the applicability of streamflow indices, derived from recorded historical observations and transposed data under the assumption of stationarity, has also become questionable. It is worth pointing out that many human related activities clearly affect the climate system. Most importantly, emissions of greenhouse gases, especially carbon dioxide and methane, are causing more heat to be trapped within earth's atmosphere. Therefore, the case for significant climate change is compelling in both the empirical observations and theoretical predictions. A warmer air mass can hold more water (i.e., warmer air has a higher saturation vapor pressure) and therefore, it is reasonable to expect higher amounts of water vapor in the air, leading to intensification of the hydrologic cycle, with impacts ranging from one region to another and from one component of the hydrologic cycle to another (Khaliq 2019). However, it is not so straightforward to consider the impacts of a changing climate when deriving streamflow indices, specifically when transposed FDCs are employed to assess hydrokinetic resource potential at various ungauged points of interest within a target region. Though openly recognized and acknowledged, the topic of non-stationary climate is not considered in this report with respect to assessment of hydrokinetic resources in Canada. The same was also assumed in similar previous studies related to hydrokinetic resource assessment. However, future studies could explore the influence of climatic changes on hydrokinetic resources in different regions of Canada.

This report is specifically focused on hydrological aspects of hydrokinetic resource assessment procedures and the description of hydraulic aspects is kept at the minimum. A significant portion of this report is devoted to reviews of previous work completed by the NRC, particularly the methods and datasets related to hydrological aspects of the hydrokinetic resource assessment procedures. For the reader's convenience, a primer on hydrokinetic power is provided in Chapter 2, while Chapters 3 to 5 of the report respectively pertain to the work reported in Phase I (Jenkinson 2010), Phase II (Jenkinson and Bomhof 2012) and Phase III (Jenkinson and Bomhof 2014) reports, completed by the NRC. Each of these chapters is concluded with a high level summary and a set of potential avenues that can be explored in the future for improving hydrologic aspects of resource assessment procedures in order to develop reliable estimates of hydrokinetic potential. Additional relevant detail, specific objectives and limitations of this report are provided below.

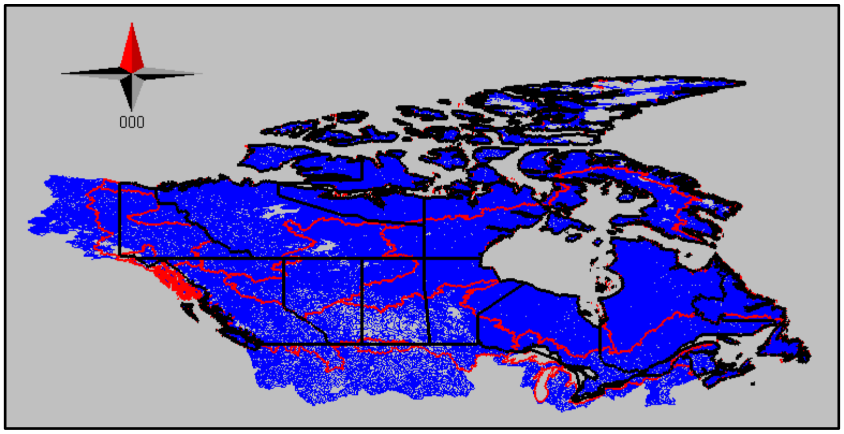

Figure 1: Canadian map showing large drainage basins and stream segments on a map of scale 1:1 million.

1.2 Objectives

The objectives of this report are to:

- Document a critical review of the hydrological aspects of the hydrokinetic resource assessment study, which was completed by the NRC during 2010 to 2014 period, in an effort to improve prediction of various relevant streamflow indices across Canada;
- Identify the areas where additional research is needed to support re-assessment of hydrokinetic resources so that additional site-specific investigations can proceed for power extraction and equipment deployment, as well as to support large scale commercial operations; and
- Carve a path forward for future research and development in order to improve assessment of hydrokinetic resources from a hydrological viewpoint and that, in turn, can inform the development of national guidelines and best practices for boosting renewable power from Canada's vast network of rivers, creeks and streams.

1.3 Organization of the Report

This report is divided into seven chapters, including this introduction chapter, and a section on references. The background information on hydrokinetic resources and previous assessment studies is provided in Chapter 1 in order to provide the reader with sufficient background on the topic. Objectives and limitations of the report are also discussed in this chapter. A short primer on hydrokinetic energy systems in provided in Chapter 2. Chapter 3 of the report provides an overview of the work completed in Phase I of the resource assessment study, previously conducted by the NRC. Review of existing literature was a significant part of this phase of the study. Here, this review is further screened to highlight important outcomes of this part of the study. Chapter 4 presents a similar overview as given in Chapter 3, but for the work completed within the scope of Phase II of the resource assessment study, wherein the main focus was on evaluating a suitable methodology using data from selected large geographic regions within Canada. Chapter 5 presents an overview of the work completed in Phase III of the study. In this phase the focus was on developing estimates of hydrokinetic resources on a national scale following the methodology validated in Phase II. Chapter 6 presents an overall synthesis of the resource assessment study and discusses some of the main findings of Chapters 3 to 5. The final Chapter 7 explores avenues of future research, and discusses potential recommendations and steps necessary to be followed for updating resource assessment study, by overcoming theoretical shortcomings identified in Chapters 3 to 5 and data reliability issues related to watershed attributes. A list of references cited is available at the end of the report. This list is derived mainly from the previously completed reports for Phase I, Phase II and Phase III of the resource assessment study. Selected new references have also been added.

1.4 Convention on the Usage of Acronyms and Other Considerations

A number of acronyms are used in this report, which are devised based on various acronyms used previously in the three related technical reports. Some of the acronyms are chapter-specific, while others are utilized throughout the report. Therefore, to facilitate easy comprehension and smooth readability, the acronyms are reintroduced in their expanded form in each chapter so that each chapter can be read independently, without referring back and forth to other chapters.

Various aspects of open channel flow play a significant role in the assessment of hydrokinetic resources. The terms like river flow, streamflow, flow rate, or simply flow, reflecting open channel flow conditions, are considered equal in terms of meanings in this report. It was necessary to state it upfront since different terms are used in many relevant studies and merely using the word “flow” does not convey true meanings of the contextual analyses. Mean monthly flow, mean annual flow and flow duration curves are collectively referred to as streamflow indices, where appropriate. These three indices of streamflow are frequently referred to in this report.

In Chapters 4 and 5, a number of different datasets have been discussed, in addition to the hydrologic datasets pertaining to recorded streamflow and channel cross-sectional information. These datasets include (1) topographic dataset; (2) soil and land use characteristics dataset; (3) surficial geology dataset; and (4) climatic dataset. A number of different attributes have been derived from each of these datasets to support various analyses related to hydrokinetic resource assessment. In this report, for simplicity reasons, the attributes derived from the above mentioned first three datasets are referred to as physiographic attributes, while all attributes together as watershed/catchment attributes. We also note here that the former convention may not be applicable in general.

1.5 Scope and Limitations

The review, analysis and discussions provided in this report are intended for use by individuals that have some basic understanding of runoff-generating mechanisms in riverine environments, methods pertaining to streamflow analysis and the statistical theory involved in time series modelling and estimation of flow duration curves or many other hydrological indices (e.g. mean monthly and annual flows) at gauged and ungauged locations at various temporal and spatial scales.

The documents and technical/scientific information considered for this report are those which were originally considered for the technical reports by Jenkinson (2010) and Jenkinson and Bomhof (2012, 2014), in addition to the findings documented in these reports. Most of these documents and information sources are publicly available. In this report, where applicable, references are also provided for obtaining additional information and details on various methods, originally described in the above mentioned three technical reports. To set the stage for assessing hydrokinetic resources across Canada, a number of Canadian studies were reviewed and included in Jenkinson (2010) and only a few international studies were mentioned. Therefore, the report by Jenkinson (2010) did not provide a broader perspective on the subject of hydrokinetic resource assessment world-wide.

The scope of this current report is limited to only hydrological investigations. The hydraulic aspects are equally important for resource assessment, but are discussed minimally in this report. For a detailed account of all hydraulic aspects related to resource assessment, the reader is referred to Jenkinson (2010) and Jenkinson and Bomhof (2012, 2014) and the references to previous studies mentioned in there. Some related information can also be found in Kirby et al. (2020). To improve quality of predicted indices of streamflow at ungauged locations for hydrokinetic resource assessment, some avenues of future research are identified based solely on the analyses presented in Jenkinson and Bomhof (2012, 2014) and the developments reported in the literature on physical and stochastic hydrology. Detailed descriptions of theoretical aspects that underpin various statistical and physical methods considered for this report are outside the scope of this report. For such descriptions, scientific articles and technical reports associated with these methods should be referred to. Information on these sources can be found in the references section of this report.

When considering a location for hydropower or other water resources assessment and development related projects, one can divide the project developmental levels into five basic stages (NRCan 2004a, Tudor Engineering Company 1991, Jenkinson 2010): (1) pre-reconnaissance; (2) reconnaissance; (3) pre-feasibility; (4) feasibility, and (5) final design. The primary difference among these stages is the degree of confidence one can have in the results obtained. The first level can be considered as involving low cost, low effort and low confidence analyses. Compared to this, the final stage involves very high confidence analyses, leading to ultimate deployment to achieve the desired design goals. Reconnaissance level may involve site visits, assessments and rankings. The pre-feasibility and feasibility levels involve detailed investigations of selected sites, followed by detailed physical studies of the same sites. The hydrokinetic resource assessment study conducted by the NRC during the 2010–2014 period and documented in Jenkinson (2010) and Jenkinson and Bomhof (2012, 2014) and reviewed here falls under the pre-reconnaissance stage. It was anticipated that various data products that were expected to result from the resource assessment study will help achieve some objectives of the reconnaissance and feasibility levels. Therefore, the review documented in this report should also be considered along the same level.

2 A Primer on Hydrokinetic Power

2.1 General

This chapter provides some basic information on hydrokinetic power and associated mathematical framework in a simplified manner. Most of the information provided here is directly taken from Jenkinson (2010) and Jenkinson and Bomhof (2012), and modified based on the results and analyses presented in Jenkinson and Bomhof (2014). It is important to note that the objective of the above mentioned studies was to characterize and quantify the energy resource specifically for hydrokinetic systems.

2.2 Hydrokinetic Systems

Hydrokinetic systems convert kinetic energy from flowing water into electricity or other forms of energy (Verdant Power 2006). There are a number of features that make hydrokinetic energy systems distinct from other hydropower systems (Verdant Power 2006, Khan et al. 2009) in that they:

- Rely on existing kinetic energy in the river/stream;
- Do not rely on artificial water-head from impoundments, or barrages;
- Do not require large civil works for implementation; and
- May operate in the river's natural pathway and do not require any flow diversions.

Hydrokinetic systems offer a number of advantages over conventional run-of-river and water storage systems, particularly with regards to environmental impact and capital costs. For example, civil works required for the development of run-of-river and water storage hydropower often represent the most significant portion of the project development and can often render a project financially unviable (NRCan 2004a). Systems not requiring a barrage or similar storages have a much lower capital cost. However, the efficiencies and power production capacity of hydrokinetic turbines is much lower than other systems. Hydrokinetic turbines have other advantages too. For example, they may be deployed on an incremental basis, as a single unit or in a clustered configuration. They also remain below the water surface and have a lower noise and vibration issues than the conventional turbines.

2.3 Hydrokinetic Power Estimation

The kinetic energy of moving water can be determined from the density of the water, the velocity at which the water travels and the cross-sectional area of the channel from where the energy will be extracted. The hydrokinetic energy can be expressed as:

$$P_K = \frac{\rho}{2} A V^3 \tag{1}$$

where P_K is the available kinetic energy, ρ is the fluid density, A is the cross-sectional area of the channel and V is the flow velocity. Hydrokinetic power is often reported as a power density (i.e. specific power, P_S) which is the power normalized to a unit cross-sectional area, i.e.

$$P_S = \frac{P_K}{A} = \frac{\rho}{2} V^3 \tag{2}$$

When considering flows in rivers, one can make a reasonable assumption that the density of water remains essentially constant, even with changes in temperature. The velocity and area are the only variables required to determine the hydrokinetic power. In fact, the velocity in a river is rarely constant and is expected to vary significantly on a daily or smaller/larger temporal scale. Determining the average energy at a location in the river requires kinetic energy over a period of time to be integrated. Assuming that the velocity will change with time, and with the same extraction area, the power equation can be written as:

$$\overline{P_K} = \frac{\int P_K(t)dt}{\int dt} = \frac{\int \frac{\rho}{2} A [V(t)]^3 dt}{\int dt} \tag{3}$$

If one is interested in the entire average kinetic energy available in a river at a given point then the area will also change with time, as the flow changes. The above equation can be re-written as:

$$\overline{P_K} = \frac{\int P_K(t)dt}{\int dt} = \frac{\int \frac{\rho}{2} A(t) [V(t)]^3 dt}{\int dt} \tag{4}$$

The total energy available at a river cross-section is strongly influenced by the velocity and the temporal variability of flow. The resource assessment study (i.e. Jenkinson 2010 and Jenkinson and Bomhof 2012, 2014) focused mainly on estimating the hydrokinetic power available in flowing water and did not consider turbine characteristics or turbine efficiency when characterizing the potential resource.

The above equations represent the energy with an average flow velocity across the cross-sectional area of interest, that being the entire river cross-sectional area or the cross-sectional area of just the turbine device. In fact the velocity may vary within the cross-sectional area itself. This is particularly the case in natural channels if the turbine cross-sectional area is not small compared to the cross-sectional area of the channel, or if the area considered is the entire cross-sectional area of the channel. Velocity within the channel cross-section will vary substantially both in the vertical and the horizontal directions. Velocity profiles along the vertical axis are generally required for any type of hydrokinetic device. As power is a function of the velocity cubed, it is important to have a complete vertical velocity profile and knowledge of where the device is going to be anchored.

In the resource assessment study, the total hydrokinetic energy in the stream was assessed using predicted values of average flow velocity across the river cross-section. Average velocity at a cross-section does not represent the average kinetic energy flux across the cross-section, as the relationship between velocity and power is non-linear. The power available will likely be some value greater than that calculated with the average velocity, depending on the nature of the velocity distribution. While the spatial variability of velocity across a river cross-section is expected to have an impact on the hydrokinetic energy assessment (e.g. due to river meandering), the complexities associated with predicting that variability at a regional scale across Canada, precluded its consideration in the resource assessment study (Jenkinson 2010).

Evaluation of the available hydrokinetic energy at a river cross-section was carried out by integrating the power over the cross-sectional area as shown above. To assess the hydrokinetic power potential along a river requires longitudinal integration of the available hydrokinetic power over the entire length of the river. Integration of this type will require assumptions as to energy extraction by installed turbines, and estimates for the allowable spacing between installed turbines along a river length. In the Phase II study, longitudinal integration of the power potential was not considered. However, this was explicitly conducted in the Phase III study.

3 Assessment of Canada’s Hydrokinetic Power Potential: Phase I – Methodology and Data Review

3.1 General

As already mentioned in Chapter 1 of this report, hydrokinetic resource assessment study was conducted in three different phases, i.e. Phase I, Phase II and Phase III. This study was led by the NRC, with financial support from NRCan. The work completed under Phase I of the study was documented in Jenkinson (2010). Most of the information presented and discussed in this chapter is derived from this source. Where applicable and where it was necessary, changes have been incorporated to improve technical/scientific quality of various methods, pertaining to estimation of river flows at ungauged locations, reviewed in Jenkinson (2010). The overall objective of the resource assessment study was to provide high level estimates of hydrokinetic power potential at regional and national levels. However, precise site-specific assessments for power development and equipment installations were not considered for this assessment study. Within the scope of Phase I, the following elements were explored, mainly from a literature review perspective and documenting general observations and guidelines for Phase II and Phase III of the study. Below, an overview of the work completed in Phase I is presented first, followed by additional details of selected components in Sections 3.2 to 3.5 of this chapter.

<u>a) River Flow Estimation Techniques:</u>

As a large part of the Canadian river and stream network is ungauged, reasonable methods from both national and international literature on ungauged hydrology were required to be reviewed, tested and refined for national scale applications. Therefore, a number of selected studies on the estimation of various characteristics of river flows at ungauged locations were reviewed, in addition to methods pertaining to transposition of flow duration curves (FDCs) from gauged to ungauged locations. Perhaps, it was envisioned at the outset of the study that FDCs can provide a reasonable estimate of the time-averaged hydrokinetic potential of a river at a given location. A possible application of a smaller set of reviewed methods in subsequent phases of the study was also envisioned. Though regulated rivers could also exhibit considerable hydrokinetic potential, the techniques on the estimation of streamflow characteristics in regulated rivers were not considered. Also, the impacts of river ice, ice jams or ice cover on the estimation of river flows at ungauged locations were not considered. In addition, the study also precluded the use of deterministic hydrologic models, often used for river flow forecasting and warning purposes, for the estimation of river flows at ungauged locations. Though expensive and time consuming, hydrologic modelling is a reasonable approach for detailed investigation of hydrokinetic resources across a given watershed.

<u>b) Channel Geometry and Slope Estimation Techniques:</u>

Estimation of channel velocity is a critical component for assessing hydrokinetic potential of a river. Consequently, it was envisaged that channel geometry and slope are also critical parameters for obtaining estimates of channel velocity at different locations within a river reach. A review of available techniques, with potential applications on regional basis, was also included in the Phase I report (Jenkinson 2010).

c) Data Uncertainty and Uncertainty Analysis Techniques:

Analysis of uncertainty from various potential sources is generally considered an integral part of almost all data analysis exercises. Therefore, methods related to estimation of uncertainty when using DEM data, streamflow records, roughness estimates, channel geometries, etc., were reviewed and discussed to support subsequent phases of the resource assessment study.

d) Regional Channel Velocity Estimation Studies:

Review of studies wherein investigators have attempted to estimate channel velocities at a regional scale was also included in Phase I of the study. An estimation of average flow velocity within a river cross-section was targeted and the impact of irregular velocity distribution, within the cross-section, on hydrokinetic assessment were not considered.

Additional detail is provided in the following sections, wherein the major focus is on hydrologic considerations (i.e. streamflow estimation techniques) as the objective of this review report is to improve streamflow estimates at ungauged locations across Canada so that more reliable assessment of the hydrokinetic potential can be obtained. Due to this specific focus, hydraulic aspects are not discussed in greater detail in this report.

3.2 Hydrologic Considerations

It is important to note that continuous time-series of streamflow are not typically employed directly for assessing hydrokinetic potential of a river at a given location. However, the frequency with which a specific streamflow is expected to be observed at the target location over a longer period of time is important for resource assessment and selection of suitable power generation equipment. The FDC, which is generally derived from continuous streamflow records, provides a graphical representation of streamflow variability and expected frequencies at a given location (see Figure 2). It is straightforward to derive FDCs at locations with continuous streamflow data. Compared to this, estimation of FDCs at ungauged locations is accomplished by a number of indirect means, including transposition from gauged locations, using empirical or regression-based statistical methods, and through hydrologic modelling. For hydrokinetic power extraction, high flows, which are generally rare, are not so useful but are critical with respect to hydrokinetic turbine design and deployment. The same is the case for extreme low flows with respect to power extraction. However, these flows are not considered critical for deployment of hydrokinetic turbines and associated equipment.

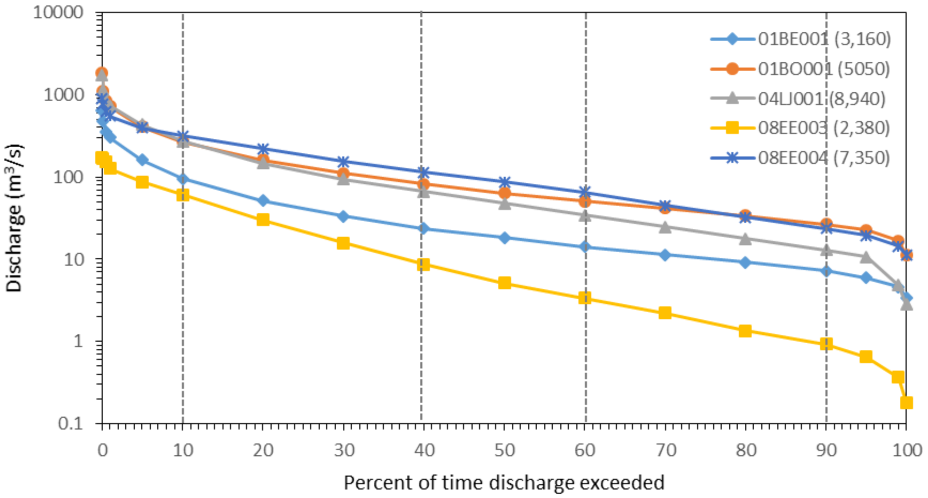

Figure 2: Flow duration curves for five selected streamflow recording stations of Environment Canada (shown in the legend) for the 1961–1990, 30-year period. Arbitrary divisions in terms of different flow regimes are also shown. Drainage area (in km^2) of the respective watershed is listed in the legend.

In determining hydrokinetic potential at the regional or national scale, it is important to have reasonable estimates of streamflow regimes at all channel reaches within the area of interest. Considering the spatial extent of Canada and the amount of data to be processed, estimates of streamflow need to be acquired in an efficient manner, with ideally minimal manual processing. Therefore, some of the techniques that were employed for estimating streamflow characteristics at ungauged locations, specifically within the context of hydrological regionalization approach, were examined in Phase I of the study. Through regionalization approaches, one could estimate desired characteristics of streamflow at all ungauged locations in an efficient and robust manner.

When estimating streamflow characteristics/indices at ungauged locations, two primary aspects are generally considered: (1) the nature of the streamflow characteristic that needs to be estimated at the target ungauged location (e.g. annual or seasonal high flows, percentiles of FDCs, annual or seasonal low flows, etc.), and (2) the selection of a suitable technique for regionalizing streamflow characteristics based on data from gauged locations. This partitioning is necessary since most of the regionalization approaches are tied with the hydrologic variable being estimated at the desired ungauged locations. Furthermore, the above mentioned second step consists of two additional independent steps, i.e. (1) delineation of homogenous regions (DHR), which classifies or groups a number of source data sites that exhibit similarity in terms of some selected features of interest (i.e. climatological, geophysical and statistical characteristics of interest), and (2) regional estimation method (REM), which is employed to transfer the required information from source data sites to the target ungauged site. Below, the DHR and REM steps are elaborated further.

3.2.1 Delineation of Homogeneous Regions

The regionalization framework for estimating streamflow characteristics at ungauged locations is based on the principle/understanding that the sites with recorded data that are more similar to that of the ungauged site within the space of selected attributes are the best possible predictors of streamflow characteristics at the target site and should therefore be included in defining a homogeneous region or neighbourhood. The attributes could be selected from observed streamflow statistics or could be derived from climatic and physiographic characteristics of the associated drainage areas. A mixture of these characteristics is often used. However, the process of identifying homogeneous regions is generally tied with the variable of interest, e.g. high flows or low flows. The regions identified homogeneous for high flow analysis may not be the same as those identified for low flow analysis. The main reason behind this disparity is that the underlying physical mechanisms that govern low flow processes may not be similar to those that govern high flow generating processes within various watersheds of a large geographic area. For the case of FDCs, considerable difficulty arises because the FDC represents the entire flow regime of a gauged watershed, ranging from low flows to high flows, as well as flows in-between these extremes. With respect to FDCs, Dingman (1978) and Searcy (1959) noted that the lower flow ranges of a FDC are controlled less by climactic drivers than by basin geology and physiography, whereas in a runoff-dominated watershed, the local climate would have a significant impact on higher flow ranges of an FDC. Consequently, regionalization of FDCs is not a trivial task. Generally, precipitation amounts and temperature and evaporation patterns can affect river flows on large regional scales, while physical properties of watersheds (i.e. geology, land use, and presence or absence of surface water bodies) can affect river flows on local scales (Homes et al. 2002).

There are many possible approaches that have been used for delineating homogeneous regions. One of the most popular and the easiest to comprehend approach is to delineate geographically contiguous homogeneous regions based on the geographic proximity concept. If an ungauged site falls within a homogeneous geographic region then the characteristics of that region as a whole (from all sites within the region) are used to estimate target streamflow characteristics at the ungauged location. According to Jenkinson (2010), Acres Consulting Services Limited (1984a) is the only study in Canada that identified hydrologic homogeneous regions at the national scale. In this study 12 hydrologic homogeneous regions were identified. This was done by first identifying a number of predefined physiographic regions within Canada and then sub-dividing them by the presence or absence of permafrost and based on the differences in regional climactic parameters (Acres Consulting Services Limited 1984b). Among other regional studies wherein a similar approach was used is Gingras et al. (1994), who identified nine homogeneous regions in Ontario and Quebec based on statistical characteristics of flood flows. In a similar manner, the Ontario Ministry of the Environment delineated homogeneous regions within the province of Ontario based on various characteristics of low and high flows (Chang et al. 2002). For instance, to predict low flows at ungauged locations, six different regions were delineated within the province, while for predicting high flows, 12 different regions were delineated. It is important to point out that there were a number of other provincial and regional studies (e.g. Moin and Shaw 1985, 1986; Loukas and Quick 1995; Wang 2000; Eaton et al. 2002, etc.) which were not reviewed in the Phase I report (i.e. Jenkinson 2010). This might had been due to time or project constraints.

Cluster analysis technique has been commonly used for delineating homogeneous regions (e.g. Leboutillier and Waylen 1993, Nathan and McMahon 1990, Tasker 1982). This technique identifies different groups or clusters of sites/stations based on similarity of statistical, climatic, geophysical and hydrologic attributes. For the case of FDCs, clustering is generally based on similarity of geophysical attributes. Once homogeneous regions are identified, desired characteristics of streamflow at ungauged locations within the identified homogeneous regions can be estimated. For marginal cases, when ungauged locations are suspected to belong to more than one cluster, a weighting scheme is generally adopted. In the literature, this approach is also referred to as fractional membership technique (Acreman and Wiltshire 1989). Based on the number of studies conducted (e.g. Acres Consulting Services Limited 1984a, Leboutillier and Waylen 1993, Natural Resources Canada 2004a, 2004b, Ottawa Engineering Limited 1997, and Tasker 1982), the use of geographically contiguous homogeneous regions seems to be the most popular approach. In certain circumstances, especially when the number of gauging stations is very limited, delineation of non-contiguous homogeneous regions is also common in the hydrologic literature. Two such techniques that have been employed in Canada are described below.

