LOW FLOW FREQUENCY STUDY FOR NEWFOUNDLAND AND LABRADOR

SHABNAM MOSTOFI ZADEH

Low Flow Frequency Study for Newfoundland and Labrador

Prepared by

© Shabnam Mostofi Zadeh

A Thesis submitted to the School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Engineering

Faculty of Engineering and Applied Science

Memorial University of Newfoundland

November 2012

St. John's Newfoundland

ABSTRACT

The objectives of this study were to quantify the characteristics of low flows in rivers of the province of Newfoundland and Labrador, and to develop equations which could be used to estimate the magnitude, frequency, duration, and spells of low flow events. These different aspects of low flows were analyzed by applying methods of flow frequency, flow duration, and flow spell analysis, respectively. Sixty hydrometric stations in the Island of Newfoundland which have more than 20 years of complete data were selected for the current low flow study. Because of the sparseness and shortness of hydrometric data in Labrador, sites with more than 15 years of data were chosen with a total of 12 stations. An L-moment based approach was applied for regional frequency analysis of annual minimum 1-day and 7-day flows for two separate homogeneous regions, Island of Newfoundland, and Labrador and it yielded prediction equations for low flows of different durations and return periods. The performance of these regional models was verified using new sets of data, and showed reliable results. Therefore, one can use these prediction models for ungauged sites in Newfoundland and Labrador. To perform regional flow duration analysis, physiographic parameters of the regions under study were regressed against quantiles of flow duration curves obtained for each hydrometric station to produce a regional model for predicting flow duration curves at any ungauged sites. Regional model of flow durations were validated successfully using a new set of data, and the results were promising. Different hydrological methodologies were applied to define flow spells for rivers in Newfoundland and Labrador, and regional models were defined to predict the annual maximum flow spell variables in Newfoundland and Labrador.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Dr. Leonard M. Lye for his continual guidance, teaching, financial support and encouragement for pursuing the Master of Engineering Program at Memorial University of Newfoundland (MUN). Special thank goes to Dr. Amir Ali Khan, Department of Water Resources, Government of Newfoundland and Labrador for providing help. Funding for the study was provided by the Government of Newfoundland and Labrador through the Institute for Biodiversity Ecosystem Science and Sustainability (IBES). I wish also to thank the School of Graduate Studies for the financial support.

My heart-felt appreciation goes to my dear husband, Vandad, for his support and sharing my aspiration during the study period and thesis preparation. I appreciate my beloved parents, for being understanding, and helping me to endure their distance from me during the period of my study.

A number of people have helped in data gathering for this study. I wish to thank Mr. Michael Colbert, Hydrological Modeling Section, Department of Environment and Conservation, Government of Newfoundland and Labrador for helping with the hydrological records, and delivering the requested data for the province of Newfoundland and Labrador.

Lastly, I wish to thank the academic staff, the support staff, and friends who have helped me in many ways during the period of my study.

Table of Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

α	Scale parameter of the distribution
α_r	Population probability weighted moment
B_4	Bias of sample regional L-kurtosis
β_r	Population probability weighted moment
C_V	Coefficient of variation
D_i	Discordancy measure
$E(x)$	Expectation of random variable x
F	Non-exceedance probability
Φ	Standard normal CDF
$F(x)$	Cumulative distribution function
Φ^{-1}	Inverse of standard normal CDF
h	4th parameter of the distribution
H	Heterogeneity measure
k	Shape parameter of the distribution
ln	natural log
λ_r	Population L-moments
l_r	Sample L-moments
μ	mean
μ_V	mean of Vs
N_{sim}	Number of simulated regions

Q Flow rate

σ Standard deviation

σ^2 Variance

σ_4 Standard deviation of sample regional L-kurtosis

σ_V standard deviation of Vs

T Return period

τ L-CV

t Sample L-CV

τ_3 L-Skewness

t_3 Sample L-Skewness

${t_3}^R$ Regional average sample L-Sk

τ_4 L-Kurtosis

t_4 Sample L-Kurtosis

${\tau_4}^{Dist}$ Distribution L-kurtosis

${t_4}^R$ Regional average sample L-Ku

t^R Regional average sample L-CV

u L-moment ratio vector

V weighted standard deviation of sample L-CVs

ξ Location parameter of distribution

$X(F)$ Quantile function of frequency distribution

Z_{Dist} Goodness-of-fit measure of the candidate distribution

7Q10	7-day consecutive low flow with ten year return period
AFDC	Annual Flow Duration Curve
AM	Annual Minimum
Eff-P	Effective Precipitation
FDC	Flow Duration Curve
L-CV	Coefficient of L-variation
L-ku	Coefficient of L-Kurtosis
NSE	Nash-Sutcliffe Efficiency
OLS	Ordinary Least Squares
PWM	Probability Weighted Moments

List of Appendices

1 Introduction

1.1 General

Stream flows naturally vary both during a year and from year to year. In the face of these variabilities, water management decisions can only be made with predicted estimates of stream flows. More importantly, design and planning of water resources projects requires the assessment of the probability of extreme hydrological events such as low or high flows. A low flow condition can be defined as a period during which the average stream flow is a minimum for the year. The characteristics and estimation of low flows are important issues in hydrologic studies such as the determination of minimum downstream flow requirement of hydropower station, estimation of available water supply for municipal and industrial uses, water quality management, determination of potential capacity for effluent dilution, assessing the impact of low flows on aquatic ecosystem, and in general for environment impact assessment studies (Govt. of Newfoundland and Labrador, 1991).

The low flow regime of a river can be analyzed in a variety of ways depending on the type of data initially available and the type of output information required (Smakhtin, 2001). Low flow studies often require that the hydrologists estimate the magnitude, frequency, duration, and spells of low flow events as different aspects of low flow analysis by applying methods of flow frequency, flow duration, and flow spell analysis. Flow frequency, flow duration, and flow spell analysis are the three main objectives of the current study.

Flow frequency analysis is traditionally based on fitting a probability distribution to the available data at a specific site of interest. This probability only gives us an idea how likely a flow is to happen in future. It is generally assumed that flow magnitude can be reliably estimated from a long period of data records. However, the available historical flow data at the site of interest are often too short to give these reliable estimates of critical flow (low or high). This condition has led hydrologists to wonder whether the estimation from one sample can be more accurate by not just using information from one sample but also from other related samples. Therefore approaches have been developed to augment the limited flow record available at a specific location by involving data from neighboring locations, the so called homogeneous hydrological region. This technique would not only improve the estimates at the site of interest with short data records, but would also provide a basis for flow estimation at any ungauged locations within that homogeneous region. The process of using data from several sites to estimate the frequency distribution is known as regional frequency analysis. This procedure can be used for estimating any flow statistics such as mean, low or high flows (Hosking and Wallis, 1997). In this study the interest is in the minimum low flow estimation, thus the outcome of the regional frequency analysis would be the low flow minimums with associated frequency of flow being equal or below this amount.

The general procedure of conducting regional frequency analysis involves the following basic steps: collecting low flow data at the gauged rivers; screening the collected data for any gross errors or any other causes that makes the data unusable; identifying homogeneous regions and testing their homogeneity; determining the regional prediction equations (growth curves or regression relations) for the homogeneous regions; and establishing the flow quantiles of interest. Estimating flow magnitudes using the regional approach has been documented for the last four decades (Hosking and Wallis, 1997).

The index flow method suggested by the USGS (Dalrymple, 1960) is the earliest and still most popular approach for regional estimation which is still in use with slight modifications over time. Regression on quantiles was suggested as an alternative approach to overcome the apparent problems associated with the original index flow method regarding its assumption about the distribution characteristics of flow data within a region. With the introduction of the L-moments approach in statistics and its application in hydrology the index flow method has been firmly re-established as a general procedure of flow frequency analysis. This approach has been used for conducting the regional low flow frequency analysis in this study.

To find out what percentage of time of a year flow in a river will be below a certain amount, it is necessary to conduct flow duration analysis using flow-duration curves. Flow-duration curves simply provide the relationship between streamflow and the percentage of time it is exceeded (Vogel and Fennessy, 1994). Flow-duration curves, as a comparison to flood or low flow frequency analysis, are derived from all the historic data available for a stream rather than just the annual lowest flow. Similar to flow frequency analysis, regional flow-duration analysis can be conducted for a region. There are different methods for regionalization of flow-duration curves, the most common being the multiple-regression approach. Where no flow measurements exists at a site, a common approach for estimating streamflow is to develop a relationship between flow measurements from a gauged river and physiographic parameters of its basin. Some of the factors which have been considered include catchment area, main channel slope, drainage density, difference in elevation, and percentage of the area covered by forest, swamps, lakes and impermeable rock. Regression equations can be developed between these factors and any streamflow indices derived from the flow-duration curves. And finally for any ungauged location within the defined region, stream flow indices can be estimated from the multiple regression equation and estimates of these physiographic factors (McMahon et al., 2004).

Environmental instream flow requirment are flows in a river that are deemed as a minimum to maintain the river ecosystem (Karim, 1995). Therefore, it is critical to have an estimate of these required instream flows, and planning for the time that streamflow goes below this amount. There are several ways that instream flow can be estimated. Methods such as percentiles of the flow duration curves, percentages of the mean annual flow, and consecutive seven-day averaged low flow with an estimated ten year return period. Selecting a method to estimate the environmental instream flow depends on the particular requirement that is being considered for the ecosystem (McMahon et al, 2004).

In order to estimate how long streamflow will be below a certain amount (instream requirement), and how large the deficit volume is, it is necessary to conduct flow spell analysis. This may be found by using the aforementioned instream minimum flows as a threshold on the sequential daily flows. Flows below these thresholds are considered as spells (IH, 1980) that may be quantified in terms of duration (in days), volume (in m^3)

and intensity of flow spell (volume divided by the duration). Therefore, flow spell analysis takes into account the sequencing of flows. Flow duration curves, in contrast, give no information on how the low flow days are distributed.

1.2 Low Flow Analysis for the Island of Newfoundland

The history of low flow estimation in the island of Newfoundland dates back to 1991, when the Government of Newfoundland and Labrador (Govt. of Newfoundland and Labrador, 1991) conducted the first study to quantify the characteristics of low stream flows and came up with set of equations to estimate low flows of various durations and return periods on ungauged streams. However, at the time of that study the present state-of-the-art regional frequency analysis techniques were not available, and the recorded data period was short.

A hydrological study of the Island of Newfoundland was performed in 1995 (Richter and Lye, 1995) to identify the key basin characteristics associated with flow measures and assessed several methods of regional subdivision for improving flow estimates at ungauged sites. A data set of 40 stations with more than 10 years of record was used in this study.

In 1997 a research study was performed on duration, volumes, and intensities of flow spells for a few rivers in Newfoundland and Labrador (Shaughnessy, 1997). Different methods of estimating environmental instream flow requirements were used as the threshold values. Again, the number of suitable gauged rivers and their record period was short in this study. A more detailed review of the aforementioned studies is given in Section 2.4.

1.3 Research Objectives

The first objective of this research is to apply the popular L-moments based index flow approach to conduct a regional frequency analysis for low flows for rivers of Newfoundland and Labrador, Canada. The L-moments and regional frequency analysis based on L-moments were introduced in the early 1990's (Hosking, 1990, Hosking and Wallis, 1993). The 1991 low flow study for the Island of Newfoundland was based on 'regression on quantiles' approach, and data records were short at that time. The Island of Newfoundland was the only region used in the 1991 study, and no research was performed on rivers of Labrador. In the present study the more efficient 'L-moments' approach will be used to conduct a regional analysis for rivers in both the Island of Newfoundland, and the Labrador region where the records are now of sufficient length for frequency analysis.

The next objective of the proposed research is the development of regional flow duration estimation equations for Newfoundland and Labrador. A regional regression approach will be used between flow indices of flow duration curves and related physiographic parameters of river basins, to produce a set of prediction equations for ungauged sites.

The final goal of this thesis is to provide a means of estimating flow spells for rivers of Newfoundland and Labrador, and to quantify duration, volume and intensity of those spells based on different instream flow requirements. This part of the study is revisiting the previous study undertaken in 1997 by using longer available record data, and additional gauged rivers for the analysis.

1.4 Outline of Thesis

The thesis is organized into three major groups of chapters:

- Introduction to the problem, overview and approaches: Chapter 1 and 2
- Main methodologies: Chapter 4, 5 and 6
- Summary and conclusions: Chapter 7

Chapter 1 covers the introduction of the topic in which the general concepts of low flow estimation including regional low flow frequency, flow duration, and flow spell analysis and their application in Newfoundland and Labrador are briefly discussed. Chapter 2 surveys the existing literature review on flow estimation methods with the particular emphasis on regionalization techniques. The main methodologies proposed, regional low flow, flow duration, and flow spell are briefly introduced in Chapter 3, and then followed up by their application for selected rivers within Newfoundland and Labrador in Chapters 4, 5 and 6, respectively. Summary and conclusions of this study can be found in Chapter 7. Finally, the computer programs that were developed for various processing of the data are presented in the appendices.

2 Literature Review

Low flows have been investigated only in the recent past few decades. This includes low flow frequency analysis, base flow separation, recession analysis, flow spell analysis, and low flow estimation at ungauged sites. Although there is a high interest in low flow studies, the mass of literature has still been relatively less compared with flood or precipitation studies. It could be a result of that low flows are viewed less destructive as floods. The characteristics and estimation of low flows are important issues in many hydrologic studies and in general for environmental impact assessment studies. Such studies often require that the hydrologists estimate the magnitude, frequency, duration, and spells of low flow events as different aspects of low flow analysis (Smakhtin, 2001). Proposing any solutions to a problem is only justifiable after a complete knowledge and understanding of the existing solution(s) to the problem at hand or problems with some similar characteristics. For this reason, this chapter reviews the developments and existing theories and methods that are relevant to low flow analysis in general. At the end, the earlier report of the Provincial Government of Newfoundland and Labrador on the low flow characteristics of the rivers in the Island, the study on relationship between flow and basin variables on the Island by Richter and Lye (1995), and also flow spell analysis research by Shaughnessy (1997) for rivers in the province are reviewed.

2.1 Low Flow Frequency Analysis

2.1.1 General

Unlike the flow duration curve which shows the proportion of time during which a flow is exceeded, a low flow frequency curve shows the proportion of years when a flow is exceeded or equivalently the average interval in years (return period or recurrence interval) that the streamflow falls below a given discharge. Figure 2-1 illustrates a typical low flow frequency curve.

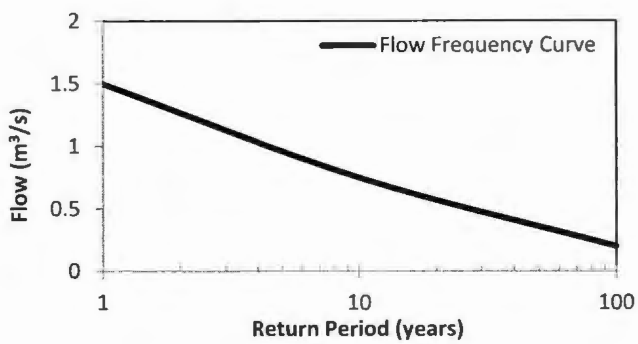

Figure 2-1 Low flow frequency curve

Low flow frequency analysis form a part of the frequency analysis of extreme events and as such has been covered in many classical hydrology text books (e.g. McMahon et al., 2004). Some authors note that the literature on low flow frequency analysis remains to be limited compared, for example, with the literature on flood frequency (e.g. Vogel and Wilson, 1996).

The existing approaches in flow estimation can broadly be divided into two sections: (1) statistically based; (2) physically based. Statistically based approaches refer to the analysis of raw data collected from a site or a region using state-of-the-art statistical tools in deriving probabilistic functions or frequency distributions pertaining to flow (low flow) quantiles. Physically based approaches essentially model the actual flow conditions in a river channel based on all the available physical theories and data. Two distinctive components can be delineated based on the theories used: hydrologic and hydraulic. Since the physically based approach is not within the scope of this study, only the statistical approach is reviewed in the following sections.

Traditional methods that dominate the statistical approach are single station flow frequency analysis and regional versions of this analysis. Flow frequency analysis is a standard procedure for the planning and design of water resources projects and other civil engineering works. It provides the probabilistic assessment of the magnitude of (flood or low) flows associated with a certain risk tolerance level. It was discussed before that one of the objectives of this study is to develop regional low flow frequency models, and to apply the mature method of L-moments for this regional analysis. Therefore, the rest of this section will be confined to some of the history of regional flow frequency analysis.

2.1.2 Regional Flow Frequency Analysis

This section will review the literature on regional flow frequency analysis under the following subheadings that constitute the general procedure of the analysis:

- Data screening;
- Delineation of homogeneous regions;
- Regional homogeneity test;
- Selection and estimation of regional frequency distribution;
- Estimation of flow magnitudes; and
- Quantile estimation accuracy assessment.

2.1.2.1 Data Screening

The first essential step of any statistical data analysis is to check that the data are appropriate for the analysis. For frequency analysis of any hydrological event, the data collected at a site must be a true representation of the quantity being measured and must be drawn from the same frequency distribution. It is also based on the assumption that the data are random, independent and homogeneous. For hydrological data errors could be due to incorrect recording, systematic changes over time (type or location of the measuring instrument), human-induced flow regulation, or any combination of these. These errors may cause data to have outliers, non-homogeneity, serial correlation, and trends which subsequently reduce the reliability of the frequency analysis based on these data.

Statistical tests for outliers and trends can be found in the literature (e.g. Kendall, 1990; Barnett and Lewis, 1994). Double-mass plots and quantile-quantile plots are some of the techniques that can be used for between-site comparisons. In addition, there are many computer software packages that can perform tests for outliers, trends, and serial correlation (e.g. Environment Canada CFA 3.1). In the context of regional frequency analysis using L-moments, Hosking and Wallis (1997) found that comparing sample L-moment ratios of different sites provide useful information. They noted L-moments of the data can reflect the incorrect data values, outliers, trends, and shift in the mean of a sample. They introduced a composite statistic based on L-moment ratios, a measure of discordancy between the L-moment ratios of a site and the average L-moment ratios of a group of similar sites, called the discordancy measure (D_i). The details on computations and interpretation of D_i statistics are given in Section 3.2.5.1.

2.1.2.2 Delineation of Homogeneous Regions

The identification of homogeneous regions is usually the most difficult stage in a regional frequency analysis, and requires the greatest amount of subjective judgment. The aim is to form groups of sites such that their frequency distributions are identical except for a site-specific scale factor (Hosking and Wallis, 1997).

Several methods have been proposed for grouping similar sites into regions and for use in the regional frequency analysis which can be roughly categorized based on the following basis:

Geographical convenience: Regions are often defined by sets of contiguous sites, based on administrative areas (e.g. FEH, 1999; Beable and McKerchar, 1982), or major physiographic sites grouping (e.g. Matalas et al, 1975). However, as Wiltshire (1986) and Acreman and Sinclair (1986) discussed, geographical proximity could not guarantee hydrological homogeneity, as some neighboring basins could be physically very different. Kachroo et al (2000) in a more recent study utilized sound judgment about the hydrological responses of the basins based on geographic information and similarity of the statistics of the observed flow data. The geographic regions they delineated were found to be hydrologically homogeneous.

Subjective partitioning: Regions can be defined subjectively by inspection of the site-characteristics, especially for small-scale studies. Therefore, the formed regions may or may not be geographically contiguous. The resulting regions from this subjective method can be objectively tested by heterogeneity measure described in Section 3.2.5.3. The Government of Newfoundland and Labrador (1991) study divided the Island of Newfoundland into three regions based on site characteristics factors. Gingras et al. (1994) is an example of subjective partitioning as well. They formed regions for annual maximum streamflow data in Ontario and Quebec by grouping the sites according to the time of year at which the largest flood typically occurred. It should be noted that the use of at-site statistics in subjective partitioning is discouraged as this might affect the validity of test of homogeneity which is usually based on the at-site data itself (Hosking and Wallis, 1997).

Objective partitioning: In this method of partitioning, the sites are assigned to one of two groups depending on whether a chosen site characteristic does or does not exceed some threshold value. This threshold value is chosen to minimize a within-group heterogeneity criterion. Wiltshire (1985) used a single measured partitioning value of one or more basin characteristics to group the basins. In an iterative fashion, the optimum size of the region would be defined by minimizing the within-group departure of these statistics. Pearson (1991a) applied similar procedure and used within-group variation of sample L-moments. Hosking and Wallis (1997) described this procedure as an effective 'objective partitioning' approach. They used it in conjunction with an efficient homogeneity test (heterogeneity measure) as defined in Hosking and Wallis (1993). Pearson (1991b) successfully applied this heterogeneity measure along with Wiltshire's (1985) partitioning criterion for regionalization of streamflow data for small drainage basins in New Zealand.

Cluster analysis: It is a standard method of statistical multivariate analysis and it is used for dividing data into groups. This method has been successfully used to form regions in regional frequency analysis. In this method, a data vector represents the characteristics of a site and the sites are grouped according to the similarity in their respective data vectors. De Coursey (1973) was the first one who applied cluster analysis to form groups of sites having similar peak flow response. Acreman and Sinclair (1986), Burn (1989), Guttman (1993), and Lim and Lye (2003) are some of examples of using this partitioning method for identifying homogeneous regions in regional frequency analysis.

Hosking and Wallis (1997) regard cluster analysis of site characteristics as the most practical method of forming regions from large data sets. However, they noted that the output of this analysis should not be considered final and it needs subjective decisions at several stages. In addition, they provided insight into the maximum and minimum size of the regions to be formed by this procedure for use with the index flow method.

2.1.2.3 Regional Homogeneity Tests

Once regions are formed based on the physical characteristics of the site, it is required to assess whether the regions are hydrologically homogeneous, so that the information obtained from the region is useful for flow frequency analysis. It should be tested whether a region is homogeneous or it needs to be divided into more regions, or whether two or more homogeneous regions are similar and so should be combined to form one homogeneous region. The hypothesis of homogeneity is based on the assumption that the at-site frequency distributions of the observed data at the sites in a homogeneous region are identical except for a site-specific scale factor. This test is constructed as a statistical significance test of the similarity of appropriately chosen statistics calculated from the distribution of at-site data. However, selection of which statistic to use and which distribution to assume for the at-site data has remained controversial for the last few decades. This test examines the similarity between the at-site distribution and hypothesized regional distribution. Some of the regional homogeneity tests in the literature are reviewed next.

Dalrymple (1960) apparently is the first published literature on a regional homogeneity test. He suggested a procedure to test homogeneity of a region for the index flow method based on the study of 10-year flood estimated from the Gumbel frequency distribution at each gauging station within the region. Wiltshire (1986 a, b) proposed the next two approaches after Dalrymple (1960) based on statistical hypothesis tests. The first approach involved testing the regional homogeneity based on the coefficient of variation (CV) of standardized annual maximum series, whereas the second approach was a distribution based procedure and used the geometry of the cumulative distribution function of the dimensionless regional parent. He concluded that the second approach is better in terms of statistical power. In order to evaluate the regional homogeneity, Wiltshire used a non-parametric jack-knife procedure to estimate the at-site distribution, unlike Dalrymple who assumed Gumbel distribution as the parent distribution at each site. Hosking and Wallis (1993) proposed the next important statistical test for homogeneity test based on the sample L-moments ratios. Chowdhury et al. (1991) suggested another statistical test based on L-moments which was more powerful than previous tests; however, the most rigorous L-moment based test of homogeneity is that of Hosking and Wallis (1993). It compares the variability of the L-moment ratios of the sites within a region with the expected variability obtained from simulation from a collection of sites with the same record length as their real world counterparts. A heterogeneity measure is then calculated based on the difference between the weighted standard deviation of the sites' L-CVs in the region and the mean of the same statistics obtained from the simulation. Hosking and Wallis (1997) used a 4-parameter Kappa distribution for their simulation. This test has been used as a standard test of homogeneity in recent years (e.g. Castellarin et al., 2001; Lim and Lye, 2003). Details of this test are discussed in Section 3.2.5.3.

2.1.2.4 Selection and Estimation of Regional Distribution

After confirming the homogeneity of a region in regional frequency analysis, a single frequency distribution is fitted to the data from several sites within that region. The candidate distributions are usually evaluated for the accuracy of the quantile estimates for each site. There are many families of distribution that might be candidates to be a regional frequency distribution. The choice of this distribution can be evaluated by considering its ability to reproduce features of data that are of particular importance in modeling. There may be a particular range of return periods for which quantile estimates are required, for example, in analysis of extreme events such as drought, quantiles of one tail of frequency distribution are of particular interest. Matalas and Wallis (1973) mentioned that the competing distributions that fit the observed data satisfactorily may differ significantly in their tails. These considerations may affect the choice of a regional frequency distribution. Therefore, 'robustness' was recognized to be the most important property of a frequency distribution employed for regional analysis.

Different regional frequency distributions were selected in several regional studies. For example, the Flood Estimation Handbook (1999) recommended an index-flow method employing the GEV distribution for a site with short period of record. Durrans and Tomic (1996) applied the log-Pearson III distribution to regional low flow frequency analysis. Chen et al. (2006) analyzed low flow frequency in South China, and selected the three-parameter lognormal distribution as the most appropriate distribution for the region.

As the purpose of regional frequency analysis is to augment the data at one site, it was possible to fit a three or more parameter distribution, more reliably. Hosking and Wallis (1997), thereby noted that distributions with three to five parameters are appropriate candidates for regional frequency analysis, because they yield less biased estimates of quantiles in the tails of the distribution. It is possible that more than one distribution fits the data adequately; in this case, the best choice would be one that provides the most robust and efficient quantile estimation. Furthermore, Hosking and Wallis suggest that the final choice of distribution should be made based on 'goodness-of-fit' tests of the candidate distributions. They provided an approach that directly involves the regional average L-moments. For a three-parameter distribution, the goodness-of-fit is judged by how closely the L-kurtosis of the fitted distribution matches its regional average counterpart of the observed data.

McCuen (1985) introduced the moment ratio diagram which is a tool to visually judge the fit of a particular data set to a theoretical distribution. The basic advantage of using this diagram is that a single diagram can visually compare the fit of several distributions for a given set of data. In the regional context, the position of regional average dimensionless moments on the diagram would give closer resemblance of the underlying regional distribution. Later on, Hosking (1990) introduced the L-moment ratio diagram. Vogel and Fennessy (1993) showed that the L-moment ratio diagrams are more accurate than the product moment diagrams in discriminating between the distribution and they proposed to replace the product moment diagram with L-moment diagram in hydrological investigations. However, Hosking and Wallis (1997) indicated that the L-moment diagrams is only a tool in selecting the candidate distributions and final distribution selection should be made using more objective test that reflects the robustness of the distribution. Details on regional frequency distribution selection are provided in Section 3.2.5.4

2.1.2.5 Estimation of Flow Magnitudes

The frequency distributions at the sites within a homogeneous region are assumed to be identical apart from a scale factor, and a probability distribution will have been chosen for fitting to each region. Several methods have been proposed for fitting a distribution to data from homogeneous regions, for example, methods based on index-flood (index-flow), station-year, and maximum likelihood procedures.

The index-flood procedure was first introduced by Dalrymple (1960), in which the observed annual peaks at each site are first standardized by dividing each data point by its sample mean (the index-flood), and then all the standardized observations are used to estimate an average dimensionless frequency distribution (growth curve). Then, the quantile for each site within the homogeneous region is calculated by multiplying the quantile estimate of the regional growth curve by the site's sample mean (the index-flood) of annual records. The procedure is called index-flood because of its first application in flood studies. However, in this study, it has been called the index-flow procedure without loss of generality. The index flow procedure is very popular among practicing hydrologists, and have been adopted in many regional frequency studies with limited modifications.

The well-known station-year method combines the rescaled data (by site-dependent scale factor) from all sites into a single sample and fits a distribution by treating the combined samples as a single random sample. This method is now rarely in use, as in many cases, it is not appropriate to treat the rescaled data as a single random sample.

An approach based on maximum likelihood estimation treats data as a statistical model that is completely specified by scale factors and unknown parameters of a regional growth curve. These parameters can be estimated by using the method of maximum likelihood. This method has been used for example by Loaiciga and Marino (1988).

As discussed before, the main goal of regional analysis is to be able to estimate the flow variables at a site where there are no records available. In this case, the index flow variable at the site of interest must be estimated in another way, as there is no flow recorded at this site. Usually, the index flow is estimated from a regionally calibrated linear or log-linear relationship between the mean or median flows and physically measurable catchment characteristics (Lim and Lye, 2003; Mostofi Zadeh et al., 2012). The US Geological Survey (Thomas, 1987; Tasker, 1987) proposed a different approach from the index flow. They estimated the flow quantile of interest at every station, and regressed these quantiles from a homogeneous region to their respective sets of significant catchment characteristics. The quantiles at the site of interest would be obtained by substituting the important catchment characteristics in the respective regional regression relations. This method has been widely used all over the world as well and it is known as 'regression on quantile' method of regional analysis. Two advantages of this method over the index flow method is that, firstly, it avoids specifying a regional average frequency distribution (the growth curve), and secondly, it uses the regression techniques that are readily understood by hydrologists. However, the introduction of L-moments has firmly re-established the index flow method as a general procedure for regional flow frequency analysis, because the extent of distribution selection and parameter estimation problem in the index flow method have been significantly reduced by using L-moments.

Hosking and Wallis (1993) provided a general framework for carrying out index flow based regional frequency analysis using L-moments. As the L-moments approach gained popularity among hydrologists, the index flow method based on L-moments has been accepted as a standard method of regional frequency analysis in recent years. Most of the applications of this methodology were in flood frequency analysis. However, some researchers have attempted to apply this mature method to low flow frequency analysis. Pearson (1995), Durrans and Tomic (1996), Tate et al. (2000), Kroll and Vogel (2002), Chen et al. (2006), Modarres (2008), and Shi et al. (2010) are some of the examples of regional low flow frequency analysis based on L-moments. The suggested L-moments algorithm by Hosking and Wallis (1997) is summarized in Section 3.2.5.4.