The Region of Influence (ROI) Approach

The ROI approach is used to identify homogeneous regions that are not necessarily geographically contiguous. The sites included in such regions share similarity within the space of selected hydrologic, climatic or geophysical attributes. Each station or site is assumed to be associated with a specific region. This technique was proposed originally by Acreman and Wiltshire (1989) for the UK, but has been used in many other parts of the world, including Canada (Burn 1990a, 1990b). Though not necessary, a weighting function is used to weight individual stations/sites depending upon a similarity/dissimilarity measure in the form of Euclidian distance, calculated for a set of attributes within the attribute space. A biggest advantage of the ROI approach is that it allows formation of homogeneous regions that can contain a large number of sites and that in turn can be useful to obtain robust estimates of low frequency quantiles, which are generally associated with high uncertainty. If a geographic homogeneous region contains a smaller number of stations/sites then the ROI approach can be useful to expand the number of neighbouring sites for that region and to obtain relatively more reliable estimates of low frequency quantiles. Tasker et al. (1996) and many others (e.g. Burn 1990a, 1990b) have used this approach for regional frequency analyses. Holmes et al. (2002) used the ROI approach to estimate FDCs at ungauged locations.

Canonical Correlation Analysis (CCA)

The CCA approach is a multivariate statistical technique that permits establishment of interrelationships between two groups of variables by determining linear combinations of one group that are most correlated to linear combinations of the second group. The CCA technique has been employed as a regionalization method for flood frequency analysis, where one group of variables represents flood characteristics and the second group represents physical and climatological characteristics of watersheds. The principle being that by knowing the second the first can be predicted (Bobée et al. 1996). For a given gauging station/site, a homogeneous region or a group of stations can be identified by examining the proximity of the site to other gauging sites within the canonical space of attributes. A chi-squared distance measure is used to identify neighbouring sites for each ungauged location. This procedure has been applied for flood frequency analysis in Quebec (Ribeiro-Correa et al. 1995) and Ontario (Ouarda et al. 2001). The applications of this method for determining FDCs at ungauged locations are relatively limited in the literature.

3.2.2 Regional Estimation Methods (REMs)

A number of REMs from the literature that had shown some promise for estimating FDCs at ungauged locations were reviewed in Jenkinson (2010). Some of these methods are described below.

Index Flood Method:

This method was proposed by Dalrymple (1963) for regional flood frequency analysis, but can also be used for other variables of interest. The principle is that the at-site flood frequency curves in a homogenous hydrologic region are identical, except a scale factor that can be described in terms of watershed characteristics (e.g. climatic, geophysical or other characteristics). First, this method requires identification of homogenous regions and then determining a standardized (or normalized) flood frequency curve, commonly known as ‘growth curve’, for each homogeneous region. In the original application of this method, delineation of geographic regions and pooling of standardized flood flows to derive the regional growth curve were considered. Derivation of site-specific flood flow indices can be described by the following relationship:

$$Q_k(T) = \mu_k Q^*(T) \tag{5}$$

where $Q^*(T)$ represents the regional growth factor corresponding to return frequency T (also known as return period or return interval), μ_k is the scale factor and $Q_k(T)$ is the estimated flood magnitude at site k. This concept from regional flood frequency analysis was borrowed by some investigators for determining FDCs at ungauged locations.

Acres Consulting Services Limited (1984a) employed this concept and normalized FDCs using the 2-year return flow that needed to be estimated at the target point of interest. The technique employed in the RETScreen (Natural Resources Canada 1984a, 1984b) application of NRCan also uses a variation of this approach. In this application, a representative FDC was determined for a region of interest (see Figure 3) and normalized using the mean annual flow. The mean annual flow was estimated at the target location by employing specific runoff values from published maps (see Figure 4) and by calculating the related drainage area at the target location.

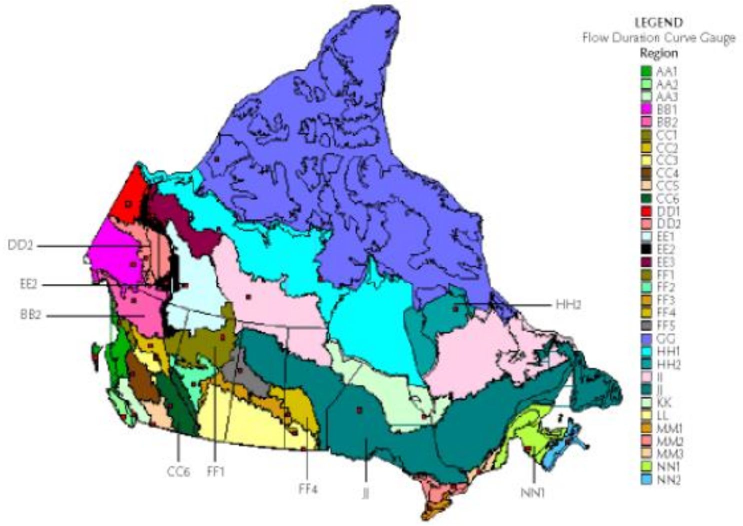

Figure 3: Map of flow duration curves employed in RETScreen decision-support system. Source: NRCan (2004a, 2004b); Jenkinson (2012).

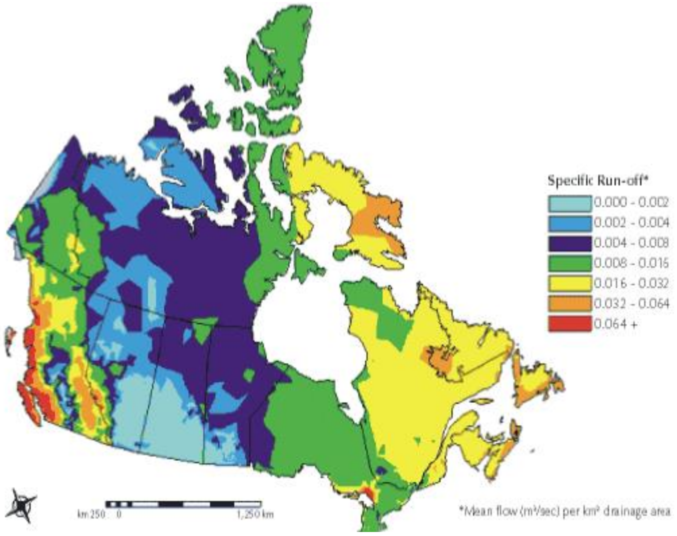

Figure 4: Map of specific runoff employed in RETScreen decision-support system. Source: NRCan (2004a, 2004b); Jenkinson (2012).

Some variants of this approach were also implemented in Smakhtin et al. (1997) and Smakhtin and Masse (2002), wherein the normalization was conducted using the mean annual flow.

Drainage Area Ratio Method:

This is one of the simplest methods for estimating streamflow at ungauged locations. In this method the streamflow from a source site is scaled based on the ratio of drainage areas, as shown below:

$$Q_u = Q_g \left(\frac{A_u}{A_g}\right)^m \tag{6}$$

where Q_u and Q_g are respectively the streamflow values for the ungauged and gauged locations, A_u and A_g are respectively the upstream drainage areas at the ungauged and gauged locations, and m is a calibration parameter that accounts for the non-linearity of the relationship. The parameter m requires calibration, but is often taken as unity for simplicity reasons (NRC-CHC 2008, Shu and Ouarda 2012). Using this method, a complete FDC curve can be generated at an ungauged site either by estimating continuous streamflow data or by estimating selected percentiles of the FDC. However, caution is generally warranted when using this method as the relationship between streamflow and drainage area is affected by a number of physiographic, climatic and other factors. The strength of the relationship drops off quickly as the drainage area ratio diverges significantly from unity (Copeland et al. 2000, McCuen and Levy 2000). A number of studies have employed this method (e.g. Mohamoud and Parmar 2006, Gulliver and Murdock 1996, NRC-CHC 2008b). Drainage area differences of more than 25–50% have been considered as the applicability limits of this method (Durand 2002, McCuen and Levy 2000). Mohamoud and Parman (2006) proposed modified drainage area ratio methods for the US Mid-Atlantic region, but their modifications seem to have very limited applications as they have never been evaluated in later studies.

Parametric Characterization of FDCs:

In parametric characterization of FDCs, FDC is assumed to be represented by analytical relationships. These relationships could be in the form of polynomial or exponential relationships. The parameters of the relationship are estimated through regional analyses. One of the first FDC regionalization and transposition studies was conducted by Quimpo (1983), who was the first to propose parametric characterization of FDCs. Though the approach is parsimonious in nature, it makes the FDC inflexible due to constraining the shape of the FDC. Applications of this method are rare in the literature. Franchini and Suppo (1996) also proposed a parametric technique for estimating the FDC by fitting a curve to three selected quantiles of the FDC. The authors proposed two possible relationships for describing the lower portion of the FDC. This technique was later extended by Castellarin et al. (2004), who considered four quantiles instead of three. This method regionalizes streamflow quantiles rather than parameters as a function of watershed attributes.

Statistical Characterization of FDCs:

Statistical characterization of FDCs involves describing the FDC in the form of a probability distribution function. Leboutillier and Waylen (1993) conducted a streamflow regionalization study in British Columbia by fitting a two-component, two-parameter lognormal mixture distribution to the FDC, resulting in a five parameter FDC. The values of each of the five parameters were clustered into seven regional clusters using a two-stage density linkage cluster analysis. The generated clusters showed distinct contiguous regions within the province of British Columbia. Averaged parameter values were then determined and representative FDCs were generated for each of the identified regions. Predictive capabilities of this method were not investigated in the literature as indicated by Jenkinson (2010).

Graphical Characterization of FDCs:

This technique was proposed by Smakhtin et al. (1997) and was further developed in Castellarin et al. (2004), as a “graphical” FDC transposition method. In this method, FDCs from gauged sites were normalized using an index flow and the regional FDC was determined by averaging percentiles of the normalized FDCs within the region. The index flow values at the ungauged sites were estimated using a linear regression technique. Shu and Ouarda (2012) stated that the distinctive characteristic of this technique was that the method made no assumptions about the shape of the FDC as is often done in parametric and statistical distribution based methods. The entire FDC at the ungauged site was derived from observed FDCs at other locations. This technique is advantageous if the entire FDC is required at the site of interest, or if a specific region of the FDC is required that cannot be easily represented by other methods. Along similar lines, Mohamoud (2008) estimated 15 percentiles of the FDC by employing step-wise regression and grouping these percentiles into low, median, and high flow ranges, with five percentiles in each, and determining unique predictors for each of the three ranges. The selection of source sites was based on pre-defined landscape classifications. Shu and Ouarda (2012) expanded these techniques, proposed by Smakhtin (1997) and Mohamoud (2008), and considered 17 different percentiles to represent the FDC. A separate regression relationship was developed for each of the 17 percentiles by employing a step-wise regression analysis and using climatic and watershed characteristics. FDCs at ungauged locations were estimated using various distance weighting schemes and employing area, positional, physiographic and climatic data from multiple sites. Logarithmic interpolation technique was used for obtaining values lying in-between any two predicted percentiles, where required. The authors found that the FDC technique outperformed area ratio method and that the inclusion of multiple source sites consistently improved predictive capability of the method.

Other Methods:

A few other methods have also been proposed for estimating FDCs at ungauged locations and these methods were found to be similar to regional estimation techniques. For example, non-linear spatial interpolation technique of Hughes and Smakhtin (1996). These authors developed a non-linear technique for infilling missing data at proximal gauges, and generating continuous streamflow time series using the FDC as a transfer function. Although this technique was developed primarily to fill-in missing data, the procedure resembles to that of FDC transposition at ungauged locations. Monthly FDCs using daily streamflow data for each calendar month for both the target and source stations were employed for constructing streamflow time series on a monthly basis. Smakhtin (1999) and Smakhtin and Masse (2000) also developed a technique for ungauged locations where the FDC at the target site was unknown. These authors suggested to normalize FDCs using an index flow, and then determining the target FDC, along with the index flow at the target location. The FDC at the source site was represented as a discharge table for fixed percentage points and then data between points was interpolated using a logarithmic interpolation technique. The source sites were weighted based on similarity of the source sites to the target location. The authors recommended that up to five sites could be used as source sites. Normalized FDCs created for ungauged locations were then de-normalized with the respective index flow determined through regional regression analyses. The authors also suggested that one should avoid direct use of drainage area and preferred the use of mean annual flow in the regression analyses. The authors also suggested 20 to 25 years of data as being adequate for applying this method. This method has also been reviewed favourably by Metcalfe et al. (2005) for generating flow regimes at ungauged locations in Ontario.

Inspired by Smakhtin (1999) and Smakhtin and Masse (2000), Shu and Ouarda (2012) suggested Regression Based Logarithmic Interpolation (RBLI) technique for generating FDCs at ungauged locations. They transposed 17 different percentiles of the FDC from source sites to ungauged locations using multiple regression based on climatic and physiographical attributes, but without normalizing the FDC as was the case in Smakhtin (1999) and Smakhtin and Masse (2000). Like the studies of Smakhtin (1999) and Smakhtin and Masse (2000), in-between percentiles of the FDCs were obtained through a logarithmic interpolation technique, when generating continuous time series of streamflow at ungauged locations. The authors evaluated the effect of considering single and multiple source sites in the RBLI approach and variants of the area ratio method for transposition of FDCs using data from Quebec. The RBLI method performed better than the area ratio method and multiple source sites option was found to show substantial improvement over the single source site option in most cases. For the case of multiple source sites, geographic distance based weighting scheme was found to perform better compared to the weighting scheme based on physiographic attributes.

3.3 Data Sources

To support Phase II of the study, a number of datasets were described in the Phase I report by Jenkinson (2010), along with recommendations on the use of these data sources. Following sections are compiled based on the information provided in that report. Where appropriate, additional information pertaining to various data sources is also discussed and the references to these data sources are kept the same as reported originally in Jenkinson (2010). However, new references may have become available overtime.

3.3.1 National Hydro Network (NHN)

The NHN is a Geographical Information System (GIS) product that includes a geometric description and a set of basic attributes that describe Canada's inland surface waters. It provides information on hydrographic features such as lakes, reservoirs, rivers, streams, canals, islands, obstacles (e.g. waterfalls, rapids, and rocks in water) and constructed elements (e.g. dams, wharves, and dikes), as well as a linear drainage network and toponymic information (i.e. geographical names) associated with hydrography. This product was available in GML (Geography Mark-up Language) and ESRI shape file formats (Canadian Council on Geomatics 2009). At the time, it was anticipated that this product can support estimation of river widths at gauging sites and that, in turn, can support estimation of river flow velocities to develop velocity duration curves.

3.3.2 HYDAT

HYDAT is Environment Canada’s (EC) database that contains information on river flow magnitudes, water levels and sediment concentrations for over 2,500 active and 5,500 discontinued hydrometric monitoring stations, located across Canada (Environment Canada 2004). This database also provides information on streamflow statistics, such as mean monthly and annual flows and extreme flow values and their dates of occurrences. This database is available online, and is also available through ECDE (Environment Canada Data Explorer) desktop application. It was anticipated that the HYDAT database will be used to develop site-specific FDCs and to validate selected methodologies in Phase II of the study.

3.3.3 Environment Canada Measurement Database

The Environment Canada’s measurement database (EC-MDB), which is maintained by Water Survey of Canada (WSC), is an MS-Access database developed for internal use by EC and is a repository of hydrometric field measurements. This database includes information on river cross sections, channel geometry and measured velocities for all completed field surveys across Canada. Such information for the province of Quebec is not available in this database. It was anticipated that this database will be useful for calibration and validation of regional techniques for estimating channel geometry related elements and for validation of velocity duration curves.

3.3.4 Canadian Digital Elevation Data (CDED)

The CDED consists of an array of ground elevations or digital elevation model (DEM) extracted from the National Topographic Database and other data sources from provinces and territories (Centre for Topographic Information 2000). The geographic resolution is 0.0001 decimal seconds (maximum 93 m pixel resolution) for the entire country. This data can be used to delineate drainage areas at gauging stations and for determining associated slopes for regional studies.

3.3.5 Soil and Land Dataset

Agriculture and Agri-Food Canada (AAFC), through Soil Landscapes of Canada Working Group and the Canadian Soil Information System (CanSIS), created a series of GIS maps that show major characteristics of soil and land for the whole country. Soil Landscapes of Canada (SLCs) were compiled at a scale of 1:1 million, and the information was organized according to a uniform national set of soil and landscape criteria based on permanent natural attributes (NLWIS 2008).

3.3.6 Canadian Climate Data

A number of climatic datasets were discussed in Jenkinson (2010) as potential sources for climatic variables required for estimating streamflow characteristics at ungauged locations and for identifying homogeneous regions for the same. Many of these datasets are discussed below.

(a) Canadian Daily Climate Data (CDCD) consists of daily temperature, precipitation, and snow depth, recorded at over 6900 active or inactive meteorological stations across Canada. It is available as part of EC's National Climate Data and Information Archive (Environment Canada 2006).

(b) AAFC has also produced a national ecological framework for Canada and introduced national ecodistricts with associated climate normals based on 1961 to 1990 data (AAFC 2008). The ecodistricts were delineated in ESRI shape files and included information on mean annual rainfall; mean annual snowfall; mean annual precipitation; average, minimum and maximum average daily temperatures; potential evapotranspiration; growing degree days; and growing season start and end dates. The biggest shortcoming of this source was that this dataset was available only for the 1961 to 1990 period at the time of the study and therefore deriving the same information for the later normal periods (e.g. 1990 to 2010) from other sources was a considerable challenge.

(c) AAFC, in collaboration with NRCan, EC and the Australian National University, also developed a 10 km gridded daily climate dataset (i.e. daily maximum temperature, daily minimum temperature and daily precipitation amount) for Canada south of the 60° North (NLWIS 2008). These grids, available in two file formats (text and GeoTIFF), were interpolated from daily EC climate station observations using a thin plate smoothing spline surface fitting method implemented within ANUSPLIN V4.3 (Hutchinson 2009).

(d) A national mean annual runoff map was produced by the Department of Fisheries and Oceans (DFO) and EC's Inland Waters Directorate in 1978 as part of the Hydrology Atlas of Canada. This map is available online through NRCan's National Atlas website. A similar map is also available through the RETScreen decision-support system of NRCan (see Figure 4; Natural Resources Canada 2004a, 2004b). The database associated with RETScreen also includes representative FDCs for various large hydrological regions of Canada (see Figure 3).

(e) EC has also developed a map of the annual mean total precipitation for the 1971 to 2000 period. This map represents average precipitation conditions across Canada.

3.4 An Overview of Selected Hydrokinetic and Hydropower Resource Assessment Studies

Many hydropower resource assessment studies generally seek to estimate the availability of potential hydropower resources to inform preliminary decision-making and later to support various phases of design and deployment. This can be done for hydropower projects of various sizes and types, but most often it has been conducted for large or medium size hydropower, low-head and run-of-river type systems. These systems involve construction of barrages or penstocks to generate an artificial hydraulic head which drives a pressurized turbine. For these systems the necessary elements that are required for assessment are the mean annual flow or the flow frequency and the hydraulic head that can be produced at the project site. With these variables, as well as assumptions about diversion or turbine efficiency, an estimate of the resource can be made at all target locations. In contrast, hydrokinetic resource assessment studies require information on river flow velocities and the frequency with which they occur (in the form of velocity duration curve) at all locations within a region of interest. However, estimation of velocity can be difficult in river systems due to difficulties involved in the estimation of reliable values of river flow, channel geometry, roughness, and slope at all locations, specifically when these locations are ungauged. In order to guide the path forward for Phase II and Phase III of the study, a number of selected hydrokinetic and hydropower related studies were reviewed in the report by Jenkinson (2010). In that report, findings of each of the reviewed studies were summarized and the approaches taken to estimate river flow and velocity, and cross-sectional geometry, roughness, and slope were also discussed. The following studies (selected only) were considered by Jenkinson (2010).

- The first Canada-wide study by UMA Group (1980) on hydrokinetic energy assessment. In addition to hydrokinetic potential in selected rivers, this study also examined tidal power potential at some Canadian coastal locations.
- The study by Miller et al. (1986) on the assessment of hydrokinetic resource potential in the United States, completed for the US Department of Energy.
- NRC-CHC (2008b) study on the development of a methodology for assessing the hydrokinetic energy contained in Canadian rivers. Like many subsequent studies on the same topic, this study was also sponsored by NRCan.
- The study by Acres (Acres 1984a), who developed a pre-feasibility methodology for assessing hydropower resources in Canada.

- The study by Tudor Engineering (1991), completed for the World Bank, Industry and Energy Department. This study provided a methodology for rapid and accurate assessment of the number, size, cost and economic feasibility of small hydro projects.
- Natural Resources Canada's RETScreen clean energy project (NRCan 2004a, 2004b).
- The study by the US Department of Energy to examine hydropower energy resources across the United States for small- and low-power hydro projects (USDOE 2004, 2006).
- The study by Kerr Wood Leidel Associates, focused on the development of a GIS based hydropower assessment system for BC Hydro and the BC Transmission Corporation – called the Rapid Hydropower Assessment Model (KWL 2008, Monk et al. 2009).
- An automatic power assessment study by Rojanamon et al. (2009), wherein the authors considered a combination of hydrology, economic, environmental and social factors when deciding the location and potential run-of-river resources as specific locations.

3.5 Recommendations for Hydrologic Investigations

After reviewing a number of selected studies and available hydrologic, physiographic and other data sources, a number of recommendations for conducting Phase II of the study were made in Jenkinson (2010). The recommendations related to hydrologic investigations are discussed here and the reader is referred to Jenkinson (2010) for recommendations on hydraulic aspects of the study.

A regionalization approach was recommended for both hydrologic and hydraulic aspects of the study. Though not explicitly described, canonical correlation analysis (CCA) was selected to support identification of homogeneous regions or neighbourhoods. This was decided due to the reason that the CCA is a multivariate approach and it allows for prediction of multiple variables that was also the target of the resource assessment study. It was anticipated that a number of regions will be selected based on hydrologic and physiographic considerations and the recommended methodologies will be validated on these regions. A research-based approach was favoured for hydrokinetic resource assessment, but nothing was indicated about the specific aspects that will be researched and where the effort will be concentrated for innovating suggested approaches. The following techniques were recommended in Jenkinson (2010) to carry out Phase II of the study:

- Estimation of MAF based on the methodology associated with the RETScreen framework of NRCan (Natural Resources Canada 2004a, 2004b) and using the CCA-based delineation of homogeneous regions. Estimation of hydrokinetic resource potential based on estimates of MAF was used previously in some studies (e.g. UMA Group 1980, USDOE 2006 and KWL 2008). This method was suggested because it is quite useful in deriving a quick estimate of the power potential.

- Estimation of FDCs at ungauged locations based on the RETScreen framework of NRCan (Natural Resources Canada 2004a, 2004b). This framework was used previously for national resource assessment studies. The underlying procedures involve estimation of FDCs at ungauged locations. Application of these methods does not require any new data to be developed because the required data sources have already been integrated in the RETScreen framework. It was suggested to validate this procedure based on HYDAT stations. It was also suggested that the RETScreen method can act as a benchmark for comparison purposes.
- Estimation of FDCs at ungauged locations based on the area ratio method as described in Shu and Ouarda (2012). The area ratio method represents a simple and quick approach for estimating streamflow at ungauged locations and it has also been used for transposition of FDCs in some previous studies (e.g. NRC-CHC 2008b). Though it was mentioned that this method will be used within the CCA-based regionalization approach that has already been used for estimating streamflow at ungauged locations in Upper Great Lakes region for the International Joint Commission, no valid reference or source was provided in the report to support this assertion.
- CCA-based transposition of graphical FDCs: Transposition of graphical FDCs was developed by Hughes and Smakhtin (1996) for generating streamflow sequences at ungauged locations. The same method was adapted by Metcalfe (2005) for transposition of FDCs and was developed further in Shu and Ouarda (2012). CCA-based regionalization approach for transposition of graphical FDCs, based on selected 17 percentiles of the FDC, was recommended for ungauged locations across Canada.

3.6 Concluding Remarks

Most of the information presented in this chapter was derived from Jenkinson (2010) and modified to overcome some scientific/technical shortcomings, where applicable. Furthermore, no attempt was made to add additional literature review to strengthen what was originally presented and discussed in Jenkinson (2010). A number of new regional studies have emerged since the completion of this work.

The goal of Phase I of the study (Jenkinson 2010) was to: (1) review selected studies and document available methodologies that could be employed for determining Canada’s hydrokinetic potential; (2) identify available hydrologic, physiographic and other related data sources; and (3) make recommendations for conducting Phase II of the study.

For estimating hydrokinetic potential at a given location within a river reach, it was suggested to use Manning’s equation for estimating channel velocity and that requires estimates of channel geometry, river flow (in the form of a FDC or MAF), channel slope and roughness. Hydrologic regionalization or other suitable approaches were suggested for estimating most of the variables required for hydrokinetic resource assessment.

Based on the reviewed literature, it was found that numerous techniques exist for regionalization of extreme river flow conditions (e.g. high and low flows), but fewer techniques exist for regionalization of FDCs. The study also identified that graphical characterization of FDCs by employing CCA within the regionalization framework as a promising approach for Phase II of the study. However, no specific study was reported wherein this potential combination was explored before. For estimating channel geometry related elements (e.g. channel width and depth) and slopes on regional basis, only a few techniques were found in the literature. Among the available studies, discharge-based estimates were quite common, however, other approaches that relate channel geometry to drainage area and physiographic and climatological characteristics of a given watershed were also noted. Estimates of channel slope at the regional scale were generally obtained from the DEM data and occasionally through power-function smoothing in low-gradient streams. Similar approaches for estimating channel roughness on regional basis were not found in the literature. Typically, the roughness values were merely assumed based on the published literature in most of the reviewed studies.

A number of national datasets were identified to support estimation of required variables for assessing national hydrokinetic resource potential. This was a significant contribution of Phase I of the study (Jenkinson 2010). Identified datasets included: climate data, hydrometric data, digital soil and land use maps, hydro network maps and digital elevation data. A database that contains cross-sectional information, and estimates of streamflow and velocity at all surveyed stations across Canada was also a significant finding. This database, which is regularly maintained by WSC, was expected to be useful for validating various selected techniques for estimating velocity duration curves and channel geometry.

For estimating streamflow characteristics (e.g. selected indices of FDC and MAF) at ungauged locations, the study recommended to use some simple methods (e.g. the area ratio method), transposition of graphical FDCs and the methods that are integrated with the RETScreen decision support system of NRCan. The study supported implementation of the area ratio method and the graphical FDC estimation technique within a regionalization context, in combination with the CCA approach. The use of physiographic and climactic datasets discussed in Section 3.3 to identify homogeneous regions and to drive regional regression relationships were recommended. In the absence of regional approaches for channel roughness, the study suggested using a range of values from the published literature.

Estimation of uncertainty was also given a significant importance. In most cases, the study recommended the use of jack-knife and bootstrap techniques for uncertainty analyses.

In summary, the report by Jenkinson (2010) made a comprehensive effort to review various techniques available in the literature and focused mainly on how to employ those techniques at regional levels by identifying various baseline hydrologic, physiographic, and climatic datasets to assess hydrokinetic resource potential across Canada. However, no attempt was made to explore and discuss how the identified techniques can be innovated or modified to support better estimates of various indices of streamflow at ungauged locations compared to the efforts made in previous studies. This was an obvious shortcoming of the study. The area ratio method was one of the recommended methods, but it has certain limitations. For example, this method is specifically suitable for ungauged locations within the same watershed and not so when applied across multiple watersheds in a larger geographic region. For identifying homogeneous regions, the focus was kept on climatic and physiographic parameters and the CCA approach. Many popular statistical approaches (e.g. L-moments based approach of Hosking and Wallis 1997) were not considered at all. Many new approaches for identifying contiguous and non-contiguous hydrologic homogeneous regions have been published in the literature. Canadian climatic regions can also be considered to derive large contiguous homogeneous regions. Many of these approaches can be explored in future studies for the estimation of streamflow indices at ungauged locations. Therefore, a more comprehensive review is needed to devise new methods for estimating FDCs at ungauged locations. The ROI approach was mentioned but was not included in the recommended approaches. The ROI approach is as good as the CCA approach for identifying non-contiguous homogeneous regions or neighbourhoods. A blended approach that can exploit strengths of both the ROI and the CCA approaches could be useful for improving estimates of streamflow indices, including FDCs, at ungauged locations across Canada.

4 Assessment of Canada's Hydrokinetic Power Potential: Phase II – Development of Datasets

4.1 General

The work conducted within the scope of Phase II of the hydrokinetic resource assessment study was documented in Jenkinson and Bomhof (2012). Most of the information provided in this Chapter is derived mainly from this source and modified where necessary. First an overview is presented here, followed by additional information in Sections 4.2 to 4.7.

Canada has a large network of rivers and streams, but the majority of this network is ungauged. In order to estimate hydrokinetic potential at ungauged river reaches a number of regionalization approaches were evaluated in this phase of the resource assessment study. Following the recommendations from Phase I, flow duration curves (FDC) were employed for this purpose. In addition to FDCs, mean monthly flows (MMFs) were also employed for comparison purposes, although this was not included in the original scope of the study, envisioned in the Phase I report. For validating the proposed methodology from Phase I, six large geographical areas were selected that represented different physiographic regions of Canada. For generating FDCs and MMFs at selected stations within each of the selected study regions, recorded streamflow data was obtained from Water Survey of Canada's (WSC) HYDAT database. Channel geometry data obtained from Environment Canada's Measurement Database (EC-MDB) were analyzed to produce estimates of channel bank-full widths and depths for locations where this data was available. This data was also used to estimate hydrokinetic potential at HYDAT gauging stations within each study region to validate adopted methodologies.

To drive hydrologic regionalization approaches, a database of physiographic and climatic attributes of watersheds upstream of the WSC gauging locations was also developed. For the development of this database, a number of datasets were processed including elevation, land cover, soils types, and climatological data. These datasets have already been described in Chapter 3. Hydrologic regionalization approaches included: (1) the RETScreen technique, (2) multiple regression (MR) alone, and (3) multiple regression in combination with canonical correlation analysis (MR-CCA). These approaches were employed to estimate FDCs and MMFs at target sites. For estimating channel velocity at desired locations, investigations pertaining to estimation of channel velocity using both hydraulic and geometric relationships were conducted, wherein estimates of water depth, channel width and slope, and a measure of roughness were also considered. For hydrokinetic power estimates, two approaches were evaluated, i.e. (1) direct regionalization of hydrokinetic potential and (2) using FDCs and MMFs in conjunction with channel geometry and continuity and Manning's equations.

Additional detail on the above analyses, investigations and data development procedures is provided in the following sections. Again, as for Chapter 3, the focus in this chapter is also on hydrologic investigations and all elements related to hydraulic modelling and analyses are addressed minimally. For information on these elements, the reader is referred to Jenkinson and Bomhof (2012).

4.2 Study Regions

In order to test the methodology, proposed in Phase I, a number of study regions was identified across Canada. Two products were used to delineate these regions: Canadian ecoregions identified by Agriculture and Agri-Foods Canada (AAFC) (AAFC 2008) and the Canadian hydrogeological regions (Sharpe et al. 2008) identified by Canadian Geological Survey (CGS). These regions are respectively shown in Figure 5 and Figure 6. Intersection of these two products resulted in a number of other new regions. Of these regions a sub-set of six regions was selected based on the number of available hydrometric stations. The selected regions are shown in Figure 7. These regions were named as: (1) British Columbia, (2) Canadian Shield, (3) Prairies, (4) North, (5) Southern Ontario, and (6) Maritimes. The naming convention partially retains the original naming conventions used in AAFC (2008) and Sharpe et al. (2008). These regions broadly represent the hydrologic variability observed within Canada.

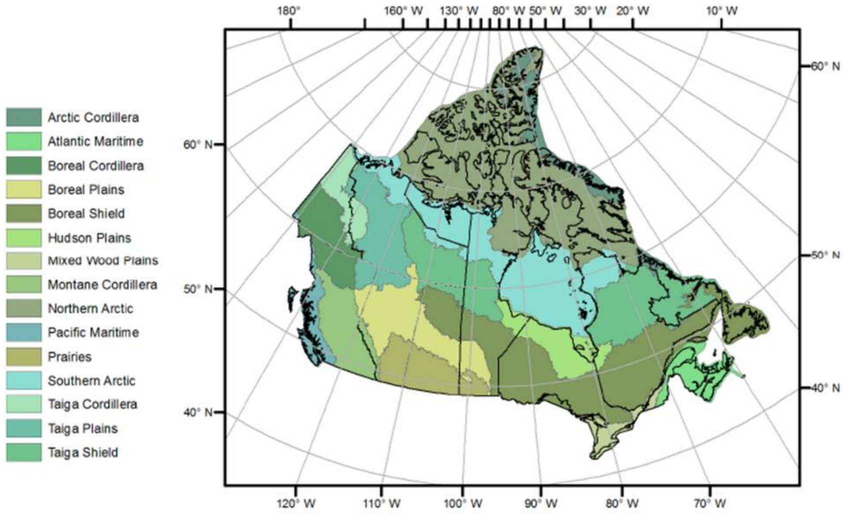

Figure 5: Canadian ecoregions identified by Agriculture and Agri-Food Canada (AAFC 2008).

The study watersheds within four of these regions (i.e. British Columbia, Canadian Shield, Prairies, and Maritimes) were delineated using Digital Elevation Model (DEM) driven techniques, while in the remaining two (i.e. North and Southern Ontario) Voronoi polygon-driven techniques were used. For each study region, hydrometric stations with upstream drainage areas located within the boundary of each region were identified first and then a subset of stations was considered based on the following criteria: (1) stations with minimum 10 years of continuous data, (2) currently active stations, and (3) unregulated (representing near natural flow conditions) stations. However, it was not clear what parameters were used to satisfy the first and second requirements (e.g. the starting and ending years).

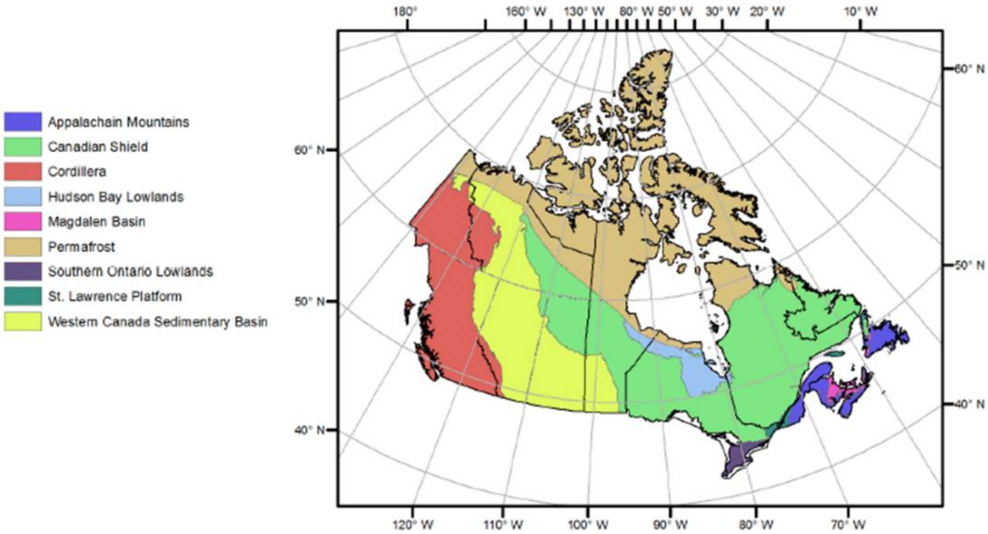

Figure 6: Canadian hydrogeologic regions identified by the Canadian Geologic Survey (Sharpe et al. 2008).

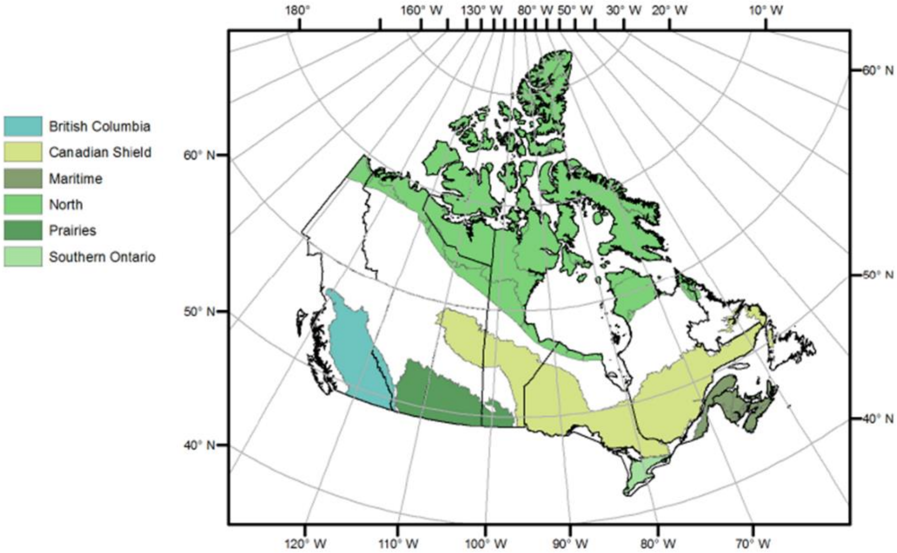

Figure 7: Study regions identified for validating the methodology, proposed in Phase I of the resource assessment study. Source: Jenkinson and Bomhof (2012).

4.3 Streamflow Indices and Hydrokinetic Dataset

For estimating hydrokinetic potential at a given location within a river reach, information on streamflow regimes over a longer period of time is generally required. Therefore, for each watershed within the identified study regions, Jenkinson and Bomhof (2012) processed continuous streamflow records from Environment Canada’s (EC) hydrometric stations to generate the following three indices: MAF, MMF and FDC. MMFs were not included in the list of identified indices at the time of Phase I study. The decision of considering MMFs for estimating hydrokinetic potential was made later during the Phase II study.

The values of MAF were obtained directly from EC’s HYDAT database. The MMF values are also available in HYDAT, but were calculated by taking the average of all streamflow values within each month of the year for the period of available record. FDCs were calculated using the procedure outlined by Searcy (1959). According to this procedure, data is sorted in a descending order and probabilities are assigned based on the Weibull plotting position formula. The same procedure has also been described in many text books on Applied Hydrology (e.g. Shaw 2004). Following the suggestions by Smakhtin et al. (1997) as discussed in Chapter 3, 17 discrete exceedance probability point values were selected to characterize the FDC (referred to herein as percentiles of the FDC). Where necessary, linear interpolation between the two closest points, on either side of the target point, were used. The considered exceedance probabilities (in percent) were: 0.01, 0.1, 0.5, 1.0, 5.0, 10.0, 20.0, 30.0, 40.0, 50.0, 60.0, 70.0, 80.0, 90.0, 95.0, 99.0, and 99.99.

In order to compare results derived on the basis of MAF, MMF and FDC, MMFs were averaged to generate an equivalent MAF and in a similar manner, FDCs were integrated to produce equivalent MAFs. It was pointed out that with the use of 17-point FDC and 12 MMFs, some data were lost. Therefore, a separate investigation was carried out to determine the level of differences and then the best possible integration method was used. The authors came up with using the Piecewise Cubic Hermite Interpolating Polynomial procedure (MathWorks 2011). It will be seen later that this procedure was not retained for Phase III of the resource assessment study.

To validate results from various regionalization approaches, estimates of hydrokinetic energy at all hydrometric stations within the four selected study regions (i.e. British Columbia, Canadian Shield, Prairies and Maritimes) were developed. The primary task involved in this estimation was the development of power duration curves (PDCs) in order to characterize total hydrokinetic energy. PDCs were developed by equating flow rates to equivalent hydrokinetic potential. For this purpose, separate relationships were developed to characterize velocity and cross-sectional area as a function of discharge based on EC-MDB. PDCs were numerically integrated to obtain time-averaged power. Additional detail on this topic can be found in Appendix B of Jenkinson and Bomhof (2012). Graphical visuals demonstrating spatial patterns and total amount of available hydrokinetic power for each of the considered study regions, are available in Jenkinson and Bomhof (2012). British Columbia was associated with the highest hydrokinetic potential, followed by Maritimes. Canadian Shield and Prairies were associated with much less hydrokinetic potential. Here in this report, it is sufficient to mention that these estimates were obtained to develop hydrokinetic dataset in order to validate results obtained from regionalization approaches.

The influence of considering FDC and MMF profiles on estimates of hydrokinetic potential was demonstrated by comparing equivalent power calculated using the MAF and the fully integrated profiles of the FDC and that of the MMF for 1063 stations from the four selected study regions. It was shown that the choice of MAF grossly underestimates the available power when compared to the choice of FDC. On average, integration of MMF profile showed a 69% increase in the estimated power at a location when compared to the MAF. Similarly, integration of FDC resulted in 200% increase in the available hydrokinetic power compared to the MAF. This investigation demonstrated that those techniques, which depend on MAFs for estimating power potential, are likely to utterly underestimate the available power. This is due to the non-linear nature of the streamflow-power relationship, as the higher flows produce more power than the lower flows. Thus, the hydrokinetic power using a mean flow value would under-represent the time-averaged hydrokinetic power available at a given location in a river reach.

4.4 Physiographic Database Development

In order to estimate streamflow indices and then the hydrokinetic potential at all stream reaches through regionalization approaches, a database of watershed characteristics (i.e. attributes reflecting soil type, geology, landforms, climate, etc.) was required to be developed. This point has already been discussed in Chapter 3 of this report. Perhaps, for convenience reasons, this database was referred to as physiographic database in Jenkinson and Bomhof (2012). Here, in this report, this database is referred to as watershed attributes database. This database was developed by processing a number of geospatial data products and sources, outlined in Chapter 3. In Phase II of the resource assessment study, this database was developed for all hydrometric stations located in four selected regions for validating the methodology that was proposed in Phase I. To start with, watershed delineations were required at all hydrometric stations.

Watershed delineations were conducted in two different ways. In the first approach, the CDED (Canadian Digital Elevation) DEM data was employed to delineate watersheds upstream of the HYDAT gauging stations. It was found that the delineation of many watersheds required substantial computational effort. Although it was appropriate to follow this approach for the validation phase of the methodology, this approach was not tractable when millions of watersheds needed to be delineated for all river reaches in Canada. Therefore, another relatively efficient approach was investigated using Voronoi drainage area polygons, derived from published river networks. These two approaches are further described below:

<u>DEM-Based Watershed Delineations</u>

For HYDAT stations located in each study region, a number of CDED DEM data files were obtained from NRCan’s GeoBase (Centre for Topographic Information 2000) and assembled into a single large DEM tile, fully covering the study region. The large DEM tile was then processed to calculate streamflow directions, and to determine natural channel locations. The upstream watersheds corresponding to all hydrometric stations were delineated by matching a point on a calculated channel with the reported location of the HYDAT station, using the “At” algorithm (Ehlschlaeger 1989). If necessary, manual corrections were applied when locating the delineation point. The delineated watersheds were used to calculate drainage areas, perimeters, centroids and other DEM-based attributes. These watersheds for two regions (British Columbia and Canadian Shield) are shown in Figure 8 and for another two regions (Prairies and Maritimes) in Figure 9. Calculated drainage areas for all watersheds were compared with those published by EC in HYDAT database. The watersheds showing significant deviations were not included in the analysis. Jenkinson and Bomhof (2012) noted that these deviations were mainly due to DEM discrepancies. However, they did not indicate the final set of stations used in the validation exercise for each of the four study regions.

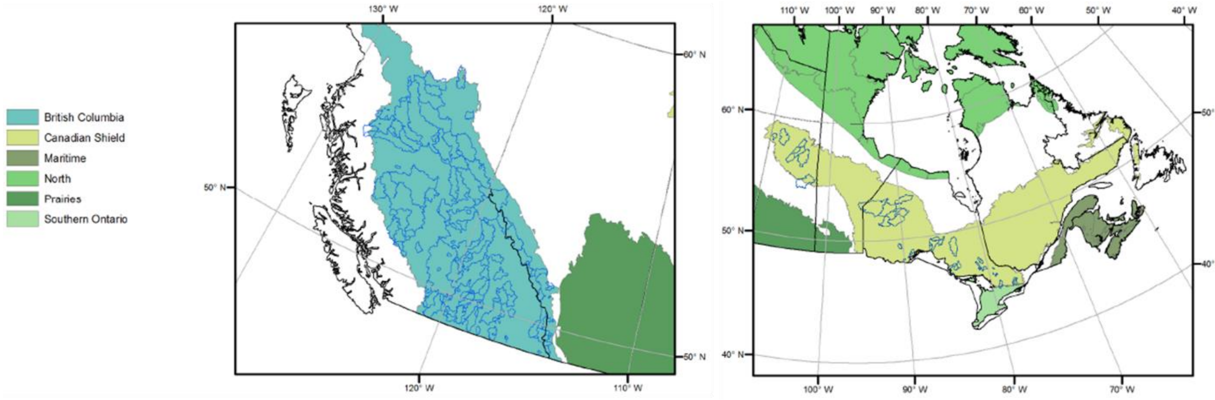

Figure 8: Delineated watersheds for two selected regions, British Columbia (left panel) and Canadian Shield (right panel). Source: Jenkinson and Bomhof (2012).

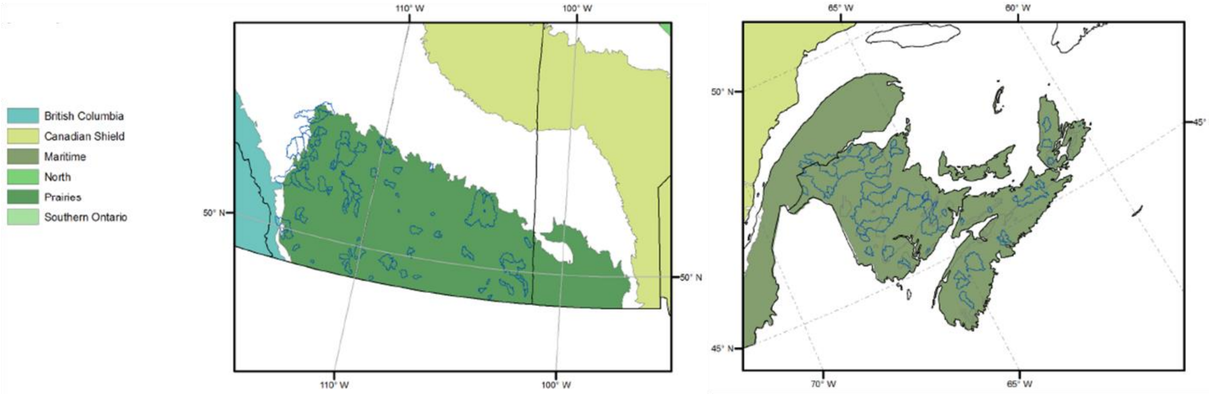

Figure 9: Delineated watersheds for two selected regions, Prairies (left panel) and Maritimes (right panel). Source: Jenkinson and Bomhof (2012).

Voronoi-Based Watershed Delineations

The best alternative for watershed delineations was to use Voronoi polygons. This approach considers watershed delineations based on distance to the nearest stream segments. Thus, each stream segment has one corresponding Voronoi polygon. This approach is not as accurate as the

DEM-based approach because it does not account for terrain elevation, however, it is computationally lot less expensive. By comparing DEM- and Voronoi-based delineations for 211 sample watersheds from the four study regions, Jenkinson and Bomhof (2012) reported that about half of these watersheds with drainage areas less than 1,150 km^2 had a relative error of 22.0%, while the other half with larger drainage areas had a relative error of only 0.8%. A comparison of DEM-based delineated watersheds with those reported in the HYDAT database resulted in a much stronger relationship, with a mean relative error of 3.1%. The errors for watersheds with smaller drainage areas were quite significant.

Following the comparison of watershed delineation methods, a total of 71 physiographic and climatic attributes were derived for each of the Voronoi polygons. These attributes were assigned by spatially superposing the polygons onto the relevant data coverages. All considered attributes can be classified into the following six groups:

- Attributes derived from the DEM data
- Attributes derived from the climatological data
- Attributes derived from the land cover data
- Attributes derived from the soils data
- Attributes derived from the surficial geology
- Attributes derived from the RETScreen database

Each of the above group, apart from the RETScreen related group, consists of multiple attributes. Description of these attributes and literature on their relevance to various streamflow characteristics were not found in the Phase II report. This information would have been very useful for many scientific reasons and to understand how various attributes influence streamflow characteristics and runoff generating mechanisms in different parts of the country.

4.5 Regionalization of Streamflow Indices

For predicting streamflow indices at ungauged locations, hydrologic regionalization based on canonical correlation analysis (CCA) was proposed by Jenkinson (2010) in Phase I of the resource assessment study. However, in general, before adopting a specific approach from a set of candidate approaches for large scale applications, such as for whole of Canada, it is important to evaluate all candidate approaches using observational data from gauged locations and that is what was accomplished by Jenkinson and Bomhof (2012) in this part of the Phase II study. This evaluation was performed separately for each of the four study regions, i.e. British Columbia, Canadian Shield, Prairies and Maritimes. The following approaches were considered in this evaluation study (see also Chapter 3 of this report):

- The RETScreen procedure for estimating FDCs,
- Multiple regression (MR) analysis for estimating FDCs and MMFs, and

- Multiple regression combined with the CCA-based delineation of homogeneous regions or neighbourhoods (MR-CCA) for estimating FDCs and MMFs.

The above three methods are further elaborated below.

4.5.1 Description of Regional Approaches

RETScreen Procedure for FDCs

The RETScreen is a decision-support system which relies on two essential components for obtaining a FDC at a given point of interest across Canada (NRCan 2004a, 2004b). These components are: (1) a map of hydrologically similar regions; each region is associated with a representative normalized FDC (see Figure 1), and (2) a specific runoff map of Canada, showing mean annual values (see Figure 3). These two components need to be used together to obtain a FDC at the point of interest. To avoid any mix up with the four selected study regions, the hydrologic regions that were an integral part of RETScreen are referred to as simply RETScreen regions hereafter. Jenkinson and Bomhof (2012) digitized associated datasets (i.e. specific runoff maps) in order to expedite application of the RETScreen procedures. For a watershed associated with a HYDAT gauging station, covering more than one RETScreen regions, a weighted average specific runoff amount (based on fractional coverages) was obtained which then was multiplied with the watershed area to obtain total mean annual runoff. If a given watershed was located entirely within a single RETScreen region, then the procedure for obtaining the mean annual runoff was straightforward. This estimation is essential for de-normalizing regional FDCs, associated with RETScreen regions. The regional FDCs stored in the RETScreen database are in the form of 21 discrete points with values at each 5% interval from zero to 100%. It is important to note that, in theoretical terms, the values at these two extremities are not defined. Each watershed was assigned to a RETScreen region based on the location of the associated HYDAT station. The ordinates of the regional FDC were de-normalized using the mean annual runoff (and not with the drainage area as mentioned in Jenkinson and Bomhof 2012). For this validation study, each FDC was converted from a 21-point FDC to a 17-point FDC. Linear interpolation was applied when the probability of exceedance values did not match directly between the 21-point and the 17-point FDCs.

MR Approach for FDCs and MMFs

It seems that Jenkinson and Bomhof (2012) applied the MR technique in each of the four study regions separately, considering all available HYDAT stations. A stepwise regression analysis was adopted and a p-value of 0.05 was used for exclusion/inclusion of a watershed attribute. In the case of FDCs, a separate MR relationship was developed for each of the 17 percentile flows in each of the study regions. In the case of MMFs, 12 values were independently modelled. For MAFs, numerical integration of the FDC and simple averaging of MMFs were used. Jenkinson and Bomhof (2012) also mentioned that a jackknife approach (Efron 1982) was used to evaluate the MR technique. However, it is unclear how the jackknife approach was used when performing a comparison with the RETScreen procedure and whether the watershed attributes were kept the same for all stations located within a study region or varied from station to station. It is also unclear if the MR relationship was redeveloped following the step-wise regression analysis for each of the considered stations. Though uncertain, the latter seems to be the likely case.

MR-CCA Approach for FDCs and MMFs

Naturally, there was a certain degree of spatial variability inherent in physiographic and hydrometric characteristics of various watersheds of the four study regions. The MR-CCA approach was considered to be beneficial in order to account for this variability. For the application of this approach, a statistically similar grouping of stations or neighbourhood was formed for a given station which in turn was used to develop the MR relationship. Some information on this approach can be found in Chapter 3 of this report. From the description provided by Jenkinson and Bomhof (2012), it is not clear how the neighbourhoods were identified for modelling both FDCs and MMFs. However, it is clear that a single neighbourhood was used for all 17 percentile flows for predicting FDCs and similarly, a single neighbourhood was used for predicting all 12 MMFs. As stated in Jenkinson and Bomhof (2012), this strategy was adopted to produce the best overall results. However, this strategy has some fundamental shortcomings which are discussed in the last section of this chapter. A parameter "alpha" that controls the size of the neighbourhood was optimized by minimizing the overall root mean square error (RMSE) and considering 25 discrete values of "alpha", i.e. alpha = 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, and 0.75. Using a zero value of alpha would be equivalent to using just the MR, i.e. applying the MR technique without considering CCA-based neighbourhoods. Unfortunately, no information was provided about the optimized alpha values and the groups of watershed attributes that were retained for MR in each of the four study regions. This information would have been very useful for many readers from a physical viewpoint.