2.2 Flow Duration Analysis

A flow duration curve (FDC) is one of the most informative methods of displaying the complete range of river discharges from low flows to flood events. It displays the relationship between streamflow and the percentage (probability) of time it is exceeded. (McMahon et al, 2004). Figure 2-2 shows a typical flow duration curve. In other words, it is the relationship between magnitude and frequency of streamflow discharges with no regards to their sequence of occurrence. The later falls mostly within the scope of flow spell analysis. Despite the wide use of FDCs in hydrological practice, the relevant literature is rather limited.

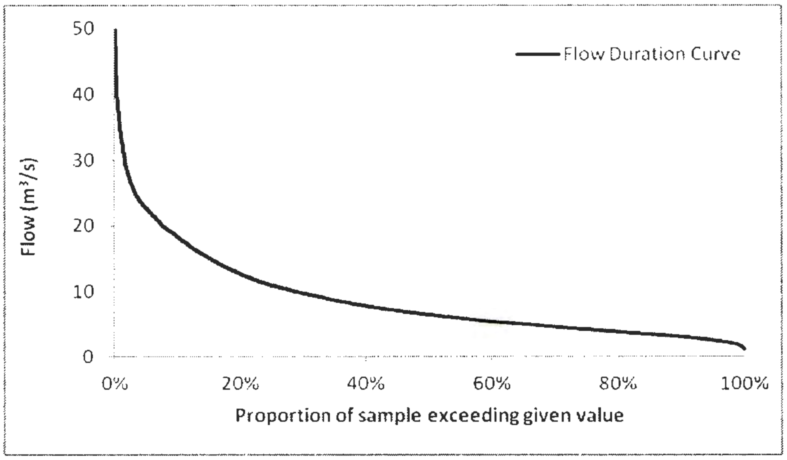

Figure 2-2 Daily flow duration curve

2.2.1 Flow Duration Curve Construction

In general, a FDC is constructed by reassembling the flow time series values in decreasing order of magnitude, assigning flow values to class intervals and counting the number of occurrences (time steps) within each class interval. Cumulative class frequencies are then calculated and expressed as a percentage of the total number of time steps in the record period. Finally, all ranked flows are plotted against their rank which is again expressed as a percentage of the total number of time steps in the record (Smakhtin, 2001).

FDC may be constructed using different time resolutions of streamflow data such as annual, monthly or daily. In addition they may also be constructed using some other time intervals, for example, from *m*-day or *m*-month average flow time series. FDCs constructed on the basis of daily flow time series provide the most detailed way of examining duration characteristics of a river. More details on construction and interpretation of FDCs are provided in some sources (e.g. Searcy, 1959; Institute of Hydrology, 1980; McMahon and Mein, 1986)

According to the period of record used for constructing FDCs, they can be divided into two major groups: (1) on the basis of the whole available record period; (2) on the basis of a portion of calendar (month or seasons). The shape and general interpretation of any FDC depend on hydrometric errors and particular period of record on which it is based. (Smakhtin, 2001)

2.2.1.1 Period of Record FDCs

These FDCs are calculated on the basis of the whole available record period of streamflow. Vogel and Fennessy, (1994) described this as Period of Record FDCs or (POR FDCs). Smakhtin et al. (1997) introduced a long-term average annual FDCs. He constructed a non-dimensional FDC for each flow gauge by dividing discharges by the long-term mean daily flow which was estimated as the average of all daily streamflows in the available record.

2.2.1.2 Monthly or Seasonal FDCs

These FDCs are constructed on the basis of all similar calendar months or all similar seasons from the whole period of record. (e.g. all Januarys, or all summers). Smakhtin et al. (1997) have used these FDCs. They may also be constructed for a particular season (e.g. Winter 2000) or a particular month (e.g. January 2009)

2.2.1.3 Annual FDCs

The period-of-record FDC represents variability and exceedance probability of flow over the available or selected period. Vogel and Fennessy (1994) introduced a different interpretation of a FDC. They constructed FDCs for individual years and treated them in the way similar to a sequence of annual flow maxima or minima. Their interpretation allows mean and median FDCs to be estimated. These median FDCs represent the exceedance probability of flow in a typical year (not wet, nor dry). These curves were demonstrated to be less sensitive to the length of the record period, especially in the area of low flows. This approach also allows constructing confidence intervals and return periods for FDCs.

2.2.2 Application

Searcy (1959) was the first to summarize a number of FDC applications including the analysis of catchment geology on low flow, hydropower and stream water quality studies. Alaouze (1989, 1991) developed procedures based on FDC, for estimation of optimal release schedule from reservoirs. Mallory and McKenzie (1993) and Pitman (1993) employed FDCs in design of flow diversions. Hughes et al. (1997) developed an operating rule model which is based on FDCs and is designed to convert the original tabulated values of estimated ecological instream flow requirements for each calendar month into a time series of daily reservoir releases. Vogel and Fennessy (1995) provided a review of numerous possible applications of FDCs in engineering practice, water resources management and water quality management.

2.2.3 Interpretation and Indices

Flow duration curves are a convenient way of portraying the flow characteristics of a stream. The shape of the FDCs summarizes the flow characteristics of a stream for comparison with other streams. The slope of a FDC reflects the catchment's response to precipitation. The low flow end of the curve is valuable for interpreting the effect of geology on low flows. If groundwater contributions are significant, the slope of the curve at the lower end tends to be flattened whereas a steep curve indicates low baseflows (McMahon, 1976). Searcy (1959) suggests that streams draining the same geologic formations will tend to have similar FDC at the low flow end.

Flow duration curves can provide a number of indices to characterize the stream for classification and regionalization purposes. Of most interest for low flow studies is the low flow section of FDCs, which may be arbitrarily defined, for example, as part of the curve with flows below the median which corresponds to the discharge equaled or exceeded 50% of time or Q50 to Q99.

2.2.4 Flow Duration Curve Estimation for Ungauged Catchments

FDC construction and calculation methods described above require adequate observed streamflow records which can only be provided for gauged catchments. However, it is often needed to predict these flow quantities for ungauged catchments. Possible approaches for flow duration curve estimation in ungauged catchments can be classified as: (i) mathematical methods, regional regression models of some low flow index with catchment physiographic characteristics; (ii) graphical methods, based on the construction of regional prediction curves.

2.2.4.1 Regional Regression Approach

The regional regression approach is perhaps the most widely used technique in low flow estimation at ungauged sites (Smakhtin, 2001). Like all the regional approaches, it normally includes several subsequent steps as follow:

- Selection of flow characteristics of regression model
- Delineation of hydrologically homogeneous regions
- Construction of regression model

The flow characteristic for the regression model, in this case, is the constructed flow duration curve for available gauged watersheds. The regionalization of streamflow characteristics in general is based on the premise that watersheds with similar geology, topography, land cover, and climate would normally have similar streamflow responses. Therefore it is important to delineate regions that have similar catchment physiographic and climatic parameters. The identification of homogeneous regions is normally required for large territories such as countries or large regions but may be skipped for smaller regions. The regression model is a relationship between the dependent low flow characteristic (quantiles of flow duration curve in this case) and independent catchment and climatic variables. Vogel and Kroll (1992) found that low flow characteristics are highly correlated with catchment area, average basin slope, and base flow recession constant. Technically, the regression model is constructed by means of a multiple regression analysis. The parameters of the regression models have been traditionally estimated using the principle of ordinary least squares (OLS). This step includes selection of the types of regression model, estimation of regression model parameters, and assessment of estimation errors. The procedure is described in many text books (e.g. Yevjevich, 1972) and can be performed using standard statistical software packages. However, it is not an easy task to uncover a true relationship between these dependent and independent variables.

Singh et al. (2001) found that the statistical approach of nondimensional quantile estimation of flow duration curve performs satisfactorily in calibration as well as in validation for large number of Himalayan catchments. Archfield et al. (2007) developed two sets of regional regression equations to estimate daily period of record FDCs at ungauged sites in Southern New England. The first method assumed an underlying probability density function (pdf) of Kappa distribution for daily streamflow whose parameter values are related to the physical characteristics of the watershed using a regression approach. The second method related flows at selected exceedance probabilities on the FDC to physical characteristics of the watershed. It was observed that FDC estimates from regression equations developed for individual exceedance probabilities had better results and led to lower mean square error than estimates of FDCs that assumed an underlying pdf. Mohamoud (2008) presented a method to predict flow duration curves and streamflow for ungauged catchments in Mid-Atlantic region, USA. 15 percentile flow points from constructed normalized FDCs were selected for each study catchment. A step-wise regression method was used to develop models for these flow percentiles using landscape and climate descriptors of study region. The method was tested by predicting the 15 percentile flows for the ungauged evaluation sites, reconstructed complete FDCs for them, and finally, evaluated the prediction performance of the method by comparing reconstructed FDCs and observed streamflow FDCs.

2.2.4.2 Regional Prediction Curve

As opposed to the estimation of a single quantile of flow duration curve for which regression model has been constructed, the regional prediction curve approach allows the range of flow indices to be estimated. In this approach, the flow duration curves from a number of gauged catchments of varying size in a homogeneous region are converted to a similar scale, superimposed, and averaged to develop a composite regional curve. To make curves from different catchments comparable, all flows are standardized by catchment area, mean or median flow or other 'index' flows. A curve for ungauged site may then be constructed by multiplying back the ordinates of a regional curve by either catchment area or an estimate of index flow, depending on how the flows for the regional curve were standardized. The index flow is estimated either by means of a regression equation or from regional maps. (Smakhtin, 2001)

The first attempt to construct the regional FDCs was by Lane and Lei (1949). They have designed the 'variability index', a measure of streamflow variability specifically related to FDC and calculated as the standard deviation of the logarithms of 5, 15, 25,..., 85 and 95 exceedance flow values. They determined the average value of variability index and correlation to this index which are dependent on the physiography of the individual ungauged river catchment.

Regional FDCs have been constructed in several different parts of world in the last few decades and will briefly be reviewed here. Singh (1971) and Dingman (1978) constructed regional FDCs in several states in USA. Quimpo et al. (1983), Mimikou and Kaemaki (1985), Wilcock and Hanna (1987), Tucci et al. (1995), Niadas (2005) constructed regional FDCs in Philippines, Greece, Northern Ireland, Brazil and Greece respectively.

Fennessy and Vogel (1990) tried a different approach for regional FDCs. They approximate only the lower half of 1-day annual FDCs by fitting log-normal distribution to it and developing a regression equation for log-normal distribution parameters with catchment characteristics. Smakhtin et al. (1997) constructed 1-day annual and seasonal regional FDCs for one of the primary drainage regions of South Africa and used them to generate a continuous daily streamflow hydrograph at ungauged sites. Franchini and Suppo (1995) proposed a methodology for regional analysis of the drought part of a flow duration curve for a limestone region. They mathematically described the lower part of flow duration curve using discharge as a function of duration, and finally using a physiographic regression to use this equation in ungauged locations. Viola et al. (2011) performed a regional analysis on flow duration curves in Sicily, Italy. They fitted a relationship between duration of wet periods and related discharges, and predicted the parameters of this relationship using a regression between them and watershed morphological data.

2.3 Flow Spell Analysis

2.3.1 Environmental Instream Flow Requirements

Environmental instream flows are defined as flows in a river that are deemed as minimum flow required for maintaining the river ecosystem (Karim et al., 1995). Three main ways in which these flows are evaluated have been described in the literature, the Habitat method developed by the Washington Department of Fisheries (Collings, 1972); the Hydraulic Rating method; and finally the Hydrological method. The last method has been recognized as the easiest method to estimate the environmental instream flow requirement so far, and therefore has been very popular (Caissie and El-Jabi, 1995). However, all the aforementioned methods fail to indicate when and how often low flows occur. For this reason, it is necessary to study the continuous low flow events and deficit volumes. Jowett (1997) performed a comprehensive review of these methods.

The Habitat method is based upon sampling river data at various cross-sections to quantify the parameters necessary for the fish species development; such as average water velocity, water temperatures, depth and sediment transfer. Some habitat features, such as depth and velocity, are directly related to flow, whereas others describe the river and surroundings. It might not be a good assumption that the flows measured at one location in a river give an accurate representation of the suitability of that flow to support aquatic life for the whole river.

The Hydraulic method relates various parameters of the hydraulic geometry of stream channels to discharge. The hydraulic geometry is based on surveyed cross-sections, from which parameters such as width, depth, velocity and wetted perimeter are determined. Because of the field and analytical work involved in this, they are more difficult to apply than the hydrological method.

The hydrological method is based on the history of flow and relies solely on the recorded or estimated flow regime of the river. There are several ways in this method to describe the environmental instream flow requirement. Some methods assume that some percentage of the mean flow is needed to maintain a healthy stream environment. Other hydrological methods recommend flows based on the flow duration curve or an exceedance probability. The choice of method to estimate the environmental instream flow depends upon the particular requirement that is being considered for the aquatic ecosystem.

2.3.2 Continuous Low Flow Events and Deficit Volumes

Prolonged streamflow below the determined environmental instream flow requirement can imply high economic, ecosystem and even human loss where rivers act as water supply systems or as inflows to hydropower. Therefore, it is necessary to be able to predict rivers droughts and have knowledge about their time of occurrence and their duration. However, neither flow duration curve nor low flow frequency distribution provides information about the length of continuous periods below a particular flow value of interest (environmental instream flow requirement), how the low flow days are distributed during a year, and also give no indication of a possible deficit of flow which is built up during a continuous low flow event. There exist ways to overcome these limitations which will be described in the next section:

2.3.2.1 Theory of Runs

A widely used approach applies the 'truncation level' or 'threshold' concept. It originated from Yevjevich's theory of runs (Yevjevich, 1967) A run in drought hydrology is defined as the number of days (months, years) when daily (monthly, annual) streamflow remains below the certain threshold flow. This threshold flow can be described as the environmental instream flow requirement. As it was mentioned before, there exist different choices for this threshold value, dictated by the objective of the drought study or the type of flow regime. For example, in the case of drought hydrology of perennial rivers, threshold flows in the range of discharges between 70-90% exceedance on flow duration curve is meaningful. But for ephemeral rivers which flow only after significant rainfall events, discharges as high as those with 20% exceedance are not defined as unreasonably high drought thresholds (Tate and Freeman, 2000). In addition, different water resources practices and water users have different water requirements and may not have a common opinion on what threshold should be used to define a drought event (Smakhtin, 2001).

The theory of runs consist of three main low flow characteristics which are the run duration; the run severity (cumulative water deficit or the negative run sums); and the run magnitude (the intensity of flow deficit over its duration).

Some drought studies focus on the longest run duration for each year of record, such runs often interpreted simply as annual hydrological droughts. Clausen and Pearson (1995), Tallaksen et al. (1997), Stahl and Demuth (1999), Modaress (2009), and Modarres and Sarhadi (2010) are some examples of these studies. Other studies take into account other associated characteristics of droughts such as the start date of the longest run (e.g. Woo and Tarhule, 1994), or the end of the run durations (e.g. Tlalka and Tlalka, 1987). Whatever drought characteristic (longest run duration, or deficits, intensities, run start date, etc) has been chosen, it will yield one value for each year of streamflow record. These data like all other frequency analyses, can be ranked, assigned a probability or return period using a plotting position formula, and plotted against the assigned return period for a given value of threshold flow.

2.3.2.2 Flow Spells

Flow spell is a similar procedure with only a different terminology developed by the UK Institute of Hydrology (IH, 1980). Flow spell analysis, in contrast with the flow duration curve which gives no information on how the low flow days are distributed, considers how long a flow below some threshold has been maintained and how large a deficit has been built up, and therefore takes into account the sequencing of flows. In this approach, the run duration becomes spell duration, and the total volume of flow that would be required to maintain the flow at a given threshold is called deficiency volume with the same descriptions as provided in the previous section. The intensity of spells is another measured which describes the suddenness of the flow spell by dividing the deficit volume over its duration of occurrence.

For the rivers which regularly fall to zero-flow condition, the frequency of duration of continuous zero-flow periods may be of interest and analyzed by common statistical procedures as described before. Continuous zero-flow periods analysis is effectively a specific case of spell analysis which indicates the likelihood of extended periods of no flow or droughts (e.g. Armentrout and Wilson, 1987).

Spell analysis can also be applicable for the periods of high flows (Prudhomme and Gilles, 1997). It may also useful for the study of even more specific events, like short-term freshes, small peaks caused by occasional rains during prolonged low flow periods, which may have important ecological implications (Smakhtin, 2001).

Nathan and McMahon (1990a) developed a method for regionalizing spell duration and deficiency volume frequency curves. They considered the data to be log-normally distributed, thus the frequency curves plot as straight lines on log-normal probability paper. Therefore, estimating only two points would define a curve. Nathan and McMahon suggest developing regional equations using multiple-regression analysis for the prediction of 2- and 50-year events. These points are plotted on log-normal probability paper and a straight line is drawn through them. In that study, catchments were first divided into hydrologically homogeneous regions and then regression techniques used to select and weight the most important variables. For multiple regression equations, they found that the most important variables were mean annual rainfall and estimated ratio of baseflow to total streamflow.

In general, flow spell analysis has been used for estimating the amount of storage needed on a catchment to maintain water supplies, and to check the representativeness of synthetically generated streamflow time series (McMahon and Mein, 1986). In the analysis of flows for ecological and water quality requirements, flow spells are more valuable than flow duration curves.

Both the run and spell analyses have been in wide use for the identification, characterization, and management of annual or multiyear hydrological droughts. Chang and Stensoon (1990), Wijayarante and Golunb (1991), Clausen and Pearson (1995),Burn and DeWit, (1996), and El-Jabi et al. (1997) are some examples of these studies.

2.4 Previous Low Flow Studies in the Province

The first study on the characteristics and estimation of low flow in the province of Newfoundland and Labrador was carried out by Government of Newfoundland and Labrador on 1991(Govt. Newfoundland and Labrador, 1991). This study had two objectives: (1) characterize low flows in streams across the Island of Newfoundland and (2) develop equations which could be used to calculate low flows of various return periods at ungauged sections of rivers. Two low flow periods were selected for analytical purposes, one during the winter season between January and March, and the other during summer season between July and September. 39 gauges were selected across the Island where the data records exceeded 8 years. Minimum N-day low flows during winter and summer were obtained based on daily flows over the period of record for these gauges to cover the duration of low flow events as an important factor in low flow estimation, where the durations, N, were 1, 7, 15 and 30 days. The Gumbel Type III distribution was fitted to the summer and winter low flow series. It was found that summer low flows are generally lower than the winter low flows and are more likely to reach zero values and exhibit high variance. The database on low flow frequency estimates was, therefore, the summer low flow frequency values. The Island by judgment was divided into three different regions in this study. These regions were hydrologically, climatically and physiographically distinct from one another. In order to have an estimate of low flows at different return periods, regression analysis was performed between frequency estimates of low flows obtained from fitted Gumbel distribution and watershed parameters. The results of the analysis indicated that magnitudes of low flows across the Island were very highly correlated with drainage area, precipitation amounts, and type and extent of land cover. Drainage area and percentage of drainage area covered by forests were the significant independent basin characteristics in the estimation of low flow magnitudes. The results of comparison between frequency estimates and predicted values from regression equations were not very promising. The percentage difference between frequency and regression estimates ranged between -50% and +50%, on the average. Some percentage differences were very high, over 100%. These occurred mostly when the frequency estimate was very close to zero. Finally, the obtained regression equations were tested on 21 watersheds for which some data were available but these data were not used in the derivation of equations. Results indicated that relatively high percentage difference between frequency and regression estimates can be expected, particularly, when the frequency estimates are close to zero. It was recommended that the study should be repeated when more stations and longer periods of data are available within the region to improve the frequency estimates and therefore the regression equations.

Richter and Lye (1995) comprehensively reviewed the history of regional analyses in the Island of Newfoundland and performed the next hydrological study on the rivers of the Island. Identification of the key basin characteristics associated with a range of flow measures was carried out in this study. Data set of 40 stations with more than 10 years of data was used. The flow measures were selected to represent average, high and low flow regimes. They defined effective precipitation (Eff-P) as the average runoff depth over a basin. Eff-P was used as a basin's hydrological input due to lack of precipitation measured in the Island. Then it was attempted to relate Eff-P to topographic and geographic variables, thus it can be estimated for any ungauged basin. The nonlinear multiple regression analysis of Eff-P showed that distance from Southwest of the Island, elevation of centroid of the basin, and fraction of barren area are the most important explanatory variables. In the next step, relationship among flow variables in three flow categories (high, low and available flow) and basin characteristics were investigated. The most explanatory variables were found to be drainage area, area controlled by lakes and swamps, fraction of barren area in the basin, and distance of the basins north and/or southwest of defined lines. Finally, grouping of the basin into regions of geographic and basin characteristics was performed. The mean annual maximum daily flow was the measure of interest for the regionalization of the basins. A detailed assessment of several methods of regional subdivision was carried out and it found that dividing the Island into regions generally improves the estimates at ungauged sites. Clustering based on the basin characteristics was reported as a promising method of regionalization.

The next study was conducted on duration, volumes, and intensities of flow spells for 20 rivers in Newfoundland and Labrador, all with at least twenty years of consecutive unregulated daily flows. (Shaughnessy, 1997) Different methods of estimating environmental instream flow requirement were used as the threshold values. Four of the threshold values were constant during the year, including 25% mean annual flow, seven day consecutive low flow with ten year return period (7Q10), and 85 and 95 percentiles of period of record flow duration curve. Tennant's method was selected as seasonally changing threshold value. And finally, 50^{th} and 90^{th} percentile of monthly flow duration curves were used as two monthly threshold values. Linear relationship between threshold values and catchment area was attempted in this study, but found to be not statistically significant at the 5% level with the exception of 7Q10 method. Moreover, a comparison of the severity of the spell periods according to the different threshold methods was undertaken for ten out of twenty rivers in terms of annual maximum duration, volumes and intensities. Linear relationships were tentatively obtained between the threshold value and annual maximum volumes and annual maximum intensities for 18 rivers. Excluding Labrador rivers, most frequently the lognormal distribution appeared to fit the annual maximum variables, particularly for annual maximum durations. Based on the time of occurrence of annual maximum spells, it was concluded that Tennant's method provides the best hydrological instream flow requirements both for fish as well as being reasonable for water abstraction. The number of suitable gauged rivers and their record period were rather short in this study. The current study will revisit this study by Shaughnessy with the additional data available so far.

2.5 Rationale of the Thesis

From the preceding review of the literature in the developments of the regional analysis in low flow frequency, flow duration curves, and flow spells, it is apparent that several approaches have been used so far. Among the most popular in recent years for regional frequency analysis of streamflows is the index flow based on L-moments which will be applied in the current study for frequency analysis of low flows in Newfoundland and Labrador. The previous studies showed that the index flow method based on L-moments has been successfully applied in regional flow frequency analyses and therefore it will be applied in the current study. The regional regression method will also be used to estimate the index flow for flow frequency analysis. The regional regression approach is the method which will be adopted in this study herein for regional analysis of flow duration curves and flow spells of the rivers of the province. The next chapter will review the methodologies of the aforementioned methods for the different aspects of low flow analysis.

3 Methodology

3.1 General

This chapter will review the methodologies applied in this study for conducting regional low flow frequency, flow duration, and flow spells analyzes. The first section will describe the index flow method of regional flow frequency analysis based on L-moments. The second section will describe the regional flow duration approach which has been adopted in this study. And finally the third section will describe the flow spell analysis procedure followed in this study.

3.2 Regional Low Flow Frequency Analysis

It has been well accepted that using a regional approach in any frequency analysis is effective in extending the available information at a site to sites within a homogeneous region, or creating information when there is no data available at a site of interest. Estimating extreme flows using a regional approach can be carried out using methods such as the index flow method and the direct regression on quantiles method. The purpose of this study is to adapt the mature method of index flow based on L-moments in regional frequency analysis which has been extensively applied for flood flows to low flows (e.g. Lim and Lye, 2003; Modarres, 2008; and Shi et al. 2010).

The definition and procedure for deriving L-moments will be discussed next followed by the stepwise procedure necessary for conducting a L-moment based regional flow frequency analysis and the procedure to estimate the index flow. All the provided procedures for regional frequency analysis in this section are largely based on Hosking and Wallis (1997) manuscript introducing L-moments based approach.

3.2.1 Probability Distribution

The fundamental quantity of statistical frequency analysis is the frequency distribution. Let X be a random variable which in this study is flow magnitude at a given time at a given site. X takes values that are real numbers. For example, suppose that observations are made at regular intervals at some site of interest. X is regarded as a random quantity or a random variable, taking any value between zero and infinity. The relative frequency with which these X value occur defines the *frequency distribution* or *probability distribution* of X and is specified by the *cumulative distribution functionF*(x), the probability that the actual value of X is at most x:.

$$F(x) = \Pr[X \leq x] \tag{3-1}$$

$F(x)$, the cumulative distribution function of the frequency distribution is an increasing function of x, and its value is always between zero and unity for all x. We only consider continuous random variables here, therefore, the inverse function of the cumulative distribution function exists and it is called the quantile function of the frequency distribution and denoted by $x(F)$. It expresses the magnitude of an event in terms of its nonexceedance probability F, that is, the value such that the probability that X does not exceed $x(F)$ is F. In engineering and environmental applications a quantile is usually expressed in terms of its *return period*. The quantile of return period T, X_T, is an event magnitude so extreme that it has probability of $1/T$ of being exceeded by any single event. For an extreme low event, in the lower tail of the frequency distribution, X_T is given by:

$$X_T = x(1/T) \tag{3-2}$$

$$F(X_T) = 1/T \tag{3-3}$$

Accurate estimation of the quantiles of the distribution of random variable for return periods of interest is the main goal of frequency analysis.

3.2.2 Moments

Moments of the distribution has been traditionally used to describe the shape of a probability distribution. The first moment is the mean, μ, which is the center of location of the distribution. The dispersion of the distribution about its center is measured by the standard deviation, σ, or the variance,σ^2, The coefficient of variation (CV), $C_v = \sigma/\mu$, describes the dispersion of a distribution as a proportion of the mean. Dimensionless higher moments can also be used such as skewness and kurtosis which are obtained by ratios of the various central moments. Detailed information on how to derive these moments for a sample data can be found in Hosking and Wallis (1997).

However, Wallis et al. (1974) found that moment estimators have some undesirable properties. These estimators can be severely biased. Therefore, inferences based on sample moments were likely to be very unreliable. A more satisfactory measure of frequency distribution is obtained from L-moments as described next.

3.2.3 L-Moments

L-moments were defined as an alternative system for describing the shape of a probability distribution. Hosking (1990) derived L-moments by modifying the probability weighted moments (PWMs) which were defined by Greenwood et al. (1979) as follow:

$$M_{p,r,s} = E[X^p\{F(X)\}^r\{1 - F(X)\}^s] \quad (3\text{-}4)$$

If a distribution has a quantile function, $x(u)$, useful special cases of the probability weighted moments are described as:

$$M_{1,0,r} = \alpha_r = \int_0^1 x(u)(1 - u)^r du \quad (3\text{-}5)$$

$$M_{1,r,0} = \beta_r = \int_0^1 x(u)u^r du \quad (3\text{-}6)$$

These equations are similar to the definition of conventional moments which can be written as:

$$E(X^r) = \int_0^1 \{x(u)\}^r du \quad (3\text{-}7)$$

These PWMs were used as the basis for estimating parameters of probability distribution in the previous studies such as Landwehr et al. (1979a,b) and Hosking and Wallis (1987). However, PWMs method suffers from difficulties in directly interpreting as a measure of scale and shape of a probability distribution. Hosking (1990) solved this problem by considering certain linear combinations of the probability weighted moments. L-moments are defined by Hosking in terms of the PWMs α and β as follow:

$$\lambda_{r+1} = (-1)^r \sum_{k=0}^{r} P_{r,k}^{*} \alpha_k = \sum_{k=0}^{r} P_{r,k}^{*} \beta_k \quad (3\text{-}8)$$

Or for a random variable X with quantile function of $x(u)$ L-moments can be described as follow:

$$\lambda_r = \int_0^1 x(u) P_{r-1}^*(u) du \qquad (3\text{-}9)$$

Where $r = 0,1,2, \ldots$ and:

$$P_{r,k}^* = (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} = \frac{(-1)^{r-k}(r+k)!}{(k!)^2 (r-k)!} \qquad (3\text{-}10)$$

$$P_r^*(u) = \sum_{k=0}^{r} P_{r,k}^* \, u^k \qquad (3\text{-}11)$$

The following equations are the first four L-moments in terms of probability weighted moments:

$$\lambda_1 = \alpha_0 = \beta_0 \qquad (3\text{-}12)$$

$$\lambda_2 = \alpha_0 - 2\,\alpha_1 = 2\,\beta_1 - \beta_0 \qquad (3\text{-}13)$$

$$\lambda_3 = \alpha_0 - 6\alpha_1 + 6\alpha_2 = 6\beta_2 - 6\beta_1 + \beta_0 \qquad (3\text{-}14)$$

$$\lambda_4 = \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \qquad (3\text{-}15)$$

L-moments ratios are dimensionless version of L-moments and they are achieved by dividing the higher-order L-moments by the scale measure, λ_2. Therefore, they are measures of the shape of probability distribution independent of its scale of measurement. L-moment ratios are defined as:

L-CV: $\tau = \lambda_2 / \lambda_1$ (3-16)

L-skewness: $\tau_3 = \lambda_3 / \lambda_2$ (3-17)

L-kurtosis: $\tau_4 = \lambda_4 / \lambda_2$ (3-18)

The first L-moment, λ_1 is a measure of central tendency and is equivalent to the mean of the distribution whereas λ_2 is the measure of dispersion. Their ratio, L-CV is termed as the L-coefficient of variation, τ; the ratio of λ_3/λ_2 is referred to as τ_3 or L-Skewness; and the ratio of λ_4/λ_2 or τ_4 is called L-kurtosis.

The L-moments are easy to interpret as they are analogous to the conventional moments. Their popularity for use in regional frequency analysis procedure is growing because they are less biased than the conventional moments and they can better discriminate among the commonly used frequency distributions (Hosking, 1990).