4.5.2 Performance Evaluation

For evaluating the performance of regression-based approaches, a jackknife or "leave one out" cross validation approach was used. For the application of this approach, each of the hydrometric stations and its associated flow (e.g. a specific percentile of the FDC) and physiographic attributes were removed in turn, and then the data from the remaining stations was used to predict the target flow at the excluded station. The values predicted this way were compared to the measured values using a number of metrics, including the Nash-Sutcliffe efficiency (NASH), RMSE, relative root mean square error (RRMSE), bias, and relative bias. Since recorded flows are not available at ungauged locations, the excluded station, where the targeted streamflow indices are being estimated, is assumed ungauged in this type of evaluation. This strategy is commonly used to evaluate various candidate approaches for their application and performance at ungauged locations within a region of interest. The word "performance" could be misleading to some readers. Here, it simply indicates how well a given approach will perform when applied under real ungauged situations.

4.5.3 Results of Regionalization Approaches

The results from RETScreen, MR and MR-CCA approaches were compared on a region-by-region basis for all study regions. HYDAT stations from multiple regions were not considered to augment the sample space for the application of both MR and MR-CCA approaches. As mentioned before, the application of the MR-CCA approach does require a group of stations or a neighbourhood to be identified for each target station. It is not clear, how this was accomplished in the validation study. The RETScreen procedure was applied for each station separately, depending upon station’s location within a RETScreen region.

Prediction of FDCs

For each of the four study regions (i.e. British Columbia, Canadian Shield, Prairies, and Maritimes), FDCs in the form of 17 percentile flows were estimated for all stations and compared with those obtained from recorded observations. The performance matrices were evaluated for each percentile flow separately for each of the three methods (i.e. RETScreen, MR and MR-CCA) and graphical comparisons were made. Jenkinson and Bomhof (2012) noted that the MR and MR-CCA approaches outperform the RETScreen procedure, but the results shown in their Figures 20 to 23 do not support this statement. For the Prairies region, they noted that all three methods performed similarly in terms of RMSE. Also, for this region, the performance in terms of RMSE was very different. Overall, the performance of the RETScreen procedure was not bad as noted by Jenkinson and Bomhof (2012). In addition, some other contradictory statements were also made with respect to performance of the three methods (not discussed here). The MR-CCA approach performed in general better than the MR approach. The performance of the three methods was very similar for the majority of the 17 percentile flows. For British Columbia, where a great degree of hydrologic variability existed within the region, the MR-CCA approach outperformed the MR approach substantially.

Additionally, a comparison between the RETScreen procedure and the MR-CCA approach was also made for MAFs, estimated from predicted FDCs and from recorded observations. MAFs from predicted FDCs were obtained through numerical integration of the FDCs, defined by 17 percentile flows. Though the RETScreen procedure showed a consistent positive bias (i.e. overestimation) for each of the four regions, considerably high values of the correlation coefficient were noted for all regions, except Prairies. The results of MR-CCA for the Canadian Shield and Maritimes regions were better than those for the other two regions, with relatively poor results for the Prairies region.

Prediction of MMFs

For predicting MMFs at gauging stations within each of the four study regions, the same procedure was used as used for the case of FDCs. Instead of 17 percentile flows, 12 MMFs were predicted using only the MR and MR-CCA approaches. As the RETScreeen procedure did not facilitate determination of MMFs, comparative evaluation was not possible for this case. As for FDCs, the MR-CCA approach also outperformed the MR approach, specifically for British Columbia, in terms of NASH and RMSE values. However, this was not so obvious for the other three regions where both methods performed equally. This is contrary to the statement made in Jenkinson and Bomhof (2012) that British Columbia and Prairies benefitted substantially from the CCA-informed approach. Compared to the MR approach, the MR-CCA approach performed much better for late summer months for British Columbia. During this period, rivers are fed by glacier melt. For Prairies, extremely marginal differences were seen for non-winter months in the results shown in Jenkinson and Bomhof (2012), contrary to what was discussed.

The results of MMFs were also compared in terms of MAFs. MAFs were obtained by averaging estimated MMFs. This comparison showed good correspondence between observed and estimated MAFs for British Columbia, Canadian Shield and Maritimes and not so for Prairies. The MAFs estimated based on MMFs were more similar to observed MAFs compared to those estimated by integrating FDCs. Compared to FDCs, the advantage of using MMFs was that they were also able to preserve the temporal variability within the annual cycle.

4.5.4 Influence of Voronoi-Based Watershed Delineations

In the above evaluations of the three methods, database of watershed attributes was developed using watershed delineations from the DEM data. Alternate delineations and associated attributes based on Voronoi polygons were also evaluated in Phase II of the study. For certain cases, Jenkinson and Bomhof (2012) noted considerable differences between the delineated watershed areas. To evaluate the impact of these differences, the MR-CCA approach was applied to the entire set of 214 watersheds for which Voronoi and DEM derived drainage areas were available. The number 214 could not be verified based on the information provided in Jenkinson and Bomhof (2012). The results of NASH and RMSE from all regions for the 17 percentile flow values were amalgamated to produce a single plot for each of the two delineation options (i.e. DEM-based and Voronoi polygon based watershed delineations). No discernible differences between the two approaches were noticed and that indicated that Voronoi polygon based delineations is a viable option instead of using computationally intensive DEM-based delineations. This might be useful in expediting the process of watershed delineations and development of physiographic and other databases for large scale applications, e.g. for whole of Canada.

4.5.5 Overall Comparison of Estimated Streamflow Indices

Overall, in terms of coefficient of correlation, RMSE and bias, the MR-CCA method performed similarly for estimating MMFs and FDCs for the Prairies, Canadian Shield and Maritimes regions. However, MMFs were better estimated than the percentile values of FDCs by the MR-CCA approach for British Columbia in terms of coefficient of correlation and RMSE values.

According to Jenkinson and Bomhof (2012), the RETScreen procedure performed poorly compared to the other two methods for estimating FDCs. However, after having a closer look, the performance of this procedure is still acceptable in terms of performance indicators shown in their analysis.

4.6 Velocity Estimates and Significant Parameters

Velocity of water in a river/channel plays a critical role in the estimation of hydrokinetic potential at a given point in a river reach. It can be obtained from observed or estimated flows using two potential methods, i.e. by employing either the Manning's equation or the continuity equation. The Manning's equation can be written as:

$$V = \frac{1}{n} S^{\frac{1}{2}} R^{\frac{2}{3}} \tag{7}$$

where n is the Manning's roughness coefficient, S is the channel slope, and R is the hydraulic radius. Assuming a rectangular channel, the hydraulic radius R can be written as:

$$R = \frac{A}{P} = \frac{WD}{W+2D} \tag{8}$$

where D is the depth of water, W is the width of the channel, P is the wetted perimeter, and A is the cross-sectional area. Knowing the depth of water, channel width and magnitude of discharge, velocity can also be obtained from the continuity equation:

$$V = \frac{Q}{A} = \frac{Q}{WD} \tag{9}$$

By combining the continuity equation with the Manning's equation and eliminating depth, the Manning's equation can be written as:

$$V = \frac{1}{n} S^{\frac{1}{2}} \left(\frac{Q}{VW + \frac{2Q}{W}} \right)^{\frac{2}{3}} \tag{10}$$

For the Phase II study, the channel geometries were assumed to be rectangular with a constant width, and no overtopping of the banks. In this part of the report by Jenkinson and Bomhof (2012), various methods for estimating channel geometry, slope and roughness at locations of HYDAT gauging stations were evaluated, along with velocity estimates for a range of discharges.

4.6.1 Channel Geometry Relationships

Two sets of channel geometry relationships were evaluated, one based on drainage area and the other based on the discharge. In the first set, channel width and depth were defined as power functions of drainage area as shown below:

$$w = aA^b \tag{11}$$

$$d = cA^f \tag{12}$$

where *w* represents channel width at the water surface and *d* represents the depth of water, and *a*, *b*, *c*, and *f* are coefficients to be determined through optimization or least squares analysis. Lower case notation for *w* and *d* is used here compared to the general notation used in the description provided above. These equations were fitted through the least squares approach using data from the EC-MDB for each study region separately and for all regions together. Reasonably strong relationships were noted for channel width for most regions, with the exception of Prairies. The relationship for channel depth was not as good as for the width, especially for Prairies and Maritimes.

In the second set, width, depth and mean velocity were defined as a power function of discharge. These relationships were proposed by Leopold and Maddock (1953), as shown below:

$$w = aQ^b \tag{13}$$

$$d = cQ^f \tag{14}$$

$$v = kQ^m \tag{15}$$

where *Q* is the channel-forming discharge, *w* and *d* are the same as defined before, *v* is the average flow velocity, and *a*, *b*, *c*, *f*, *k*, and *m* are coefficients, to be estimated through least squares analysis. Two functional restrictions are applicable to the coefficients, i.e.

$$ack = 1 \tag{16}$$

$$b + f + m = 1 \tag{17}$$

Channel widths and depths were calculated for each station from the EC-MDB for each percentile of the 17-point FDC using the robust regression relationships. These relationships were described in Appendix D of Jenkinson and Bomhof (2012). The following text is derived from these authors' observations:

"After some investigations and comparisons with some established methods, a robust regression approach was used to develop power relationships among the depth and width data and corresponding flows taken from the EC-MDB. This approach was similar to the ordinary regression analysis except that a weighting factor was applied to each data point depending on its distance from the mean (MathWorks 2011). The 'Huber' weighting was applied in the development of these relationships. This approach was further refined by considering data points with flows above the MAF. The regional regression was found to be a better method for predicting bankfull widths on a regional level. However, a large amount of scatter was noticed for flows below the MAF threshold."

For each study region, width vs. flow and depth vs. flow curves were created for each percentile flow following the robust regression relationships by Jenkinson and Bomhof (2012). In their report, these curves were shown only for the 30% exceedance probability level. The values of the coefficients *b* and *f* compared well with those provided in Park (1977), based on a number of studies worldwide, and in Allen et al. (1994) for the US.

In order to further validate the generated widths and depths for each percentile flow, velocity was calculated using the continuity equation and then an average value was obtained. Averaged velocity values were compared with measured values from the EC-MDB for all four study regions together. In the comparison, estimated velocities showed a good correspondence with measured velocities.

4.6.2 Estimation of Channel Slope and Roughness

Estimates of channel slope were required to calculate in-stream velocities from estimated streamflow. This investigation involved estimation of all local slopes at the outlet of each watershed. Following a previous study (i.e. NRC-CHC 2008), Jenkinson and Bomhof (2012) used a reach length of 750 m to estimate channel slope, but due to errors in the DEM data, zero slopes were encountered, especially for many watersheds in the Prairies region. Consequently, the reach length was increased progressively by 250 m intervals starting at 750 m until encountering a non-zero slope. This slope was retained for evaluation purposes. According to an additional investigation by Jenkinson and Bomhof (2012), it was not possible to estimate slope as a function of drainage area and therefore DEM-based approach was picked.

Jenkinson and Bomhof (2012) also note that no information was found to estimate channel roughness, through functional relationships, and therefore, a uniform value of the roughness coefficient was assumed (i.e. $n = 0.04$) for all watersheds in the four study regions. They also mentioned that this value was adjusted through a sensitivity analysis, but did not explain how this was accomplished.

4.6.3 Comparison of the Manning’s Equation and the Continuity Equation

The relative accuracy of the Manning’s equation and the continuity equation was also evaluated in terms of power estimates. Channel widths and depths were calculated using the regional equations, discussed above, with flows from historical FDCs. In the case of Manning’s equation, streamflow magnitude, width, estimated slope, and uniform channel roughness were used to calculate channel velocity and specific power. Power was calculated using cross sectional areas derived from calculated widths and depths. In the case of continuity equation, streamflow, width and depth were used to calculate velocity and specific power. In this case, power was calculate in the same way as for the case of Manning’s equation. Final power estimates were obtained by integrating the resulting PDCs at each station and compared with corresponding observed values. The continuity equation procedure resulted in much better estimates than the Manning’s equation procedure. This discrepancy was attributed to inaccuracies associated with estimated values of slope and roughness.

4.7 Hydrokinetic Power Estimates

Due to the focus of the present report on hydrological investigations, only a brief description of the methods used for estimating hydrokinetic power is provided here. Additional insights on the performance of these methods, and how they were applied for power calculations, can be found in Jenkinson and Bomhof (2012).

Estimates of hydrokinetic power were obtained using two different approaches: (1) direct regionalization of the hydrokinetic potential and (2) using the hydraulic flow equations–first by estimating channel geometry and then the hydrokinetic potential. In both cases, estimated FDCs and MMFs, as described in detail in the previous section, were employed and estimates of time-averaged hydrokinetic power were compared to the time-averaged hydrokinetic potential obtained from observed data, separately for each of the four study regions (i.e. British Columbia, Canadian Shield, Prairies and Maritimes). Based on the information provided in Jenkinson and Bomhof (2012), though not explicitly mentioned, most likely the observed power was estimated by converting the observed FDC to an equivalent PDC and then numerically integrating the latter.

In the second approach, power was estimated once by employing the continuity equation and then by employing the Manning’s equation to calculate channel velocity. The latter involved the usage of the Manning’s roughness coefficient and channel slope as well. The power estimates were obtained separately based on both FDCs and MMFs for each of the four study regions and compared with corresponding observed estimates.

4.8 An Overview of the Study and Outcomes

In this chapter the work conducted within the scope of Phase II of the resource assessment study is reviewed, with specific emphasis on hydrological investigations. A specific methodology based on the hydrologic regionalization concept for estimating streamflow indices at ungauged locations was proposed in Phase I. This methodology was evaluated for estimating FDCs and MMFs at ungauged locations within the regionalization framework and assuming gauged locations as ungauged sites following the leave-one-out cross validation approach. For the evaluation, four large regions were selected, i.e. British Columbia, Canadian Shield, Prairies and Maritimes, and a dataset of watershed attributes, representing climatic, geologic, land use and soil types, was also developed. Development of watershed attributes was a significant undertaking. Multiple regression combined with canonical correlation analysis (MR-CCA) and just the multiple regression (MR), without CCA, were evaluated. It was concluded that in most cases the MR-CCA approach performed better than the MR approach. Another established approach (i.e. the RETScreen procedure of NRCan 2004a, 2003b) was also evaluated. According to Jenkinson and Bomhof (2012), this approach did not perform as good as the other two approaches and resulted in overestimation of FDCs. However, this statement could not be verified based on the results provided in the Phase II report. In comparison to FDCs, MMFs were better simulated in most regions, especially in British Columbia. The following recommendations were made based on the results of the study:

- It was recommended to replace the RETScreen procedure with the regionalization approach based on CCA-based delineation of homogeneous regions or groupings of stations for estimating FDCs for resource assessment studies.
- It was recommended to employ the multiple regression approach for channel geometry in combination with the continuity equation for estimating hydrokinetic energy from the national river network.
- Due to errors associated with small watersheds, delineated using Voronoi polygons, noted in the study, high-resolution datasets were recommended.
- Estimation of water depth based on power form channel geometry relationships were recommended for turbine deployment purposes.

4.9 General Observations and Remarks

Some important aspects that were expected to be taken care of in the Phase II study are described below. Some of these points needed additional explanations and clarifications, with support from the published literature. In addition, at certain places in the Phase II report, it was also found that some important information was either missing or incomplete. There were also some issues noticed with the usage of statistical terminology in the Phase II report.

- The title of the Phase II report is somewhat misleading and it does not justify with the work that was accomplished in this phase. The report not only covers development of physiographic and climatic datasets, it also covers validation of the proposed methodology using data from hydrometric stations.
- For each of the four study regions, hydrometric stations with upstream drainage areas located within the boundary of each study region were identified first and then a subset of stations was considered based on the following criteria: (1) stations with minimum 10 years of continuous data, (2) currently active stations, and (3) unregulated stations, representing near natural flow conditions. However, it was not clear what parameters were used to satisfy the first and second requirements (e.g. the starting and ending years of recorded data and the targeted time window).
- For the evaluation of the RETScreen procedure, regional FDCs were converted from 21-point FDCs to 17-point FDCs to validate the methodology that was proposed in Phase I. When the points did not directly match between these FDCs, a linear interpolation between two neighbouring points was used. Due to nonlinear nature of the probability scale, it was better to perform this interpolation within the logarithmic domain.
- It is difficult to assume specific runoffs as constant values, integrated with the RETScreen procedure, e.g. a map prepared based on data from the 1961–1990 period cannot be assumed as representative of the 1991–2020 period, unless the differences are minor and lie within the

range of natural variability. This was a fundamental shortcoming of the RETScreen procedure and that needed to be discussed in the report.

- It is also not clearly stated in the report why the proposed methodology was evaluated using data from gauged locations, while the focus of the hydrokinetic resource assessment study was on ungauged locations across Canada. Though underlying procedures may be clear to those individuals who are familiar with ungauged hydrology, it was necessary to make this point clear for a general reader.
- The usage of terms like “validation stations” and “validation datasets” was not so convincing. A different terminology should have been used to refer to gauging stations, where observational records are maintained. The observational records are also used for model development purposes. This type of terminology is generally suitable in situations where the distinction between “calibration/training” and “validation” phases of a study are made prior to undertaking an evaluation of a given modelling approach.
- A database of 71 different watershed attributes was developed in Phase II of the study. Undoubtedly, this was a huge undertaking and Jenkinson and Bomhof (2012) must be applauded for this achievement. The same database was used in the evaluation of two regionalization approaches, one involved region-wise multiple linear regression and the other also involved the same regression approach, but informed by canonical correlation analysis to form homogeneous neighbourhoods based on selected similarity measures. No discussion is available on the number, type and definitions of watershed attributes that were retained and used for each of the two regression approaches. This was necessary to understand the influence of various attributes on regression relationships and uncovering their role in predicting various parts of the FDC and the annual cycle of MMFs. Generally, these aspects are very insightful from a physical viewpoint, since the streamflow generating mechanisms are not the same throughout the year and also they are not the same for high and low flow conditions.
- For a regression modelling exercise, where a number of explanatory variables is involved as was the case of this study, it is important to screen those variables a priori in order to avoid inclusion of mutually correlated variables. Without such a screening, one can end up proposing or adopting an over-fitted relationship wherein some variables are merely duplicating some other variables. No discussion is available in the Phase II report regarding this aspect in order to avoid spurious attributes being included in the multiple regression relationships. The step-wise regression procedure used in the analysis for exclusion/inclusion of an attribute does not guarantee that the attributes are mutually uncorrelated. This was a fundamental shortcoming noticed in the applications of multiple regression method. In addition, regression diagnostics were not at all discussed in the report. These diagnostics are readily available in almost all computer packages (e.g. Matlab, R package, SAS, etc.).
- A single optimal neighbourhood was used in the estimation of all percentiles of the FDC and the same was the case for all MMFs. It is very difficult to justify this choice based on physical mechanisms involved in the generation of high and low flows. Similarly, the

climatic and hydrological mechanisms responsible for streamflow generation during, for example, April to May period are very different from those dominant in September to October period in different parts of Canada. It is important that statistical inferences should not contradict physical principles and understandings. It was not possible to comment on these aspects further because the list of attributes was not available in the Phase II report.

5 Assessment of Canada’s Hydrokinetic Power Potential: Phase III – Resource Estimation

5.1 General

The work conducted within the scope of Phase III of the hydrokinetic resource assessment study was documented in Jenkinson and Bomhof (2014). This was the third and the final report of the resource assessment study. Most of the information provided in this chapter is derived mainly from this source and modified where necessary to overcome some shortcomings observed in the description of methodological aspects of various approaches. First an overview is presented below, followed by additional information in Sections 5.2 to 5.4. Following specific recommendations from Phase I and Phase II of the resource assessment study, a nation-wide assessment of the hydrokinetic potential was undertaken in Phase III. The main objectives of this phase of the study were:

- To develop national physiographic and climatological datasets, including stream networks, to drive hydrologic regionalization approach, which was validated in Phase II, and then to estimate hydrokinetic potential at the national scale;
- To expand and process hydrometric datasets much beyond the level considered for Phase II to support development of hydrokinetic data products;
- To develop suitable techniques for accounting the effects of diversions, regulations and control structures on the hydrokinetic estimates;
- To calculate estimates of theoretically extractable energy from all stream reaches across Canada; and
- To evaluate and summarize uncertainties in the hydrokinetic estimates.

In this third phase of the study, multiple linear regression (MR) in combination with canonical correlation analysis (CCA) was employed to estimate selected percentiles of flow duration curves (FDCs) for all river reaches delineated in the Canadian National Atlas vector data (Natural Resources Canada 2008). Various percentiles of FDCs estimated this way were combined with those obtained from the recorded data to improve their relevance to existing conditions. Estimates of hydrokinetic potential for all stream reaches were obtained using the channel flow-geometry relationships, originally proposed by Leopold-Maddock (1953), developed based on Environment Canada’s measurement database (EC-MDB). The estimated total hydrokinetic energy potential was approximately 710 GW with a 95% uncertainty interval of 430– 2200 GW. It is important to note that this assessment was of pre-reconnaissance nature and this point has already been indicated upfront in Chapter 1. Additional description of a number of components, including data sources used, regionalization approach, Canadian drainage regions, power estimation procedures, integration of observed flows from HYDAT with the estimated flows, uncertainty assessment strategy, and channel flow-geometry relationships, is provided below. In each case, only a summarized information is provided and more detailed information can be found in Jenkinson and Bomhof (2014). It is important to point out that a few equations from Chapters 2 and 4 are intentionally repeated in Section 5.2 in order to maintain the flow of information provided and to demonstrate various inter-relationships.

5.2 Methodology

5.2.1 Data Sources

A number of different datasets were employed for hydrokinetic resource assessment, particularly for the estimation of selected percentiles of FDCs at all stream reaches; identification of homogeneous regions; and development of channel geometry relationships. These datasets were considered in the Phase II report (Jenkinson and Bomhof 2012) and have already been discussed in Chapter 3 of this report and therefore are not discussed here again. However, some important information on specific features of these datasets that was found insightful from an application and assessment perspective is retained here. In addition, the methodology adopted for estimating hydrokinetic potential across Canada is also described.

HYDAT

This is EC's repository of streamflow records for over 2,500 active and 5,500 discontinued hydrometric stations across Canada. Data from this repository was used to develop FDCs at all selected gauging sites that were used to identify homogeneous regions or neighbourhoods for estimating selected percentiles of FDCs at all ungauged locations within the hydrologic regionalization framework. The underlying procedures have already been described in Chapters 3 and 4 of this report.

Environment Canada Measurement Database (EC-MDB)

The EC-MDB is a repository of hydrometric field measurements. It includes information on cross-sectional areas, channel geometry elements and velocity of water for conducted field surveys up to the year 1997. Data was available only until 1997 at the time of the study. Selected data from this repository was used to characterize hydrokinetic power at each gauging location and a part of that was used in validating the proposed methodology in Phase II. This repository contains information for 1,861 stations, which are part of HYDAT, and does not contain data from Quebec. For the Phase III study, stations with 5 or more records were considered and that criterion resulted in 1,174 stations. Additional discussion on this aspect is provided later in this chapter.

Digital Stream Network

The National Atlas of Canada provides a vectorized representation of streams and lakes in Canada in ESRI shapefile format at the 1:1 million scale (Natural Resources Canada 2008).

There are approximately 365,000 river segments contained in this dataset, with associated drainage areas as small as 0.1 km^2. The data was divided by major drainage basins (see Figure 10). The stream network had complete connectivity and contained critical attributes that were used in the study. These attributes were: (1) unique identifiers for each stream segment; (2) upstream and downstream connecting nodes for each stream or lake segment; and (3) an identifier representing whether a segment is a river, lake, coastline or inland border. It has already been mentioned in Chapter 4 of this report that delineating a drainage area for each stream segment was computationally very expensive so Voronoi polygon based delineations using ESRI ArcGIS (ESRI 2010) were adopted (Jenkinson and Bomhof 2012). These delineations were used to develop physiographic and other required attributes.

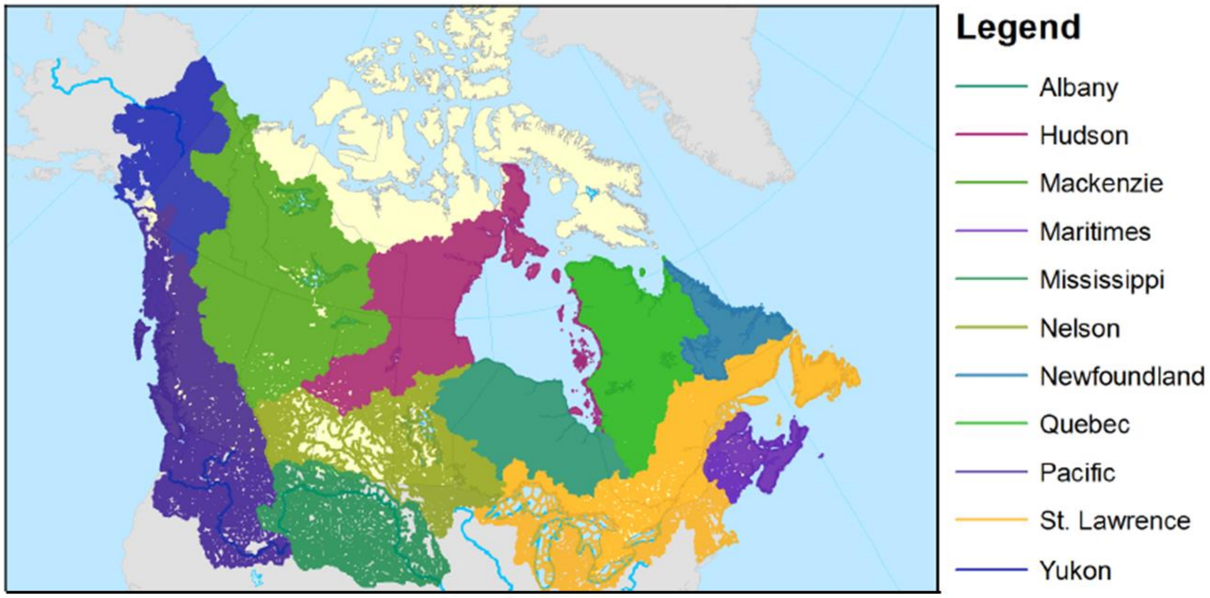

Figure 10: Major drainage basins, with stream segments. Figure adopted from Jenkinson and Bomhof (2014).

Physiographic and Climatic Datasets

The AAFC Soil Landscapes of Canada Working Group and the Canadian Soil Information System (CanSIS) created a series of GIS coverage maps that show various major characteristics of soil and land for the entire country (AAFC 2000). Soil characteristics, including hydraulic conductivity, drainage and soil water retention were derived from this database.

Land cover data was obtained from Natural Resources Canada via Geobase (Centre for Topographic Information 2009). Land use variables were classified into 45 different land cover categories. For the Phase III study, these categories were generalized into 10 unique land classes.

Surficial geology data, obtained from Earth Sciences Information Centre via GeoGratis (Turner et al. 2003), was used to obtain information on the Canadian surficial geology, separated into 26 classes. These classes are also identified by a more generic classification based on origin and grain size.

Climatic data was obtained from AAFC (2008) that included temperature (annual minimum, annual maximum and annual mean values) and precipitation (annual rainfall, annual snowfall and total precipitation amount), interpolated from EC's weather stations, for the 1961–1990 period.

A number of physiographic and climate attributes were developed from the above mentioned datasets for drainage areas of individual streams and employed in regression modelling and identification of homogeneous regions or neighbourhoods. All of the attributes were not individually explained in the Phase III report and the same was also the case with the Phase II report. Description of the attributes had been very insightful for many interested readers.

Streamflow Calibration Dataset

The streamflow calibration dataset was created to drive regionalization techniques. This dataset consisted of selected indices of streamflow that were calculated from observed records from the HYDAT database. Streamflow indices included: FDCs, mean monthly flows (MMFs) and mean annual flows (MAFs). The FDC was divided into 17 percentile values, following the studies by Mohamoud (2008) and Shu and Ouarda (2012). Shu and Ouarda (2012) found that the 17 discrete points adequately characterize a FDC, while making no prior assumptions about the shape of the FDC. The advantage of MMFs is that they provide a good indication of seasonal patterns which are not captured in the FDC. Gauging stations were carefully filtered and only those that satisfied the following criteria were considered, i.e. (1) stations that had 10+ years of data, (2) stations with recent observations (1995 or later), and (3) stations that were unregulated. Due to the limited number of watershed delineations and the time consuming nature of delineating new watersheds from the DEM data, it was not possible to include all HYDAT stations that satisfied the above three point criteria. Consequently, Jenkinson and Bomhof (2014) retained 889 stations that represented every region and province in Canada, except Quebec.

5.2.2 Prediction of Streamflow Indices

As prescribed above and also in Chapter 4, FDCs were characterized in the form of 17 discrete percentile values. These percentiles of DFCs and values of MMFs were required to be estimated for every river reach identified in the Digital Stream Network. These estimates were obtained using the multiple linear regression approach combined with CCA (referred to as MR-CCA hereafter). In the Phase III report, detailed procedure was described only for FDCs and not for MMFs. The MR-CCA approach was also evaluated in the Phase II study, already explained in Chapter 4 (cf. Jenkinson and Bomhof 2012).

Stepwise regression was used to determine the most significant attributes for all selected percentile flows of the FDC. The attributes that were found significant for at least 5 of the 17 percentile values were designated as key attributes and the same were used in CCA. The CCA approach has already been described in Chapter 4. This procedure was also used in previous studies to identify hydrologically similar groups of stations or neighbourhoods, for a given set of target locations. For example, the CCA was used for flood frequency analysis in Quebec (Ribeiro-Correa et al. 1995) and Ontario (Ouarda et al. 2001).

In the Phase III study, the CCA approach was used to determine groups of stations that were similar in the attribute space to the target stream reaches, where the percentile flows were to be estimated. According to Jenkinson and Bomhof (2014), only a small number of attributes was available for CCA. This does not seem to match with the information provided in Chapter 4 based on Jenkinson and Bomhof (2012) where 71 different attributes were mentioned. Of the 17 percentile flows, values corresponding to 0.1, 0.5, 1.0, 5.0, 30.0, and 50.0% exceedance probabilities were considered for CCA. This smaller set of percentile flows was selected to best match the higher flow range in the FDC and to account for the fact that larger flows have a much greater contribution to the available hydrokinetic potential.

From the description provided in Jenkinson and Bomhof (2014), it appears that the selected percentiles of FDCs were estimated at all river reaches separately for each of the eleven large drainage regions, shown in Figure 10. In order to identify a group of neighbouring stations for a target station through the CCA approach, a parameter alpha needed to be optimized. This parameter controls the number of stations to be included in the neighbourhood of a given station. The alpha value was identified on a regional basis (by major drainage basins) and was determined by a trade-off between having minimum errors and having enough neighbours to estimate target flows at all stream reaches. If alpha is kept very small (e.g. 0.0001), nearly every station from a group of source stations is likely to be included in the neighbourhood, leading to dissimilar groups of stations. If alpha is large (e.g. 0.7), the size of the neighbourhood becomes too small, leading to situations where additional analyses could no longer be performed and having many river reaches without estimated target flow values. For example, having fewer predictor variables than the number of parameters to be estimated for a regression relationship. The MR-CCA approach was applied for a range of alpha values (i.e. from 0 to 0.75, with multiple steps) for each drainage region, estimated flows at selected points were compared with measured values at nearby HYDAT stations. Through evaluating RMSE values and the number of reaches with sufficient neighbouring stations, optimum values of alpha were determined. These values of alpha are shown in Table 3 for all regions.

After finding the best neighbouring stations, target percentile flows of the FDC were obtained through multiple regression using the best combination of variables for each percentile flow, determined from the analysis described above. Uncertainty was taken into account by determining 95% prediction intervals, compared to the confidence intervals.

Table 1: Calibrated values of the parameter alpha for 11 large drainage regions of Canada (from Jenkinson and Bomhof 2014).

Drainage Region	Alpha Parameter	Drainage Region	Alpha Parameter
Albany	0.0050	Newfoundland	0.0001
Hudson	0.0001	Quebec	0.0010
Mackenzie	0.0001	Pacific	0.0010
Maritimes	0.0050	St. Lawrence	0.0010
Mississippi	0.0000	Yukon	0.0001
Nelson	0.0075		

5.2.3 Augmenting Predicted Flows with Measured Flows

The MR-CCA technique was developed to estimate targeted percentiles of FDCs assuming that the stream reach is unregulated, and physiographic and climatological attributes for the associated drainage area are adequately available. According to Jenkinson and Bomhof (2014), two challenges emerged in the application of this technique:

- Lack of harmonization between the physiographic and climatological data that affect river flows for many international watersheds; and
- Non-compliance to the natural flow requirements for the MR-CCA technique due to existence of partially or completely regulated rivers by dams or other control structures.

These challenges were overcome by enhancing estimated flows based on observed data from HYDAT stations, where available. If a HYDAT gauging station was found upstream of the stream reach under consideration, the corresponding estimated flow based on the relevant physiographic and climatic attributes was removed from the MR-CCA results. This resulted in estimated flows only for the incremental drainage area, not including the drainage area associated with the gauging station. Then, the recorded flow from the HYDAT gauging station was added to the incremental value. Through this strategy, the above mentioned issues were resolved to some extent. Streamflow indices could be predicted for rivers with drainage areas partially in the US, at least downstream of the location of the first HYDAT gauging station. The observed/recorded data reflected the impact of regulation. According to Jenkinson and Bomhof (2014), “the added benefit of this strategy was that the relative confidence on the estimated flows was increased, as the uncertainty in the predicted values was mitigated through the addition of measured flows where available”. Apart from this statement, no discussion is available in their report on “how in fact this reduction of uncertainty was possible”. For this strategy, HYDAT stations with at least 5 years of continuous daily data since 1990 were considered and only the data from 1990 onward was used. This requirement allowed for 1,400 HYDAT stations (out of approximately 6,500) to be considered to augment estimated flows. This criteria also ensured that historical flows did not contain both regulated and non-regulated regimes. From this statement, it was difficult to understand which river flow regime was actually preserved through this strategy. The decision of using data from 1990 onward creates considerable inconsistencies among various physiographic and climatic attributes as many of those pertain to 1960–1990 time period (cf. Jenkinson and Bomhof 2012). This raises serious questions about the utility of the data and derived information. In addition, merely five or a few more years of data adopted for the Phase III study does not reflect long-term behavior of FDCs.

In applying the above strategy, Jenkinson and Bomhof (2014) also noted some discrepancies with the National Atlas of Canada data, i.e. considerable differences in the accumulated drainage areas obtained from the network and HYDAT database. They also noted that this problem was specifically encountered in regions with complicated drainage networks or in areas where diversions were constructed. In these cases, the National Atlas of Canada network was manually adjusted or the observations from HYDAT stations were not considered.

5.2.4 Predicting Power from Flow Estimates

Available power in a stream can be expressed as a total potential energy for hydropower generating purposes or as a hydrokinetic potential. The total potential energy can be expressed as:

$$P_T = \gamma H Q \tag{18}$$

where P_T is the total power, H is the change in the elevation head, γ is the specific weight of water, and Q is the volumetric flow rate. The relationship for hydrokinetic power is given by:

$$P_K = \frac{1}{2}\rho A v^3 \tag{19}$$

where P_K is the hydrokinetic power, ρ is the density of water, A is the channel cross-sectional area and v is the average velocity of water (see also Chapter 2). For hydrokinetic power estimates, information on both the velocity and cross-sectional area of a channel is required. In some cases, cross-sectional area-averaged power, or specific power (P_S) is considered in evaluating the hydrokinetic potential of a river reach at a given point of interest. This is accomplished by dividing the P_K by the channel cross-sectional area, i.e.

$$P_S = \frac{P_K}{A} = \frac{1}{2}\rho v^3 \tag{20}$$

Knowing estimates of stream discharge (or flow rate), velocity of water can be computed using the Manning's equation or the continuity equation. The Manning's equation requires estimates of channel slope and roughness, which are quite challenging to estimate. Therefore, due to lack of reliable information on channel slope and roughness for the national river network considered in the Phase III study, the Manning's equation was not used. In short, Jenkinson and Bomhof (2012) favoured the use of the continuity equation, instead of the Manning's equation. For the application of the continuity equation, channel widths and cross-sectional areas were estimated for each percentile flow of the FDC using the downstream hydraulic geometry relationships, described in Bomhof (2013). Downstream hydraulic geometry was proposed by Leopold and

Maddock (1953), who demonstrated power form relationships between channel geometric elements (i.e. width and depth), velocity and flow rate.

In the Phase III study, “at-a-station” flow-geometry relationships for 428 gauging stations were derived from the EC-MDB described in Chapters 3 and 4. Figure 11 illustrates development of these relationships for the HYDAT station 01AJ003, i.e. the relationship between the channel width, depth, velocity and the streamflow. Only the records greater than the MAF were employed in the development of these relationships. It was observed in Phase II of the study that for lower flow rates the flow-geometric relationships considerably diverge from the expected linear behaviour on a log-log plot. Additional information for developing these relationships is provided in Chapter 3 of this report, which, in turn, was derived from Jenkinson (2010). With the at-a-station relationships defined, regional relationships were derived for large scale applications.

With estimates of geometric elements available for each stream segment, channel velocity was obtained from the estimated flow rate:

$$v = \frac{Q}{A} = \frac{Q}{wd} \tag{21}$$

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

As with estimated flows, there is a considerable degree of uncertainty with the estimated widths and cross-sectional areas, obtained using the procedures described above. Jenkinson and Bomhof (2014) described a procedure to evaluate uncertainty in estimated geometric elements. Unfortunately, due to lack of sufficient details and theoretical justifications, it was very difficult to comprehend their approach and therefore no additional information on uncertainty assessment is provided in this section.

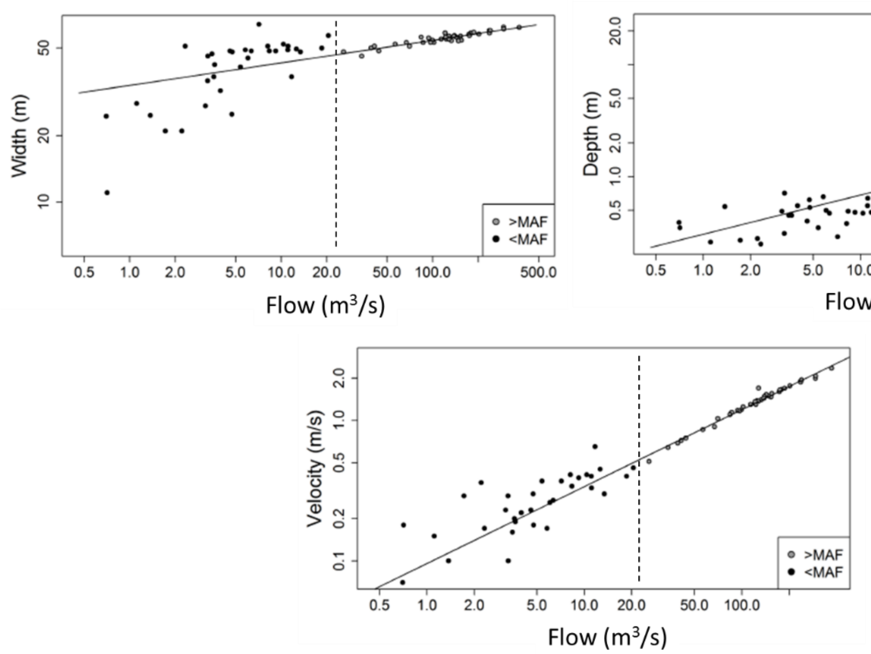

Figure 11: Channel geometric elements (width, depth and velocity) as a function of flow rates for HYDAT station 01AJ003. The vertical dotted line divides the flow rates greater and less than the MAF. Figure modified from Jenkinson and Bomhof (2014).

5.2.5 Total Available Power

The amount of extractable energy at a given point in a stream reach is limited by: (1) temporal variations in streamflow; (2) extractable energy by the turbine; (3) turbine cross-sectional area; (4) hydraulic influence of other turbines or structures in the vicinity of the deployment site; and (5) turbine efficiency. For the resource assessment study, the overall scope was restricted to determining the total amount of energy (in theoretical terms) available in all stream reaches (Jenkinson and Bomhof 2014). From the above list, only the first item was considered in the resource assessment study and other items were not considered since they depend on turbine characteristics and therefore were out of the scope of resource assessment study.

In determining the total amount of energy, it was assumed that all hydrokinetic energy is extractable at all times. In practice, this is not feasible due to considerable variations in extreme low and high flow conditions. In addition, there could be other situations such as ice or debris that can impact turbine use, efficiency and deployment. Because all of these factors are site-specific, it was assumed that the hydrokinetic energy is extractable over the whole range of streamflow variations. Generally speaking, the ability to extract energy during all flow conditions is an engineering challenge for the turbine design engineers and not strictly a natural limitation (Jenkinson and Bomhof 2014).

The maximum extractable energy in a stream reach depends on the number of sites used within the reach to extract power. The stream/river requires a certain length of the reach to ’recover’ the velocity downstream of the deployment location. The UMA Group (UMA 1980) provided some guidelines on determining this length of the river based on a modified form of the Manning's equation, wherein the slope term was replaced by the ratio of the available hydraulic head (i.e. the elevation head) and the unknown length of the river reach. The same procedure can also be applied in the hydrokinetic case by replacing the elevation head term with the velocity head term. Additional information about this procedure is available in the report by the UMA Group (UMA Group 1980).

Following the guidelines and assumptions from the UMA Group study (UMA Group 1980), potential hydrokinetic power for each stream reach was integrated over its length (Jenkinson and Bomhof 2014). Power estimates were summed up and grouped by province and drainage region. Jenkinson and Bomhof (2014) also noted that the majority of stream reaches was processed to obtain total power, with the exception of some reaches that could not be processed due to the restrictions imposed by the MR-CCA approach.

Two separate techniques were employed to estimate the total available hydrokinetic power in a stream reach. The first technique involved using a combination of the flow rate and the elevation head, assuming that the maximum theoretically extractable hydrokinetic energy is equivalent to, and cannot be higher than, the total hydropower potential available. The second approach involved estimating the velocity head and directly using the flow rate and channel geometry estimates. These two approaches are further elaborated below.

When integrating power estimates for channels using the flow rate and elevation head, $P = \gamma HQ$ relationship was employed; H was calculated as the elevation drop along a channel from the available DEM data and Q was the mean annual volumetric flow rate determined by integrating the predicted FDC for the channel. However, it is unclear from the description provided in Jenkinson and Bomhof (2014) if the MAF was derived from the entire FDC or just from the selected percentile values corresponding to five exceedance probabilities (i.e. 0.1, 0.5, 1.0, 5.0, 30.0, and 50.0%) of the FDC. Equation (22) summarizes the calculation procedure considering Q as a function of the exceedance probability f of the FDC:

$$P_T = \gamma H \int_0^1 Q(f) df \tag{22}$$

Hydrokinetic power from channel velocity was determined using equation (19), modified to account for the variability in stream velocity and cross-sectional area, as a function of FDC exceedance probability:

$$P_K(f) = \frac{1}{2}\rho A(f)[v(f)]^3, \ \ \rho = \gamma/g \tag{23}$$

Combining equations (19) and (21) and $\rho = \gamma/g$, $P_K(f)$ can be re-expressed as in the following equation:

$$P_K(f) = \gamma\left[\frac{v(f)^2}{2g}\right]v(f)A(f) = \gamma H_v(f)Q(f) \tag{24}$$

The value of the velocity head H_v was calculated assuming energy loss follows the Manning's equation, which in addition to estimates of flow rate Q at a particular exceedance probability, also requires estimates of the hydraulic radius at that flow rate, i.e.

$$H_v(f) = \frac{v(f)^2 Ln^2}{R(f)^{4/3}} \tag{25}$$

Channel cross sections were assumed rectangular in the Phase III study. Following this assumption, the hydraulic radius and cross-sectional area can be written as:

$$R(f) = 2d(f) + w(f) \tag{26}$$

and

$$A(f) = w(f)d(f) \tag{27}$$

In determining the hydrokinetic power as defined in equation (24), the flow rate and channel width and depth, must be estimated for each exceedance probability. The total hydrokinetic power is then determined through integration as:

$$\begin{aligned} P_K &= \int_0^1 P_K(f)df \\ &= \gamma Ln^2 \int_0^1 \frac{v(f)^2}{[2d(f)+w(f)]^{\frac{4}{3}}} Q(f)df \\ &= \gamma Ln^2 \int_0^1 \frac{Q(f)^3}{w(f)^2 d(f)^2 [2d(f)+w(f)]^{4/3}} df \end{aligned} \tag{28}$$

It should be noted that in obtaining the hydrokinetic estimate from velocity, equation (28), there were many more estimated variables than using only the flow rate and elevation drop, equation (22). This resulted in compounding the uncertainty associated with the estimated hydrokinetic potential. Integration of the power over the range of exceedance probabilities in both cases was done using the trapezoidal rule for numerical integration.

5.2.6 Uncertainty in Flow and Power Estimates

To characterize uncertainty in the estimated hydrokinetic power for each stream reach, the uncertainty was thought of as arising from estimated flow rates and channel geometry elements. Jenkinson and Bomhof (2014) described a Monte Carlo approach to address this uncertainty. It is difficult to comprehend their approach due to inadequate explanation. In addition, some serious statistical issues were also noted with the approach and therefore no additional insights are offered here, except some general observations. They adopted 1,000 Monte Carlo simulations for MMFs and FDCs to derive 95% uncertainty bands. Uncertainty was assumed only in the flow rates and not in the channel geometric elements, contrary to what was described in the report. For each simulation, channel geometric elements were described as a function of flow rate. These relationships were derived from the data available in the EC-MDB database (Bomhof 2013). For applying equation (28), channel geometric elements were obtained from the following relationships:

$$w(f) = a_w[Q(f)]^{b_w} \tag{29}$$

and

$$d(f) = c_d Q(f)^{f_d} \tag{30}$$

The coefficients a_w, b_w, c_d and f_d were derived from data available in the EC-MDB (Bomhof 2013). The above equations are similar to those described previously in Chapter 4, except subscripts are introduced here to differentiate the coefficient f_d from the placeholder f that represents the exceedance probability for the FDC. For each channel reach, 1,000 Monte Carlo simulated FDCs (in the form of five selected percentile flows) were used to determine associated channel width and depth using the above two equations. In reality, contrary to the claim by Jenkinson and Bomhof (2014), uncertainty associated with the channel flow-geometry relationships were not accounted for because these relationships were developed through a separate process and were associated with a different set of uncertainties. In other words, the uncertainty inherent in flow-geometry relationships was not considered.

5.3 Results

Following the methodology and data sources described in Section 5.2, Jenkinson and Bomhof (2014) developed two sets of outputs to meet the study objectives. First was a national dataset that included stream velocities and hydrokinetic potential for each stream reach and the second was a power map that included (theoretically extractable) maximum power associated with each stream reach. The maps of hydrokinetic energy were designed for examining locations with a high power potential. The underlying data was not integrated regionally for estimating regional resources. Compared to this, extractable power product was integrated over different regions and large drainage basins to determine the respective resource potential.

Extractable power was estimated by two different methods, as outlined in Section 5.2: (1) using the hydropower potential estimates (in terms of elevation head) as described by equation (22), and (2) using estimates of channel geometry to directly estimate hydrokinetic potential as described by equation (23). These estimates represented national resources and did not account for the energy that was being already extracted through existing dams and other structures.

5.3.1 National Geospatial Datasets: Estimates of Flow Rate, Velocity and Hydrokinetic Power

Following the steps described in Section 5.2, estimates of flow rate, velocity and hydrokinetic potential were generated for each stream reach within the national river network to develop a national database. Power estimates were summarized and demonstrated in the form of spatial maps (not included here in this report) for different exceedance probabilities. These maps showed large areas in northern Canada with no power estimates. This was due to the lack of available data. The resolution of the available satellite data from northern Canada was too coarse to support reliable estimates of physiographic and other attributes required for the MR-CCA analysis. Secondly, some areas in the Prairie Provinces were also not processed, due to the presence of non-contributing areas (PFRA 2008). According to Jenkinson and Bomhof (2014), it was difficult to predict FDCs based only on physiographic and climatic attributes in these areas. In order to increase accuracy of predicted flow rates in the Nelson River basin, which also includes the prairies region, the value of the parameter alpha for the CCA approach was required to be relatively high which in turn increased the number of stream reaches, with insufficient neighbours for the application of the MR-CCA method.

5.3.2 Influence of CCA-Based Neighbourhood Identification on Estimated Flows

To demonstrate the usefulness of CCA-based delineation of neighbourhoods and then estimating target streamflow indices, Jenkinson and Bomhof (2014) used the notion of normalized prediction intervals. This demonstration was conducted by predicting and comparing percentiles of the FDC corresponding to 50% exceedance probability for three cases: (1) without using CCA-based neighbourhoods (i.e. by assuming the parameter alpha to be equal to zero), (2) using neighbourhoods identified with parameter alpha = 0.0001, and (3) using neighbourhoods identified with parameter alpha = 0.0075. In the comparison, a reduction in the prediction intervals was noticed for the entire country. In the third case, larger areas of the stream network emerged with “no estimate” in different parts of the country, especially in Manitoba and the North. This highlighted the trade-off between the need of having the estimated flows available all over the network to support development of a national database and the required high accuracy of predicted flows. Through this comparison, Jenkinson and Bomhof (2014) demonstrated the value of using CCA-based neighbourhoods.

5.3.3 Regional Hydrokinetic Power and Hydropower Estimates

Integrated or total power available for each stream reach was calculated both as hydropower (as a surrogate for hydrokinetic power) and solely as hydrokinetic power, separately for each province/territory and for each of the large drainage basins, considered in the study.

<u>Hydropower Estimates</u>

The total hydropower potential, along with uncertainty intervals, were developed by province and by region based on large drainage basins. The results were shown in the form of bar graphs (on a logarithmic scale), as well as in the form of tables, showing the median values and the corresponding 2.5% and 97.5% point estimates. There is an estimated 710 GW of total potential hydropower in Canada with the most power concentrated in British Columbia (220 GW), followed by Quebec (120 GW). As parts of northern Canada were not considered due to insufficient data, it was speculated by Jenkinson and Bomhof (2014) that there is a possibility that Nunavut and the Northwest Territories may have a higher power potential.

Hydrokinetic Estimates

The total hydrokinetic potential, along with the uncertainty intervals were developed by province and by region based on large drainage basins and the results were shown in the form of bar graphs (on a logarithmic scale), as well as in the form of tables, showing the median values and corresponding 2.5% and 97.5% point estimates. There is an estimated 340 GW of potential hydrokinetic power in Canada with the most power concentrated in the Northwest Territories and Quebec (respectively 73 GW and 74 GW), followed by British Columbia (45 GW), Ontario (42 GW) and Manitoba (28 GW). The degree of uncertainty in the estimates was found to be much higher for this technique than that noticed with the hydropower technique.

Comparison of Regional Power Estimation Techniques

A comparison between both methods showed that the prominent difference between the two methods was in the form of greater degree of uncertainty associated with the hydrokinetic power estimates, compared with the hydropower estimates. In addition to regional comparisons, this disparity was also demonstrated at the national scale. The median estimates in both cases were found similar. The greater level of uncertainty associated with hydrokinetic estimates was due to the compounding influence of the additional parameters (i.e. channel width and depth) that were required to be estimated for the hydrokinetic case. In theoretical terms, the maximum total hydrokinetic potential should be comparable to the hydropower potential. The results of the study satisfy this concept as the range of hydropower estimates were contained within the range of hydrokinetic estimates. At the end, Jenkinson and Bomhof (2014) recommended using that method which is relatively more precise for quantifying the national resource.

5.4 Overview and Outcomes

Hydrokinetic resource assessment study was conducted in three separate phases. Analyses, investigations, and outcomes from each of the three phases were documented in three separate technical reports. In this chapter, the work undertaken in Phase III and documented in Jenkinson and Bomhof (2014), is reviewed from a hydrological standpoint. The methodology adopted in Phase III was driven mainly by numerous recommendations that came out of Phase I and Phase II and was strengthened further in Phase III for application across whole of Canada.

The results of Phase III study suggest that Canada has a massive resources of hydropower, even when estimates of uncertainty are considered. The expected value of total energy is approximately 710 GW, with the lower end of the uncertainty interval at 430 GW. This very large number represents all energy and that in fact cannot be extracted from the system due to a number of reasons, e.g. many locations being not even be feasible for energy extraction. At the country scale, British Columbia and Quebec were found to have the largest overall hydropower resources. Based solely on the hydrokinetic energy assessment procedures, it was found that the total energy potential is approximately 340 GW, with the lower end of the 95% uncertainty interval at 29 GW. Jenkinson and Bomhof (2014) note that due to the uncertainties associated with channel geometry related elements, the uncertainty intervals for the latter case were very large.

Most of the data products related to hydrokinetic resources were developed to analyze and investigate resource locations across the country. Jenkinson and Bomhof (2014) offered some recommendations for practitioners on the use of these products, including the following list (the text is modified from the original source where appropriate for technical and scientific reasons):

- It is useful to consider closely the uncertainty (or prediction) intervals for all relevant data as some Canadian regions are associated with a higher degree of confidence in various estimates compared to other regions, stemming from the nature of underlying regression relationships and supporting datasets.
- It was found that the greatest degree of uncertainty in the hydrokinetic power estimates was due to higher uncertainties in the estimates of channel geometry related elements. If a region is associated with high energy potential, the estimates of channel geometry related elements could also be obtained from the areal-photography or satellite imagery of the region. Contrary to this recommendation, the investigations of these data products conducted in Phase II did not support this recommendation.
- Incorporation of design limitations for turbine deployment can be done by filtering developed geospatial datasets. For example, filtering estimates of channel depth, width and flow rate.
- Inclusion of feasibility factors like proximity to the electrical grid or communities could readily be done with the standard GIS software, with integrated spatial analysis tools.