3.2.4 Sample L-moments

L-moments have been determined for some of the known probability distributions, but it is necessary to estimate these L-moments from a finite sample of data. This estimation can be made based on a sample of size n arranged in ascending order. Let the order sample be $x_{1:n} \leq x_{2:n} \leq \cdots \leq x_{n:n}$. The following will be an unbiased estimator of probability weighted moment β_r:

$$b_r = n^{-1} \sum_{j=r+1}^{n} \frac{(j-1)(j-2)\ldots(j-r)}{(n-1)(n-2)\ldots(n-r)} x_{j:n} \quad (3\text{-}19)$$

Analogously to Eq. (3-12) to (3-15), the sample L-moments can be defined as follows, and based on these, L-moment ratios can be calculated.

$$l_1 = b_0 \quad (3\text{-}20)$$

$$l_2 = 2b_1 - b_0 \quad (3\text{-}21)$$

$$l_3 = 6b_2 - 6b_1 + b_0 \quad (3\text{-}22)$$

$$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \tag{3-23}$$

3.2.5 Steps in Regional Frequency Analysis

When the data are available at a large number of sites and the quantile estimates are sought at each site, regional frequency analysis using an index flow procedure based on the L-moments approach can be performed, and will involve the following steps.

In regional frequency analysis procedure, the following notations have been used. Suppose that there are N sites in the study region with sample size of $n_1, n_2, \dots, n_N$ respectively. The sample L-moment ratios at site i are denoted by $t^{(i)}, t^{(3)}$ and $t^{(4)}$. The regional average L-moment ratios are then given by:

$$\bar{t} = \frac{\sum_{i=1}^{N} n_i t^{(i)}}{\sum_{i=1}^{N} n_i} \quad \text{and} \quad \bar{t}_r = \frac{\sum_{i=1}^{N} n_i t_r^{(i)}}{\sum_{i=1}^{N} n_i} \quad \text{where } r = 3, 4, \dots \tag{3-24}$$

3.2.5.1 Data Screening

Hosking and Wallis (1997) introduced a discordancy measure, D_i, to identify grossly discordant sites from the whole group of sites. Discordancy is measured in terms of the L-moments of the sites' data. Hosking and Wallis (1997) stated that incorrect data values, outliers, trends, and shifts in the mean of sample can all be reflected in the L-moments of samples and therefore in discordancy measure. By using this test, sites with gross error will be screened out from the others.

Let $u_i = \left[t^{(i)}, t_3^{(i)}, t_4^{(i)}\right]^T$ be the vector containing the L-moment ratios of sites under study, L-CV, L-sk, and L-ku respectively for site i. Let:

$$\bar{u} = \sum_{i=1}^{N} u_i / N \tag{3-25}$$

$\bar{u}$ is the unweighted regional average. The discordancy measure for site i is then defined as follows:

$$D_i = \frac{1}{3}N(u_i - \bar{u})^T A^{-1}(u_i - \bar{u}) \tag{3-26}$$

Where, A, defines the matrix of sums of squares and cross-products as follow:

$$A = \sum_{i=1}^{N}(u_i - \bar{u})_i\,(u_i - \bar{u})^T \tag{3-27}$$

Hosking and Wallis (1997) stated that site i should be declared as discordant if D_i is large. Based on the above definition of the discordancy measure, this large value depends on the number of sites under study, N. They suggested some critical values based on the number of sites in the group, and suggested that a site be regarded as discordant if its D_i value exceeds the critical value provided in Table 3-1.

Table 3-1 Critical values for discordancy measure, D_i (after Hosking and Wallis, 1997)

Number of Sites	7	8	9	10	11	12	13	14	≥15
Critical Value	1.917	2.140	2.329	2.491	2.632	2.757	2.869	2.971	3

The sites having high D_i values are either removed from the set of data, or moved to a different region. This decision depends upon the physical reasons associated with the apparent discordancy. The above procedure for calculating the discordancy measure can be performed by writing a program in MATLAB (Appendix A-1).

3.2.5.2 Delineation of Homogeneous Regions

Section 2.1.2.2 discussed the possible ways of delineating hydrologic homogeneous regions. One of them was the use of subjective judgment based on at-site characteristics.

Hosking and Wallis (1997) mentioned that region formation based on subjective judgment is suitable for small-scale studies, and it needs to objectively be tested later for heterogeneity.

In the current study, a subjective delineation was adopted for the Island of Newfoundland and also Labrador. Based on subjective judgment, the whole Island of Newfoundland was considered to be one homogeneous region, and Labrador as another separate homogeneous region. In addition, the delineated homogeneous regions on the Island of Newfoundland in the 1991 study were also considered. An objective test using the L-moment based heterogeneity measure discussed in the following section will be used to confirm the delineated regions.

3.2.5.3 Regional Homogeneity Test

Once groups of hydrological homogeneous regions have been identified, it is desirable to assess whether these regions are hydrologically homogeneous and meaningful. It is necessary to test whether the proposed region is accepted as a homogeneous region and whether two or more homogeneous regions are identical so that they can be merged together and form a single region. If a region is acceptably homogeneous, it is assumed that all sites within that region have the same L-moment ratios population, and if there is any difference between these measures it is attributed to sampling variability.

In other words, the hypothesis of the homogeneity test is that the at-site frequency distributions are identical except for a site-specific scale factor (Hosking and Wallis, 1997). This heterogeneity measure has been developed by Hosking and Wallis (1997) and it is based on the study of the sites' L-CVs.

Suppose that the selected region has N sites, with site i having record length of n_i. The weighted standard deviation of the at-site sample L-CVs is given by:

$$V = \left[\sum_{i=1}^{N} n_i \left(t^{(i)} - t^R\right)^2 / \sum_{i=1}^{N} n_i\right]^{1/2} \tag{3-28}$$

It is necessary to calculate the regional average L-CV, L-sk, and L-ku denoted by t^R, t_3^R, and t_4^R as described in Eq. (3-24). Then, a four parameter kappa distribution with the quantile function as described in Eq. (3-29) is fitted to the regional weighted average L-moment ratios, 1, t^R, t_3^R, and t_4^R. Detail of this distribution fitting can be found in Hosking and Wallis (1997). A MATLAB code has been developed in this study (Appendix A-2) to perform this task and estimate the parameters of the kappa distribution based on regional weighted average L-moment ratios. The quantile function of the kappa distribution is given by:

$$\mathrm{x}(F) = \xi + \alpha\{1 - [(1 - F^h)/h]^k\}/k \tag{3-29}$$

Where ξ, α, h, and k are parameters of the distribution. After estimating the parameters of the kappa frequency distribution, a large number of simulation of homogeneous kappa regions, N_{sim}, say for example 10000, are then simulated, each region having N sites with the exact same record length as their real counterparts. A larger number of simulations, N_{sim}, will give more reliable values of μ_V and σ_V. These simulated regions are homogeneous and no correlations exist between them. For each of these simulated regions, the weighted standard deviation, V as described in Eq. (3-28) is then calculated.

After completing all the simulations, the mean, μ_V, and standard deviation, σ_V of the N_{sim} values of V are calculated. Then the heterogeneity measure, H, is given by:

$$H = \frac{(V - \mu_V)}{\sigma_V} \tag{3-30}$$

H is a measure of the departure of V from similar statistics obtained from the simulation of a large number of realizations for a region. Hosking and Wallis (1997) suggested considering the region as "acceptably homogeneous" if $H < 1$, "possibly heterogeneous" if $1 \leq H \leq 2$, and "definitely heterogeneous" if $H > 2$. Robson and Reed (1999) provided a more relaxed criterion by suggesting that region could be considered heterogeneous if $2 \leq H \leq 4$, and strongly heterogeneous if $H > 4$. The MATLAB program code developed in Appendix A-3 will perform this homogeneity test.

3.2.5.4 Selection and Estimation of Regional Distribution

As discussed earlier, the aim of regional frequency analysis is to fit a single frequency distribution to data from several sites within a homogeneous region. The region might be slightly heterogonous in reality, and the chosen distribution may not necessarily fit the data well. Therefore the aim is to find a distribution that will yield accurate quantile estimates for each site.

There are several families of distribution that can be considered for fitting to a regional data set. Their suitability as a regional frequency distribution should be evaluated somehow. Several methods are available in the literature for testing the goodness of fit of a distribution to a set of data. L-moment ratios diagram and Hosking and Wallis goodness of fit test based on L-kurtosis are two of these methods selected for this study. A L-moment ratio diagram which is a plot of L-sk vs. L-ku for the candidate distribution will help to select the best candidate distribution based on the position of the regional weighted average L-moments on this diagram. The goodness of fit test introduced by Hosking and Wallis (1997) is a hierarchy of statistical tests that is more powerful to discriminate among the candidate distributions.

3.2.5.4.1 L-moments Diagram

L-moments have been calculated for many common distributions. A convenient way of representing the L-moments of different distributions is the L-moment ratio diagram whose axes are L-skewness and L-kurtosis. A two-parameter distribution plots as a single point on this diagram, three-parameter distributions as a line, and distribution with more than three-parameters generally cover two-dimensional areas on the graph (Hosking and Wallis, 1997). Figure 3-1 shows the L-moment ratio diagram.

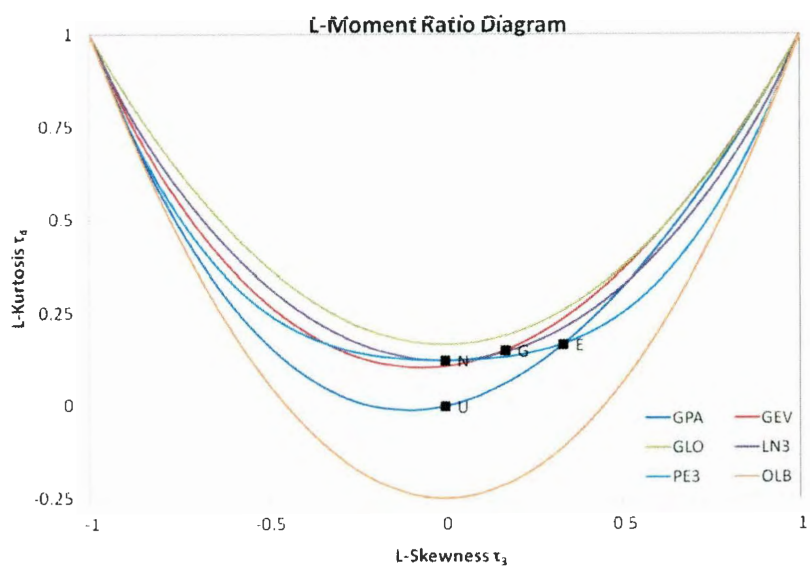

Figure 3-1 L-moment ratio diagram, key to distributions: E-exponential, G-gumbel, N-normal, U-uniform, GPA, generalized pareto, GEV-generalized extreme value, GLO-generalized logistic, LN3-lognormal, PE3- Pearson type III.

It was mentioned by Hosking and Wallis that it is convenient to express L-ku as a function of L-sk. This relationship has been stated by polynomial approximations for many common distributions reported by Hosking and Wallis (1997). The following equation represents the polynomial approximation form. For obtaining the coefficient A_k for commonly used distributions refer to Hosking and Wallis (1997).

$$\tau_4 = \sum_{k=0}^{8} A_k \tau_3^k \qquad (3\text{-}29)$$

Regionally weighted average L-sk and L-ku point is plotted on L-moment ratio diagram. The position of this point indicates the most appropriate candidate regional distribution.

3.2.5.4.2 Hosking and Wallis goodness of fit test

Hosking and Wallis (1997) stated that the goodness of fit test will judge how well the L-kurtosis of the candidate distribution match the regional weighted average L-kurtosis of the observed data which was corrected for sampling bias.

Suitable candidate three-parameter distributions are the generalized logistic (GLO), generalized extreme-value (GEV), generalized Pareto, lognormal, and Pearson type III. Each candidate distribution will be fitted to the regional average L-moments, and the L-kurtosis of fitted distribution will be calculated and denoted as τ_4^{DIST}. Then the same procedure as the heterogeneity test can be followed to fit a kappa distribution to the regional average L-moments, and to simulate of a large number of kappa regions. For each simulated region, the regional average L-ku, τ_4^m, is calculated. The goodness of fit measure for each candidate distribution is then given by following equation:

$$Z^{DIST} = (\tau_4^{DIST} - t_4^R + B_4)/\sigma_4 \quad (3\text{-}30)$$

Where B_4 is the bias of t_4^R, and σ_4 is the standard deviation of t_4^R defined as follow:

$$B_4 = \sum_{m=1}^{N_{sim}} \left(t_4^{R(m)} - t_4^R\right)/N_{sim} \quad (3\text{-}31)$$

$$\sigma_4 = \left[(N_{sim} - 1)^{-1} \left\{\sum_{m=1}^{N_{sim}} \left(t_4^{R(m)} - t_4^R\right)^2 - N_{sim}\, {B_4}^2\right\}\right]^{1/2} \quad (3\text{-}32)$$

The candidate distribution is declared as adequate fit if Z^{DIST} is sufficiently close to zero. Hosking and Wallis (1997) suggested a reasonable criterion being $|Z^{DIST}| \leq 1.64$.

3.2.5.5 Flow Quantile Estimation

Once the delineated region has been shown to be homogeneous, and a suitable distribution has been identified, the index flow procedure can be applied to estimate flows. The index flow procedure is a convenient way of pooling summary statistics from different data samples. The key assumption in index flow procedure is that the frequency distributions of all sites in a homogeneous region are identical, except for a site-specific scale factor, the index variable.

Suppose that data are available at N sites, with site i having sample size of n_i and observed data Q_{ij}, $j = 1,2,..,n$. Let $Q_i(F), 0 < F < 1$, be the frequency distribution quantile function at site i. Then the quantile of non-exceedance probability of $F, Q_i(F)$ of the site i can be written as:

$$Q_i(F) = \mu_i q(F) \quad (3\text{-}35)$$

Where, μ_i is the index flow variable, and $q(F)$ is regional growth factor, a dimensionless quantile function common and constant to every site. The index variable can be estimated by $\mu_i = \overline{Q}_i$, the sample mean of the annual low flow data at site i.

3.2.5.5.1 Estimation of Index flow

Based on the index flow procedure, for estimating a T-year return period flow quantile at ungauged sites, an estimate of the index flow or sample mean of annual low flow data is required. Since observed flow data are not available at ungauged sites, the at-site mean cannot be computed. In such a situation, it is necessary to establish a relationship between the mean annual low flow of the gauged catchments within the homogeneous region, and their pertinent physiographic and climatic characteristics to obtain an estimate of the mean annual low flow. For this purpose a non-linear regression based on the least squares method between the site characteristics and the index flow of the corresponding sites in the region is carried out. The regression model usually has the following form.

$$\overline{Q} = \alpha_0 A_1^{\alpha 1} A_2^{\alpha 2} \ldots A_n^{\alpha n} \qquad (3\text{-}36)$$

Where A_1, A_2,...,A_n are the site characteristics, α_0, α_1, ...,α_n are the model parameters, ε_0 is the additive error term and n is the number of site characteristics.In this study it was assumed that the climatic characteristics are identical throughout the region, and among physiographic specifics, catchment size was used to establish a relationship with the magnitude of discharge. Minitab statistical software package was used to perform the regression analysis based on least squares method.

3.2.5.5.2 Estimation of Regional Growth Curve

The parameters of the regional growth curve which is identical to the selected regional distribution can be estimated by pooling the information available from all the sites within the homogeneous region. Hosking and Wallis (1997) suggested the following procedure to estimate the parameters of regional growth curve.

- The first four unbiased L-moments and their ratios should be computed separately for each site within the homogeneous region.
- The average L-moment ratio weighted proportionally to the record length of each site should be obtained.
- The parameters of selected regional distribution for homogeneous region should be estimated using the regional average L-moment ratios. These estimations can be performed by using the provided relationships between the L-moments and the parameters of some distributions by Hosking and Wallis (1997). It should be noted that regional weighted average L-moments, 1, t^R, t_3^R, and t_4^R should be inserted as the L-moments of selected regional distribution in order to obtain the parameters of the selected distribution.
- Plot the quantile function $q(F)$ of the regional frequency distribution estimated in step (iii) versus the return period. The resulting curve is the regional low flow growth curve for the region. Figure 3-2 presents a typical regional growth curve.

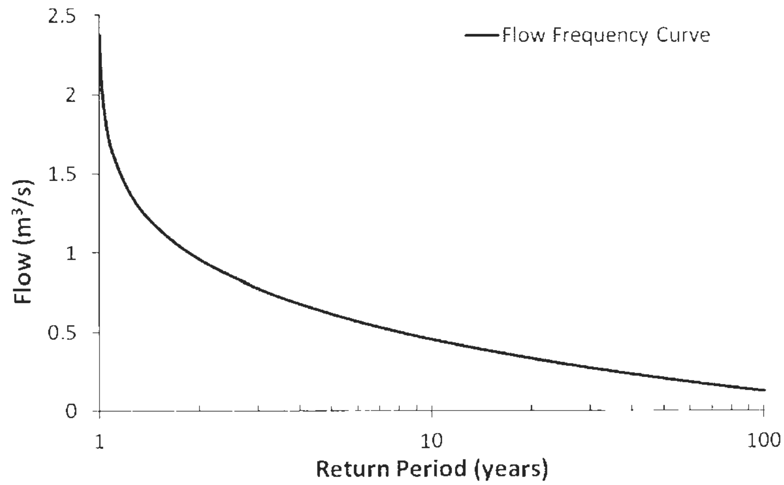

Figure 3-2 Typical regional growth curve

3.3 Regional Flow Duration Analysis

3.3.1 Constructing Flow Duration Curves

In general, a FDC is constructed by reassembling the flow time series values in decreasing order of magnitude, assigning flow values to class intervals and counting the number of occurrences (time steps) within each class interval. Cumulative class frequencies are then calculated and expressed as a percentage of the total number of time steps in the record period. Finally, all ranked flows are plotted against their rank which is again expressed as a percentage of the total number of time steps in the record (Smakhtin, 2001). As it was discussed earlier, according to the period of record used for constructing FDCs, they can be divided to two major groups: (1) on the basis of the whole available record period; (2) on the basis of a portion of calendar (month, seasons or year). Two types of FDCs were studied for rivers in Newfoundland and Labrador; period of record

FDCs and Annual FDCs. Details of constructing these FDCs are described in the next sections.

3.3.1.1 Period of Record FDC

A flow duration curve (FDC) provides the percentage of time (duration) a daily or monthly (or some other time interval) streamflow is exceeded over a historical period for a particular river basin. FDC may also be viewed as the complement of the cumulative distribution function of the considered streamflow (Vogel and Fennessy, 1994).

The non-parametric procedure described for example by Vogel and Fennessy (1994) can be used to construct FDCs based on streamflow observation consisting of two main steps: (a) observed streamflows Q_i, $i = 1,2, \dots, N$, are ranked to produce a set of ordered streamflows $Q_{(i)}$, $i = 1,2, \dots, N$, where N is the same length, and $Q_{(1)}$ and $Q_{(N)}$ are the largest and the smallest observations respectively; (b) each ordered flow observation, $Q_{(i)}$is then plotted against its corresponding duration D_i which is generally dimensionless and coincides with P_i an estimate of the exceedance probability of the flow observation, Q_i. If the Weibull plotting position is used, the exceedance probability is:

$$P_i = P(Q > Q_i) = \frac{i}{N+1} \qquad (3\text{-}37)$$

3.3.1.2 Annual FDC

A different approach proposed by Vogel and Fennessy (1994) is an annual interpretation of flow duration curves. This interpretation considers n FDCs for n individual years of records (AFDCs), each one constructed analogously to the FDCs, described in the previous section, using only hydrometric information collected in a calendar or water year. Then one can treat those n annual FDCs in much the same way one treats a sequence of annual maximum or annual minimum streamflows. For each exceedance probability P, the median value of discharge is computed. The AFDC is actually a plot of these median values against their exceedance probability. This median AFDC represents the distribution of streamflow in a 'typical' or median hypothetical year and is not affected by the observations of abnormally wet or dry periods during the period of record. And this is the significant difference between period of record FDC and median AFDC. The period of record FDC is highly sensitive to the particular period of record whereas the median AFDC is not (Vogel and Fennessy, 1994).

3.3.2 Flow Duration Curves Regional Regression

Scarcity of streamflow data is a common problem in many watersheds discussed in many studies. Therefore the regionalization of FDCs appears to be an essential operative tool when dealing with ungauged river basins or those with short streamflow record. Hence the development of regional FDCs for estimation of FDCs at ungauged river basins or the enhancement of empirical FDCs constructed for gauged streams where only limited amount of hydrometric information is available is necessary.

As discussed in Section 2.2.4, there are two main approaches of regional FDC analysis; one based on a graphical method which is predicting the whole FDC by fitting a distribution to it, and the other, predicting some flow quantiles of FDC using physiographic parameters of the region by means of regression. Studies showed better performance of the second method in predicting FDC. For this reason, the latter was adopted as the approach for regionalization of FDCs in Newfoundland and Labrador for the current study.

3.3.2.1 Regression Model

The common form of relationship between flow magnitudes (quantiles of flow duration curve) and physical parameters of a gauging river has the following form:

$$Q = a \times (Var1)^b \times (Var2)^c \times \ldots \qquad (3\text{-}38)$$

Where $Var1$, $Var2$,... representing the basin characteristics and physical parameters of it. It is necessary to find out which physical parameters are important, and what are those equation constants $(a, b, c\ldots)$. These can be achieved by conducting a regression between physical characteristics of basins as independent variables or predictors and flow quantiles of FDC. However, taking natural log of both sides of the above equation will yield the following equation for which a simple multiple linear regression can be easily performed.

$$Ln(Q) = Ln(a) + b \times \ln(var1) + c \times \ln(Var2) + \cdots \qquad (3\text{-}39)$$

The regression equation can be obtained by using statistical software packages (e.g. Minitab). By having these flow quantile prediction equations in hand, it is only necessary to obtain an estimate of physiographic parameters of a site with no hydrologic data, to estimate the percentiles of FDC. For regression equations developed in natural log space, bias correction factors were estimated by Smearing Estimator (Duan, 1983) and applied to the final regression equations.

3.3.2.2 FDC Quantiles

Some sets of flow quantiles of FDC associated with selected exceedance probabilities were chosen for this study:

- High flows: $Q_{0.01}$, $Q_{0.05}$, $Q_{0.1}$, $Q_{0.15}$, $Q_{0.2}$
- Median flows: $Q_{0.25}$, $Q_{0.3}$, $Q_{0.4}$, $Q_{0.5}$, $Q_{0.6}$
- Low flows: $Q_{0.7}$, $Q_{0.8}$, $Q_{0.9}$, $Q_{0.95}$, $Q_{0.99}$

These selected percentiles of FDC represent all the flow ranges of FDC from the high flows end to the low flow ends. The values of these percentiles were obtained for all the gauged rivers in Newfoundland and Labrador for their both period of record FDC and AFDC. Then they have been regressed against physiographic parameters of respective watersheds.

3.3.2.3 Physiographic Parameters

The possible significant site characteristics for river basins in Newfoundland and Labrador include: drainage area; fraction of lake area; fraction of forest area; fraction of swamp area; fraction of barren area; fraction of lake and swamp area; fraction of area controlled by lakes and swamps, lake and swamp factor, length of main channel, elevation difference of main channel, slope of main channel, drainage density, and shape factor. Definition of all these physiographic parameters and how to extract them is available in 1989 regional flood frequency report of Gov. of Newfoundland and Labrador. These parameters were extracted for some of the hydrometric gauges of Newfoundland in that report and were adopted in the current study. The physical parameters for remaining gauges that were not available in that report were extracted using ArcGIS software.

3.4 Regional Flow Spell Analysis

3.4.1 Defining Flow Spell

Flow spell analysis considers how long a low flow (below some threshold) has been maintained and how large a deficit has been built up, and takes into account the sequencing of flows (McMahon et al., 2004). A graphical description of the method is shown in Figure 3-3. Two main measures are obtained directly from a flow spell graph: spell duration, and deficiency volume. And the third measure, intensity, is derived by dividing these two measures (as described in section 2-3).

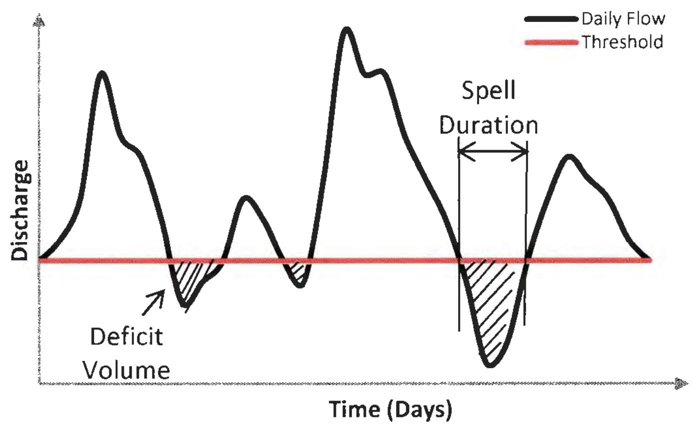

Figure 3-3 General diagram of defining flow spells

Daily flows are serially correlated and therefore it is expected that flow spells would follow one another during the dry months of year which indicates the dependence of the present spell on the previous one. In order to estimate the recurrence interval of these events, probability distributions must be used and they are subject to the condition of independent data value. This can be achieved by considering only the maximum annual duration, volume and intensity of flow spells. Annual frequency refers to the proportion of years in which deficit volume or spell duration is exceeded. To analyze the spell data using this method, the longest spell duration, largest deficit volume below a given threshold, and largest intensity is found for each year. These are known as annual spell maxima.

It should be noted that the annual maximum spell duration is not necessarily the same event as that of the annual maximum volume, and the start dates might not be the same. If this happens then the annual maximum intensity must be calculated by taking the annual maximum volume and dividing it by its duration of spell not the annual maximum duration which would yield a higher intensity.

3.4.2 Environmental Instream Requirement as Threshold

3.4.2.1 Percentiles of FDC and AFDC

As discussed earlier in Chapter 2, a certain percentile of flow duration curve can be used as environmental instream flow requirements, for example, Q_{85} or Q_{95}. However, the percentiles derived based on period of record flow duration curve are more sensitive to extreme low flows than other environmental instream flow requirements methods, even though a period of record more than minimum recommended ten years may have been taken (Shuaghnessy, 1997). To overcome this issue, Q_{85} and Q_{95} percentiles of annual flow duration curves are also adopted in this study as instream flow requirements. These percentiles can be used as a constant threshold value for environmental instream flow requirement throughout a year.

3.4.2.2 Percent of Mean Annual Flow

The mean annual flow (MAF) is based on complete years of record data. It is calculated by first finding the mean flow of each year of data, and then the mean flow of these means, by summing and dividing these means by the number of complete years. This type of instream requirement is less sensitive to extreme low flows that the traditional thresholds based on period of record FDC method. Two different percentage of MAF was selected in this study as threshold values to calculate the flow spells, Tennant's method and 25% MAF.

3.4.2.2.1 Tennant's Method

Tennant's method (1976) is easy to estimate and implement, it takes into account seasonal variability of flow, and it reduces the weight given to extreme streamflows as compared to POR FDCs. Because of these advantages, Tennant's method is now widely used in some parts of US (Caissie and El-Jabi, 1995).

Tennant performed a study on the change in percentiles of widths, depths and velocities to the reduction in MAF over a ten year period for 58 rivers in Montana, Wyoming and Nebraska regions of US. He concluded that aquatic habitat conditions were similar on streams carrying similar MAFs. Afterward, some studies carried out in 21 other states of US and confirmed this theory (Karim, 1995). Tennant then defined recommended flows during summer and winter months according to different river conditions that are necessary to be maintained or enhanced. Table 3-2 provides the information. The excellent river condition is used as environmental instream flow requirement in most of the studies as well as the current study.

Table 3-2 Tennant's Method (adopted from McMahon et al., 2004)

River Condition	Recommended Minimum Flow (%MAF)	
	Oct to Mar	Apr to Sept
Flushing or maximum	200%	200%
Optimum range	60 to 100%	60 to 100%
Outstanding	40%	60%
Excellent	30%	50%
Good	20%	40%
Fair or degraded	10%	30%
Poor or minimum	10%	10%
Severe degradation	<10%	<10%

3.4.2.2.2 25% MAF

This method is also called the modified Tennant's method. Similar to the percentile of the period of record FDC, the threshold value is held constant throughout the year for this method, regardless of season. This threshold is widely used throughout Atlantic Canada since a fixed percentage of MAF is best suited to water abstraction systems whose intake structures corresponds to a specific stream water elevation (Caissie and El-Jabi, 1995).

3.4.2.3 7Q10

This method was adopted by Chiang and Johnson (1976), and like the MAF methods recognizes that daily flows are serially correlated and so a yearly value ensures more independent data. This method is different from all other methods because it uses this independent data to obtain plotting position and estimate return period of events using probability plotting position (IH, 1980).

A minimum of 20 years of data is recommended for this method (IH, 1980). Here, the water year is defined as the year from 1^{st} of January to the 31^{st} of December. First, it is necessary to find the lowest 7day moving average flows for each year of data, 7Q. These values then are fitted to a distribution function to estimate the low flows having 10-year return period.

The $7Q10$ method is limited where the gauging instrument is faulty and there are missing data, since the yearly 7 day low flow is calculated based on only complete years of data. Thus it suffers the same problem as the MAF method in addition to the complexity of correlating few data points to a probability distribution function. This criterion was found to underestimate the minimum flow throughout the year for aquatic biota (Bovee, 1982) which explains why this method is primarily used for maintaining water quality parameters in rivrers not sufficient for aquatic life.

For each of the thresholds discussed above, a Microsoft Excel macro was developed in this study to calculate all the flow spells (including their start and end date, duration, volume, and intensity) for each year of streamflow data based on these threshold values, and then to find the maximum flow spells variables (duration, volume, and intensity) for each year.