The resource assessment study could be expanded and improved by examining the following research questions (Jenkinson and Bomhof 2014):

- A small number of percentile points of the FDC were employed in the CCA approach for delineating prospective neighbourhoods. Additional experimentation can be done to determine the best combination of various percentile points of FDCs and physiographic and climatological attributes to delineate most robust neighbourhoods.
- The neighbourhoods based on the CCA approach were established for each major drainage basin within Canada. Other methods for determining neighbourhoods could be investigated.
- Additional techniques could be developed to rapidly refine hydrokinetic resource estimates with the addition of new constraining datasets, including channel geometry. However, how

this can be accomplished was not discussed. Channel geometry has already been flagged as one of the challenging factors in the study.

- There is limited research available on hydraulic effects of kinetic turbine installations in river reaches and therefore, lab experiments and field testing should be conducted to validate computer modelling efforts.

5.5 General Observations and Remarks

The following points were observed when reviewing the Phase III report of the resource assessment study. It is important to note that some of these points are quite critical for gauging the quality of all selected approaches and various datasets used to support national assessment of hydrokinetic resources. These points are also important to appraise the real value of various data products that came out of the hydrokinetic resource assessment study.

- For assessing national hydrokinetic resources, 11 large drainage regions covering most of Canada were considered. After delineating homogeneous neighbourhoods for each of these regions using the CCA approach, selected percentiles of FDCs were estimated through MR. For the application of MR, no information was provided regarding how many stations from the available number of stations within each region were considered. In addition to this, it is also unclear how many attributes were retained for each region and how many of these were actually used in the MR analysis. Though not explicitly stated, it looks like a constant neighbourhood was used for each of the 11 drainage regions. If this was the case, then it was important to discuss the gauging stations which were not included in the neighbourhoods. This had shed some light on the regional or localized behaviour of certain hydro-meteorological phenomena from an explanatory viewpoint and to explore existing physical linkages to understand any peculiar behaviour.
- It is unclear from the description provided if the same attributes, that were identified for the CCA approach and for forming the set of key attributes, were also used in the MR analysis.
- It looks like Jenkinson and Bomhof (2014) used a constant neighbourhood for each region, but different from one region to the other. From a physical viewpoint, it was not a good idea to use a constant neighbourhood for each drainage region. There is certainly some degree of non-homogeneity within these large drainage regions, and it is very likely to have a few number of completely irrelevant stations within the neighbourhood of a given target station. Therefore, some thoughts should have been put into adopting variable neighbourhoods within a region depending upon the target stream reach.
- It was stated in the report by Jenkinson and Bomhof (2014), that “once the best neighbourhoods were found for all regions, each of the flow estimates (i.e. FDC, MMF, and MAF) were predicted using the MR approach …”. In fact, no independent regression analysis was performed for MAFs. These flows were generated from estimated FDCs and MMFs (mentioned in the Phase III report, though not reported) through numerical

integrations. It is important to mention that no independent analysis was conducted for MMFs in the Phase III study. Therefore, there is no reason to surprise the reader with irrelevant information. Regression analyses were carried out only for selected percentiles of FDCs.

- As with estimated flows, there is a considerable degree of uncertainty associated with the estimated channel widths and cross-sectional areas, obtained using the flow-geometry relationships. Jenkinson and Bomhof (2014) described a Monte Carlo simulation based procedure to evaluate uncertainty in the estimated channel geometric elements. Unfortunately, due to lack of sufficient details and theoretical justifications, it was difficult to comprehend their approach and to discuss how their approach can be assessed and improved.
- It is mentioned in the methodology section of the Phase III report that various data products that will emerge from this study (e.g. estimates of channel velocity, geometry and flow) will allow for any limitations to be overcome in post-processing using the GIS or similar tools. However, no guidance was provided on how this will be made feasible when an investigator will screen these datasets for site selection and deployment purposes.
- When estimating total power through integration for a given stream reach, the integration limits suggest that the power was calculated over the entire range of the FDC. However, only five selected percentiles were estimated in the whole process. Therefore, it is unclear how the integration for calculating total power for a given stream reach was conducted. In real situations and also from the nature of the Phase II validation study, one would expect the stream depth and width to vary spatially across the entire reach and that, in fact, was not the case since the depth and width were taken as functions of flow rate, which was estimated for the entire stream reach. Thus, spatial variations in depth, width and velocity were not taken into account. This point has not been explained explicitly in the Phase III report. Power assessments using such assumptions also have serious implications for turbine deployment purposes, since a stream reach associated with a considerable hydrokinetic potential may not actually be suitable for turbine deployment. Thus, the users of the generated data products need to be aware of this aspect that could lead to spurious positive signals for available power potentials.
- When estimating integrated power for a stream reach using the flow rate, elevation head and the mean annual volumetric flow rate, determined by integrating the predicted FDC for the reach, it is unclear if the MAF was derived from the entire FDC or just from the five estimated percentile flows corresponding to five selected exceedance probabilities (i.e. 0.1, 0.5, 1.0, 5.0, 30.0, and 50.0%). In the former case, it is also unclear how extremal points of the FDC were estimated, while in the latter case, it is possible that the real power potential may have been underestimated inadvertently.

6 An Overview and Summary of the Resource Assessment Effort

The review of the resource assessment study presented in Chapters 3 to 5 is interlinked and therefore, here, a general overview of the entire study is presented first, followed by a concise summary of each of these three chapters. This will facilitate an easy comprehension of the entire effort that spanned over almost five years from 2010 to 2014 period, financial supported by Natural Resources Canada (NRCan).

6.1 An Overview

Due to increasing interests in renewable energy globally, importantly from untapped hydrokinetic resources, NRCan commissioned a multiyear study to assess Canada's hydrokinetic potential in an effort to boost renewable energy resources for domestic and industrial consumption. The ultimate aim of the study was to develop resource databases, which governments and regional entities could use for investment planning and decision-making, as well as to support private market development. The National Research Council Canada (NRC) led this effort through an inter-departmental agreement between NRCan and the NRC and completed an assessment of hydrokinetic resources at the large regional and national levels. This assessment was completed in three different phases and outcomes from each of these phases were documented in three separate technical reports (i.e. Jenkinson 2010 and Jenkins and Bomhof 2012, 2014). The major objective of Phase I of the study was to conduct a review of various methods and data sources and to identify suitable techniques for conducting an assessment of hydrokinetic potential at both regional and national levels in subsequent phases of the study (i.e. Phase II and Phase III). In Phase II, before conducting the national assessment, a couple of selected methods were evaluated using data from the national hydrometric network, maintained by Water Survey of Canada, for four selected large geographic regions (see Figure 7) in order to guide the country-wide assessment. Based on the outcomes and recommendations from both Phase I and Phase II of the study, a nation-wide assessment of the hydrokinetic potential for theoretical energy extraction was carried out in Phase III, which was also the concluding phase of the study. The work conducted within each of the three phases of the study is respectively reviewed in Chapter 3, Chapter 4 and Chapter 5 of this report.

6.2 Summary

In general, for hydrokinetic resource assessment in a given region of interest, hydrologic investigations and hydraulic modelling are ideally required and both are required to be carried out in tandem. Hydrologic investigations pertain to estimation of desired streamflow indices (e.g. monthly and annual mean flows, percentiles of flow duration curves, etc.) at all potential points of interest within the wider region of interest, while hydraulic modelling part pertains to the estimation of river flow velocities at all points of interest. The methods and data sources pertaining to hydrologic investigations are reviewed in this report, while the hydraulic modelling aspects are not assessed. Where applicable, shortcomings are highlighted and recommendations are made for additional research in order to obtain improved estimates of various streamflow indices, which play a crucial role in estimating hydrokinetic resources. The main findings of the resource assessment study, which are reviewed and documented in Chapters 3 to 5, are discussed below.

In Chapter 3, review of existing literature pertaining to transposition of streamflow indices from gauged to ungauged locations within the realm of ungauged hydrology and various aspects of hydraulic modelling associated with hydrokinetic resource assessment, originally documented in Jenkinson (2010), is presented. A significant amount of literature was available on the former topic even at the time the review was conducted, but only selected studies were considered. Compared to this, hydraulic aspects were addressed adequately. The study by Jenkinson (2010) identified that graphical characterization of flow duration curves (FDCs) in combination with canonical correlation analysis (CCA) within the regionalization framework was a promising approach for estimating streamflow indices at ungauged locations. However, no credible study was noted where such a combination was explored previously, specifically for FDCs. Within the regionalization context, FDCs were not explored in the literature as much as some other indices of streamflow (e.g. extreme low or high flows). For estimating channel geometry related elements (e.g. channel width, depth and velocity), discharge-based approaches were found quite common, however, other approaches that relate channel geometry to drainage area and physiographic and climatological characteristics of a given watershed were also noted. A number of national datasets were identified to support estimation of required variables for assessing national hydrokinetic resource potential. This was a significant contribution of Phase I of the study. Identified datasets included in the study were: climate data, hydrometric data, digital soil and land use maps, hydro network maps and digital elevation data. A database that contained cross-sectional information, and estimates of streamflow and velocity at all surveyed stations across Canada was also a significant finding. This database was used to validate various selected techniques for estimating velocity duration curves and channel geometry in other phases of the study. In short, the effort by Jenkinson (2010) focused mainly on how to employ existing techniques at regional levels by identifying various hydrologic, physiographic and climatic datasets to be used in hydrokinetic resource assessment study across Canada. However, no attempt was made to explore and discuss how the identified techniques could be innovated or improved further.

In Chapter 4, the work completed within the scope of Phase II of the resource assessment study is presented. In this phase, the main focus was on validating a suitable methodology based on hydrologic regionalization concept, identified in Phase I, using observational data from hydrometric stations located within four selected large geographic regions of Canada. Most of this work was in the form of a proof of concept so that the verified concepts and methodologies can be taken forward to complete Phase III of the study. For evaluation purposes, a dataset of watershed attributes, representing climatic, geologic, land use and soil types, was developed. This was undoubtedly a significant undertaking. Simple multiple linear regression combined with the CCA approach (MR-CCA) and just the multiple regression (MR), without the consideration of CCA, were evaluated. It was concluded that in most cases the MR-CCA approach performed better than the MR approach alone. Another established approach (i.e. the RETScreen procedure of NRCan (2004a, 2004b)) was also evaluated. In comparison to FDCs, MMFs were found to be better simulated in most of the selected geographic regions. Finally, Jenkinson and Bomhof (2012) recommended to replace the RETScreen procedure with the CCA-based approach for delineating neighbourhoods or groups of stations with similar physiographic and climatic attributes for estimating selected percentiles of FDCs, and MR-based approach for channel geometry, in combination with the continuity equation, for estimating hydrokinetic energy within the entire national river network. Estimation of water depth based on power form channel geometry relationships were recommended for turbine deployment purposes. For applying the MR and MR-CCA approaches, a database of 71 different physiographic and climatic attributes was developed. However, no discussion was provided on the type and definitions of various attributes that were used in each of these two regression based approaches. This was necessary to understand the influence of various attributes on regression relationships and uncovering their role in predicting various parts of the FDC as well as the annual cycle in terms of MMFs. In addition to this, a number of other shortcomings were also noted in the analyses conducted in Phase II of the study. These are described in detail in Section 4.9 of Chapter 4 in this report.

In Chapter 5, an overview of the work completed in Phase III of the study from Jenkinson and Bomhof (2014) is presented. In this phase, Jenkinson and Bomhof (2014) focussed on developing estimates of hydrokinetic resources on a national scale, following the methodology and concepts identified in Phase I and validated in Phase II. The results obtained in Jenkinson and Bomhof (2014) suggest that Canada has a massive resources of hydropower, even when considering some estimates of uncertainty. The expected amount of total energy was estimated to be approximately 710 GW, with the lower end of the uncertainty interval at 430 GW. This very large value represents all energy and some of that even cannot be extracted due to a number of constraints, e.g. many locations are just not feasible for energy extraction. At the country level, British Columbia and Quebec were found to have the largest overall hydropower resources. Based solely on the hydrokinetic energy assessment, it was found that the total energy potential is approximately 340 GW, with the lower end of the 95% uncertainty interval at 29 GW. Jenkinson and Bomhof (2014) noted that due to the uncertainties associated with channel geometry related elements, the uncertainty intervals for the latter case were found to be very large. They offered some recommendations for practitioners on the use of the resulting data products and improving estimates of hydrokinetic potential including: (1) experimenting with different combinations of various percentiles of the FDC and physiographic and climatological attributes for delineating neighbourhoods through the CCA approach, and (2) developing other techniques than the one used in their study to rapidly refine hydrokinetic resource estimates with the addition of new constraining datasets, including channel geometry. Apart from the above mentioned findings and estimates, a number of assumptions and theoretical considerations were not adequately explained in Jenkinson and Bomhof (2014) and reader was left with numerous unanswered questions. Some of these issues have been discussed in Section 5.5 of this report, where additional information can be found.

7 Future Considerations and Research Avenues

7.1 Background

For hydrokinetic resource assessment in a given region of interest, detailed information on both hydrologic and hydraulic aspects for the entire regional river/stream network is generally required. This information is derived mostly through hydrologic investigations combined with hydraulic modelling. In hydrologic investigations, which is the main focus of this report, long-term streamflow records play a critical role in deriving information on some useful indices of streamflow regimes. The streamflow indices that are often considered for hydrokinetic resource assessment are the mean monthly flows (MMFs), mean annual flows (MAFs) and flow duration curves (FDCs). The use of FDCs is more common compared to the other two indices. As already explained before in Chapter 1, a FDC establishes a time-independent relationship between streamflow magnitudes and their frequencies of occurrence. It is well known that the majority of the Canadian river network is ungauged and recorded observations are available more frequently in southern parts of the country and much less so for northern regions. In the absence of recorded observations, information on streamflow indices at ungauged locations is obtained through indirect means, e.g. by transposing known or processed information from gauged to ungauged locations, following a credible methodology.

Hydraulic modelling component pertains to obtaining information on channel velocities or velocity duration curves, which are analogous to FDCs, at various points of interest along a river reach, after establishing streamflow-geometry relationships. As the hydrokinetic power at a given point in a river reach is proportional to cube of the velocity of moving water, reliable estimates of various streamflow indices are very important.

For the estimation of streamflow indices at ungauged locations, numerous techniques abound in the literature. These techniques include drainage area ratio methods (e.g. Mohamoud and Parman 2006), parametric characterization of FDCs (e.g. Castellarin et al. 2004), graphical characterization of FDCs (e.g. Castellarin et al. 2004; Smakhtin and Masse 2000; and Shu and Ouarda 2012), and various variants of the regression framework, developed mainly on the basis of hydrologic homogeneous regions or groups of watersheds/gauging stations with similar attributes of interest (e.g. FSR 1975; Haan 2002). An important feature of these techniques is that one seeks regression relationships between streamflow indices (e.g. MMFs) and watershed attributes (e.g. drainage area) from gauged locations and then transposes those to target ungauged locations, with known attributes. Recent research has shown that if a set of nearest-neighbours (i.e. gauged sites with watershed attributes similar to those of the target ungauged site within a geographic space or within the attribute space) can be identified then the reliability of estimated streamflow indices can be improved at the ungauged location (e.g. Burn 1990b; Zrinjti and Burn 1994; Tasker et al. 1996; Ouarda et al. 2000, 2001; Cavadias et al. 2001; Eng et al. 2005; Shu and Ouarda 2012). It is important to note that there is no consensus on the use of geographic space or attribute space for defining hydrologic similarity and therefore it remains an open question in statistical hydrology. Furthermore, lack of available gauged locations within a given geographic region remains a serious issue in achieving hydrologic similarity solely within the geographic space (Khaliq et al. 2015).

As described in Chapters 3 to 5 and summarized in Chapter 6, selected percentiles of FDCs were estimated at all target stream reaches by employing multiple regression (MR) in combination with the canonical correlation analysis (CCA) approach. The CCA approach was used to identify nearest-neighbours and then the MR was used to estimate desired percentiles based on a set of physiographic and climatic attributes. The estimated percentiles of FDCs were used to estimate time-averaged hydrokinetic potential for all stream reaches. Altogether, 71 different physiographic and climatic attributes were considered. These attributes included geographic, geologic, climatic, soil, and land cover characteristics. A number of shortcomings pertaining to statistical applications and reliability of various attributes are identified in the analyses documented by Jenkinson and Bomhof (2012, 2014) and some of these have already been highlighted in Chapters 4 and 5. These shortcomings can be overcome by adopting some well proven approaches and performing systematic analyses with a rational thinking, as explained and echoed in the following sections.

7.2 Supporting Datasets and Watershed Attributes

A number of different datasets were considered for the hydrokinetic resource assessment study. Most of these datasets have already been described in Chapters 3 to 5 in different contexts. For developing a set of watershed attributes, Jenkinson and Bomhof (2012) considered Canadian digital elevation data, soil characteristics data, land cover data, surficial geology data, and climatic data. The watershed attributes were derived for over 365,000 river segments from the Canadian digital stream network (Natural Resources Canada 2008), which were divided by major drainage basins (see Figure 10). The drainage area for every river segment was delineated by Jenkinson and Bomhof (2012) based on Voronoi tessellation of the vectorized stream segments using Esri ArcGIS.

Though derivation of various watershed attributes was truly a considerable undertaking, no information on the derived attributes was documented in the associated technical reports (i.e. Jenkinson and Bomhof 2012, 2014). As a consequence, description of these attributes was not provided in Chapters 3 to 5 of this report. Subsequent to the resource assessment study, the same datasets and derived attributes from NRC’s archives were also considered in Khaliq et al. (2015), who estimated MMFs at all stream reaches of two large hydrologic regions of Canada (i.e. Pacific region on the west coast and Maritimes region on the east coast) within the framework of a study sponsored by Environment Canada. These authors provided some information on the nature and character of 43 of the all physiographic and climatological characteristics, originally considered in Jenkinson and Bomhof (2012). These characteristics are briefly explained below and their definitions are provided in Table 2:

- From the Canadian digital elevation data, MinElev, MaxElev, MeanElev, MedElev, and StdDevElev were derived.
- DrainageIndex, KSAT, and KP0 were derived from soil characteristics data.
- Land use variables, derived from land use maps, were generalized into 10 unique land classes by Jenkinson and Bomhof (2012) from 45 different land cover categories. From these classes, Lakes, Barren, Developed, Shrublands, Wetlands, Grasslands, Croplands, and Forest were considered.
- Canadian surficial geology was separated into 26 classes from the available maps. These classes were lumped together into 9 different categories by Jenkinson and Bomhof (2012), i.e. GeoGlaciers, GeoLakeMud, GeoLakeSand, GeoMud, GeoPeat, GeoRock, GeoSandGravel, GeoTill, and GeoWater.
- Climatic indicators for the 1961 to 1990 period were derived from a national dataset and included AnnPrecRain, AnnPrecSnow, AnnPrecTotal, MinAnnTemp, MaxAnnTemp, MeanAnnTemp, PE, GSS, GSE, GSL, GDD0, GDD5, GDD10, and GDD15.

As mentioned above, additional description in terms of definitions of these climatic indicators is provided in Table 2.

Table 2: Watershed attributes considered for finding nearest-neighbours and estimation of MMFs.

Attribute	Description	Attribute	Description
DrainageArea	Drainage area [km^2]	MeanAnnTemp	Mean annual temperature [°C]
Perimeter	Watershed perimeter [m]	PE	Potential evapotranspiration [mm]
CentroidLat	Latitude of the centroid of the watershed [°N]	GSS	Growing season start date [Julian day]
CentroidLong	Longitude of the centroid of the watershed [°E]	GSE	Growing season end date [Julian day]
MinElev	Minimum elevation of the watershed [m]	GSL	Length of growing season [days]
MaxElev	Maximum elevation of the watershed [m]	GDD0	Growing degree days above 0°C [.]
MeanElev	Mean elevation of the watershed [m]	GDD5	Growing degree days above 5°C [GDD]
MedElev	Median elevation [m]	GDD10	Growing degree days above 10°C [GDD]
StdDevElev	Standard deviation of elevation [m]	GDD15	Growing degree days above 15°C [GDD]
Lakes	Proportion of watershed containing lakes [.]	DrainageIndex	Scaled drainage index [categorical]
Barren	Proportion barren land [.]	KSAT	Saturated hydraulic conductivity [m/day]
Developed	Proportion developed land [.]	KP0	Soil permeability
Shrublands	Proportion shrublands [.]	GeoGlaciers	Geological class glaciers [.]
Wetlands	Proportion wetlands [.]	GeoLakeMud	Geological class lake mud [.]
Grasslands	Proportion grasslands [.]	GeoLakeSand	Geological class lake sand [.]

Croplands	Proportion crop lands [.]	GeoMud	Mud type geological classes [.]
Forest	Proportion of watershed with forests [.]	GeoPeat	Geological class peat [.]
AnnPrecRain	Amount of mean annual rain [mm]	GeoRock	Rock type geological classes [.]
AnnPrecSnow	Amount of mean annual snow [mm]	GeoSandGravel	Combined sand and gravel classes [.]
AnnPrecTotal	Total mean annual precipitation [mm]	GeoTill	Geological class thick and continuous till [.]
MinAnnTemp	Minimum annual temperature [°C]	GeoWater	Geological class water [.]
MaxAnnTemp	Maximum annual temperature [°C]		

The above mentioned attributes were identified for individual stream segments as well as for 898 hydrometric stations, with near natural flow regimes, from Environment Canada’s HYDAT network. These attributes were employed for identifying hydrological neighbourhoods based on the CCA approach and then developing regression relationships for estimating selected percentiles of FDCs at all ungauged streams across Canada for the resource assessment study. It is important to note that according to the CCA approach each target stream reach is associated with its own neighbourhood or a group of nearest-neighbours. Thus, it is reasonable to expect that the size as well as the formation of the neighbourhoods will vary spatially from one location to the next within a target region of interest.

Screening of Attributes

Altogether, 71 different physiographic and climatic attributes were considered for the resource assessment study. However, having a larger set of attributes does not guarantee that the resulting regression relationships or groups of nearest-neighbours will bear higher degree of reliability because it is very likely that many of the attributes are mutually correlated. Therefore, all attributes need to be screened individually through pair-wise plots and on the basis of Variance Inflation Factors (VIFs) (Eng et al. 2005; Fox 2008). To have statistically meaningful attributes, it is important to consider just one attribute from a pair of strongly correlated attributes in order to avoid the influence of multicollinearity, which may lead to irrational regression coefficients. For example, Figure 12 suggests that watershed drainage area and watershed perimeter are highly correlated. Thus, it will be illogical to consider both the drainage area and perimeter in a regression relationship. Consistent with the literature on ungauged hydrology, drainage area is generally preferred in regression relationships and the same should be adopted for hydrokinetic resource assessment studies as well. Pair-wise plots of many other groups of attributes, shown in Figure 13 to Figure 17, suggest that a number of attributes are highly correlated with each other.

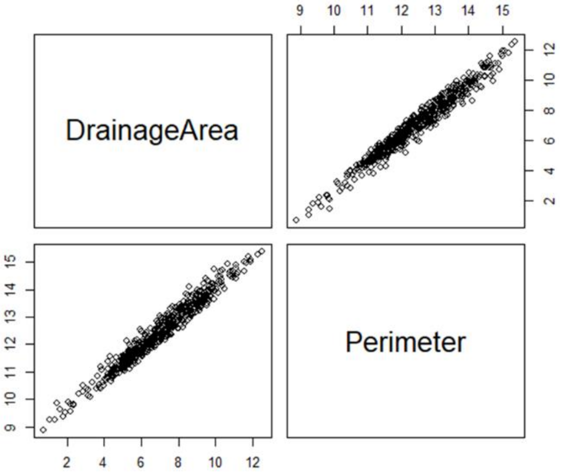

Figure 12: Correlation between watershed drainage area and perimeter (for log-transformed values).

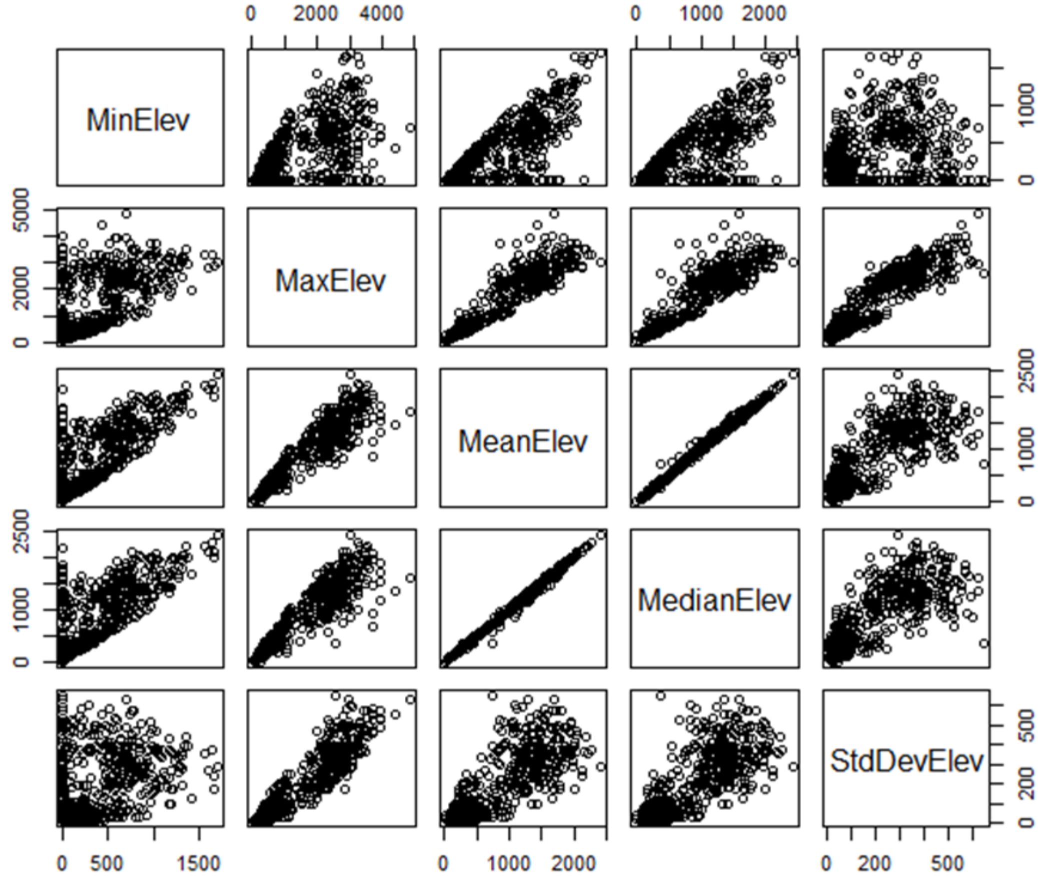

Figure 13: Pair-wise correlation plots for topographic attributes.