3.4.3 Predicting Flow Spells

Based on the above discussion, to ensure the independence of spell periods it is only necessary to fit the probability distribution to annual maximum flow spell variables. Fitting a probability distribution to annual maximum spell variables allows an estimation of the x-year spell event in terms of its duration, volume and intensity. First of all it is assumed that a probability distribution exists that will fit the data, where the data is defined as the annual maximum flow spell variables for a particular threshold method.

This can be achieved by using Minitab statistical software to fit different probability distributions to data, and choose which fits the data better.

If the data fits then a comparison between the methods of threshold estimation and spell variables can be undertaken. A relationship between the defining parameters of the fitted distribution and physiographic parameters of the watershed is then sought. In addition, direct relationships between catchment drainage area and threshold values, and then threshold values and annual maximum spells can also be obtained to predict these spell variables in any ungauged sites within the study region. Further discussion will be provided in Chapter 6.

3.5 Study Area and Data

The study area is the province of Newfoundland and Labrador, the most easterly province of Canada. The streamflow data are available through HYDAT, Environment Canada for rivers in the province up to 2010. The criterion for selecting rivers for study was at least 20 years of complete streamflow record on unregulated rivers. This lead to the selection of 60 gauged stations in the Island of Newfoundland. However, only 8 gauges in Labrador region met this criterion. Therefore, rivers with at least 15 years of data in record were selected for further studies in Labrador which gave a total of 12 rivers. Table 3-3 and 3-4 lists all the gauging stations within Newfoundland and Labrador respectively used in this study along with their information and sample sizes provided by HYDAT. It should be noted that only complete years of data were selected in this study with no attempt to extend the data record. The remainder of this study refers to these stations using their ID numbers. Figures 3-4 and 3-5 illustrate the location of these gauges on the map.

Table 3-3 Selected Hydrometric Stations in Newfoundland (HYDAT database)

ID	Station Num.	Station Name	Start Year	End Year	Drainage Area (km^2)
1	02YA001	STE. GENEVIEVE RIVER NEAR FORRESTERS POINT	1969	1996	306
2	02YA002	BARTLETTS RIVER NEAR ST. ANTHONY	1986	2010	33.6
3	02YC001	TORRENT RIVER AT BRISTOL'S POOL	1960	2010	624
4	02YD002	NORTHEAST BROOK NEAR RODDICKTON	1980	2010	200
5	02YE001	GREAVETT BROOK ABOVE PORTLAND CREEK	1984	2010	95.7
6	O2YG001	MAIN RIVER AT PARADISE POOL	1986	2010	627
7	O2YJ001	HARRYS RIVER BELOW HIGHWAY BRIDGE	1968	2010	640
8	O2YK002	LEWASEECHJEECH BROOK AT LITTLE GRAND	1952	2010	470
9	02YK004	HINDS BROOK NEAR GRAND LAKE	1956	1979	529
10	02YK005	SHEFFIELD BROOK NEAR TRANS CANADA	1973	2010	391
11	02YK008	BOOT BROOK AT TRANS-CANADA HIGHWAY	1986	2010	20.4
12	02YL001	UPPER HUMBER RIVER NEAR REIDVILLE	1928	2010	2110
13	02YL004	SOUTH BROOK AT PASADENA	1983	2010	58.5
14	02YL005	RATTLER BROOK NEAR MCIVERS	1985	2010	17
15	02YL008	UPPER HUMBER RIVER ABOVE BLACK BROOK	1988	2010	471
16	02YM001	INDIAN BROOK AT INDIAN FALLS	1956	1979	974
17	02YM003	SOUTH WEST BROOK NEAR BAIE VERTE	1980	2010	93.2
18	02YM004	INDIAN BROOK DIVERSION ABOVE BIRCHY LAKE	1990	2010	238
19	02YN002	LLOYDS RIVER BELOW KING GEORGE IV LAKE	1981	2010	469
20	02YO006	PETERS RIVER NEAR BOTWOOD	1981	2010	177
21	02YO008	GREAT RATTLING BROOK ABOVE TOTE RIVER	1984	2010	773
22	02YO012	SOUTHWEST BROOK AT LEWISPORTE	1989	2010	58.7
23	02YQ001	GANDER RIVER AT BIG CHUTE	1950	2010	4450
24	02YQ005	SALMON RIVER NEAR GLENWOOD	1987	2010	80.8
25	02YR001	MIDDLE BROOK NEAR GAMBO	1959	2010	275
26	02YR002	RAGGED HARBOUR RIVER NEAR MUSGRAVE	1978	1997	399
27	02YR003	INDIAN BAY BROOK NEAR NORTHWEST ARM	1981	2010	554
28	02YS001	TERRA NOVA RIVER AT EIGHT MILE BRIDGES	1951	1984	1290
29	02YS003	SOUTHWEST BROOK AT TERRA NOVA PARK	1968	2009	36.7
30	02YS005	TERRA NOVA RIVER AT GLOVERTOWN	1985	2010	2000

Table 3-3 Continue Selected Hydrometric Stations in Newfoundland (HYDAT database)

ID	Station Num.	Station Name	Start Year	End Year	Drainage Area (km^2)
31	02ZA002	HIGHLANDS RIVER AT TRANS-CANADA	1982	2010	72
32	02ZB001	ISLE AUX MORTS RIVER BELOW HIGHWAY	1963	2010	205
33	02ZC002	GRANDY BROOK BELOW TOP POND BROOK	1982	2010	230
34	02ZD002	GREY RIVER NEAR GREY RIVER	1969	2010	1340
35	02ZE001	SALMON RIVER AT LONG POND	1944	1965	2640
36	02ZE004	CONNE RIVER AT OUTLET OF CONNE POND	1990	2010	99.5
37	02ZF001	BAY DU NORD RIVER AT BIG FALLS	1950	2010	1170
38	02ZG001	GARNISH RIVER NEAR GARNISH	1959	2009	205
39	02ZG002	TIDES BROOK BELOW FRESHWATER POND	1978	1996	166
40	02ZG003	SALMONIER RIVER NEAR LAMALINE	1980	2009	115
41	02ZG004	RATTLE BROOK NEAR BOAT HARBOUR	1981	2009	42.7
42	02ZH001	PIPERS HOLE RIVER AT MOTHERS BROOK	1953	2009	764
43	02ZH002	COME BY CHANCE RIVER NEAR GOOBIES	1961	2009	43.3
44	02ZJ001	SOUTHERN BAY RIVER NEAR SOUTHERN BAY	1977	2009	67.4
45	02ZJ002	SALMON COVE RIVER NEAR CHAMPNEYS	1983	2009	73.6
46	02ZJ003	SHOAL HARBOUR RIVER NEAR CLARENVILLE	1986	2009	106
47	02ZK001	ROCKY RIVER NEAR COLINET	1948	2009	301
48	02ZK002	NORTHEAST RIVER NEAR PLACENTIA	1979	2009	89.6
49	02ZK003	LITTLE BARACHOIS RIVER NEAR PLACENTIA	1983	2009	37.2
50	02ZK004	LITTLE SALMONIER RIVER NEAR NORTH HARB	1983	2009	104
51	02ZL004	SHEARSTOWN BROOK AT SHEARSTOWN	1983	2009	28.9
52	02ZL005	BIG BROOK AT LEAD COVE	1985	2009	11.2
53	02ZM006	NORTHEAST POND RIVER AT NORTHEAST POND	1954	2009	3.63
54	02ZM008	WATERFORD RIVER AT KILBRIDE	1974	2009	52.7
55	02ZM009	SEAL COVE BROOK NEAR CAPPAHAYDEN	1980	2009	53.6
56	02ZM016	SOUTH RIVER NEAR HOLYROOD	1983	2009	17.3
57	02ZM018	VIRGINIA RIVER AT PLEASANTVILLE	1984	2009	10.7
58	02ZM020	LEARY BROOK AT PRINCE PHILIP DRIVE	1986	2009	17.8
59	02ZN001	NORTHWEST BROOK AT NORTHWEST POND	1966	1996	53.3
60	02ZN002	ST. SHOTTS RIVER NEAR TREPASSEY	1985	2009	15.5

Table 3-4 Selected Hydrometric Stations in Labrador (HYDAT database)

ID	Station Num.	Station Name	Start Year	End Year	Drainage Area (km^2)
1	02XA003	LITTLE MECATINA RIVER ABOVE LAC FOURMONT	1978	2010	4540
2	03NF001	UGJOKTOK RIVER BELOW HARP LAKE	1979	2010	7570
3	03OC003	ATIKONAK RIVER ABOVE PANCHIA LAKE	1972	2010	15100
4	03OE003	MINIPI RIVER BELOW MINIPI LAKE	1979	2010	2330
5	03PB002	NASKAUPI RIVER BELOW NASKAUPI LAKE	1978	2010	4480
6	03QC001	EAGLE RIVER ABOVE FALLS	1966	2010	10900
7	03QC002	ALEXIS RIVER NEAR PORT HOPE SIMPSON	1978	2010	2310
8	02XA004	RIVIERE JOIR NEAR PROVINCIAL BOUNDARY	1980	1996	2060
9	03NG001	KANAIRIKTOK RIVER BELOW SNEGAMOOK LAKE	1979	1996	8930
10	03OB002	CHURCHILL RIVER AT FLOUR LAKE	1955	1971	33900
11	03OE010	BIG POND BROOK BELOW BIG POND	1994	2010	71.4
12*	03OE001	CHURCHILL RIVER ABOVE UPPER MUSKRAT FALLS	1948	2010	92500

* Only unregulated period of data was used in this study (1954-1970)

Figure 3-4 Location of hydrometric stations in Newfoundland

Figure 3-5 Location of hydrometric stations in Labrador

4 Low Flow Frequency Analysis and Results

In this chapter the analysis and results of the regional low flow frequency analysis based on L-moments approach will be presented for the rivers of the province of Newfoundland and Labrador. The methodology described in section 3.2 was applied step by step. This analysis used the annual minimum 1-day (1-day AM) and 7-day (7-day AM) flows of selected rivers in Newfoundland and Labrador (refer to section 3.5 for selection criteria). These annual minimum values were extracted from the Environment Canada's HYDAT database available online. Then L-statistics were calculated for each of selected rivers which are the basis for the rest of analyses. The regional approach was then validated using other sets of data.

4.1 Data Screening: Discordancy measure

The discordancy measures (D_i's) were computed for the sites in the study region to find out whether any sites were grossly discordant from the other sites. If the D_i statistic for a site is more than the determined critical value, the data at such site have to be examined for possible problems. For the present study, the whole Island of Newfoundland and Labrador are assumed as two separate regions, and L-statistics of rivers in these regions were examined for overall gross errors for both 1-day and 7-day minimum annual flow data sets. The computation was carried out using a MATLAB program, *Discordany.m* (Appendix A-1). A Microsoft Excel worksheet captures the data file from the MATLAB program with a $N \times 3$ matrix of L-moment ratios, τ, τ_3, and τ_4 for each of the site within the group, where N is the number of stations in the respective group. The names of gauging stations, their record lengths, mean of annual minimum flows, L-moment ratios of data, and computed D_i values at each station for Labrador and Newfoundland are presented in Table 4.1 and 4.2 respectively for 1-day AM and 7-day AM. The computed L-moment ratios for group of sites in Labrador and Newfoundland are plotted in Figures 4.1 and 4.2 respectively for 1-day AM, and in Figures 4.3 and 4.4 for 7-day AM.

D_i values range from 0.08 to 2.8 and 0.35 to 2.14 in 1-day AM dataset and from 0.04 to 2.97 and 0.39 to 1.78 in 7-day AM dataset for Newfoundland and Labrador respectively. The high D_i values always warrant a careful scrutiny of the data at the respective stations. However, one can observe that the critical value of discordancy measure test for 60 sites in Newfoundland and 12 sites in Labrador were not exceeded at any of the sites within their groups. These D_i values in Newfoundland region are actually quite far from the critical values given the relatively large number of sites (60). Therefore, data within these two groups are not discordant and they are suitable for applying the regional low flow frequency using their L-moments. In addition, Figures 4.1(a to d) illustrate that no combination of L-moment ratios seems to be discordant with the pattern of other sites in the group.

4.2 Identifying Homogeneous Regions

After finding no discordant site in the group of sites in the two areas under study, it is rational that in the first attempt is to check whether they belong to one homogeneous region. Then if this was not the case, division of region into small groups should be considered. The Hosking and Wallis Homogeneity test outlined in section 3.2.5.3 was applied for these two regions.

Table 4-1 Statistics summary of gauging stations in Labrador

			1-day minimum annual flow					7-day minimum annual flow				
ID	Station Num.	Sample Size (years)	Mean flow (m^3/s)	L-CV	L-sk	L-ku	D_i	Mean flow (m^3/s)	L-CV	L-sk	L-ku	D_i
1	02XA003	30	15.58	0.1182	0.1362	0.1829	0.81	15.80	0.1188	0.1414	0.1642	0.56
2	03NF001	31	14.15	0.1927	0.1450	0.1530	1.08	14.39	0.1932	0.1505	0.1606	0.92
3	03OC003	16	58.31	0.1163	0.2150	0.2533	1.17	60.82	0.1164	0.1973	0.2291	0.88
4	03OE003	28	11.97	0.1623	0.0227	0.1430	0.88	12.10	0.1626	0.0272	0.1419	1.74
5	03PB002	29	18.95	0.1644	0.0801	0.0421	0.62	19.23	0.1652	0.0838	0.0439	0.47
6	03QC001	40	30.05	0.2201	0.1051	0.2166	1.60	30.48	0.2210	0.1065	0.2097	1.33
7	03QC002	33	5.63	0.1698	0.0632	0.0375	0.64	5.78	0.1753	0.0771	0.0517	0.39
8	02XA004	15	3.79	0.1208	0.0927	0.0136	1.39	3.83	0.1201	0.0848	0.0062	1.53
9	03NG001	17	23.31	0.1208	0.2265	0.2106	0.81	23.49	0.1200	0.2316	0.2090	1.09
10	03OB002	15	189.60	0.1507	0.1486	0.0921	0.51	190.51	0.1521	0.1449	0.0800	0.66
11	03OE010	17	0.146	0.2254	-0.0832	0.1155	2.14	0.167	0.2498	-0.0278	0.0475	1.78
12	02OE001	17	444	0.1610	0.1127	0.2091	0.35	448	0.1619	0.1028	0.1983	0.64

Table 4-2 Statistics summary of gauging stations in Newfoundland

			1-day minimum annual flow					7-day minimum annual flow				
ID	Station Num.	Sample Size (years)	Mean flow (m^3/s)	L-CV	L-sk	L-ku	D_i	Mean flow (m^3/s)	L-CV	L-sk	L-ku	D_i
1	02YA001	27	2.628	0.2049	0.1733	0.2068	0.96	2.784	0.2042	0.1322	0.1884	0.54
2	02YA002	25	0.065	0.3317	0.2012	0.2321	0.79	0.072	0.3254	0.1842	0.2158	0.67
3	02YC001	51	3.602	0.2045	0.0649	0.0930	0.31	3.787	0.1931	0.0821	0.0854	0.48
4	02YD002	31	0.378	0.3172	0.2689	0.2419	1.12	0.420	0.3129	0.2465	0.2192	0.87
5	02YE001	27	0.382	0.3562	0.3067	0.2291	1.41	0.445	0.3654	0.3391	0.2745	2.28
6	O2YG001	25	2.718	0.1771	0.1261	0.0968	1.22	3.019	0.1790	0.1428	0.0981	1.16
7	O2YJ001	42	4.571	0.1829	0.1061	0.0873	0.97	5.121	0.1801	0.1283	0.0562	1.46
8	O2YK002	48	2.832	0.1714	0.0579	0.2255	1.24	3.052	0.1745	0.0502	0.2002	0.90
9	02YK004	22	2.990	0.2786	-0.0853	0.0949	1.87	3.239	0.2651	-0.0853	0.0823	1.47
10	02YK005	38	1.879	0.2103	-0.0101	0.0378	0.64	2.040	0.2154	0.0268	0.0593	0.34
11	02YK008	25	0.016	0.3763	0.0758	0.0344	1.86	0.022	0.3652	0.0436	0.0166	2.04
12	02YL001	72	7.990	0.2304	0.0828	0.1759	0.21	8.977	0.2241	0.1154	0.1826	0.28
13	02YL004	28	0.192	0.1595	0.0574	0.1267	0.68	0.215	0.1562	0.1156	0.1036	1.22
14	02YL005	26	0.013	0.3644	0.2628	0.1061	1.55	0.019	0.3512	0.2262	0.0980	1.26
15	02YL008	23	2.409	0.2137	0.0793	0.0978	0.27	2.641	0.2255	0.0984	0.0772	0.36
16	02YM001	41	2.885	0.2197	0.0372	0.0592	0.40	3.136	0.2173	0.0441	0.0655	0.30
17	02YM003	31	0.090	0.4489	0.2701	0.1333	2.30	0.113	0.4207	0.2345	0.1183	2.23
18	02YM004	21	0.909	0.2753	-0.0118	-0.0027	1.35	1.017	0.2720	0.0162	-0.0009	1.15
19	02YN002	30	2.674	0.1516	0.0557	0.2639	2.22	2.929	0.1507	0.0622	0.2170	1.41
20	02YO006	30	0.402	0.2069	-0.0060	0.0943	0.30	0.459	0.2033	0.0311	0.1267	0.19

Table 4.2 Continue Statistics summary of gauging stations in Newfoundland

			1-day minimum annual flow					7-day minimum annual flow				
ID	Station Num.	Sample Size (years)	Mean flow (m^3/s)	L-CV	L-sk	L-ku	D_i	Mean flow (m^3/s)	L-CV	L-sk	L-ku	D_i
21	02YO008	27	1.665	0.2658	0.0254	0.1590	0.51	1.993	0.2631	0.0291	0.2036	0.95
22	02YO012	22	0.159	0.3125	0.0343	0.1080	0.66	0.184	0.3032	0.0190	0.0868	0.68
23	02YQ001	61	21.277	0.2441	0.0195	0.1329	0.25	22.481	0.2417	0.0241	0.1387	0.22
24	02YQ005	24	0.106	0.3744	0.1365	0.1871	1.17	0.131	0.3327	0.1254	0.2469	1.42
25	02YR001	51	1.057	0.2995	0.0398	0.0528	0.67	1.143	0.2991	0.0322	0.0398	0.82
26	02YR002	20	0.811	0.4347	0.1546	0.0621	2.54	0.917	0.4361	0.1598	0.0656	2.66
27	02YR003	30	2.295	0.2712	-0.0844	0.0038	1.84	2.485	0.2698	-0.0988	-0.0233	2.22
28	02YS001	29	6.700	0.2589	0.0360	0.1330	0.19	7.291	0.2578	0.0213	0.1335	0.29
29	02YS003	42	0.063	0.3560	0.2439	0.1558	0.84	0.079	0.3202	0.1927	0.1514	0.44
30	02YS005	26	9.608	0.2240	0.0288	0.0438	0.54	10.426	0.2188	-0.0026	0.0296	0.62
31	02ZA002	29	0.249	0.1453	-0.1016	0.1810	2.26	0.286	0.1272	-0.1218	0.2234	2.97
32	02ZB001	48	0.792	0.2280	0.2118	0.2092	1.04	0.955	0.2354	0.2637	0.2484	1.73
33	02ZC002	29	0.877	0.2482	0.2252	0.1800	0.93	1.079	0.2345	0.1838	0.1543	0.55
34	02ZD002	27	4.905	0.2136	0.0081	0.1484	0.37	5.858	0.2076	0.0009	0.1572	0.51
35	02ZE001	21	19.055	0.3421	0.1589	0.1353	0.42	20.335	0.3334	0.1518	0.1501	0.44
36	02ZE004	21	0.182	0.3790	0.2777	0.2122	1.22	0.219	0.3625	0.2784	0.1844	1.29
37	02ZF001	58	8.743	0.2070	0.0184	0.1789	0.60	9.416	0.2107	0.0381	0.1844	0.52
38	02ZG001	51	1.145	0.2431	0.0136	0.2318	1.67	1.285	0.2454	0.0216	0.2350	1.49
39	02ZG002	20	1.087	0.2502	-0.1093	0.0031	1.91	1.216	0.2411	-0.1463	0.0076	2.26
40	02ZG003	30	0.261	0.3133	0.2087	0.2501	0.93	0.345	0.2744	0.1308	0.2085	0.39

Table 4.2 Continue Statistics summary of gauging stations in Newfoundland

			1-day minimum annual flow					7-day minimum annual flow				
ID	Station Num.	Sample Size (years)	Mean flow (m^3/s)	L-CV	L-sk	L-ku	D_i	Mean flow (m^3/s)	L-CV	L-sk	L-ku	D_i
41	02ZG004	29	0.134	0.2215	-0.0738	0.1517	1.50	0.173	0.2229	-0.0718	0.1506	1.36
42	02ZH001	57	2.605	0.2777	0.0775	0.1705	0.25	2.898	0.2767	0.0793	0.1527	0.14
43	02ZH002	40	0.103	0.3361	0.2127	0.1664	0.50	0.128	0.3177	0.1663	0.1356	0.34
44	02ZJ001	33	0.107	0.3910	0.2520	0.1482	1.22	0.131	0.3762	0.2337	0.1683	1.13
45	02ZJ002	27	0.273	0.3215	0.1385	0.2722	1.72	0.313	0.2972	0.1048	0.2702	1.65
46	02ZJ003	24	0.198	0.3007	0.1597	0.1857	0.24	0.235	0.2942	0.1531	0.2025	0.37
47	02ZK001	60	1.142	0.2809	0.1790	0.1539	0.25	1.377	0.2919	0.1838	0.1343	0.35
48	02ZK002	31	0.435	0.2311	0.1621	0.1437	0.48	0.523	0.2291	0.1542	0.1096	0.51
49	02ZK003	27	0.214	0.1391	0.1420	0.0767	2.69	0.239	0.1375	0.1577	0.1198	2.01
50	02ZK004	27	0.472	0.1966	0.0913	0.2020	0.63	0.546	0.2048	0.0315	0.1081	0.17
51	02ZL004	27	0.098	0.2470	0.0960	0.0270	1.09	0.112	0.2296	0.0416	0.0272	0.66
52	02ZL005	25	0.044	0.3113	0.2360	0.2235	0.75	0.049	0.3027	0.2656	0.2246	1.07
53	02ZM006	56	0.008	0.2495	0.1205	0.2190	0.52	0.009	0.2555	0.1506	0.1737	0.19
54	02ZM008	36	0.277	0.1437	-0.0841	0.0479	1.15	0.324	0.1407	-0.1012	-0.0092	1.76
55	02ZM009	30	0.339	0.2376	0.0776	0.1062	0.08	0.399	0.2302	0.0513	0.1102	0.05
56	02ZM016	27	0.087	0.2344	0.0456	0.0768	0.21	0.100	0.2331	0.0690	0.1098	0.04
57	02ZM018	26	0.084	0.1297	0.0465	-0.0006	2.79	0.098	0.1327	0.0620	-0.0261	2.76
58	02ZM020	24	0.112	0.1617	-0.0774	0.0550	0.95	0.128	0.1667	-0.0166	0.0984	0.56
59	02ZN001	28	0.469	0.1792	0.0430	0.1075	0.41	0.530	0.1926	0.0223	0.1037	0.26
60	02ZN002	25	0.095	0.2064	0.0066	0.0743	0.33	0.117	0.2021	0.0153	0.0995	0.22

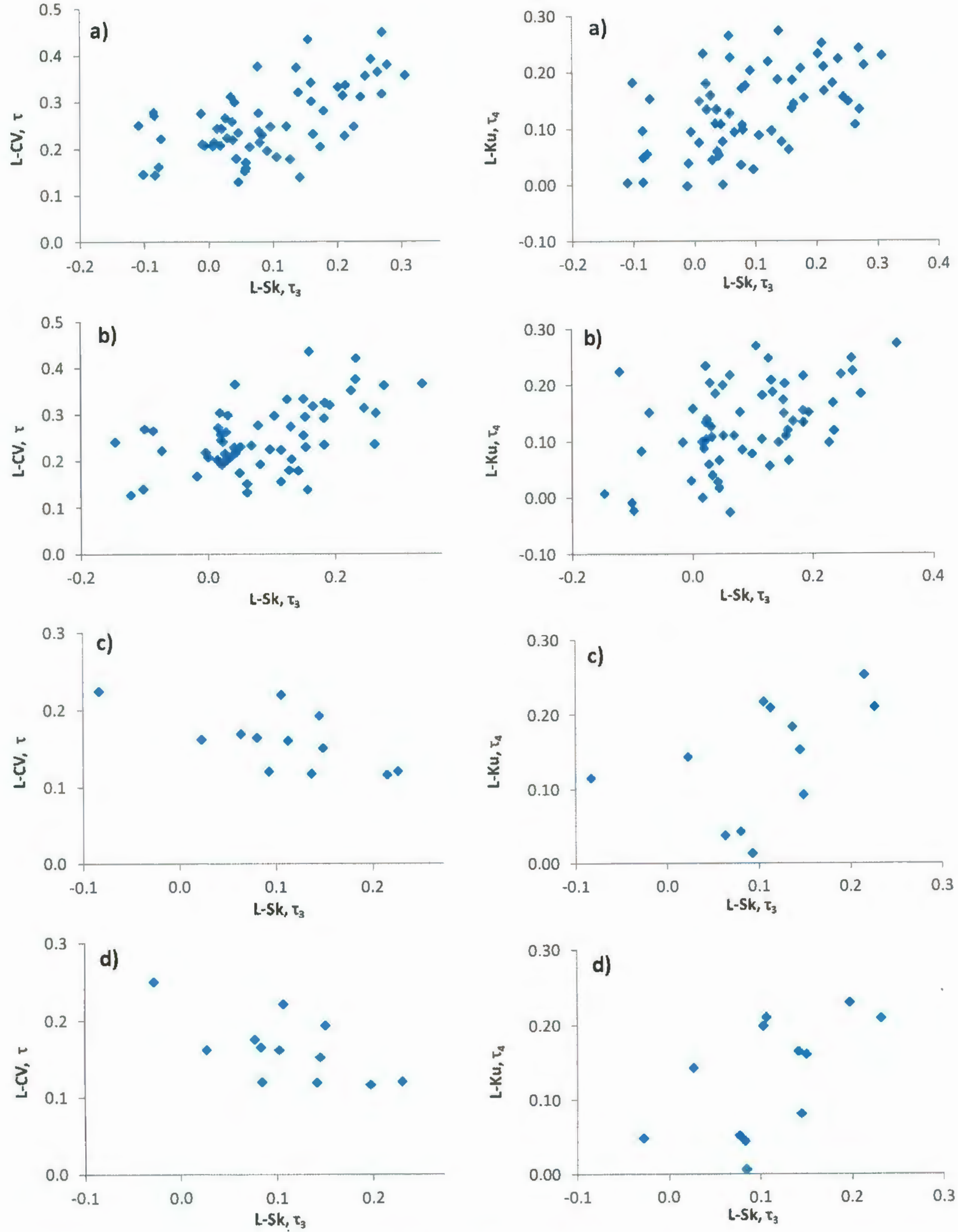

Figure 4-1 L-moment ratios in Newfoundland (a: 1-day; b: 7-day) and in Labrador (c: 1-day; d: 7-day)

Table 4.3 shows the regional average L-moment ratios for Newfoundland and Labrador both as separate regions along with their computed V, the weighted standard deviation of the at-site sample L-CVs.

Table 4-3 Weighted regional average L-statistics and weighted regional standard deviation

Region.	L-CV τ_R	L-sk τ_R^3	L-ku τ_R^4	V
Newfoundland (1-day AM)	0.25390	0.09150	0.13916	0.0707
Newfoundland (7-day AM)	0.24904	0.09017	0.13417	0.0659
Labrador (1-day AM)	0.16595	0.10153	0.13922	0.0362
Labrador (7-day AM)	0.16843	0.10667	0.13172	0.0392

A four parameter kappa distribution was then fitted to the regional average L-moment ratios of each region. The parameters of this distribution were estimated using the MATLAB code, *kappa distribution*, in Appendix A-2. Then large number of kappa regions (10000) were simulated using the *Heterogeneity test* code in Appendix A-3.

The inputs to the simulation code were kappa distribution parameters, ε, α, κ and h for the proposed region; number of sites in the proposed region, N and available record length at each site, n; and finally the weighted standard deviation of at-site sample L-CVs, V. The Heterogeneity program executes the following tasks. It generates 10000 regions from kappa distribution having the same number of sites each having the same record length as the real sites under study. Then it computes the L-CV for each site in the simulated region followed by computing the regional average L-CV weighted by the record length at each site. The weighted standard deviation of these at-site L-CVs then is computed for each of the simulated regions. And finally the overall mean, μ_V and standard deviation, σ_V are calculated for the simulated regions. Finally, the heterogeneity measure, H described in Eq. (3-30) is determined.

Table 4-4 Kappa distribution parameters and heterogeneity measures

Region.	Kappa Distribution Parameters				μ_V	σ_V	Heterogeneity measure H
	ξ	α	k	h			
Newfoundland (1-day)	0.84869	0.36061	0.06741	-0.1773	0.030	0.025	1.6
Newfoundland (7-day)	0.84039	0.36799	0.08797	-0.11778	0.029	0.026	1.4
Labrador (1-day)	0.89474	0.23804	0.06213	-0.1452	0.017	0.015	1.3
Labrador (7-day)	0.87671	0.25828	0.08920	-0.0348	0.017	0.013	1.5

The estimated four parameters of the kappa distribution for each region and the computed heterogeneity measure are presented in Table 4.4. The heterogeneity measure for Newfoundland and Labrador regions determined as 1.6 and 1.3 respectively for 1-day AM and 1.4 and 1.5 for 7-day AM indicates that these two regions are "possibly heterogeneous" under defined criteria by Hosking and Wallis (1997), but homogeneous as described by Robson and Reed (1999). Since heterogeneity measures for Newfoundland and Labrador regions are both less than the critical value of 2, one can conclude that these two regions can be considered as homogeneous, and there is no need to further divide the regions into smaller areas.

4.3 Identification of Regional Frequency Distribution

Once the homogeneous regions have been delineated, an appropriate distribution has to be selected as the regional frequency distribution. In this section, the results of a step-wise procedure outlined in section 3.2.5.4 employed for choosing the regional distributions are presented for Newfoundland and Labrador regions. The L-kurtosis based goodness-of-fit test was applied to the candidate distributions in order to select the best one. Then the L-moment ratio diagram was used as a graphical tool to confirm the choice of candidate distribution.

Hosking and Wallis's L-kurtosis based goodness-of-fit test was applied to the candidate distributions. This test compares regionally weighted average L-kurtosis corrected for the sampling bias with that of the candidate distributions. A MATLAB program, *goodness-of-fit* for carrying out this procedure was developed and is given in Appendix A-4. The bias and standard deviation of the regional L-kurtosis were estimated from the simulated kappa regions (see Table 4.5 for the regional kappa distribution parameters) as 0.0002 and 0.0081; and 0.0003 and 0.00080 for Newfoundland region 1-day and 7-day AM respectively, and 0.0001 and 0.0224; and -0.0003 and 0.0228 for Labrador region 1-day and 7-day respectively.

Table 4-5 L-Kurtosis based goodness-of-fit measure

Region		LN3	GLO	GEV	PE3	GPA
Newfoundland (1-day)	τ_4^{DIST}	0.129	0.174	0.124	0.125	0.026
	$\lvert Z^{DIST} \rvert$	1.198*	4.295	1.831	1.735	13.96
Newfoundland (7-day)	τ_4^{DIST}	0.129	0.173	0.124	0.125	0.026
	$\lvert Z^{DIST} \rvert$	0.592*	4.968	1.254*	1.122*	13.55
Labrador (1-day)	τ_4^{DIST}	0.131	0.175	0.127	0.126	0.030
	$\lvert Z^{DIST} \rvert$	0.369*	1.617*	0.542*	0.602*	4.865
Labrador (7-day)	τ_4^{DIST}	0.133	0.178	0.131	0.127	0.035
	$\lvert Z^{DIST} \rvert$	0.492*	1.466*	0.595*	0.774*	4.785

* These fits are acceptable

Table 4.6 presents the L-kurtosis τ_4^{DIST} of the candidate distributions fitted to the regional average L-moment ratios and the computed goodness-of-fit measure, Z^{DIST}. It is observed that all the candidate distributions except generalized Pareto are acceptable for Labrador datasets as their $|Z^{DIST}|$ value is smaller than critical value of 1.64. However, the lognormal distribution is the most appropriated with minimum $|Z^{DIST}|$ value. The only candidate distribution which passed the goodness-of-fit measure criteria for Newfoundland 1-day AM is the three-parameter lognormal distribution, and again the best fitted distribution for Newfoundland 7-day AM dataset is the three-parameter lognormal distribution.

In addition, the L-moment ratio diagram is also a very effective, simple and quick tool for regional frequency distribution. Figure 4.3 indicated that the points representing the regional average L-moment ratios, τ_3^R=0.09150 and τ_4^R=0.13916 for Newfoundland 1-day AM, τ_3^R=0.09017 and τ_4^R=0.13417 for Newfoundland 7-day AM, τ_3^R=0.10153 and τ_4^R=0.13922 for Labrador 1-day AM, and τ_3^R=0.10667 and τ_4^R=0.13922 for Labrador 7-day AM, lie close to the lognormal distribution, which supports the results of the goodness-of-fit test. Based on these tests it can be concluded that three parameter-lognormal distribution is the best distribution to represent the regional model for both Newfoundland and Labrador regions.

4.4 Regional Estimation using Index-flow Procedure (Regional Growth Curve)

Once the regions have been shown to be homogeneous, and suitable distribution has been identified for each region, the index flow procedure can then be applied to estimate the regional flows. As it was discussed in section 3.2.5.5, the key assumption in the index flow procedure is that the frequency distributions of all sites in a homogeneous region are identical, except for a site-specific scale factor, the index variable.

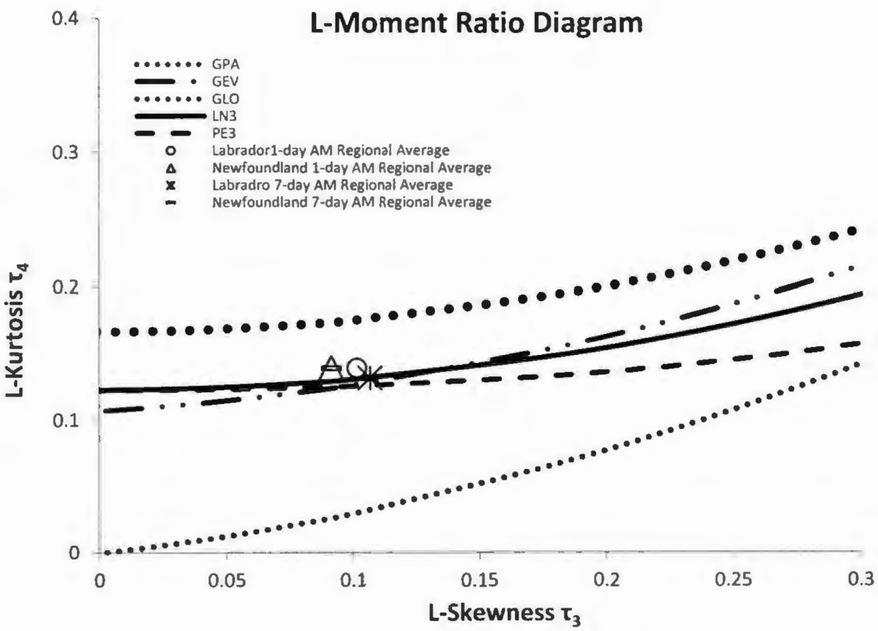

Figure 4-2 L-moment ratio diagram and regional averages

Now, the regional growth curve, $q(F)$ in Eq. (3-35), can be developed based on the best fitted distribution, three-parameter lognormal distribution, to the regional data. The quantile function of the lognormal distribution can be defined as:

$$[1] \quad q(F) = Q_T/Q_{mean} = \begin{cases} \xi + \alpha k^{-1}[1 - exp\,\{-k.\Phi^{-1}(F)\}] & \text{if } k \neq 0 \\ \xi + \alpha.\Phi^{-1}(F) & \text{if } k = 0 \end{cases} \qquad (4\text{-}1)$$

Where Φ is the cumulative distribution function of the standard normal distribution, Q_T is the low flow quantile, and Q_{mean} is the at-site mean of annual minimum discharges. The case $k = 0$ corresponds to the normal distribution.

The T-year return period of the regional growth factor is defined by Eq. (4-1), when F is replaced by $1/T$. Using the regional average L-moment ratios, the parameters of the lognormal distribution can be estimated. Hosking and Wallis (1997), page 197, provide the details on parameter estimation for three-parameter lognormal distribution. In this study a MATLAB program code, *Parameters of lognormal distribution* developed to perform this task, described in Appendix A-5. Table 4.7 provides the estimated lognormal distribution parameters for Newfoundland and Labrador regions.

Table 4-6 lognormal distribution parameters

Region	ξ	α	k
Newfoundland (1-day)	0.9581	0.4433	-0.1876
Newfoundland (7-day)	0.9594	0.4351	-0.1849
Labrador (1-day)	0.9696	0.2888	-0.2083
Labrador (7-day)	0.9676	0.29264	-0.2189

Figures 4.4 and 4.5, and 4-7 and 4-8 illustrate the estimated T-year regional growth factor, $q(F)$ along with observed values of Q_T/Q_{mean} for sample data at each site in Newfoundland and Labrador, respectively. The empirical distribution for estimating return periods for at-site date is obtained by using Cunnane plotting position formula $p_i(j) = (j - 0.4)/(n_i + 0.2)$ (Cunnane, 1978) for the jth ordered observation of site having n_i data. The horizontal axis using Z-values for standard normal distribution was transformed so that a normal distribution would plot as a straight line. It can be seen that the estimated return periods have reasonable agreement with the empirical values for all the sites both within Newfoundland and Labrador. Figures 4-3 and 4-6 in summary compares the differences between regional models of Newfoundland and Labrador for 1-day and 7-day AM, respectively. The lognormal regional model of Newfoundland in both cases shows a steeper line than the Labrador regional model.

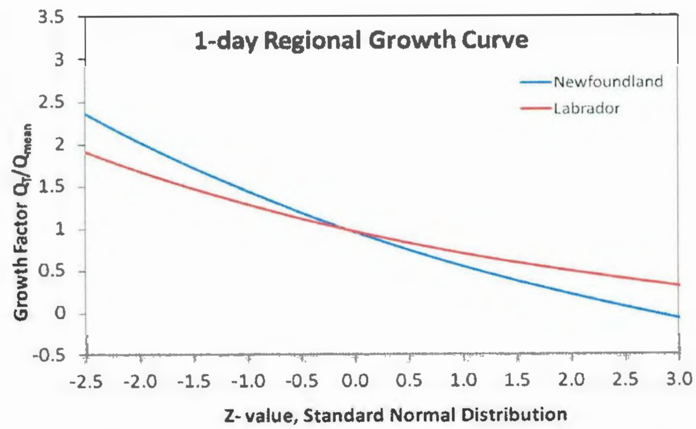

Figure 4-3 Regional comparison between fitted lognormal distributions 1-day

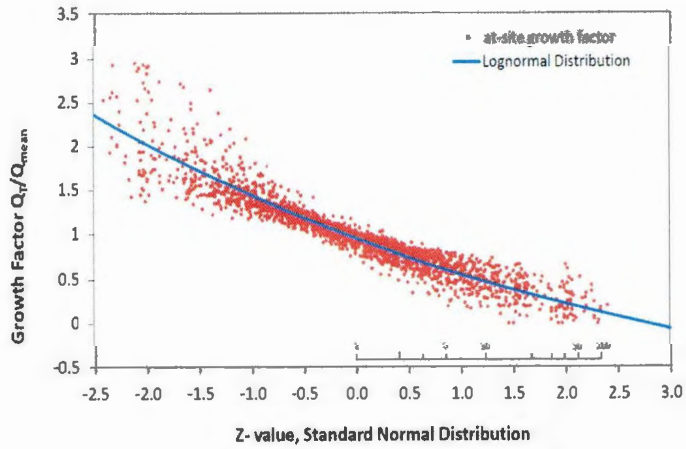

Figure 4-4 Regional comparison between at-site and fitted lognormal distribution, Newfoundland 1-day

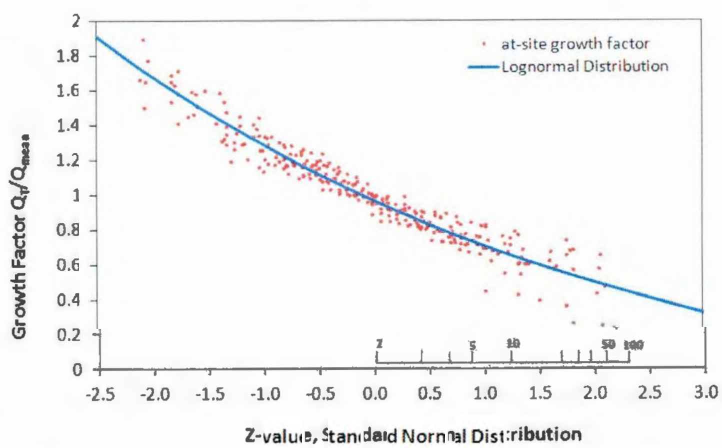

Figure 4-5 Regional comparison between at-site and fitted lognormal distribution, Labrador 1-day

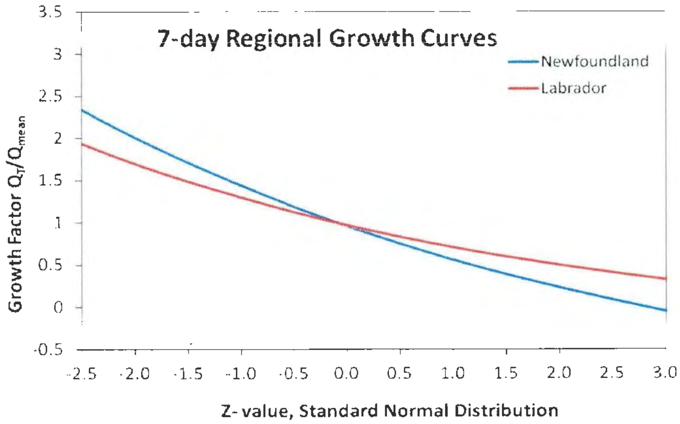

Figure 4-6 Regional comparison between fitted lognormal distributions 7-day

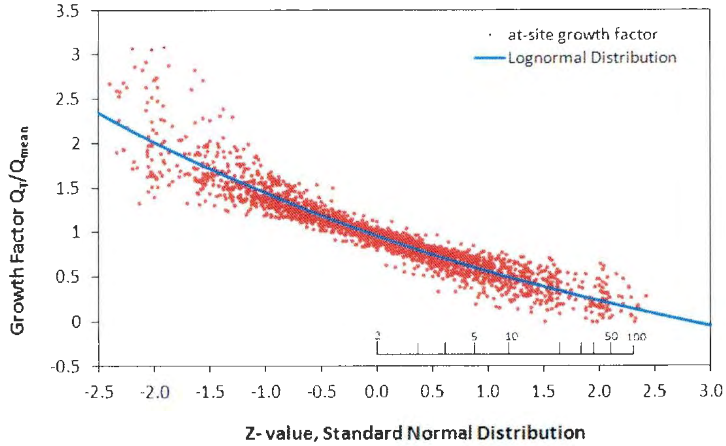

Figure 4-7 Regional comparison between at-site and fitted lognormal distribution, Newfoundland 7-day

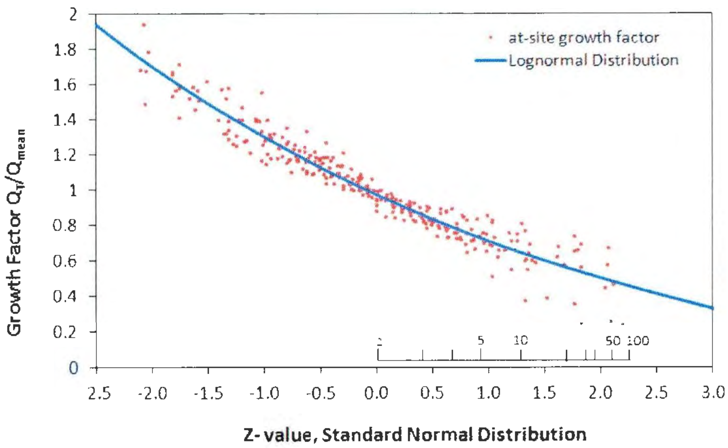

Figure 4-8 Regional comparison between at-site and fitted lognormal distribution, Labrador 7-day

4.5 Low Flow Estimation for Ungauged Sites

Based on the index flow procedure, for estimating a T-year return period minimum annual flow at any ungauged sites, an estimate of the mean annual minimum flow as index flow is required. Since observed flow data are not available at ungauged sites, at-site mean cannot be computed. In such a situation, it is necessary to establish a relationship between the mean annual minimum flow of gauged catchments within the homogeneous region and their pertinent physiographic and climatic characteristics to obtain an estimate of the mean annual minimum flow.

Unlike the previous low flow frequency study (Government of Newfoundland and Labrador, 1991) which used drainage area, precipitation amounts and land cover types as explanatory variables, in this study, it was assumed that the climatic characteristics are identical throughout the regions, and among the physiographic specifics, catchment size was used to establish a relationship with the magnitude of discharge. The drainage area data and mean annual minimum flow are available in Table 3.3 and 3.4, and 4.1 and 4.2 respectively for sites within Newfoundland and Labrador. Using the least-squares method the relationships are as follows in Table 4-7 for Newfoundland 1-day AM and 7-dayAM, and Labrador 1-day AM and 7-dayAM:

Table 4-7 mean annual minimum flow prediction equations

Region	Equation		R^2
Newfoundland (1-day)	$Q_{mean} = 0.0021\ A^{1.1067}$	(4-2)	0.91
Newfoundland (7-day)	$Q_{mean} = 0.0027\ A^{1.0848}$	(4-3)	0.92
Labrador (1-day)	$Q_{mean} = 0.0011\ A^{1.1225}$	(4-4)	0.97
Labrador (7-day)	$Q_{mean} = 0.0013\ A^{1.1075}$	(4-5)	0.97

Where Q_{mean} is mean annual minimum flow (m^3s^{-1}), and A (km^2) is catchment area. The coefficient of determination calculated from log-transformed data is R^2=0.91, R^2=0.92, R^2=0.97 and R^2=0.97 for Newfoundland and Labrador respectively which is quite satisfactory. Figure 4.9 illustrates these relationships.

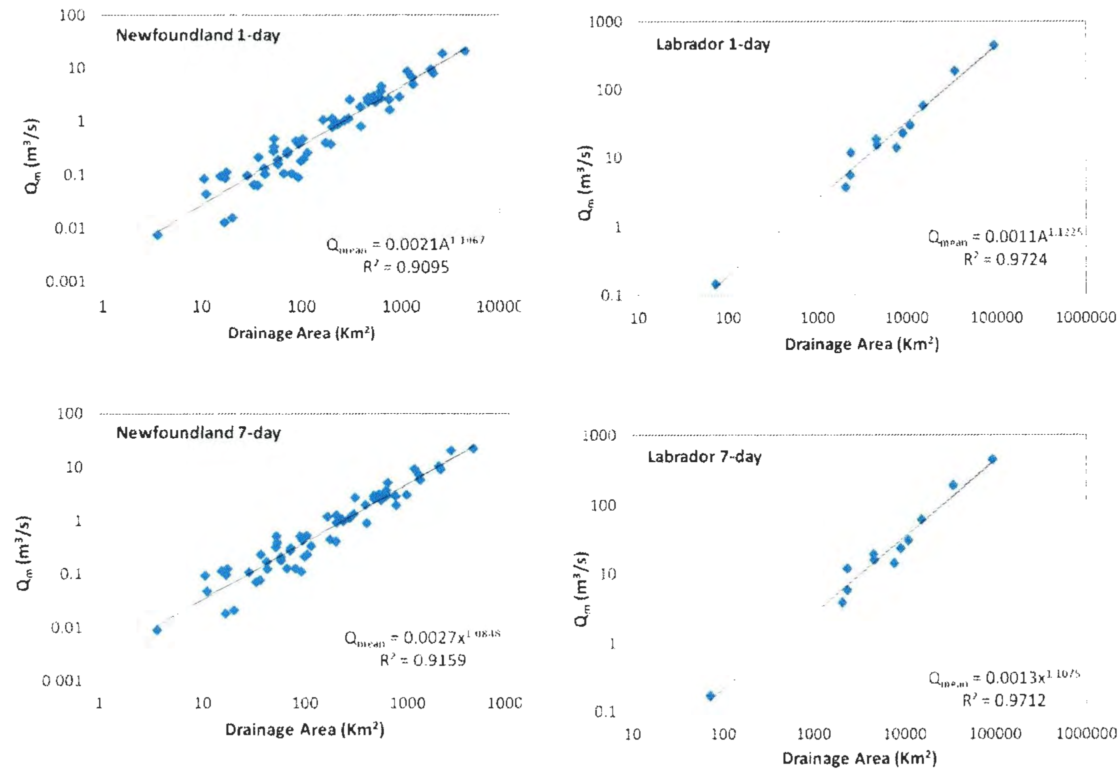

Figure 4-9 Regression of index flow with basin areas in Newfoundland and Labrador

It should be noted that the three-parameter lognormal distribution has no explicit form of the quantile function. Numerical iterations, such as Newton-Raphson method is needed to obtain an estimate of the quantile function (Hosking, 1996). For this reason a MATLAB code, *Quantiles of lognormal distribution* in Appendix A-6, was developed to perform this task and compute the quantiles of lognormal distribution. Finally, the minimum low flow estimate at return period T, Q_T for Newfoundland and Labrador 1-day and 7-day respectively can be then written as:

Table 4-8 minimum low flow prediction equations

Region	Equation	
Newfoundland (1-day)	$0.0021 \times \Phi^{-1}\{(0.1876)^{-1}Ln[1+0.1876(T^{-1}-0.9581)]/0.4433\} \times A^{1.1067}$	(4-6)
Newfoundland (7-day)	$0.0027 \times \Phi^{-1}\{(0.1849)^{-1}Ln[1+0.1849(T^{-1}-0.9594)]/0.4451\} \times A^{1.0848}$	(4-7)
Labrador (1-day)	$0.0011 \times \Phi^{-1}\{(0.2083)^{-1}Ln[1+0.2083(T^{-1}-0.9696)]/0.2888\} \times A^{1.1225}$	(4-8)
Labrador (7-day)	$0.0013 \times \Phi^{-1}\{(0.2189)^{-1}Ln[1+0.2189(T^{-1}-0.9676)]/0.2926\} \times A^{1.1075}$	(4-9)

Equation (4-6) to (4-9) derived from the quantile functions given in Eq. (4-1) and the relationships given by Eq. (4-2) to (4-5). One can then use the above equations to estimate the annual minimum 1-day and 7-day flow at any ungauged catchment within the studied regions, once the catchment area is known.

4.6 Verification of Results

Ten new hydrometric sites and four new sites have been selected in Newfoundland and Labrador regions respectively to verify the accuracy of previously defined regional growth models. Table 4-9 and 4-10 give information about these stations. Figure 4-10 to 4-13 illustrate good agreements between the observed growth factor and their respective regional estimated values. The Nash-Sutcliffe efficiency (NSE) is also computed and presented in Table 4-9 and 4-10 to give a numerical value for the comparisons. A NSE of one corresponds to a perfect match of modeled data to the observed data (Nash and Sutcliffe, 1970). NSE-1 and -2 refer to 1-day AM and 7-day respectively.

Table 4-9 Selected sites for verification of Newfoundland regional models

ID	Station Num.	Station Name	Sample Size	Drainage Area (km^2)	NSE-1	NSE-2
1	02YD001	Beaver Brook near Roddickton	19	237	0.88	0.89
2	02YF001	Cat Arm River above Great Cat Arm	12	611	0.93	0.92
3	02YH001	Bottom Creek near Rocky Harbour	12	33.4	0.93	0.90
4	02YJ003	Pinchgut Brook at outlet of Pinchgut	11	119	0.92	0.93
5	02YK003	Sheffield River at Sheffield Lake	10	362	0.94	0.92
6	02YK007	Glide Brook below Glide Lake	13	112	0.89	0.94
7	02YO007	Leech Brook near Grand Falls	12	88.3	0.87	0.85
8	02YP001	Shoal Arm Brook near Badger Bay	15	63.8	0.98	0.97
9	02YQ004	Northwest Gander River near Gander	15	2200	0.89	095
10	02ZA003	Little Codroy River near Doyles	15	139	0.92	0.91

Table 4-10 Selected sites for verification of Labrador regional models

ID	Station Num.	Station Name	Sample Size	Drainage Area (km^2)	NSE-1	NSE-2
1	03NE001	Reid Brook at outlet of Reid Pond	12	75.7	0.83	0.84
2	03OD007	East Metchin River	12	1750	0.89	0.87
3	03OE011	Pinus River	12	779	0.96	0.94
4	03PB001	Naskaupi River at Fermount Lake	13	8990	0.94	0.92

Based on the results it can be seen that the observed and model growth factors for both Labrador regions at sites 2, 3, and 4 have better agreement than at site 1. Site 1 has the smallest drainage area among the sites and this may mean that the prediction model may not be well calibrated for very small drainage areas because of the limited available data for very small catchments. The model predictions for Newfoundland sites in both models in overall have a quite satisfactory agreement with their respective observed values.

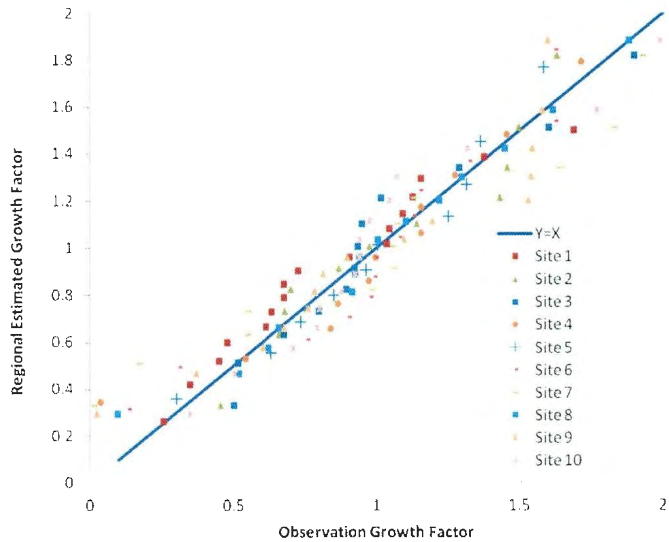

Figure 4-10 Observed and regional estimated growth factor, Newfoundland 1-day AM verification sites

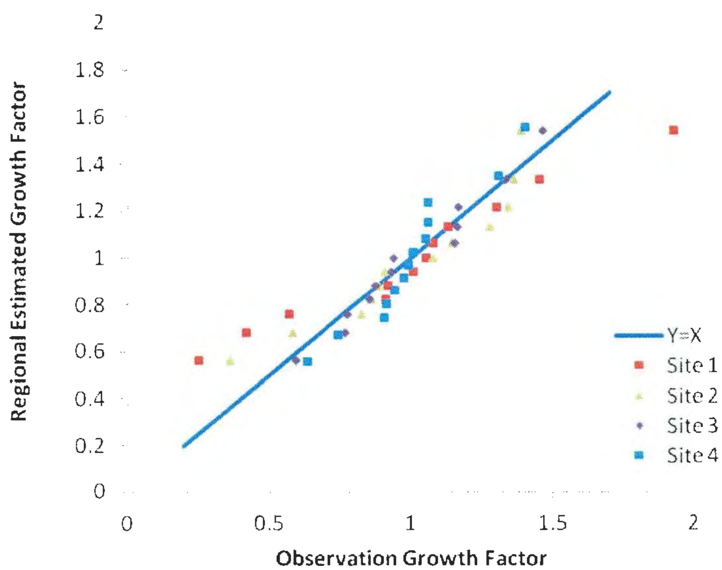

Figure 4-11 Observed and regional estimated growth factor, Labrador 1-day AM verification sites

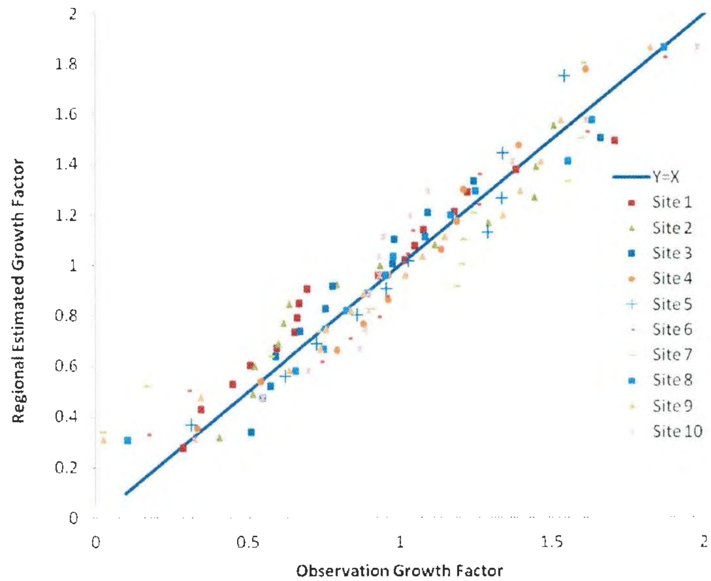

Figure 4-12 Observed and regional estimated growth factor, Newfoundland 7-day AM verification sites

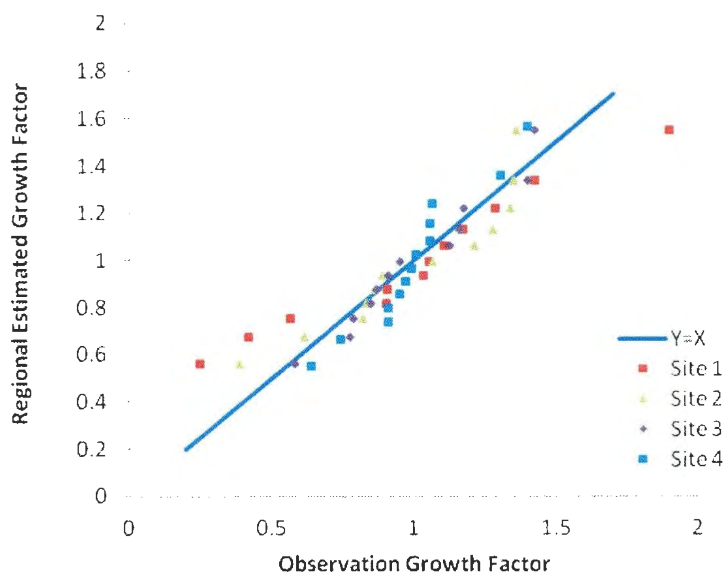

Figure 4-13 Observed and regional estimated growth factor, Labrador 7-day AM verification sites

5 Flow Duration Analysis and Results

This chapter presents the results of flow duration analyses for rivers within Newfoundland and Labrador region. The methodology described in section 3.3 was adopted for this study. The approach presented differs from other flow duration curve approaches in the scientific literature in some ways. First, it employs extensive landscape descriptors, whereas many earlier studies used fewer descriptors. Second, many of previous studies predicted only a few percentile flows representing low flows, or they used the same set of parameters to estimate the complete flow duration curve. This study identifies the relationship between landscape descriptors and 15 percentile flows, ranging from high to low flows. The objective of this chapter is to develop and test a regional regression method to predict flow duration curves and also annual flow duration curves for ungauged catchments within Newfoundland and Labrador.