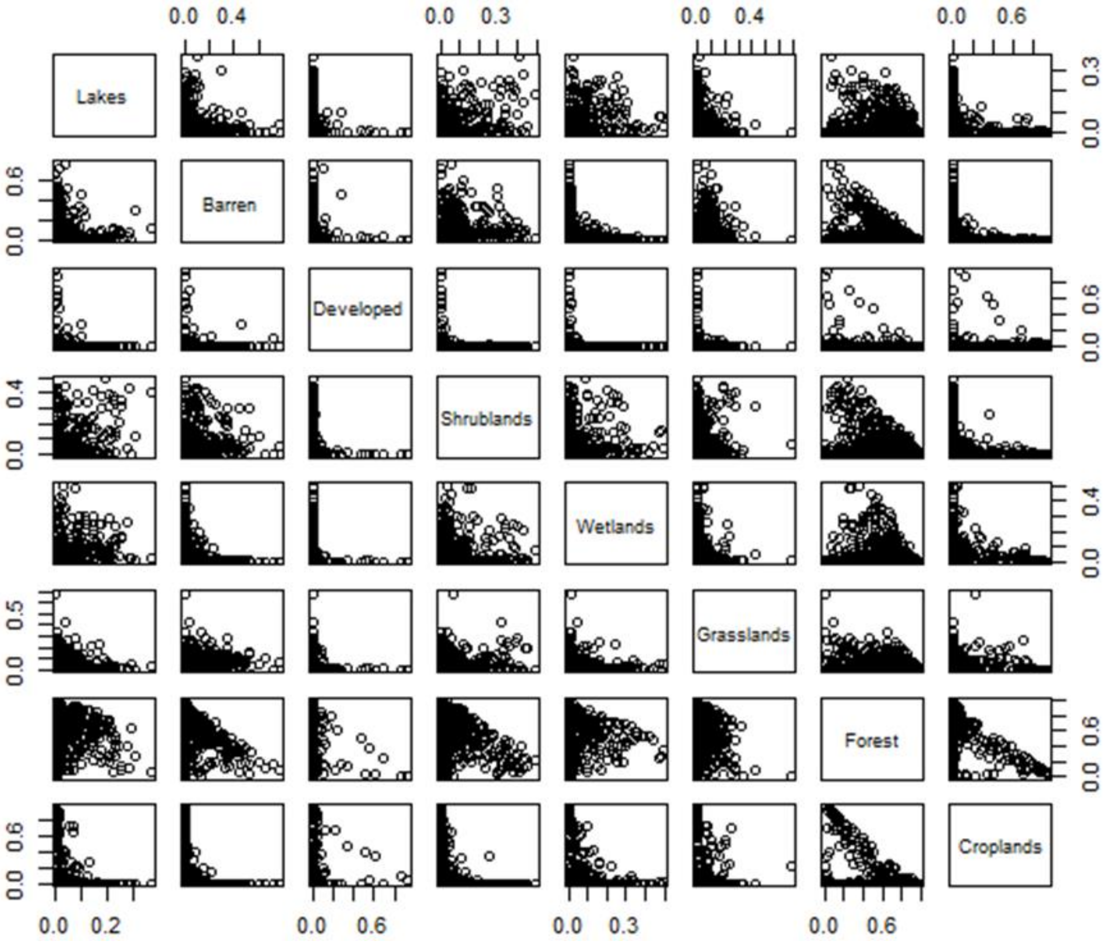

Figure 14: Pair-wise correlation plots for land use related attributes.

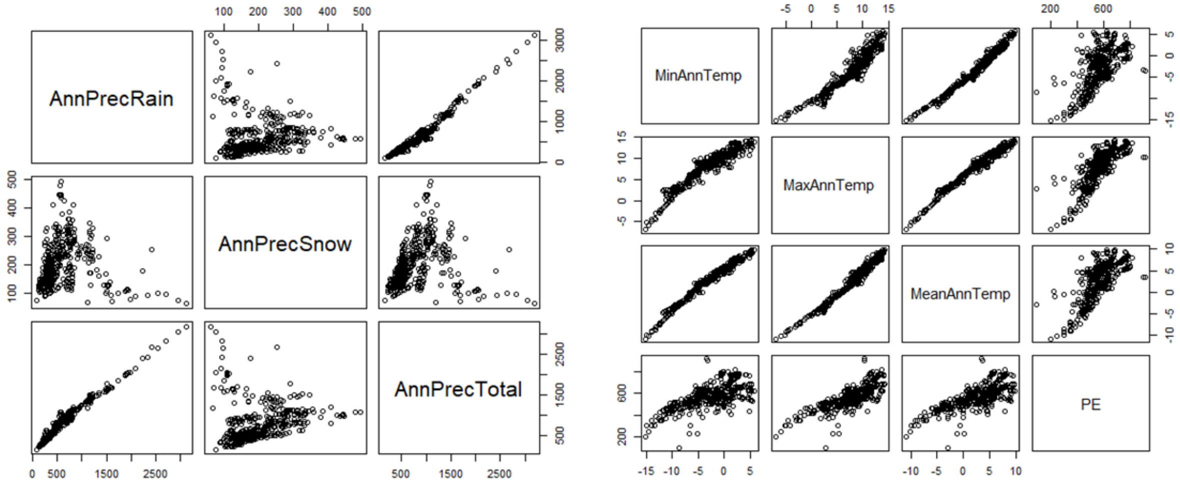

Figure 15: Pair-wise correlation plots for precipitation and temperature related climatic attributes.

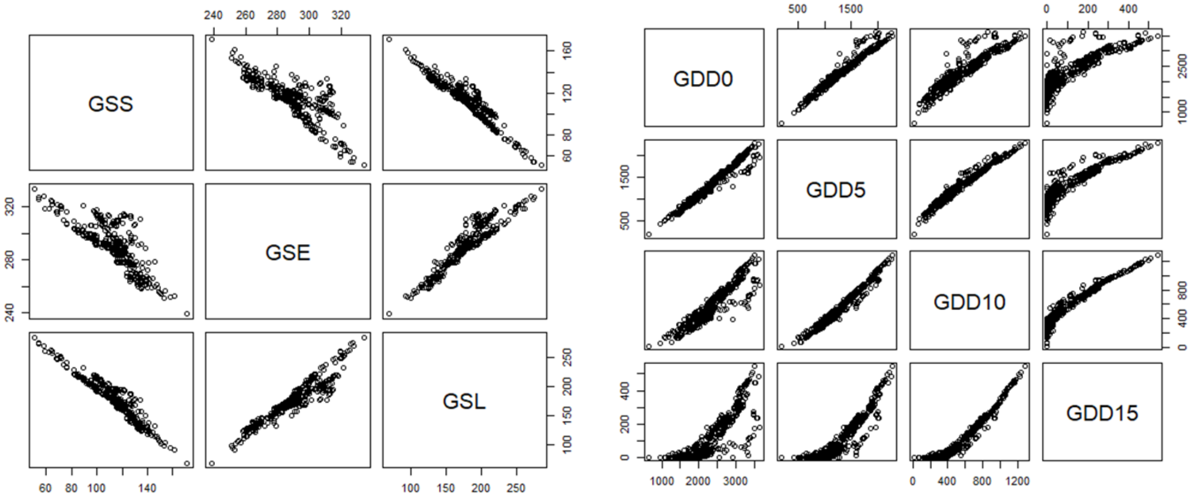

Figure 16: Pair-wise correlation plots for growing season and degree day related climatic attributes.

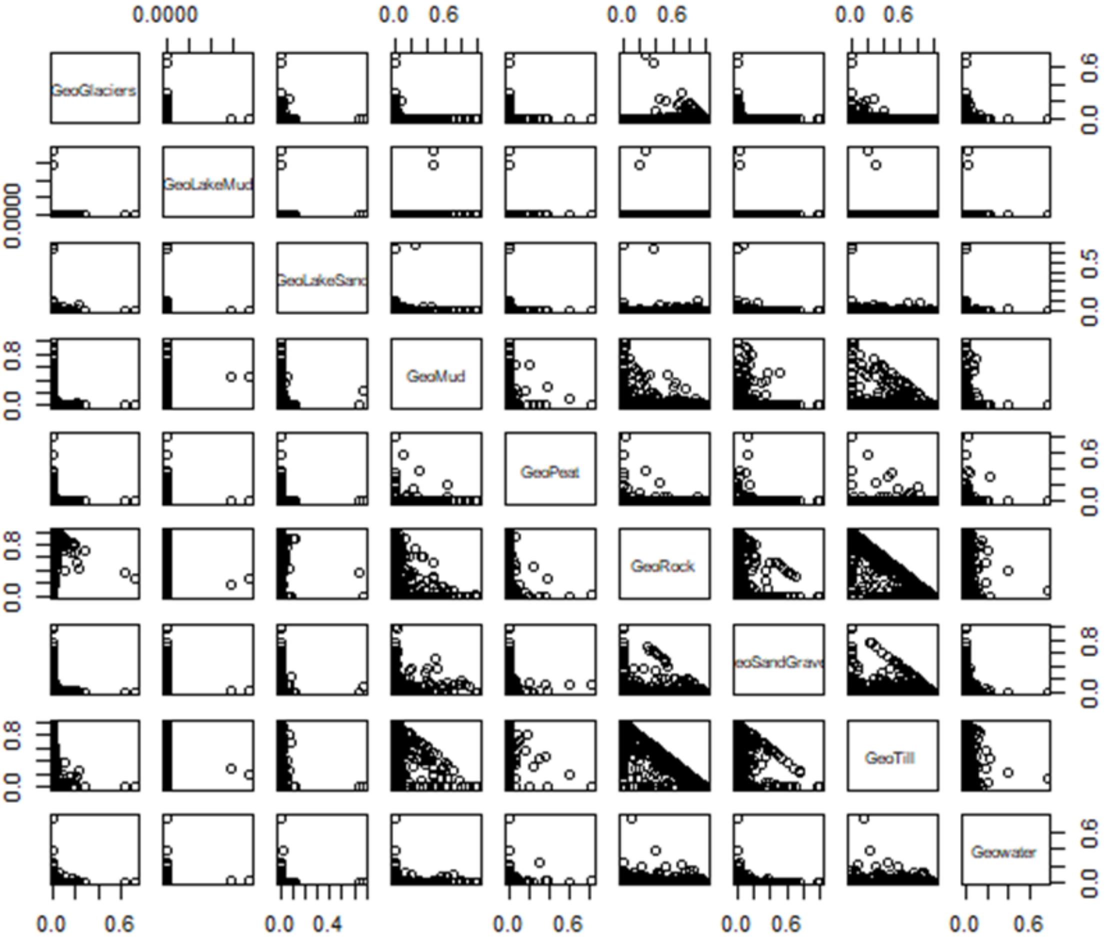

Figure 17: Pair-wise correlation plots for surficial geology related attributes.

The VIF approach can expedite the process of selecting independent attributes from a given set of specific attributes. Following Eng et al. (2005), the VIF is given as:

$$VIF = \frac{1}{1-R_{VIF}^2} \tag{31}$$

where R_{VIF}^2 is the coefficient of determination obtained when the predictor variable of interest (i.e. a selected attribute) is regressed on the remaining predictor variables. A high correlation among the predictor variable of interest and the other variables will result in a large value of R_{VIF}^2 and that, in turn, will lead to a large value of VIF, and vice versa for low correlations. According to Montgomery et al. (2001), a value of the VIF greater than 5 to 10 would be indicative of significant multicollinearity. For stringent requirements a smaller cut-off threshold can also be used. For example, Khaliq et al. (2015) used a cut-off value of 2 in their study on the estimation of MMFs. Some information on this aspect is also available in Fox (2008).

Following the results of pair-wise plots and VIFs, Khaliq et al. (2015) dropped CentroidLon from their analysis as this attribute was found to be highly correlated with some of the climatic and land cover related attributes. In addition, due to data reliability and interpolation related issues, KSAT, KP0 and GeoLakeMud were found unreliable and therefore were also dropped from their analysis. From the climate related attributes, AnnPrecRain, MaxAnnTemp, MeanAnnTemp, GSE, GSL, GDD0, GDD5, GDD15, and Forest were also dropped due to the presence of high within-group cross correlations. Similarly, GeoMud, GeoPeat, and GeoRock related to surficial geology; MaxElev, MeanElev and MedElev related to topography; and the watershed Perimeter were dropped as well. The final set contained only 23 attributes, which were used in various analyses related to estimation of MMFs. It is important to mention here that the above noted attributes were not excluded from the analyses presented in Jenkinson and Bomhof (2014) and therefore the reliability of their results could have been compromised.

After screening all attributes and identifying a set of independent attributes, it is reasonable to develop region-wise or country-wide MR relationships for each of the selected indices of streamflow (e.g. selected percentiles of FDCs or MMFs for each of the 12 months). For developing these relationships, it is important to avoid observational records with zero flow values. When finalizing these relationships, it is also important to retain only those attributes that are statistical significant at a chosen significance level, which is commonly taken as 5%. Inclusion or exclusion of an attribute in the regression relationship can be guided through step-wise regression technique. Following these guidelines, Khaliq et al. (2015) developed country-wide regression relationships for each of the 12 MMFs and found that only DrainageArea, AnnPrecSnow, MinElev, Wetlands and GeoTill are the most frequent predictors, meaning that these attributes were statistically significant for the majority of the cases. From these predictors, DrainageArea was the most important predictor as it was able to explain a greater portion of the variability in MMFs. An analysis of standardized partial correlation coefficients (McCuen 2003) also confirmed this. In such situations, naturally one would like to explore the contribution of additional attributes in predicting desired streamflow indices. Therefore, a comparison of three relationships based on DrainageArea only, most frequent predictors, and all significant predictors (which can vary depending upon the selected percentile value of the FDC or the MMF) can be performed. Through a similar analysis, Khaliq et al. (2015) showed that additional predictors did help in explaining some of the unaccounted for variability of MMFs.

7.3 Identification of Neighbourhoods or Nearest-Neighbours

In the context of regional frequency analysis, many approaches are available in the literature for identifying neighbourhoods or nearest-neighbours for estimating target variables of interest at an ungauged location. From these approaches, canonical correlation analysis (CCA) approach was favoured by Jenkinson (2010) and the same approach was later used in Jenkinson and Bomhof (2012, 2014) when estimating selected indices of FDCs for the hydrokinetic resource assessment study. Another comparable approach is based on the concept of region of influence (ROI) and that has also been used in many studies world-wide. Both approaches are described in Chapter 3 of this report. Here, the CCA approach is discussed once again from a different perspective.

The CCA approach finds groups of similar watersheds by correlating a group of streamflow characteristics (i.e. one set of, so called, dependent variables) with watershed attributes (i.e., a second set of variables, obtained through data screening procedures). More specifically, the CCA approach simplifies such a multidimensional dataset so that all of the original variables are represented by new canonical variables, which are made from linear combinations of the original normalized variables such that the correlation of the canonical variables is maximized. If the correlation between the canonical variables is high then it is assumed that one set of variables will be useful for estimating the other set of variables and vice versa (Cavadias et al. 2001). For the case of resource assessment study, one set of variables could be MMFs or selected percentiles of FDCs, and the other set could be all watershed attributes. In order to find nearest-neighbours for a target ungauged site, the location of the site is determined in the canonical space based on site’s attributes, and nearest-neighbours are identified using the Mahalanobis distance measure and an extreme upper quantile of the chi-squared distribution (taken as a cut-off value) corresponding to a selected exceedance probability ‘alpha’. As already explained in Chapter 4, smaller (larger) values of alpha would lead to more (less) nearest-neighbours in the neighbourhood of a target ungauged site. A graphical demonstration of the influence of parameter alpha is provided in Figure 18. In practice, there is a trade-off between achieving a higher degree of similarity, with having only a small number of neighbours in the neighbourhood, and the desired robustness of the relationships derived on the basis of those neighbours. The value of the parameter alpha can be optimized through a cross validation approach by evaluating a set of assumed alpha values. In the cross validation approach, a station from the group of available stations is systematically removed and the CCA approach is applied to the remaining stations to find nearest-neighbours for the removed station. The neighbours found so are used to develop regression relationship to estimate the target flow at the removed station. The estimated flows at all stations are assessed by calculating an assessment criterion (e.g. root mean square error). Graphical plots of the chosen assessment criterion against the range of alpha values can be used to select a suitable value of the parameter alpha for each of the target streamflow indices and those could be all 12 MMFs or selected indices of FDCs or any other indices of interest.

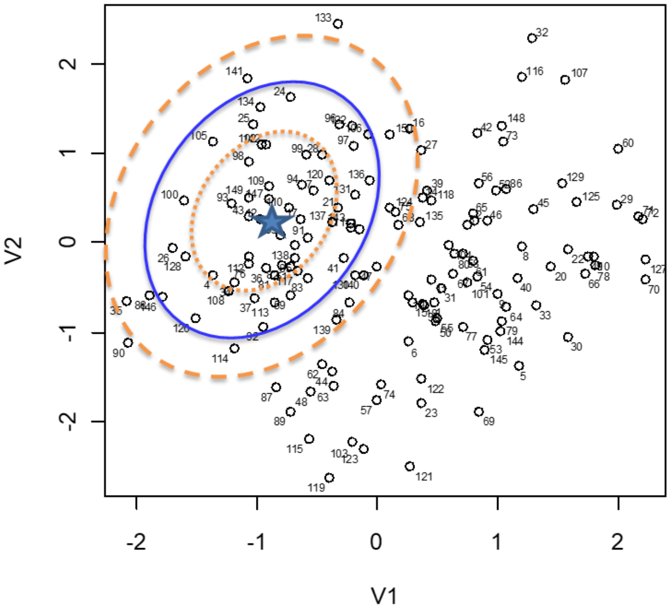

Figure 18: A visual demonstration of the influence of parameter alpha on the size of the neighbourhood. Asterisk shows the target ungauged location within the canonical space and three ellipses correspond to different values of parameter alpha (smaller values of alpha correspond to larger ellipses).

Following the attribute screening procedure and development of a baseline MR relationship, it is possible to identify some attributes that are common across all MMFs or selected percentiles of FDCs. For example, one could decide to retain only those attributes which are found significant for at least (say) six of the 12 MMFs or for 5 out of (say) 7 selected percentile flows of the FDC. This arbitrary criterion can help in reducing the undue noise due to relatively less influential attributes, with the aim to achieve a higher degree of similarity within the neighbourhoods. Since not all MMFs or selected percentile values of the FDCs will have the same number of significant variables in the regression relationships, the above approach is also helpful in overcoming such non-uniform scenarios.

7.4 Estimation of Streamflow Indices at Ungauged Locations

Following the screening of attributes and identification of neighbourhoods, the next step is to estimate the required indices of streamflow (i.e. selected percentiles of FDCs or MMFs) through regression relationships as was done previously in Jenkinson and Bomhof (2012, 2014). These relationships could be developed in a linear or nonlinear manner. The former procedure is opted for the current discussion in order to remain consistent with previous hydrokinetic resource assessment studies.

Weighting of Nearest-Neighbours

In the neighbourhood of a target ungauged site, it is very likely that some neighbours are relatively more similar to the target site than others in the attribute space and therefore it is reasonable to adopt a weighting scheme such that more similar neighbours will receive higher weight than less similar neighbours in the estimation of streamflow indices. In the case of CCA-assisted neighbourhoods, Mahalanobis distance measure of each neighbour from the target site can be used to weight various neighbours for developing MR relationships. The following weighting function, which has some similarity to the one used in some earlier studies (e.g. Burn 1990b) can be used:

$$w_i = 1, \text{if } d_i \leq d_L \text{ else } w_i = 1 - \left(\frac{d_i - d_L}{d_U - d_L}\right)^{\eta} \tag{32}$$

where w_i is the weight and d_i is the Mahalanobis distance measure of the *i*th neighbour in the neighbourhood of a target site; d_L and d_U are respectively the lower and upper thresholds and η is called the weighting exponent. It is possible to select d_L from a smaller group of percentiles (e.g. 5th, 10th, 20th and 25th) of the Mahalanobis distance measure and d_U is generally taken as the maximum value of the distribution of Mahalanobis distance measure, which in turn depends on the value of the parameter alpha. This weighting function ensures higher weights to be assigned to the closest neighbours and lower weights to the distant neighbours, with rapidly decaying values for smaller values of η (e.g. 0.05). For real world applications, all weights need to be normalized.

Transformation of Attributes

In addition to the above considerations, it is also important to investigate which transformation of a given attribute is relatively more suitable for the overall regression relationship. For example, the drainage area and the MMFs are highly correlated with each other in the logarithmic domain (see Figure 19). Thus, it will be advisable to regress logarithmically transformed MMFs on to the logarithmically transformed drainage areas. In addition to the logarithmic transformation, square root, cube root, or other suitable transformations can be used. For the case of MMFs or percentiles of FDCs, a simple way of identifying a suitable transformation is to regress logarithmically transformed indices against the selected attribute, separately for each of the selected transformations and select the one that produces the highest value of the coefficient of determination.

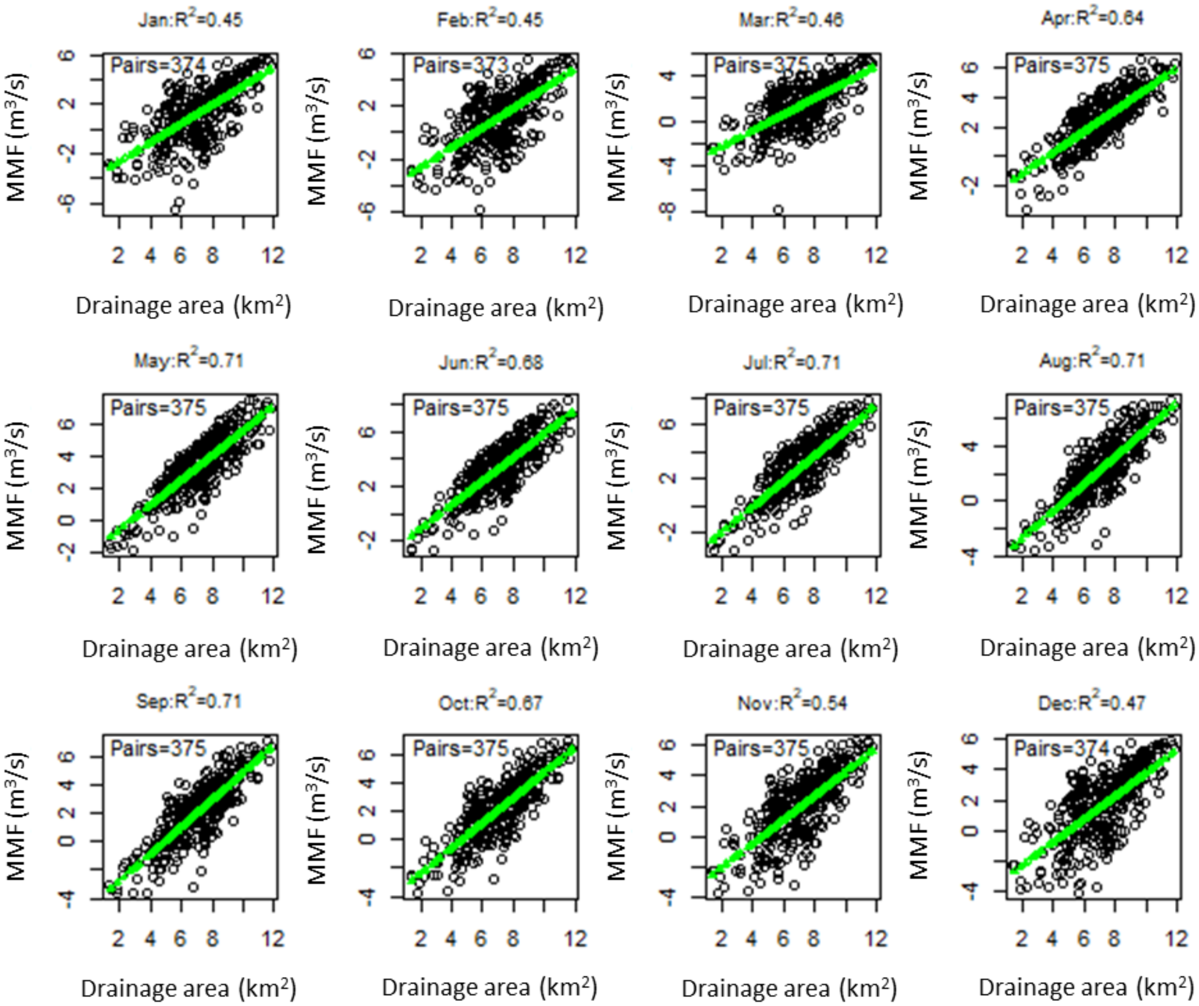

Figure 19: Relationships of MMFs and watershed drainage areas for HYDAT stations for the 1961–1990 period, plotted on a log-log scale. Green dots represent the linear regression results. The values of the coefficient of determination are shown at the top of all panels that suggest relatively stronger (weaker) relationships for ice-free (ice-dominated) months. Number of available non-zero data pairs used in the plots is shown inside each panel.

Following the above choices, MMFs or selected percentiles of FDCs can be estimated at all ungauged locations within a region of interest. In this whole process, the goal should be to improve the preliminary MR model (i.e. the regional or the country-wide MR model) for each target site based on the neighbourhood identified using the CCA approach (Ouarda et al. 2001; Spence and Saso 2005) and by suitably weighting each of the neighbours within the neighbourhood and also by selecting suitable transformations for various attributes. A comparison of four different approaches is shown in Figure 20, where it can be seen that the neighbourhood-based approach performs better than the other three approaches.