The sequential data analysis approach described in section 3.3 was used and consists of the following steps: (1) construct FDCs and AFDCs from daily streamflow time series data for each catchment under study; (2) determine the 15 selected flow percentiles from both FDCs and AFDCs; (3) identify landscape descriptors that are the best percentile flow predictors using a step-wise regression method; (4) build regional models to predict percentile flow for the study area; (5) test the model by predicting the 15 percentile flows for the ungauged evaluation sites and reconstruct the complete FDC and AFDC for the ungauged evaluation sites; and finally (6) evaluate the prediction performance of the method by comparing reconstructed FDCs and AFDCs with their observed ones. The following sections in this chapter are described these analyses and the results.

5.1 Percentiles of FDCs and AFDCs

Using the methodologies introduced in section 3.3.1 FDCs and AFDCs were constructed for all the hydrometric stations under study in Newfoundland and Labrador. As discussed before, traditional period of record FDC leads to steady state or long-term probabilistic statements concerning streamflow exceedance and it will change by adding each year of data. However, AFDC has been shown to be quite useful for making probabilistic statements about typical (neither wet, nor dry) years of data and it is not affected by the observation of abnormally wet or dry periods during the period of record. For this reason, both POR FDCs and AFDCs were studied in this research. Selected sets of flow quantiles of FDC and AFDC described in section 3.3.2.2, for high, median, and low flows were determined for all the gauged rivers under study.

5.2 Physiographic Parameters

Drainage area, fraction of lake area, fraction of forest area, fraction of swamp area, fraction of barren area, fraction of lake and swamp area, fraction of area controlled by lakes and swamps, lake and swamp factor, length of main channel, elevation difference of main channel, slope of main channel, drainage density, and shape factor were introduced as the most important physical parameters of catchments in Newfoundland in the 1989 regional flood frequency report of Gov. of Newfoundland and Labrador. These significant physiographic parameters were extracted for the study area and provided in Table 5.1.

Table 5-1 Physiographic database

ID	Station Number	DA	FA	SW	FL	L+S	AB	ACLS	LSF	LAF	Length Main R	ELEV DIFF	Slope	DD	SF
		Km^2	-	-	-	-	-	-	-	-	(Km)	(m)	%	(km^{-1})	-
1	02YA001	306	0.64	0.14	0.22	0.35	0.01	0.96	1.78	1053	38.9	88	0.23	0.54	1.48
2	02YA002	33.6	0.4	0.03	0.13	0.16	0.44	0.99	1.91	652	13.2	150	1.14	0.91	1.64
3	02YC001	624	0.33	0.04	0.13	0.17	0.5	0.99	1.91	175	48.3	479	0.99	0.76	1.45
4	02YD002	200	0.83	0.04	0.13	0.17	0.01	0.99	1.9	484	38.3	270	0.7	0.93	1.65
5	02YE001	95.7	0.49	0.06	0.06	0.12	0.39	0.88	1.82	134	24.5	700	2.86	0.75	1.64
6	02YG001	627	0.78	0.06	0.07	0.13	0.09	0.63	1.55	18.3	31.9	375	1.18	1.3	1.83
7	02YJ001	640	0.79	0.09	0.06	0.14	0.07	0.75	1.67	141	60	509	0.85	1.12	1.81
8	02YK002	470	0.55	0.06	0.1	0.16	0.29	1	1.92	274	54.9	561	1.02	0.63	2.32
9	02YK004	529	0.35	0.24	0.12	0.36	0.29	0.95	1.77	666	49.3	320	0.65	0.64	1.78
10	02YK005	391	0.68	0.08	0.1	0.17	0.15	0.94	1.85	590	38.1	378	0.99	0.19	1.98
11	02YK008	20.4	0.75	0.22	0.02	0.24	0.01	0.65	1.5	0	10.1	137	1.35	1.28	1.47
12	02YL001	2110	0.74	0.06	0.05	0.11	0.15	0.75	1.68	50	118.8	678	0.57	0.79	1.56
13	02YL004	58.5	0.94	0.01	0.01	0.02	0.05	0.08	1.06	0	13.2	130	0.99	1.34	1.54
14	02YL005	17	0.91	0.08	0.02	0.1	0	0.46	1.39	0	8.2	244	2.98	1.05	1.1
15	02YL008	471	0.58	0.01	0.07	0.08	0.34	0.99	1.95	0	48.5	393	0.81	0.57	1.9
16	02YM001	974	0.79	0.07	0.09	0.16	0.05	0.88	1.8	36.4	65	290	0.45	0.45	1.88
17	02YM003	93.2	0.91	0.07	0.05	0.11	0	0.56	1.49	0	18.6	107	0.58	0.68	1.67

ID	Station Number	DA	FA	SW	FL	L+S	AB	ACLS	LSF	LAF	Length Main R	ELEV DIFF	Slope	DD	SF
		Km^2	-	-	-	-	-	-	-	-	(Km)	(m)	%	(km^{-1})	-
18	02YM004	243.8	0.48	0.093	0.134	0.227	0.294	0.918	1.80	218.1	23.66	116	0.490	0.472	7.469
19	02YN002	469	0.23	0.06	0.12	0.18	0.63	1	1.91	371	57.3	166	0.29	1.37	2.15
20	02YO006	177	0.83	0.13	0.03	0.16	0.02	0.97	1.89	0	42.7	190	0.45	0.8	1.93
21	02YO008	823	0.73	0.19	0.05	0.24	0.03	0.55	1.4	0	69	221	0.32	0.69	1.8
22	02YO012	58.7	0.8	0.08	0.12	0.2	0	0.67	1.55	128	22.7	134	0.59	0.54	1.87
23	02YQ001	4400	0.76	0.08	0.09	0.17	0.07	0.91	1.82	277	133.8	297	0.22	0.45	2.08
24	02YQ005	80.8	0.85	0.11	0.04	0.15	0	0.87	1.79	0	22.5	372	1.65	1.09	1.78
25	02YR001	267	0.75	0.07	0.18	0.24	0.01	0.98	1.83	881	49.3	177	0.36	0.26	1.93
26	02YR002	399	0.68	0.16	0.17	0.33	0	0.96	1.79	65.1	42	95	0.23	0.74	1.68
27	02YR003	554	0.7	0.13	0.2	0.33	0	0.9	1.8	307	52.4	136	0.26	0.68	1.72
28	02YS001	1290	0.55	0.21	0.09	0.3	0.15	0.92	1.76	138	105	207	0.2	0.73	2.35
29	02YS003	36.7	0.84	0.14	0.02	0.16	0	1	1.92	0	11.2	143	1.28	0.64	1.43
30	02YS005	2000	0.61	0.23	0.13	0.36	0.03	0.93	1.74	113	128.8	274	0.21	0.35	2.12
31	02ZA002	72	0.82	0.01	0.04	0.05	0.13	0.43	1.39	0	20.4	460	2.26	1.15	1.72
32	02ZB001	205	0.08	0.06	0.07	0.13	0.78	0.6	1.52	0	33.3	444	1.33	0.72	2.09
33	02ZC002	230	0.2	0.01	0.05	0.06	0.82	0.34	1.3	38.4	28.9	360	1.24	0.96	1.84
34	02ZD002	1340	0.04	0.16	0.04	0.2	0.75	0.63	1.51	0	60	310	0.52	0.15	5.31
35	02ZE001	2640	0.35	0.02	0.14	0.16	0.5	1	1.92	619	100.4	122	0.12	0.36	1.75
36	02ZE004	99.7	0.6	0.34	0.05	0.39	0.01	1	1.81	0	18.7	109	0.58	1.38	1.52

ID	Station Number	DA	FA	SW	FL	L+S	AB	ACLS	LSF	LAF	Length Main R	ELEV DIFF	Slope	DD	SF
		Km^2	-	-	-	-	-	-	-	-	(Km)	(m)	%	(km^{-1})	-
37	02ZF001	1170	0.32	0.05	0.18	0.24	0.44	0.96	1.84	401	68.1	282	0.41	0.61	2.15
38	02ZG001	205	0.26	0.01	0.09	0.1	0.63	0.96	1.91	202	44.7	370	0.83	0.55	2.45
39	02ZG002	166	0.37	0.04	0.09	0.13	0.49	0.92	1.82	588	26.7	221	0.83	1.35	1.84
40	02ZG003	115	0.16	0.06	0.07	0.13	0.73	0.92	1.85	42.8	24.5	136	0.55	1.55	1.62
41	02ZG004	42.7	0.34	0.03	0.14	0.16	0.46	0.92	1.83	123	10	107	1.07	1.62	1.53
42	02ZH001	764	0.11	0.48	0.18	0.66	0.23	0.91	1.57	17.4	50.9	207	0.41	0.71	1.67
43	02ZH002	43.3	0.4	0.02	0.08	0.1	0.5	0.92	1.87	20.8	17	110	0.65	1.11	1.66
44	02ZJ001	67.4	0.82	0.06	0.1	0.16	0.03	0.86	1.78	89.3	16	128	0.8	1.24	1.64
45	02ZJ002	73.6	0.74	0.06	0.13	0.19	0.07	0.82	1.72	436	18	137	0.76	1.11	1.33
46	02ZJ003	106	0.65	0.1	0.07	0.17	0.18	0.68	1.58	166	25.1	250	0.99	0.66	1.66
47	02ZK001	301	0.51	0.02	0.1	0.12	0.37	0.58	1.49	8.79	45.2	165	0.37	0.96	1.95
48	02ZK002	89.6	0.48	0.16	0.15	0.31	0.24	0.81	1.64	278	26.9	200	0.74	1.11	1.91
49	02ZK003	37.2	0.86	0.11	0.02	0.13	0.01	0.34	1.24	0	14.6	228	1.56	1.16	1.48
50	02ZK004	104	0.23	0.38	0.08	0.46	0.31	0.91	1.67	116	28.5	236	0.83	1.5	1.85
51	02ZL004	28.9	0.7	0	0.04	0.04	0.27	0.39	1.36	0	13.4	122	0.91	1.14	1.73
52	02ZL005	11.2	0.39	0.03	0.07	0.1	0.51	1	1.95	272	8.7	211	2.43	1	1.52
53	02ZM006	3.9	0.75	0.17	0.04	0.21	0.04	1	1.89	265	2.6	64	2.44	1.04	1.24
54	02ZM008	52.6	0.53	0.012	0.007	0.019	0.447	0.023	1.0	0	11.15	152	1.363	0.779	2.455
55	02ZM009	53.6	0.38	0.01	0.12	0.14	0.51	1	1.93	193	14.9	133	0.89	1.13	1.37

ID	Station Number	DA	FA	SW	FL	L+S	AB	ACLS	LSF	LAF	Length Main R	ELEV DIFF	Slope	DD	SF
		Km^2	-	-	-	-	-	-	-	-	(Km)	(m)	%	(km^{-1})	-
56	02ZM016	17.3	0.22	0.05	0.06	0.11	0.68	0.9	1.84	148	8.7	259	2.98	1.01	1.4
57	02ZM018	14.82	0.34	0.042	0.025	0.067	0.598	0.179	1.12	11.98	6.94	165	2.378	0.735	0.950
58	02ZM020	19.02	0.73	0.010	0.003	0.012	0.258	0.032	1.02	0	5.4	139	2.574	0.941	1.222
59	02ZN001	53.3	0.09	0	0.13	0.13	0.79	1	1.94	132	14.6	93	0.63	1.09	2.06
60	02ZN002	15.5	0.88	0	0.12	0.12	0	0.82	1.75	512	10.3	23	0.22	1.03	1.53
61	02XA003	4478	0.89	0.016	0.064	0.080	0.029	0.602	1.55	0	274.6	329	0.120	0.436	2.680
62	02XA004	2056.6	0.81	0.096	0.059	0.155	0.031	0.578	1.48	0	96.8	162	0.167	0.420	1.818
63	03NF001	7307.3	0.46	0.005	0.103	0.108	0.432	0.829	1.77	0	193.2	452	0.234	0.386	2.272
64	03NG001	8926.0	0.69	0.042	0.089	0.131	0.177	0.987	1.92	0	280.1	393	0.140	0.406	2.196
65	03OC003	15884.5	0.70	0.130	0.147	0.277	0.026	1.000	1.86	270	291.2	259	0.089	0.320	2.063
66	03OE003	2219.0	0.84	0.026	0.126	0.152	0.000	1.000	1.92	366	106.5	151	0.142	0.314	1.921
67	03OE010	70.7	0.93	0.006	0.064	0.070	0.000	0.994	1.96	115	27.5	128	0.466	0.663	1.791
68	03PB002	4540.9	0.81	0.023	0.147	0.170	0.019	0.974	1.89	0	174.1	298	0.171	0.398	1.939
69	03QC001	10705.0	0.73	0.088	0.084	0.173	0.093	0.849	1.76	0	252.8	428	0.169	0.425	1.989
70	03QC002	2312.0	0.88	0.050	0.030	0.080	0.037	0.304	1.24	0	81.0	437	0.539	0.541	1.624

ID	Station Number	DA	FA	SW	FL	L+S	AB	ACLS	LSF	LAF	Length Main R	ELEV DIFF	Slope	DD	SF
		Km^2	-	-	-	-	-	-	-	-	(Km)	(m)	%	(km^{-1})	-
71	02YD001*	237	0.81	0.04	0.05	0.08	0.11	0.73	1.68	0	40.6	328	0.81	0.34	2.23
72	02YF001*	611	0.69	0.05	0.08	0.13	0.18	1	1.93	0	30.2	250	0.83	0.58	1.86
73	02YG002*	224	0.83	0.06	0.09	0.15	0.02	0.96	1.88	299	26.4	255	0.96	0.45	1.84
74	02YJ003*	119	0.86	0.05	0.05	0.1	0.04	1	1.95	290	16.6	164	0.99	1.73	1.54
75	02YK003*	362	0.67	0.07	0.11	0.18	0.15	1	1.91	688	37	351	0.95	0.43	1.85
76	02YK007*	112	0.87	0.09	0.04	0.13	0	0.98	1.91	132	26.8	234	0.88	1.28	1.61
77	02YO007*	88.3	0.7	0.24	0.04	0.28	0.02	0.73	1.57	0	23.1	272	1.18	0.74	1.52
78	02YP001*	63.8	0.88	0.07	0.06	0.13	0	0.79	1.72	119	20	113	0.56	0.88	1.62
79	02YQ004*	2150	0.66	0.25	0.06	0.31	0.03	0.44	1.22	0	104.2	265	0.25	0.45	1.63
80	02ZA003*	139	0.66	0.07	0.04	0.11	0.16	0.73	1.66	131	25.2	450	1.78	1.46	1.68
81	03NE001*	75.5	0.09	0	0.137	0.137	0.769	1.000	1.93	310	17.7	412	2.332	0.380	1.330
82	03OD007*	1776.0	0.70	0.066	0.157	0.223	0.080	0.937	1.82	126	145.2	366	0.252	0.401	2.036
83	03OE011*	800.2	0.63	0.139	0.143	0.282	0.087	0.946	1.80	122	93.0	105	0.113	0.372	1.910
84	03NE002*	24.9	0.73	0	0.126	0.126	0.142	0.897	1.83	333	9.8	60	0.615	0.430	1.385

* Used only for the verification of results.

DA= Drainage area; FA=fraction of forest area, SW= fraction of swamp area; FL= fraction of lake area; AB=fraction of barren area, L+S=fraction of lake and swamp area; ACLS= fraction of area controlled by lakes and swamps; LSF=lake and swamp factor; LAF=lake attenuation factor, DD=drainage density; SF= shape factor

5.3 Sets of Regression Models

To develop sets of regressions that will estimate FDC and AFDC, 15 selected flow quantiles were regressed against the basin characteristics listed in Table 5.1.

Minitab software was used to perform these regressions. Natural-log transformations were taken of the basin characteristics and also flow quantiles to linearize the relation between the two. All the coefficients and selected variables in the final regression equations were significantly different from zero at the 0.05 significance level. Basin physiographic parameters included in the final equation had variance-inflation factor less than 2. Residuals of all the regression models were normally distributed, and they successfully passed all the diagnostics tests required for regression models. R-squared adjusted was close to the R-squared predicted values for all the models. For regression equations developed in log-space, bias correction factors were estimated by the Smearing Estimator to eliminate the retransformation bias of predicted data. Table 5.2 and 5.3 provides these prediction models, adjusted and predicted R-squared and correction factors for flow quantiles of FDCs and AFDCs for the region under study, Newfoundland and Labrador. Drainage area was the physical parameter with the most influence in all the models. The other selected parameters are also important as they represent the type of land cover (forest, barren, etc.), the drainage potential of watershed and effect of large lakes. One can observe from the closeness of the R-squared values to unity that these prediction models have good performance. However, it is necessary to investigate their accuracy by applying them on a new set of data which were not used in the construction of the models, and compare their prediction performance with the actual observed values.

Table 5-2 Sets of regression equations for FDC quantiles in Newfoundland and Labrador

Quantile	Prediction Equation	Correction Factor	R^2-adjusted	R^2-predicted
$Q_{0.01}$	Ln Q0.01 = - 1.60 + 0.993 Ln DA - 0.236 Ln FA + 0.148 Ln SL + 0.264 Ln DD - 0.115 Ln L+S - 0.0425 Ln LAF	1.020	0.985	0.982
$Q_{0.05}$	Ln Q0.05 = - 2.52 + 1.02 Ln DA - 0.221 Ln FA + 0.256 Ln DD - 0.0990 Ln L+S + 0.117 Ln SL	1.017	0.99	0.99
$Q_{0.1}$	Ln Q0.1 = - 2.87 + 1.02 Ln DA - 0.228 Ln FA + 0.244 Ln DD + 0.131 Ln SL - 0.0586 Ln SW	1.015	0.99	0.988
$Q_{0.15}$	Ln Q0.15 = - 3.38 + 1.02 Ln DA - 0.253 Ln FA + 0.265 Ln DD - 0.0590 Ln SW + 0.124 Ln SL + 0.0470 Ln LAF	1.015	0.99	0.987
$Q_{0.2}$	Ln Q0.2 = - 3.67 + 1.02 Ln DA - 0.268 Ln FA + 0.260 Ln DD - 0.0602 Ln SW+ 0.0623 Ln LAF + 0.105 Ln SL	1.016	0.989	0.987
$Q_{0.25}$	Ln Q0.25 = - 3.54 + 0.978 Ln DA - 0.220 Ln FA + 0.224 Ln DD + 0.0354 Ln AB+ 0.0638 Ln LAF - 0.0473 Ln SW	1.016	0.989	0.987
$Q_{0.3}$	Ln Q0.3 = - 3.76 + 0.981 Ln DA - 0.217 Ln FA + 0.0387 Ln AB + 0.219 Ln DD + 0.0756 Ln LAF - 0.0460 Ln SW	1.017	0.989	0.987
$Q_{0.4}$	Ln Q0.4 = - 3.95 + 0.985 Ln DA + 0.0547 Ln AB - 0.185 Ln FA + 0.0941 Ln LAF+ 0.206 Ln DD	1.021	0.986	0.984
$Q_{0.5}$	Ln Q0.5 = - 4.29 + 0.986 Ln DA + 0.0616 Ln AB + 0.114 Ln LAF - 0.180 Ln FA + 0.202 Ln DD	1.026	0.983	0.98
$Q_{0.6}$	Ln Q0.6 = - 4.61 + 0.978 Ln DA + 0.0655 Ln AB + 0.137 Ln LAF - 0.188 Ln FA+ 0.205 Ln DD	1.035	0.977	0.973
$Q_{0.7}$	Ln Q0.7 = - 4.90 + 0.969 Ln DA + 0.0715 Ln AB + 0.151 Ln LAF - 0.189 Ln FA + 0.199 Ln DD	1.046	0.968	0.963
$Q_{0.8}$	Ln Q0.8 = - 4.87 + 0.944 Ln DA + 0.119 Ln AB + 0.147 Ln LAF	1.059	0.962	0.955
$Q_{0.9}$	Ln Q0.9 = - 5.43 + 0.968 Ln DA + 0.120 Ln AB + 0.155 Ln LAF	1.073	0.955	0.948
$Q_{0.95}$	Ln Q0.95 = - 5.87 + 0.994 Ln DA + 0.127 Ln AB + 0.154 Ln LAF	1.095	0.946	0.938
$Q_{0.99}$	Ln Q0.99 = - 6.62 + 1.04 Ln DA + 0.148 Ln AB + 0.140 Ln LAF	1.204	0.904	0.891

Table 5-3 Sets of regression equations for AFDC quantiles in Newfoundland and Labrador

Quantile	Prediction Equation	Correction Factor	R^2-adjusted	R^2-predicted
$Q_{0.01}$	Ln Q0.01 = - 1.50 + 0.985 Ln DA - 0.253 Ln FA - 0.0604 Ln LAF + 0.165 Ln SL + 0.275 Ln DD - 0.122 Ln L+S	1.024	0.982	0.978
$Q_{0.05}$	Ln Q0.05 = - 2.53 + 1.03 Ln DA - 0.203 Ln FA + 0.244 Ln DD + 0.155 Ln SL - 0.0656 Ln SW	1.016	0.989	0.987
$Q_{0.1}$	Ln Q0.1 = - 2.85 + 1.02 Ln DA - 0.230 Ln FA + 0.243 Ln DD + 0.129 Ln SL - 0.0555 Ln SW	1.015	0.99	0.987
$Q_{0.15}$	Ln Q0.15 = - 3.39 + 1.02 Ln DA - 0.254 Ln FA + 0.270 Ln DD - 0.0550 Ln SW + 0.123 Ln SL + 0.0523 Ln LAF	1.016	0.989	0.986
$Q_{0.2}$	Ln Q0.2 = - 3.45 + 0.979 Ln DA - 0.283 Ln FA + 0.265 Ln DD + 0.0526 Ln LAF - 0.0504 Ln SW	1.019	0.987	0.986
$Q_{0.25}$	Ln Q0.25 = - 3.34 + 0.972 Ln DA - 0.205 Ln FA + 0.225 Ln DD + 0.0618 Ln LAF + 0.0420 Ln AB	1.019	0.987	0.986
$Q_{0.3}$	Ln Q0.3 = - 3.56 + 0.978 Ln DA + 0.0477 Ln AB - 0.190 Ln FA + 0.218 Ln DD + 0.0731 Ln LAF	1.020	0.986	0.983
$Q_{0.4}$	Ln Q0.4 = - 3.97 + 0.988 Ln DA + 0.0520 Ln AB - 0.180 Ln FA + 0.215 Ln DD + 0.0954 Ln LAF	1.022	0.986	0.983
$Q_{0.5}$	Ln Q0.5 = - 4.30 + 0.988 Ln DA + 0.0555 Ln AB + 0.113 Ln LAF - 0.182 Ln FA + 0.209 Ln DD	1.027	0.982	0.979
$Q_{0.6}$	Ln Q0.6 = - 4.60 + 0.978 Ln DA + 0.0630 Ln AB + 0.137 Ln LAF - 0.179 Ln FA + 0.204 Ln DD	1.034	0.977	0.973
$Q_{0.7}$	Ln Q0.7 = - 4.83 + 0.961 Ln DA + 0.0722 Ln AB + 0.152 Ln LAF - 0.170 Ln FA + 0.199 Ln DD	1.047	0.966	0.961
$Q_{0.8}$	Ln Q0.8 = - 4.82 + 0.935 Ln DA + 0.115 Ln AB + 0.148 Ln LAF	1.060	0.958	0.952
$Q_{0.9}$	Ln Q0.9 = - 5.40 + 0.964 Ln DA + 0.113 Ln AB + 0.164 Ln LAF	1.068	0.957	0.95
$Q_{0.95}$	Ln Q0.95 = - 5.85 + 0.991 Ln DA + 0.109 Ln AB + 0.177 Ln LAF	1.080	0.953	0.946
$Q_{0.99}$	Ln Q0.99 = - 6.97 + 1.06 Ln DA + 0.253 Ln LAF + 0.106 Ln AB - 0.209 Ln ACLS	1.095	0.95	0.944

5.4 Verification of Results

To test the performance of the prediction models, another set of data was selected which consists of 10 sites from the Island of Newfoundland region, and 4 sites from the Labrador region. The physiographic parameters of these stations were also provided in Table 5.1, sites 71 to 84, but they were not included in the regression equation development.

The prediction equations provided in Tables 5.2 and 5.3 were used to estimate the flow quantiles of FDC and AFDC for these selected sites based on their measured physiographic parameters. Correction factors were applied to adjust the retransformed natural-log estimated flow quantiles. These estimated values then have been compared to the actual observed flow quantiles. Figures 5.1 and 5.2 compare these two. One can observe that the model performance for low flows is much better than high flows. In addition, Table 5.4 provides the Nash-Sutcliffe Efficiency (NSE) for the validation sites. Low NSE values in some cases (e.g 02YD001) despite the good agreement in low flow end of graph might be the effect of comparison based on the combined flow quantiles. The differences between estimated and observed FDCs for sites 11 to 14 are higher than other site. These sites belong to Labrador region, and it might mean that the models have not been very well calibrated for this region because of lack of enough data. In general, the performances of the models are reasonable, and can be used for future predictions.

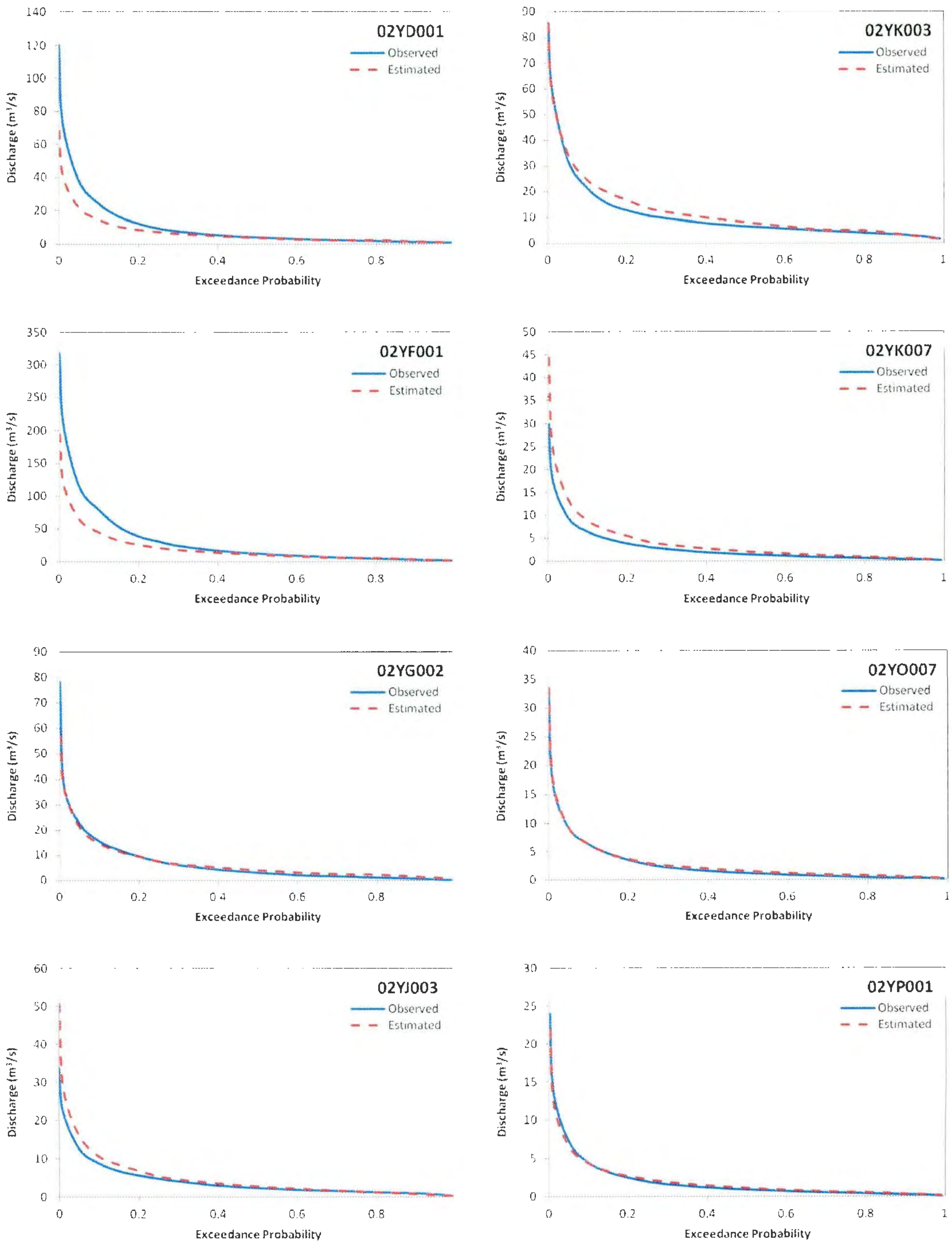

Figure 5-1 Comparison of observed and estimated FDCs for validation sites

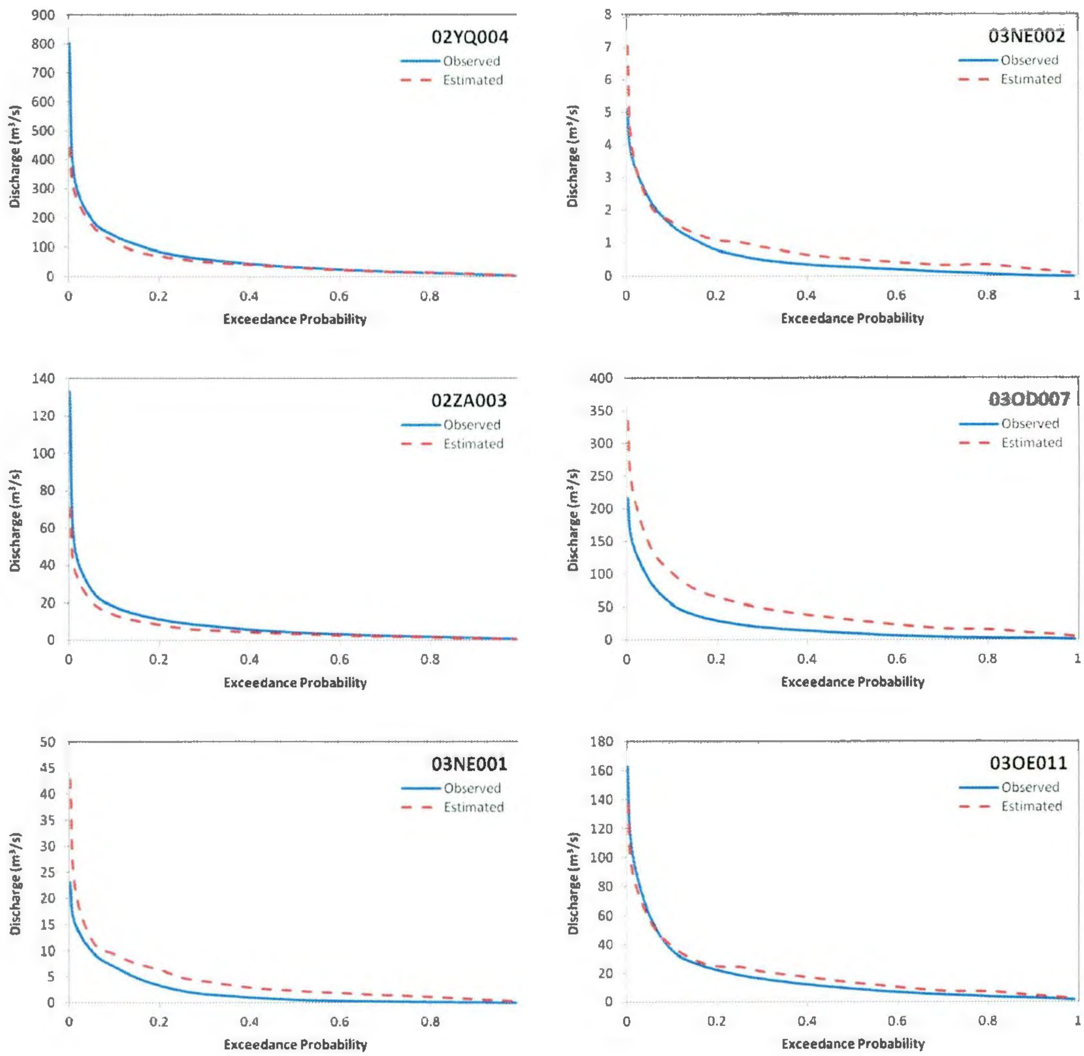

Figure 5-1 continue: Comparison of observed and estimated FDCs for validation sites

Table 5-4 NSE values for FDC and AFDC predictions in validation sites

ID	Station Num.	NSE-1 FDC model	NSE-2 AFDC model	ID	Station Num.	NSE-1 FDC model	NSE-2 AFDC model
1	02YD001	0.74	0.73	8	02YP001	0.99	0.97
2	02YF001	0.77	0.75	9	02YQ004	0.78	0.95
3	02YG002	0.93	0.99	10	02ZA003	0.75	0.85
4	02YJ003	0.72	0.86	11	03NE001	0.65	0.63
5	02YK003	0.99	0.98	12	03OD007	0.57	0.62
6	02YK007	0.70	0.67	13	03OE011	0.97	0.94
7	02YO007	0.99	0.95	14	03NE002	0.86	0.93

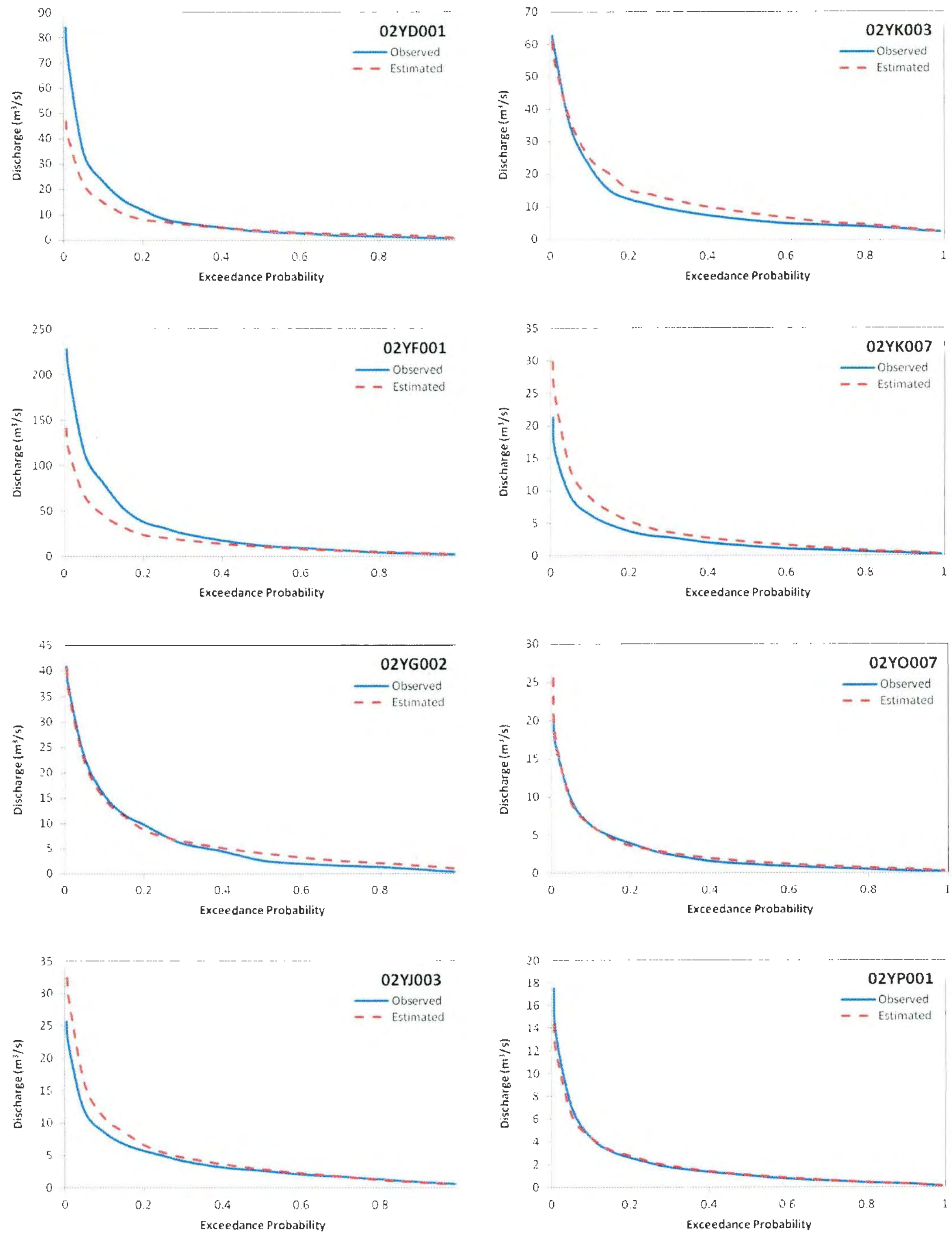

Figure 5-2 Comparison of observed and estimated AFDCs for validation site

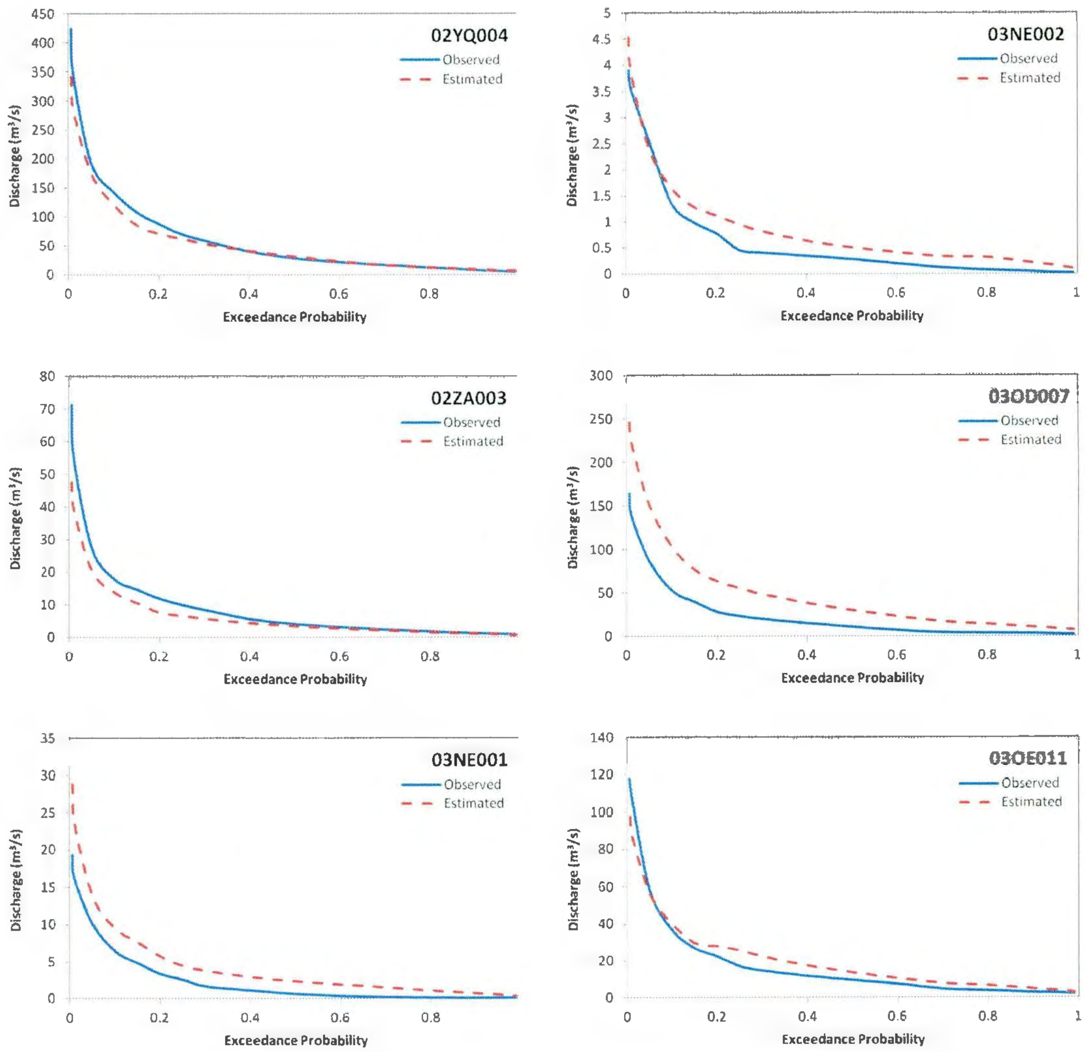

Figure 5-2 continue: Comparison of observed and estimated AFDCs for validation site

6 Flow Spell Analysis and Results

This chapter presents the analysis and results of comparing 3 major types of hydrologically based methods for instream flow evaluation, (*i*) Percentiles of FDC and AFDC (Q85 and Q95); (*ii*) Percentage of mean annual flow (25% MAF and Tennant's method); and (*iii*) the statistical low flow frequency method (7Q10, 7-day low flow having a 10-year return period). Regional models to predict annual maximum spells are sought in this study. Details on the methods of estimating each instream flow requirement were provided in Section 3.4.2.

6.1 Instream Flow Threshold Values

Tables 6-1 and 6-2 present the calculated threshold values of the above mentioned instream flow requirements Labrador and Newfoundland respectively.

Table 6-1 Results of thresholds (m^3/s) obtained for rivers in Labrador

ID	25% MAF	30% MAF*	50% MAF**	FDC Q85	FDC Q95	AFDC Q85	AFDC Q95	7Q10
1	23.34	28.01	46.68	19.30	15.20	18.80	15.68	11.81
2	39.85	47.82	79.70	18.83	12.80	18.36	13.74	8.37
3	77.04	92.45	154.08	78.00	60.50	75.90	62.15	45.82
4	14.08	16.89	28.16	14.90	10.90	15.16	13.06	7.83
5	22.49	26.99	44.98	23.20	16.90	22.12	19.50	12.67
6	63.33	75.99	126.65	37.90	24.50	39.15	31.66	15.95
7	13.00	15.60	25.99	7.50	5.09	7.53	5.84	3.67
8	10.73	12.88	21.46	5.10	3.85	5.16	4.11	2.87
9	46.05	55.27	92.11	28.00	22.40	28.06	22.92	17.63
10	179.76	215.71	359.51	210.00	159.00	217.00	196.20	131.35
11	0.43	0.51	0.85	0.24	0.15	0.25	0.21	0.08
12	406.58	487.90	813.16	513.00	399.00	504.00	456.00	295.69

* Tennant's method: Threshold for October-March period

** Tennant's method: Threshold for April-September period

Table 6-2 Results of thresholds (m^3/s) obtained for rivers in Newfoundland

ID	25% MAF	30% MAF*	50% MAF**	FDC Q85	FDC Q95	AFDC Q85	AFDC Q95	7Q10
1	2.205	2.645	4.409	3.450	2.540	3.612	3.070	1.612
2	0.343	0.412	0.686	0.150	0.073	0.140	0.096	0.025
3	6.115	7.338	12.230	5.820	3.750	5.830	4.306	2.234
4	1.358	1.629	2.715	0.732	0.385	0.777	0.459	0.160
5	1.174	1.409	2.349	0.975	0.435	0.933	0.512	0.150
6	7.083	8.500	14.166	5.200	3.470	5.346	3.586	1.910
7	6.585	7.902	13.170	8.190	5.600	8.493	5.962	2.814
8	4.533	5.439	9.065	4.700	3.260	4.975	3.534	1.872
9	4.114	4.937	8.228	5.100	3.430	5.006	3.744	1.501
10	2.766	3.320	5.533	3.200	2.010	3.374	2.538	1.096
11	0.128	0.153	0.255	0.052	0.024	0.058	0.034	0.005
12	20.375	24.450	40.751	16.700	10.100	16.670	11.000	4.754
13	0.450	0.540	0.899	0.370	0.250	0.391	0.278	0.145
14	0.124	0.149	0.248	0.053	0.025	0.062	0.031	0.006
15	6.658	7.989	13.315	4.890	2.980	5.154	3.062	1.423
16	4.839	5.807	9.678	4.840	3.200	5.200	3.880	1.694
17	0.653	0.783	1.305	0.285	0.120	0.352	0.174	0.022
18	1.611	1.934	3.223	1.850	1.200	1.806	1.374	0.442
19	5.174	6.209	10.348	4.830	3.370	5.059	3.804	2.360
20	1.133	1.359	2.265	0.823	0.520	0.880	0.599	0.257
21	5.504	6.605	11.009	4.030	2.321	4.054	2.750	0.929
22	0.386	0.463	0.772	0.379	0.201	0.396	0.254	0.065
23	30.226	36.271	60.452	35.100	20.600	38.720	27.160	10.254
24	0.611	0.733	1.221	0.365	0.172	0.416	0.216	0.056
25	1.698	2.038	3.396	1.890	0.900	1.956	1.352	0.400
26	2.382	2.858	4.764	1.730	0.509	1.728	0.914	0.145
27	3.439	4.127	6.878	4.230	2.078	4.453	3.121	0.976
28	9.182	11.019	18.364	11.800	7.116	12.700	8.996	3.233
29	0.259	0.311	0.518	0.185	0.095	0.198	0.118	0.045
30	12.504	15.005	25.008	16.900	10.275	18.280	11.420	5.540

Table 6-2 continue: Results of thresholds (m^3/s) obtained for rivers in Newfoundland

ID	25% MAF	30% MAF*	50% MAF**	FDC Q85	FDC Q95	AFDC Q85	AFDC Q95	7Q10
31	0.688	0.826	1.376	0.530	0.356	0.560	0.386	0.199
32	3.397	4.077	6.795	1.850	1.080	1.779	1.104	0.665
33	3.562	4.274	7.123	2.230	1.270	2.284	1.382	0.554
34	13.958	16.750	27.916	9.742	6.301	10.160	7.184	3.004
35	21.438	25.726	42.876	26.900	16.600	30.780	21.000	6.726
36	0.846	1.016	1.693	0.549	0.283	0.587	0.309	0.074
37	9.983	11.979	19.965	15.300	9.430	16.430	11.660	4.892
38	2.227	2.672	4.453	2.551	1.440	2.614	1.690	0.523
39	2.015	2.418	4.030	2.400	1.350	2.368	1.852	0.553
40	1.226	1.471	2.451	0.834	0.400	0.853	0.518	0.143
41	0.528	0.634	1.056	0.430	0.237	0.437	0.271	0.077
42	6.325	7.590	12.649	5.400	3.028	5.546	3.606	1.066
43	0.478	0.574	0.956	0.331	0.161	0.348	0.177	0.045
44	0.539	0.647	1.078	0.319	0.121	0.332	0.167	0.036
45	0.644	0.773	1.289	0.675	0.345	0.708	0.447	0.103
46	0.818	0.981	1.636	0.690	0.312	0.750	0.345	0.093
47	2.815	3.378	5.630	2.830	1.420	2.919	1.751	0.572
48	1.015	1.218	2.031	1.100	0.600	1.172	0.672	0.278
49	0.398	0.478	0.797	0.340	0.247	0.340	0.255	0.171
50	1.322	1.586	2.644	1.000	0.595	1.046	0.714	0.288
51	0.224	0.268	0.447	0.207	0.118	0.218	0.149	0.059
52	0.107	0.128	0.213	0.099	0.053	0.100	0.056	0.020
53	0.034	0.041	0.068	0.020	0.010	0.019	0.011	0.004
54	0.549	0.659	1.099	0.518	0.352	0.551	0.392	0.224
55	0.731	0.878	1.463	0.757	0.434	0.775	0.521	0.205
56	0.177	0.213	0.355	0.196	0.114	0.213	0.119	0.050
57	0.134	0.161	0.268	0.152	0.107	0.155	0.110	0.071
58	0.200	0.240	0.401	0.198	0.138	0.211	0.145	0.082
59	0.784	0.941	1.568	0.885	0.581	0.910	0.669	0.308
60	0.204	0.244	0.407	0.204	0.126	0.220	0.149	0.066

* Tennant's method: Threshold for October-March period

** Tennant's method: Threshold for April-September period

Considering the large area under study flow variation among the rivers is expected. Rivers 23 and 53 have the highest and the lowest threshold values respectively among rivers in Newfoundland. Rivers 12 and 11 have the highest and the lowest threshold values respectively among rivers in Labrador.

6.2 Comparison of Estimated Flows at Different Thresholds

In terms of comparing the estimated flows at different thresholds, the Tennant's method is easily compared to the 25% MAF which is the commonly used method in Atlantic Canada (Caissie and El-Jabi, 1995) as they both use a fixed percentage of MAF (Tennant's method is on average equal to 40% MAF). As compared to the 25% MAF method, the Tennant's method exceeds the 25% MAF by 20% and 100% for the periods October-March and April-September. However, comparison of the estimated flows obtained from other methods is not as straightforward. Therefore, a percentage difference between estimated flow from each method under investigation and 25% MAF method is calculated for all the hydrometric station. Figures 6-1 and 6-2 illustrate these comparisons in the form of a boxplot for all the gauges in Newfoundland and Labrador respectively. The y-axis represents a percentage difference between estimated flow for the compared method and the 25% MAF. 0% represents the complete agreement between the estimated flows of the two methods. One can observe from these graphs that the estimated flows for 7Q10 method show a significant underestimation in contrast to the 25% MAF and have the lowest threshold values both for Newfoundland and Labrador. Boxplots of the estimated flows for FDC Q85 and AFDC Q85 show quite the same difference with the

25% MAF. The estimated flows for AFDC Q95 have slightly smaller percentage difference to the 25% MAF than FDC Q95.

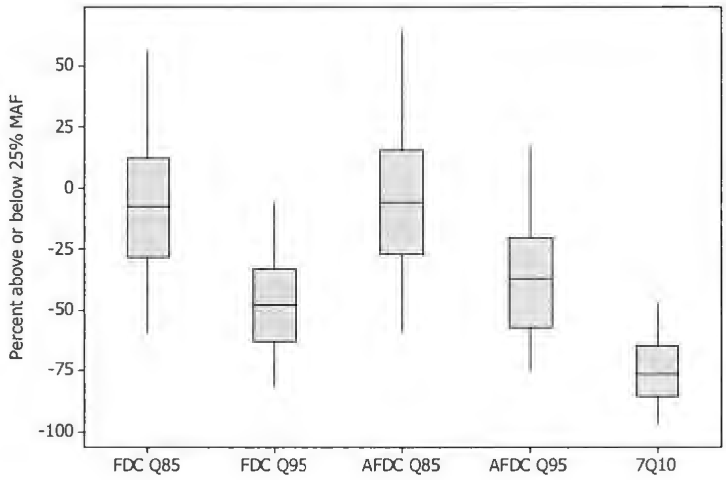

Figure 6-1 Comparison of the estimated flows for different threshold methods with 25% MAF for Newfoundland

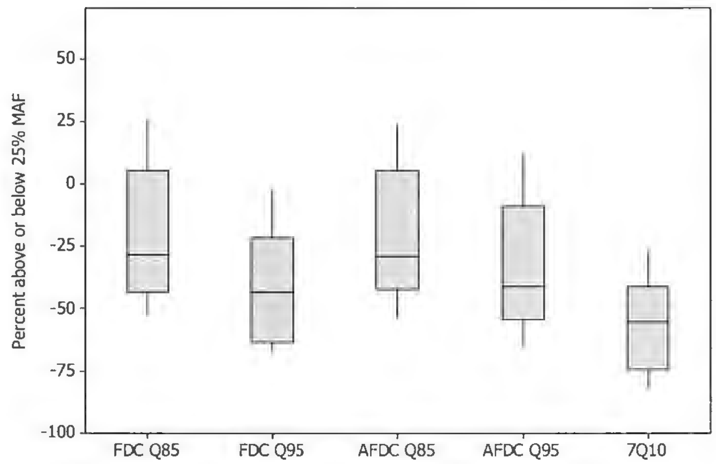

Figure 6-2 Comparison of the estimated flow for different threshold methods with 25% MAF for Labrador

Following the instream flow technique comparative study, an analysis of water availability was carried out by calculating the probability of occurrence of the instream flows. This study determines the percentage of the time that the discharge in the river is greater than the instream flow requirement calculated previously.

Using FDC Q85 and Q95, minimum instream flow requirement would be available 85% and 95% of the time respectively by definition. Tables 6-3 and 6-4 show the results of this analysis for other hydrologically based methods for Labrador and Newfoundland hydrometric stations respectively.

One can observe from the provided tables that the Tennant's method (taken as 40% MAF) on average provides the lowest probability of exceedance followed by 25% MAF method for both Newfoundland and Labrador stations as expected. The method with the highest probability of exceedance calculated by the flow duration analysis is the 7Q10 method for both Newfoundland and Labrador. Probability of exceedance for 7Q10 method is in the range of 99% of time. This means that using this particular technique as water abstraction regulation could in fact allow removing available streamflow 99% of time which leaves the required instream flow for other usages only 1% of time.

Based on the provided results, instream flows calculated using the Tennant's method on average are available 68% and 66% of the time for Newfoundland and Labrador respectively. The 25% MAF provides 81% and 77% available instream flow for Newfoundland and Labrador respectively. AFDC Q85 and Q95 showed similar results as their period of record FDC matches.

Table 6-3 Probability of exceedance by flow duration analysis for Labrador

ID	25% MAF	Tennant	AFDC Q85	AFDC Q95	7Q10	ID	25% MAF	Tennant	AFDC Q85	AFDC Q95	7Q10
1	76.7	64.9	86.1	93.7	99.5	7	67.9	54.3	84.8	91.7	99.3
2	64.1	58.3	85.7	93.1	99.2	8	64.0	56.9	84.5	93.4	99.8
3	85.4	76.6	86.2	94.3	99.6	9	66.6	60.2	84.8	94.1	99.6
4	87.8	76.2	84.2	90.1	99.1	10	90.8	77.8	83.2	87.8	99.6
5	86.4	71.7	86.8	91.4	99.4	11	65.4	52.1	84.1	89.4	98.7
6	69.6	61.1	83.8	90.6	98.5	12	94.5	83.2	85.9	91.2	98.9

Table 6-4 Probability of exceedance by flow duration analysis for Newfoundland

ID	25% MAF	Tennant	AFDC Q85	AFDC Q95	7Q10	ID	25% MAF	Tennant	AFDC Q85	AFDC Q95	7Q10
1	97.4	87.8	83.0	90.1	99.5	31	75.5	60.9	83.0	93.6	99.5
2	65.9	55.3	86.4	92.0	99.3	32	68.5	56.3	86.0	94.5	99.2
3	83.4	69.5	85.0	92.6	99.2	33	70.6	55.4	84.4	93.9	99.8
4	71.4	60.2	83.7	92.9	99.5	34	74.3	59.5	83.9	92.5	99.5
5	81.1	69.9	85.8	93.5	99.4	35	90.6	78.2	81.7	90.8	98.8
6	75.2	60.8	84.1	94.3	99.8	36	72.8	59.5	83.3	94.1	99.6
7	91.6	75.3	83.8	93.9	99.8	37	94.3	80.5	82.6	91.6	99.2
8	86.0	69.9	83.0	93.2	99.4	38	87.9	75.1	84.4	92.7	99.1
9	90.1	75.3	85.5	93.0	99.3	39	89.5	74.6	85.4	91.0	99.3
10	89.2	74.6	83.2	91.2	99.4	40	75.8	63.5	84.5	92.3	99.4
11	64.7	53.3	83.2	91.3	99.3	41	80.1	66.3	84.6	93.3	99.5
12	80.0	66.3	84.9	93.6	99.5	42	81.4	69.2	84.4	92.8	99.2
13	77.7	61.7	83.1	92.8	99.7	43	77.1	63.5	84.0	94.1	99.6
14	64.0	52.8	81.7	93.3	99.6	44	75.3	64.1	84.4	92.4	99.2
15	76.8	63.5	83.6	94.6	99.6	45	85.9	73.7	84.0	92.1	99.2
16	85.0	69.3	82.6	91.1	99.6	46	81.6	68.8	83.3	94.0	99.4
17	64.3	52.8	80.8	92.2	99.4	47	85.2	71.4	84.5	92.7	99.4
18	89.2	75.0	85.7	92.9	99.6	48	86.7	73.0	83.4	93.5	99.7
19	82.6	66.8	83.4	92.4	99.2	49	78.5	59.6	84.9	94.2	99.8
20	76.0	62.6	83.2	92.5	99.2	50	77.0	63.2	83.8	92.2	99.8
21	77.1	63.8	84.9	92.7	99.0	51	82.9	67.9	84.0	91.9	99.4
22	84.3	71.3	83.8	92.4	99.4	52	83.2	68.2	85.0	94.2	99.5
23	88.8	75.5	82.2	91.0	98.8	53	70.2	58.5	85.2	93.1	99.2
24	72.2	60.5	82.3	92.7	98.7	54	83.1	66.5	83.0	92.8	99.5
25	87.2	76.6	84.3	91.0	98.7	55	86.0	71.1	84.3	92.7	99.5
26	80.0	69.1	85.0	91.5	98.7	56	87.8	70.9	82.4	94.4	99.5
27	90.0	76.9	83.5	91.3	98.8	57	89.0	70.9	84.3	94.5	99.6
28	90.7	77.4	82.5	91.1	99.0	58	84.5	67.8	82.8	94.0	99.3
29	75.6	61.5	83.3	93.0	98.9	59	88.7	70.4	83.8	92.7	99.3
30	92.2	79.3	82.8	93.5	99.0	60	85.0	68.9	82.6	92.4	99.5

Based on the above discussion, it can be concluded that the Tennant's method provides the best and the most similar degree of protection of aquatic sources as instream flows under natural flow condition are available. The 7Q10 method clearly results in the lowest instream flows and should probably not be used as instream flow technique for rivers in Newfoundland and Labrador.

6.3 Regionalization of Flow Spells

Instream flow threshold values obtained in section 6.1 for each river can be used to estimate the flow spells, in terms of its duration (days), volume (Mm^3), and intensity (m^3/day). This will yield in a number of flow spells for each year should the instream flow goes below the defined threshold value. As it was discussed earlier, the annual maximum flow spells are the most concerned spells during the year. Thus, the regionalization of flow spells will be based on the annual maximum spell variables.

In terms of regionalization, some hydrologically based instream flow assessment methods can be applied on a regional basis using regression analysis. Regional regression equations can be obtained by linking instream flow thresholds to physiographic characteristics of watersheds such as drainage area, and then linking threshold values to annual maximum flow spells.

6.3.1 Regional Prediction of Threshold Values

Linear relationships between threshold values (m^3/s) of different methodologies and their respective drainage area (km^2) for all the rivers in Newfoundland and Labrador were obtained. Graphs 6-3 and 6-4 present these linear relationships and Table 6-5 and 6-6 provide the prediction equations along with their R-squared values which show a strong relationship between thresholds and drainage areas.