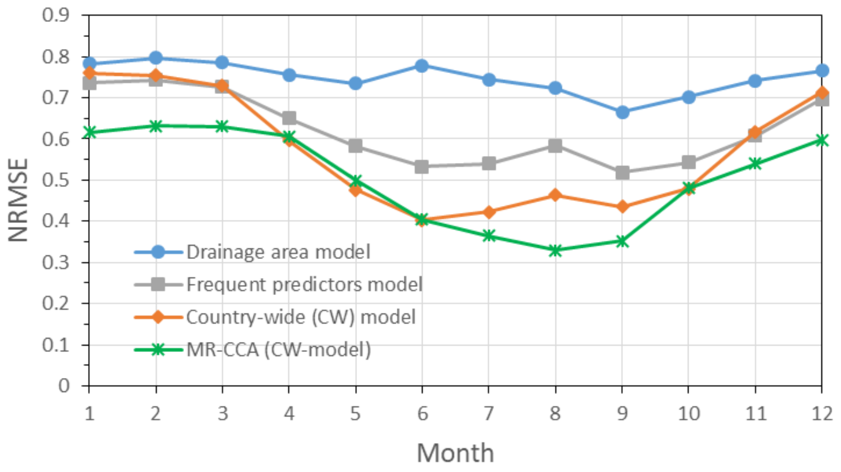

Figure 20: A comparison of four different regression-based approaches for estimating MMFs at HYDAT stations for the 1961–1990 period in terms of normalized root mean square error (NRMSE). These approaches include: (1) regression of MMFs and watershed drainage area; (2) regression of MMFs and four frequent predictors (i.e. DrainageArea, GeoTill, MinElev, and Wetlands); (3) regression of MMFs and significant predictors, obtained through step-wise regression (country-wide regression model); and (4) same as in (3) but for the CCA-based neighbourhoods.

Other Considerations

When developing and finalizing regression relationships, it is important to examine regression diagnostics in order to verify if the underlying theoretical assumptions are satisfied. For example, normality of residuals, homogeneity of variance, independence of residuals, absence of outliers, uncorrelated predictors, etc. It used to be difficult to do so, but not anymore. Almost all statistical packages (such as Matlab, SAS, R package, Minitab, etc.) provide ready to use tools to produce these diagnostics. In some situations when the values of certain predictors are close to zero, it is likely that coefficients of these predictors may end up being equal to undefined flags (e.g. NA in R package). Such attributes need to be explicitly removed from the regression relationships. In some software packages, this issue is also flagged as ‘rank deficient problem’. It is also a good practice to check finiteness of the p-value of the model. When analysing large amounts of data, as was the case of hydrokinetic resource assessment, it is possible that such problems can go unnoticed.

In the case of MMFs or FDCs, it is likely to have zero values for certain months of the year or for certain percentiles of the FDC. In these situations, some investigators tend to replace zero values with an arbitrary small number, perhaps with the intent to have larger datasets. The same was done in Jenkinson and Bomhof (2014). These arbitrary values affect regression diagnostics and therefore concerted efforts should be made to evaluate the influence of such substitutes on the quality of the regression model. This can be accomplished by excluding those datasets when calibrating a regression model.

In modelling MMFs or percentiles of FDCs as a function of watershed attributes, it is also useful to look at standardized regression coefficients. This is important in the case of both MMFs and

FDCs, since all watershed attributes do not share the same scale, i.e. they are at different scales. Irrespective of the original scale, a standardized regression coefficient of 1 for a given attribute means that an increase in its value of 1 standard deviation will produce a corresponding 1 standard deviation increase in the dependent variable. Consequently, if an attribute A has a larger absolute value of the standardized regression coefficient than the attribute B, then the attribute A has a stronger relationship with the dependent variable (i.e. MMFs or percentiles of FDCs). For any attribute, the standardized regression coefficient is equal to 'the product of its coefficient and standard deviation divided by the standard deviation of the dependent variable'. Absolute values of all standardized regression coefficients are ≤ 1. Some investigators express this assessment in terms of model rationality (McCuen 2003).

As discussed before, a reduced model consisting of most important attributes would still be useful in situations where the condition on the required minimum number of independent pairs of data is not satisfied. This situation is often encountered in the neighbourhood based regression relationships. Regarding the choice of the reduced model, the above analysis of standardized regression coefficients can be very insightful.

7.5 Final Remarks

To conclude this chapter, a number of technical and scientific aspects related to hydrokinetic resource assessment studies have been identified and discussed below, particularly from the viewpoint of hydrological investigations. These aspects should be viewed as suggestive only as many of them may need to be verified based on observational records. The discussion is driven mainly from the reviews presented in Chapters 3 to 5 and summarized in Chapter 6 of this report.

Most of the physiographic and climatic attributes were generated through automated computer-based approaches and that perhaps was the only efficient way to process multiple data sources for the entire country. A closer look at many of the processed datasets revealed that there has been lapses in ensuring the quality and reliability of derived attributes. For example, it is difficult to expect that a delineated watershed for a certain river reach will be associated with zero mean annual precipitation under Canadian climatic conditions, irrespective of the drainage area of the watershed. It was also found that for certain attributes, indicators of missing values were not taken care of and some special flags got interpolated, perhaps unintentionally, through automated approaches. Such omissions and mistakes can be quite serious, specifically for smaller watersheds. Therefore, in addition to the statistical aspects discussed above in this chapter to improve regression relationships, it is also important to carefully inspect quality and reliability of processed datasets through detailed granular analyses, before using them in regression relationships. The reliability of any model, including regression relationships, depends on the quality and accuracy of the underlying data used for calibrating the model.

When one desires to estimate various streamflow indices at ungauged locations by identifying neighbourhoods or nearest-neighbours within the attribute space, each ungauged location is expected to have its own neighbourhood. Therefore, even within the same larger homogeneous hydrologic region, there could be a smaller group of observation stations, from where the known information is generally drawn, that is relatively more similar to the location in question compared to the rest of the stations of the region. Consequently, the notion of spatially varying neighbourhoods is appealing than using a constant neighbourhood for all ungauged locations within the geographic boundaries of a larger hydrologic region. This possibility can be explored in future studies on hydrokinetic resource assessment.

In the multiple linear regression framework, adopted in previous hydrokinetic resource assessment studies, (1) relevance of predictor variables and their selection procedures; (2) inter-dependence of predictor variables; (3) how the predictor variables are introduced in the regression relationships (e.g. in their original form or using the log-transformation or square root transformation, etc.); and (4) tests of diagnostics play a crucial role in developing reliable and theoretically defensible relationships. These aspects were given little attention in previous resource assessment studies, discussed in Chapters 4 and 5 of this report. Guidelines on these aspects are readily available in many text books on applied statistics (e.g. Montgomery 2001; Helsel and Hirsch 2002; McCuen 2003; and Walpole et al. 2011). Another important aspect that is often neglected when developing multiple regression relationships concerns the number of available independent datasets and the number of unknown regression parameters. Ideally, the former should be considerably larger than the latter to develop sound relationships. McCuen (2003) recommended to have independent number of datasets more than four times the number of unknown regression parameters. This aspect is very important for neighbourhood-based regression relationships, since it is very unlikely to find sufficient number of nearest-neighbours at least for certain ungauged locations. In those situations, it is better to try a relationship with a smaller number of attributes. For example, by selecting the three most important attributes that are able to explain the majority of the variability of the dependent variable than using all (say 10) significant attributes. This will avoid having some locations with indeterminate estimates. In the case of MMFs, Khaliq et al. (2015) found that even just using the drainage area as a predictor could be beneficial to obtain estimates of MMFs for locations with insufficient number of nearest-neighbours. However, the reader is cautioned that this suggestion may not work under all situations.

Hydraulic aspects of the resource assessment study are not reviewed in detail in this report and therefore the suggestions made are merely of general nature. In determining the resource potential of a given river reach, rectangular cross-sectional shape was assumed in the resource assessment study perhaps due to the difficulties involved in determining spatially varying cross-sections. Though variable cross-sectional shape is an inherent property of natural channels, literature supports that assuming a rectangular cross-section for wide open channels is a reasonable assumption (Chow 1959). The influence of variations in the cross-sectional shape can be evaluated through carefully designed controlled laboratory experiments. The results of such experiments can be generalized to develop adjustment factors that can be applied in different regions to overcome the influence of variations in cross-sectional shapes. However, the feasibility of such laboratory experiments from a financial viewpoint needs to be determined in the light of high costs of such experiments and the advantages gained over the simplistic approach of using rectangular cross-sections.

In the hydrokinetic resource assessment study conducted by Jenkinson and Bomhof (2014), stochastically generated streamflow indices were mixed with deterministic estimates from observation stations. This was specifically done to overcome the issue of regulated flows and boundary watersheds along the Canada-US border. This is quite dangerous from a statistical viewpoint and it is difficult to find any sound justification for this operation. Furthermore, the mixing was performed only when there was five or more years of observational records available from 1990 onwards. This was also not a reasonable approach since (say) five to ten years of data cannot be considered as representative of 30 years of data used in deriving other attributes. Estimation of streamflow indices should have been attempted exclusively within the stochastic framework and deterministic estimates should have been used to validate stochastically generated values. It is necessary to point out here that the climatic variables used in the regression relationships were from the 1961–1990 period, while deterministic indices of streamflow were from a different time window.

In the literature on hydrologic regionalization in the context of ungauged hydrology, a number of different methods are available and these can be used to delineate neighbourhoods or nearest-neighbours when estimating unknown streamflow indices at ungauged locations from the corresponding known indices available at gauged locations. The region of influence approach (ROI) is quite common, while the CCA-based approach has also been used in some studies. The principles that underpin these approaches allow consideration of gauged locations from adjoining or distantly located hydrologic regions or geographic areas. Thus, geographic or political boundaries are not considered a limitation for applying these approaches. However, some investigators do object such definition of neighbourhoods due to considerable differences in associated atmospheric mechanisms that influence regional climate and local weather patterns. In order to reconcile both school of thoughts, perhaps it is useful to consider larger climatic or hydrologic regions and apply the ROI or the CCA approach to identify nearest-neighbours within the same larger region. Such an approach is advantageous from climatological, hydrological and statistical viewpoints and also ensures to some extent the physical proximity of the target location and nearest-neighbours. Canada has been divided into 11 large climatic regions (e.g. Plummer et al. 2006; Mladjic et al. 2011) and those regions can be used as a basis to develop both ROI and CCA-based approaches. For certain situations, the results from both ROI and CCA-based approaches could be very different and therefore, it will be useful to apply both approaches together within the same climatic region. This will help in reaping the benefits of both approaches and ultimately to have better estimates of streamflow indices at ungauged locations within a target region of interest.

The objective of any resource assessment study is to obtain as reliable estimates of hydrokinetic potential as possible. Therefore, it will be rather logical to process natural and regulated rivers separately. Compared to natural river reaches, regulated river reaches could exhibit a persistent pattern of variations in streamflow regimes from one year to the other. Such persistent flow patterns will also be useful in identifying suitable time windows where it will be more feasible to extract hydrokinetic energy. To uncover existence of such persistent patterns over longer periods, evaluation of annualized FDCs could be very insightful. It is very likely that the yearly patterns of flow variations in regulated rivers could be the same and therefore some investigation can directly go into identifying the time periods when flow ranges in regulated watersheds are at or above the minimum flow and velocity requirements for operating hydrokinetic turbines. To materialize this possibility, separate datasets can be developed to identify those river reaches where such a potential is available.

In future studies, investigations should be focused on both selected indices of FDCs and MMFs for hydrokinetic resource assessment. In particular for the latter case, we need to identify those months where in fact resource extraction is practically feasible. For example, open water months where the possibility of ice and ice jamming conditions is negligible. There is no point in evaluating the resource potential for the entire year, while knowing that at certain times of the year, resource extraction is not only impossible but could also be negligible. It is important to mention that it is hard to find any literature on the performance of hydrokinetic turbines under iced conditions. This could be an important research problem for future studies. As the selected indices of FDCs have already been used in previous resource assessment studies, it will be useful to include MMFs for open water months in resource assessment studies as an alternative surrogate procedure. It will also be useful to superpose remotely sensed information about flow rate and velocities on top of the velocity/flow rate maps derived from combined hydrologic and hydraulic studies to identify concomitant regions where hydrokinetic potential could possibly be quite significant.

Lastly, to improve quality of hydrokinetic resource databases, as already mentioned, it is important to improve quality of various physiographic and climatic attributes and reliability of estimated streamflow indices, and then to identify regions which are more suitable for hydrokinetic resource extraction. Additionally, to enhance the value of hydrokinetic resource databases, it will also be useful to develop different classifications of flow and velocity ranges, depending upon various parameters of hydrokinetic turbines. This process will help identify geographic regions that satisfy such classifications for further evaluation and feasibility studies.

8 References

AAFC, 2008. A national ecological framework for Canada - Canadian ecodistrict climate normals (1961-1990). Agriculture and Agri-Food Canada (AAFC), http://sis.agr.gc.ca/cansis/nsdb/ecostrat/climate_normals_1961-90.html.

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

Acreman M, Wiltshire S, 1989. The regions are dead. Long live the regions. Methods of identifying and dispensing with regions for flood frequency analysis. IAHS-AISH publication 187:175–188.

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

Acres Consulting Services Limited, 1984b. Hydrologic design methodologies for small scale hydro at ungauged sites – phase I: Study documentation report. Technical report, Environment Canada, Inland Waters Directorate.

Allen PM, Arnold JG, Byars BW, 1994. Downstream channel geometry for use in planning-level models. Water Resources Bulletin 30(4): 663–671.

Bobee B, Mathier L, Perron H, Trudel P, Rasmussen PF, Cavadias G, Bernier J, Nguyen V-T-V, Pandey G, Ashkar F, Ouarda TBMJ, Adamowski K, Alila Y, Daviau JL, Gingras D, Liang GC, Rousselle J, Birikundavyi S, Ribeiro-Corra J, Roy R, Pilon PJ, 1996. Presentation and review of some methods for regional flood frequency analysis. Journal of Hydrology 186 (1-4): 63–84.

Bomhof J, 2013. Estimating flow, hydraulic geometry, and hydrokinetic power at ungauged locations in Canada. MSc thesis, University of Ottawa, Ottawa, Ontario.

Burn DH, 1990a. An appraisal of the "region of influence' approach to flood frequency analysis. Hydrological Sciences Journal 35(2): 149–165.

Burn DH, 1990b. Evaluation of regional flood frequency analysis with a region of influence approach. Water Resources Research 26 (10): 2257–2265.

Canadian Council on Geomatics, 2009. National Hydro Network (NHN), http://www.geobase.ca/geobase/en/data/nhn/index.html.

Castellarin A, Galeati G, Brandimarte L, Montanari A, Brath A, 2004. Regional flow-duration curves: reliability for ungauged basins. Advances in Water Resources 27(10): 953–965.

Cavadias GS, Ouarda TBMJ, Bobee B, Girard C, 2001. A canonical correlation approach to the determination of homogeneous regions for regional flood estimation of ungauged basins. Hydrological Sciences Journal 46(4): 499–512.

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

Chang C, Ashenhurst F, Damaia S, Mann W, 2002. Ontario Flow Assessment Techniques Version 1.0 User's Manual. Northeast Science and Information Section, Ontario Ministry of Natural Resources. NESI Technical Manual TM-011.

Chow VT, 1959. Open Channel Hydraulics. McGraw-Hill Book Company, New York.

Copeland RR, Biedenharn DS, Fischenich JC, 2000. Channel-forming discharge. Technical Report ERDC/CHL CHETN-VIII-5, US Army Corp of Engineers (USACE).

Dalrymple T, 1960. Flood frequency analysis. US Geological Survey Water Supply Paper 1543.

Dingman SL, 1978. Synthesis of flow-duration curves for unregulated streams in New Hampshire. Journal of the American Water Resources Association 14(6): 1481–1502.

Durand N, Bourban SE, Crookshank N, 2002. Development of a Toolkit to Estimate the Concentration of Substances Released into Rivers and Streams - Scientific Literature Review. Technical report, Canadian National Research Council - Canadian Hydraulics Centre.

Eaton B, Church M, Ham D, 2002. Scaling and regionalization of flood flows in British Columbia, Canada. Hydrological Processes 16(16): 3245–3263.

Efron B, 1982. The jackknife, the bootstrap and other resampling plans. In CBMS-NSF regional conference series in applied mathematics, volume 38. Siam.

Ehlschlaeger CR, 1989. Using the "At" search algorithm to develop hydrologic models from digital elevation data. In: Proceedings of the International Geographic Information System (IGIS) Symposium, Baltimore, MD, pp 275–281.

Eng K, Tanker GD, Milly PCD, 2005. An analysis of region-of-influence methods for flood regionalization in the Gulf-Atlantic rolling plains. Journal of American Water Resources Association, 41: 135–143.

Environment Canada, 2004. HYDAT Surface Water and Sediment Data. CD-Rom.

Environment Canada, 2006. Canadian daily climate data. CD-Rom.

ESRI, 2010. ArcGIS Help Library - ArcGIS 10. ESRI.

Fox J, 2008. Applied Regression Analysis and Generalized Linear Models, 2nd Edition, Sage Publications.

Franchini M, Suppo M, 1996. Regional analysis of flow duration curves for a limestone region. Water Resources Management 10(3): 199–218.

FSR, 1975. Flood Studies Report (FSR): Hydrological Studies. Institute of Hydrology, United Kingdom, pp 1248.

Gingras D, Adamowski K, Pilon PJ, 1994. Regional Flood Equations for the Province Of Ontario and Quebec. Journal of the American Water Resources Association 30(1): 55–67.

Gulliver JS, Murdock RU, 1993. Prediction of river discharge at ungauged sites with analysis of uncertainty. Journal of Water Resources Planning & Management - ASCE 119(4): 473–487.

Haan CT, 2002. Statistical Methods in Hydrology. Iowa State Press, Ames, Iowa.

Helsel DR, Hirsch RM, 2002. Statistical Methods in Water Resources. United States Geological Survey. Available at: http://water.usgs.gov/pubs/twri/twri4a3/.

Holmes MGR, Young AR, Gustard A, Grew R, 2002. A region of influence approach to predicting flow duration curves within ungauged catchments. Hydrology and Earth System Sciences 6(4): 721–731.

Hughes DA, Smakhtin V, 1996. Daily flow time series patching or extension: A spatial interpolation approach based on flow duration curves. Hydrological Sciences Journal 41(6): 851–871.

Hutchinson MF, 2009. ANUSPLIN Version 4.3. Australian National University, http://fennerschool.anu.edu.au/publications/software/anusplin.php.

Jenkinson W, 2010. Assessment of Canada's Hydrokinetic Power Potential. Phase I Report: Methodology and Data Review. NRC Report No. CHC-TR-070. National Research Council of Canada, Ottawa, ON, Canada.

Jenkinson W, Bomhof J, 2012. Assessment of Canada's Hydrokinetic Power Potential. Phase II Report: Methodology Validation. Technical report. National Research Council of Canada, Ottawa, ON, Canada.

Jenkinson W, Bomhof J, 2012. Assessment of Canada's Hydrokinetic Power Potential. Phase III Report: Resource Estimation. NRC Report NO. OCRE-TR-2015-007. National Research Council Canada, Ottawa, ON, Canada.

Khaliq MN, Jenkinson W, Bomhof J, Serrer M, Klyszejko E, 2015. Estimation of mean monthly flows at ungauged locations in the Maritimes and Pacific hydrologic regions. Canadian Society of Civil Engineering Conference, Montreal, Quebec, April 29–May 2.

Khaliq MN, 2019. An Inventory of Methods for Estimating Climate Change-Informed Design Water Levels for Floodplain Mapping. NRC Report No. OCRE-2019-TR-011. National Research Council Canada, Ottawa, ON.

Khan MJ, Bhuyan G, Iqbal MT, Quaicoe JE, 2009. Hydrokinetic energy conversion systems and assessment of horizontal and vertical axis turbines for river and tidal applications: A technology status review. Applied Energy 86(10):1823–1835.

KWL, 2008. Computer Modelling for Water – Finding Run-of-River Hydroelectricty Potential in B.C. Kerr Wood Leidal Associates, Canadian Consulting Engineers, pp 16–17.

Leboutillier DW, Waylen PR, 1993. Regional variations in flow-duration curves for rivers in British Columbia, Canada. Physical Geography 14(4): 359–378.

Leopold LB, Maddock T, 1953. The hydraulic geometry of stream channels and some physiographic implications. United States Government Printing Office.

Loukas A, Quick MC, 1995. Comparison of six extreme flood estimation techniques for ungauged watersheds in coastal British Columbia. Canadian Water Resources Journal 20(1): 17–30.

MathWorks, 2011. MATLAB Manual - R2011a.

McCuen RH, Levy BL, 2000. Evaluation of peak discharge transposition. J. Hydrologic Engineering 5(3): 278–289.

McCuen RH, 2003. Modelling Hydrologic Change. CRC Press, Taylor & Francis Group, London.

Metcalfe RA, Chang C, Smakhtin V, 2005. Tools to support the implementation of environmentally sustainable flow regimes at Ontario's waterpower facilities. Canadian Water Resources Journal 30(2): 97–110.

Miller G, Franceschi J, Lese W, Rico J, 1986. The Allocation of Kinetic Hydro Energy Conversion Systems (KHECS) in USA Drainage Basins: Regional Resource and Potential Power. Technical Report NYU/DAS 86-151, New York University - Department of Applied Science. Prepared for the US Department of Energy.

Mladjic B, Sushama L, Khaliq MN, Laprise R, Caya D, Roy R, 2011. Canadian RCM projected changes to extreme precipitation characteristics over Canada. Journal of Climate 24: 2565–2584.

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

Mohamoud YM, Parmar RS, 2006. Estimating streamflow and associated hydraulic geometry, the Mid-Atlantic region, USA. Journal of the American Water Resources Association 42(3): 755–768.

Moin SMA, Shaw MA, 1985. Regional Flood Frequency Analysis for Ontario Streams, Volume 1: Single Station Analysis and Index Method. A study funded under the Canada/Ontario Flood Damage Reduction Program, Environment Canada.

Moin SMA, Shaw MA, 1986. Regional Flood Frequency Analysis for Ontario Streams, Volume 2: Multiple Regression Method. A study funded under the Canada/Ontario Flood Damage Reduction Program, Environment Canada.

Monk R, Joyce S, 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.

Montgomery DC, Peck EA, Vining GG, 2001. Introduction to Linear Regression Analysis (3rd edition). John Wiley and Sons, Inc., New York, New York, pp 641.

Nathan RJ, McMahon TA, 1990. Identification of homogeneous regions for the purposes of regionalisation. Journal of Hydrology 121(1-4): 217–238.

Natural Resources Canada, 2004a. Clean Energy Project Analysis: RETScreen Engineering & Cases Textbook, Chapter Small Hydro Project Analysis. Natural Resources Canada.

Natural Resources Canada, 2004b. RETScreen Software Online User Manual: Small Hydro Project Model. Technical Report, Natural Resources Canada.

Natural Resources Canada, 2008. Atlas of Canada: 1,000,000 national frameworks data, hydrology - drainage network.

NLWIS, 2008. Gridded Climate Data. The National Land and Water Information Service (NLWIS), http://www4.agr.gc.ca/AAFC-AAC/displayafficher.do?id=1227620138144&lang=eng.

NRC-CHC, 2008a. St. Lawrence River Currents – A Potential Source of Renewable Energy. Technical Report CHC-TR-53, National Research Council – Canadian Hydraulics Centre.

NRC-CHC, 2008b. Methodology for the Assessment of Hydraulic Kinetic Energy in Rivers. Technical Report CHC-CTR-075, National Research Council Canada – Canadian Hydraulics Centre.

Ottawa Engineering Limited, 1997. Hydrological Method - RETScreen. Prepared for Natural Resources Canada.

Ouarda TBMJ, Hache M, Bruneau P, Bobée B, 2000. Regional flood peak and volume estimation in Northern Canadian Basin. Journal of Cold Regions Engineering 14(4): 176–191.

Ouarda TBMJ, Girard C, Cavadias GS, Bobée B, 2001. Regional flood frequency estimation with canonical correlation analysis. Journal of Hydrology 254(1-4): 157–173.

Quimpo RG, Alejandrino AA, McNally TA, 1983. Regionalized flow duration for Philippines. Journal of Water Resources Planning & Management - ASCE 109(4): 320–330.

Park CC, 1977. World-wide variations in hydraulic geometry exponents of stream channels: An analysis and some observations. Journal of Hydrology 33(1-2): 133–146.

PFRA, 2008. PFRA Watershed Project: Areas of Non-Contributing Drainage.

Plummer DA and Coauthors, 2006. Climate and climate change over North America as simulated by the Canadian RCM. Journal of Climate 19: 3112–3132.

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

Rojanamon P, Chaisomphob T, Bureekul T, 2009. Application of geographical information system to site selection of small run-of-river hydropower project by considering engineering/economic/environmental criteria and social impact. Renewable and Sustainable Energy Reviews 13(9): 2336-2348.

Searcy JK, 1959. Flow duration curves. Water Supply Paper 1542-A, US Geological Survey.

Sharpe DR, Russell HAJ, Grasby SE, Wozniak PRJ, 2008. Hydrogeological regions of Canada: Data Release. CDROM, Open File 5893.

Shaw E, 2004. Hydrology in Practice. Routledge, Abingdon, Oxon, pp 592.

Shu C, Ouarda TBMJ, 2012. Improved methods for daily streamflow estimates at ungauged sites. Water Resources Research 48(2): 1-15.

Smakhtin VY, 1999. Generation of natural daily flow time-series in regulated rivers using a non-linear spatial interpolation technique. River Research and Applications 15(4): 311–323.

Smakhtin VY, Masse B, 2000. Continuous daily hydrograph simulation using duration curves of a precipitation index. Hydrological Processes 14(6): 1083–1100.

Smakhtin VY, Hughes DA, Creuse-Naudin E, 1997. Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa. Hydrological Sciences Journal 42(6):919–936.

Spence C, Saso P, 2005. A hydrological neighbourhood approach to predicting streamflow in the Mackenzie Valley. In: Prediction in Ungauged Basins: Approaches for Canada’s Cold Regions. Spence C, Pomeroy JW, Pietroniro A (Eds.). Canadian Water Resources Association, pp 21–44.

Tasker GD, 1982. Comparing methods of hydrologic regionalization. Water Resources Bulletin 18(6): 965–970.

Tasker GD, Hodge SA, Barks CS, 1996. Region of influence regression for estimating the 50-year flood at ungauged sites. Journal of the American Water Resources Association 32(1): 163–170.

Tudor Engineering, 1991. A methodology for regional assessment of small scale hydro power. Industry and Energy Department Working Paper No.44. World Bank, Washington D.C.

Turner R, Clague J, 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. National Research Council Canada – Canadian Hydraulics Centre.

USDOE, 2004. Water Energy Resources of the United States with Emphasis on Low Head/Low Power Resources. Technical Report DOE/ID-11111, US 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.

Verdant Power, 2006. Technology Evaluation of Existing and Emerging Technologies. Technical Report NRCan-06-01071, Natural Resources Canada.

Walpole RE, Myers RH, Myers SL, Ye K, 2011. Probability & Statistics for Engineers & Scientists, 9th Edition. Prentice Hall, Boston, pp 812.

Wang Y, 2000. Development of Methods for Regional Flood Estimation in the Province of British Columbia. PhD Thesis. Department of Forest Resources. University of British Columbia, British Columbia, Canada, pp 214.

Zrinjti Z, Burn DH, 1994. Flood frequency analysis for ungauged sites using a region of influence approach. Journal of Hydrology 153: 1–21.