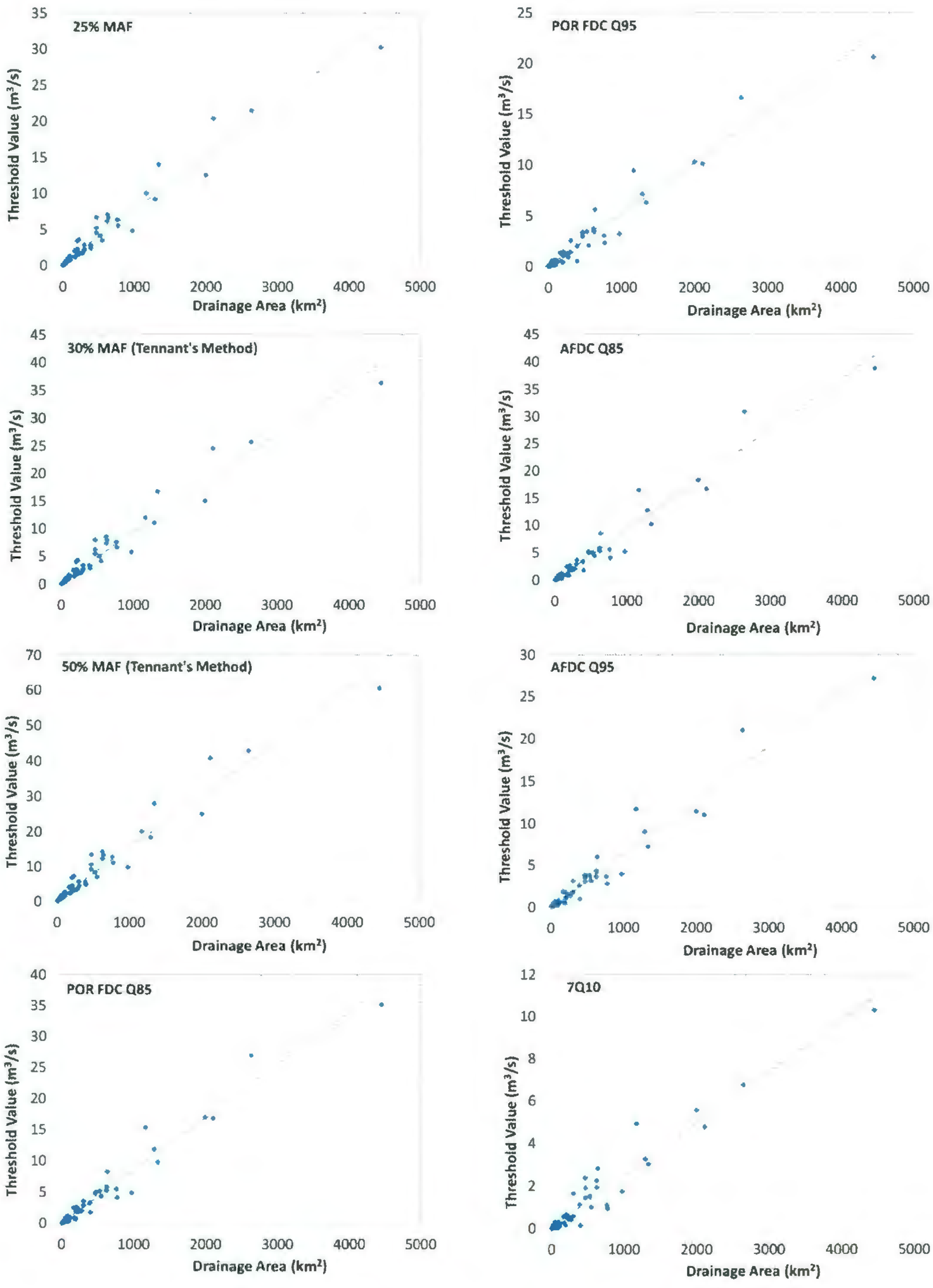

Figure 6-3 Threshold values as a function of drainage areas of Newfoundland stations

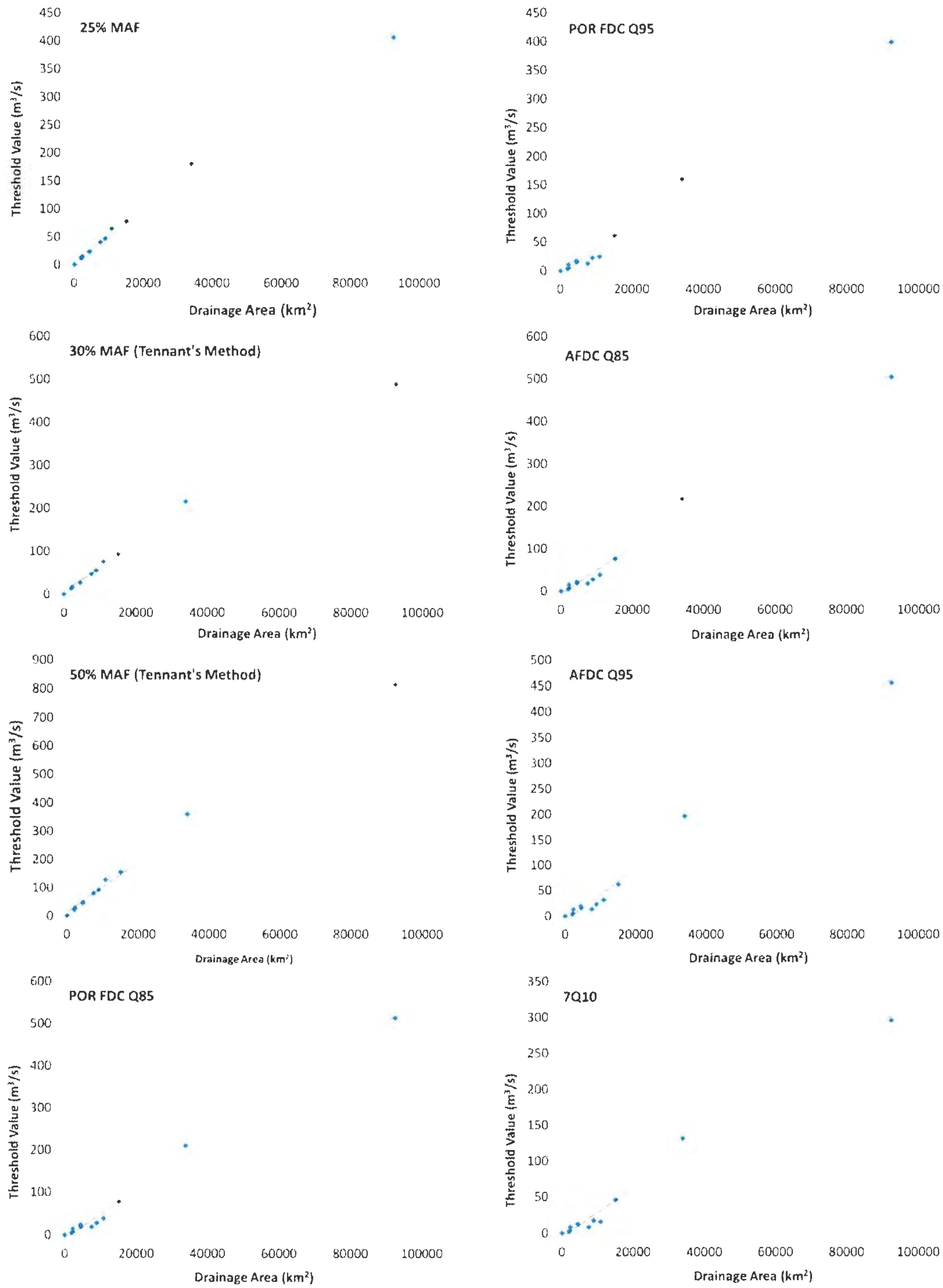

Figure 6-4 Threshold values as a function of drainage areas of Labrador stations

Table 6-5 Relationship between thresholds and drainage areas in Newfoundland

Threshold Method	Equation		R^2
25% MAF	$Q_{Threshold} = 0.0074\mathrm{DA} + 0.4089$	(6-1)	0.954
30% MAF	$Q_{Threshold} = 0.0089\mathrm{DA} + 0.4907$	(6-2)	0.954
50% MAF	$Q_{Threshold} = 0.0148\mathrm{DA} + 0.8178$	(6-3)	0.954
FDC Q85	$Q_{Threshold} = 0.0084\mathrm{DA} + 0.0639$	(6-4)	0.961
FDC Q95	$Q_{Threshold} = 0.0051\mathrm{DA} + 0.0583$	(6-5)	0.948
AFDC Q85	$Q_{Threshold} = 0.0092\mathrm{DA} - 0.053$	(6-6)	0.955
AFDC Q95	$Q_{Threshold} = 0.0064\mathrm{DA} - 0.0502$	(6-7)	0.951
7Q10	$Q_{Threshold} = 0.0024\mathrm{DA} + 0.0428$	(6-8)	0.936

Table 6-6 Relationship between thresholds and drainage areas in Labrador

Threshold Method	Equation		R^2
25% MAF	$Q_{Threshold} = 0.0044\mathrm{DA} + 6.6091$	(6-9)	0.994
30% MAF	$Q_{Threshold} = 0.0053\mathrm{DA} + 7.9309$	(6-10)	0.994
50% MAF	$Q_{Threshold} = 0.0089\mathrm{DA} + 13.218$	(6-11)	0.994
FDC Q85	$Q_{Threshold} = 0.0057\mathrm{DA} - 7.8602$	(6-12)	0.993
FDC Q95	$Q_{Threshold} = 0.0044\mathrm{DA} - 7.1835$	(6-13)	0.993
AFDC Q85	$Q_{Threshold} = 0.0056\mathrm{DA} - 7.0830$	(6-14)	0.991
AFDC Q95	$Q_{Threshold} = 0.0051\mathrm{DA} - 8.2746$	(6-15)	0.990
7Q10	$Q_{Threshold} = 0.0033\mathrm{DA} - 4.8185$	(6-16)	0.987

6.3.2 Regional Prediction of Annual Maximum Spell Variables

Next, a linear relationship between threshold values (m^3/s) and the obtained mean of annual maximum flow spell variables for all the rivers in Newfoundland and Labrador was obtained. Graphs 6-5 to 6-8 present these linear relationships and Tables 6-7 to 6-10 provide the prediction equations along with their R-squared values which show a satisfactory relationship (high R^2 values) exists between thresholds and mean of annual maximum volume and intensity.

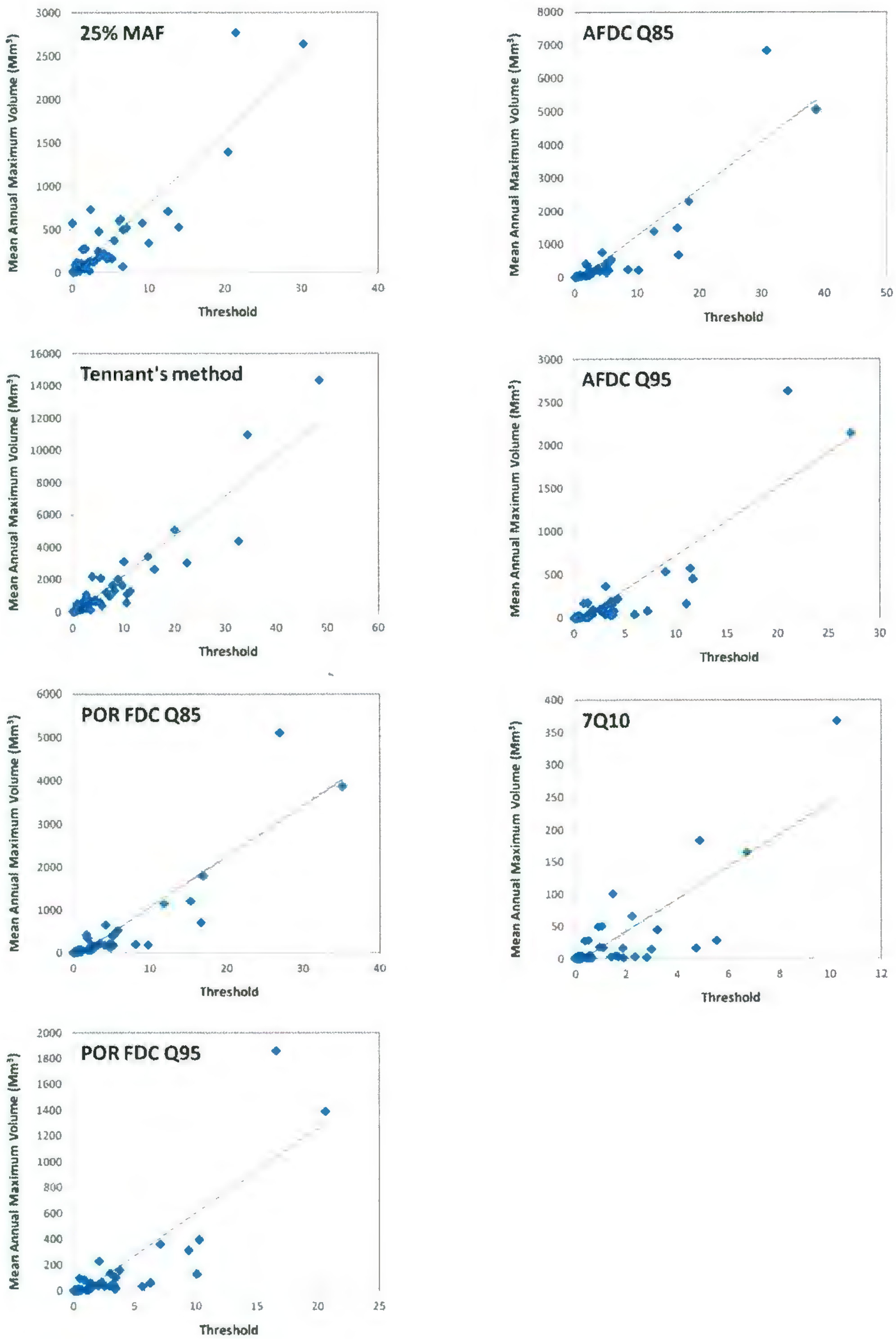

Figure 6-5 Relationship between threshold value and mean of annual maximum volumes for Newfoundland

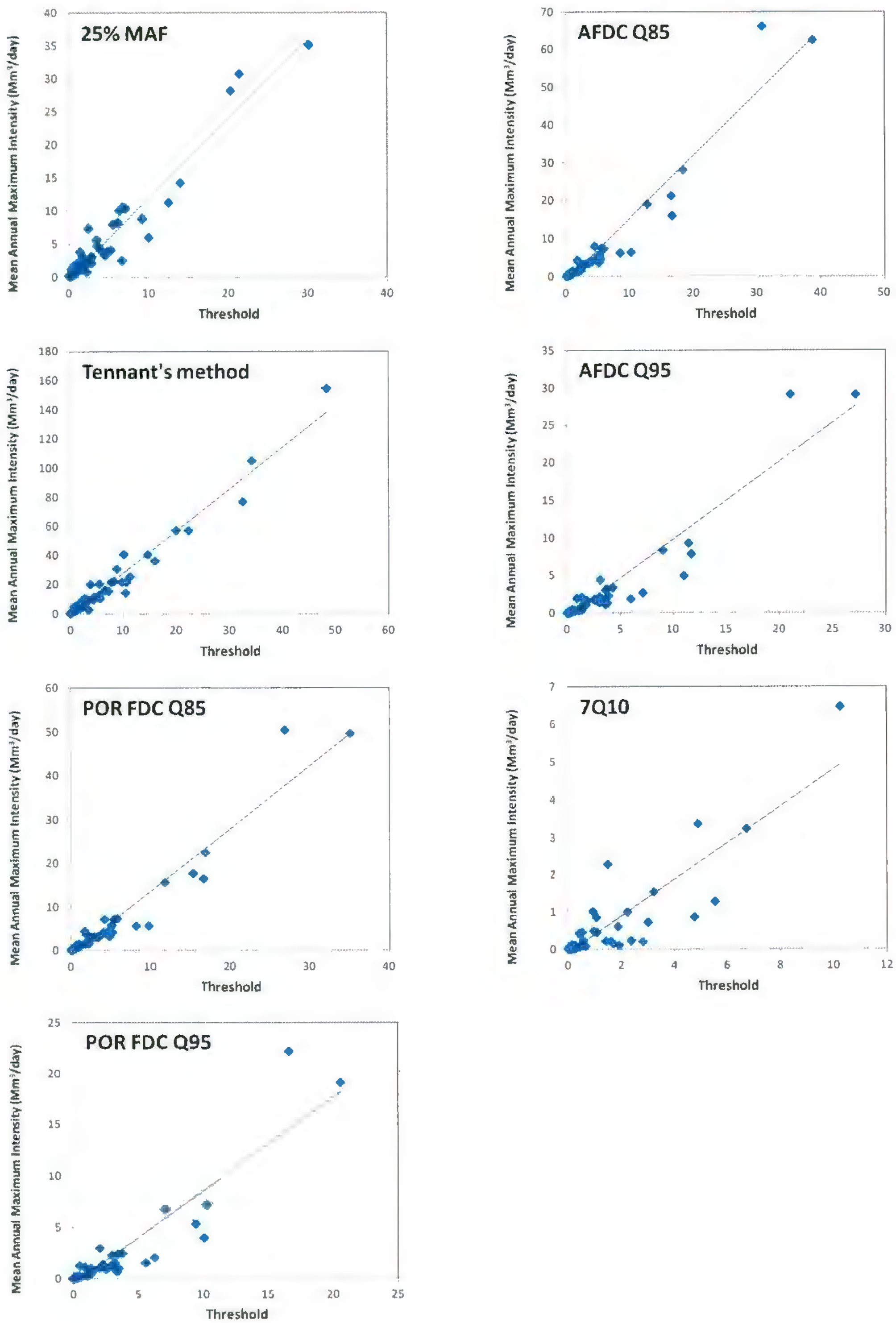

Figure 6-6 Relationship between threshold value and mean of annual maximum intensity for Newfoundland

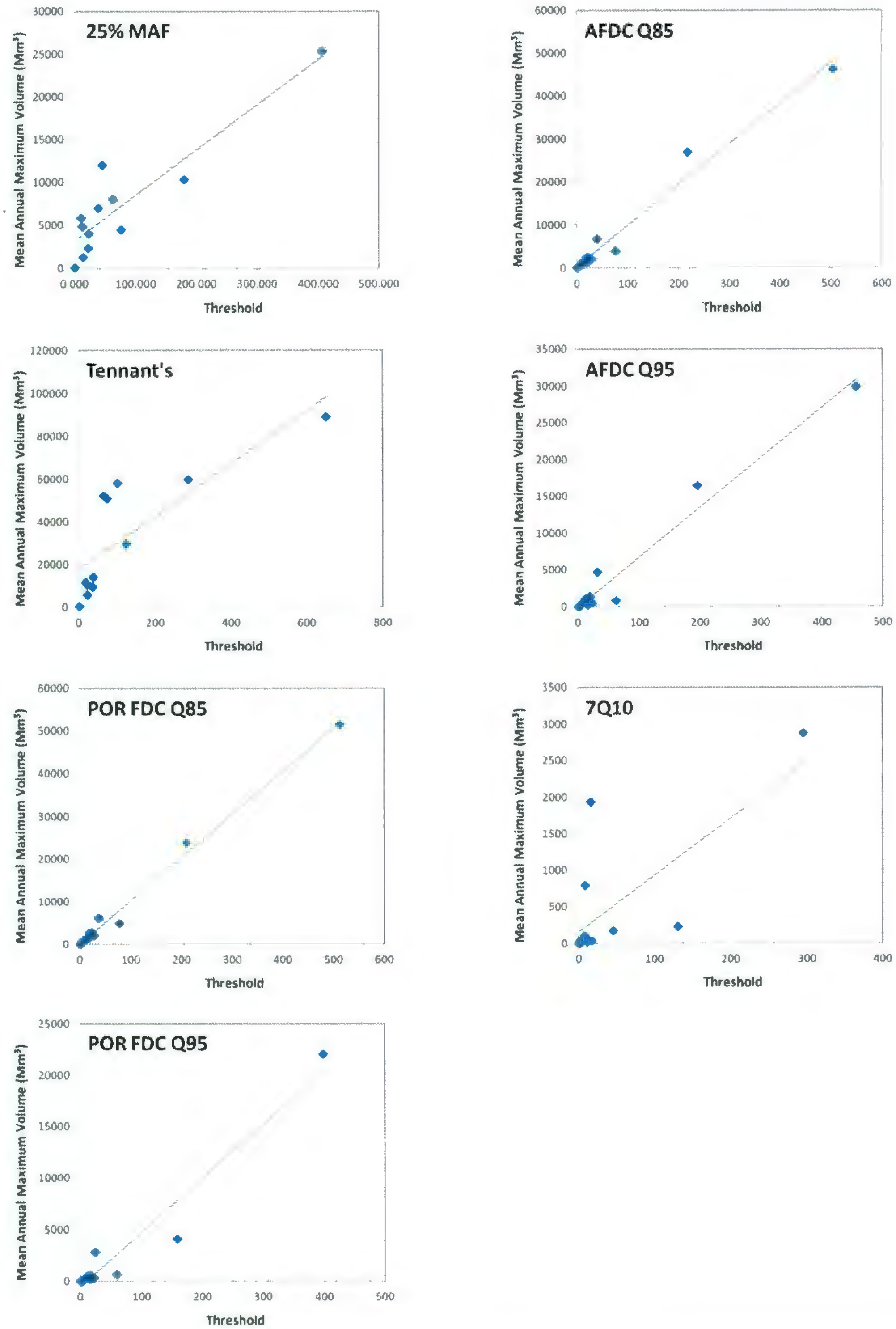

Figure 6-7 Relationship between threshold value and mean of annual maximum volume for Labrador

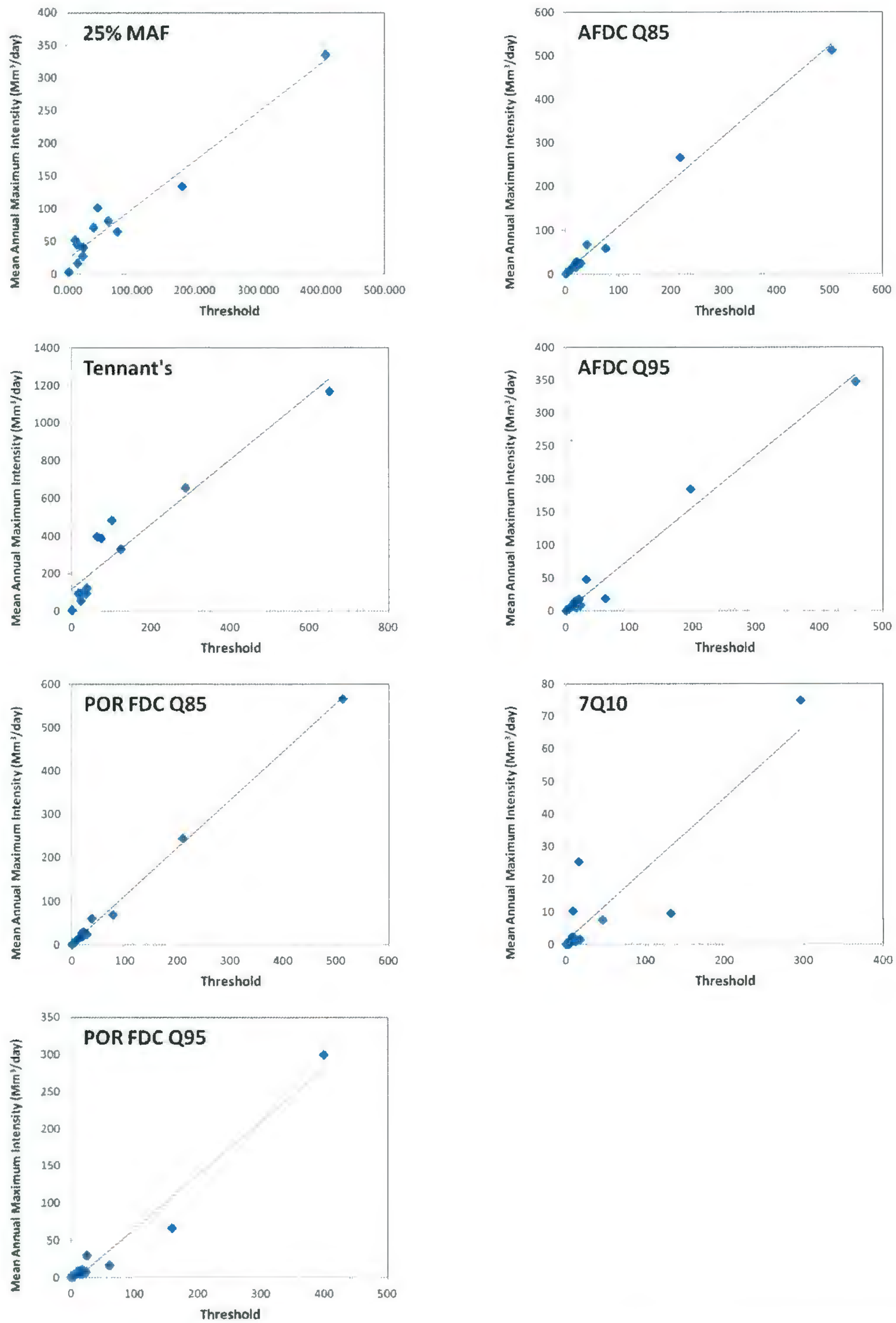

Figure 6-8 Relationship between threshold value and mean of annual maximum intensity for Labrador

Table 6-7 Relationship between thresholds and mean annual maximum volume in Newfoundland

Threshold Method	Equation		R^2
25% MAF	$V_{max} = 81.91Q_{\text{Threshold}} - 21.49$	(6-17)	0.809
Tennant's	$V_{max} = 246.98Q_{\text{Threshold}} - 223.35$	(6-18)	0.866
FDC Q85	$V_{max} = 118.15Q_{\text{Threshold}} - 121.94$	(6-19)	0.814
FDC Q95	$V_{max} = 65.528Q_{\text{Threshold}} - 52.549$	(6-20)	0.739
AFDC Q85	$V_{max} = 142.8Q_{\text{Threshold}} - 173.69$	(6-21)	0.824
AFDC Q95	$V_{max} = 79.288Q_{\text{Threshold}} - 72.56$	(6-22)	0.792
7Q10	$V_{max} = 24.967Q_{\text{Threshold}} - 7.42$	(6-23)	0.680

Table 6-8 Relationship between thresholds and mean annual maximum volume in Labrador

Threshold Method	Equation		R^2
25% MAF	$V_{max} = 52.95Q_{\text{Threshold}} + 31878.9$	(6-24)	0.827
Tennant's	$V_{max} = 124.19Q_{\text{Threshold}} + 17603$	(6-25)	0.650
FDC Q85	$V_{max} = 101.33Q_{\text{Threshold}} + 34.278$	(6-26)	0.991
FDC Q95	$V_{max} = 52.399Q_{\text{Threshold}} - 52.549$	(6-27)	0.940
AFDC Q85	$V_{max} = 95.027Q_{\text{Threshold}} + 327.31$	(6-28)	0.972
AFDC Q95	$V_{max} = 67.868Q_{\text{Threshold}} - 73.921$	(6-29)	0.966
7Q10	$V_{max} = 7.7814Q_{\text{Threshold}} + 160.07$	(6-30)	0.524

Table 6-9 Relationship between thresholds and mean annual maximum intensity in Newfoundland

Threshold Method	Equation		R^2
25% MAF	$I_{max} = 1.195Q_{\text{Threshold}} - 0.0236$	(6-31)	0.931
Tennant's	$I_{max} = 2.873Q_{\text{Threshold}} - 0.7739$	(6-32)	0.963
FDC Q85	$I_{max} = 1.432Q_{\text{Threshold}} - 0.838$	(6-33)	0.934
FDC Q95	$I_{max} = 0.907Q_{\text{Threshold}} - 0.489$	(6-34)	0.853
AFDC Q85	$I_{max} = 1.653Q_{\text{Threshold}} - 1.243$	(6-35)	0.930
AFDC Q95	$I_{max} = 1.037Q_{\text{Threshold}} - 0.648$	(6-36)	0.896
7Q10	$I_{max} = 0.489Q_{\text{Threshold}} - 0.092$	(6-37)	0.779

Table 6-10 Relationship between thresholds and mean annual maximum intensity in Labrador

Threshold Method	Equation		R^2
25% MAF	$I_{max} = 0.743Q_{\text{Threshold}} + 24.565$	(6-38)	0.946
Tennant's	$I_{max} = 1.716Q_{\text{Threshold}} + 117.07$	(6-39)	0.893
FDC Q85	$I_{max} = 1.107Q_{\text{Threshold}} + 0.871$	(6-40)	0.997
FDC Q95	$I_{max} = 0.718Q_{\text{Threshold}} - 6.629$	(6-41)	0.961
AFDC Q85	$I_{max} = 1.037Q_{\text{Threshold}} + 3560$	(6-42)	0.989
AFDC Q95	$I_{max} = 0.785Q_{\text{Threshold}} - 0.132$	(6-43)	0.978
7Q10	$I_{max} = 0.220Q_{\text{Threshold}} + 1.06$	(6-44)	0.796

By mean of these two steps regression set one can have an estimate of both the instream flow threshold value and mean of annual maximum flow spells. In addition, it is more useful to estimate the mean of annual maximum flow spell based on the instream flow threshold value than the size of drainage area.

It should be noted that no direct relationship existed to predict the annual maximum durations for rivers in Newfoundland and Labrador. However, this could be estimated by dividing the estimated values for annual maximum volume and intensity. The results obtained from Labrador prediction models should be used with caution as they were obtained based on only 12 data points.

The worst prediction equation among different instream flow methodologies belongs to the 7Q10 method for all the prediction sets. In summary, there is a stronger relationship between threshold values and mean of annual maximum intensity than mean of annual maximum volume for both Newfoundland and Labrador.

6.3.3 Regional Prediction of Probability Distribution

In an attempt to extrapolate the results, probability distributions were fitted to annual maximum flow spells for each hydrometric station. It was tried to regress the parameters of the fitted probability distribution for each hydrometric station against the physiographic parameters of the dependent watershed. However, no relationship was found to describe the parameters of the governing probability distributions using the known physiographic parameters of watershed. Therefore, regionalization of the flow spell variables is not possible using this methodology.

7 Summary of the Results

7.1 General

Low flow studies require hydrologists to estimate the magnitude, frequency, duration, and spells of low flow events as different aspects of low flow analysis. Therefore the three main objective of this study were the analyses of low flow frequency, flow durations and flow spells in a regional scale for the rivers in Newfoundland and Labrador.

Flow frequency analysis is traditionally based on fitting a probability distribution to the available data at a specific site of interest. In this study a regional approach for conducting low flow frequency analysis based on L-moment theory has been used for the rivers in Newfoundland and Labrador.

To find out what percentage of time within a year flow in a river will be below a certain amount, flow duration analysis using flow duration curves was conducted. Flow-duration curves simply provide the relationship between a given streamflow and the percentage of time it is exceeded. Flow-duration curves, in comparison to low flow frequency analysis, are derived from all the historic data available for a stream rather than just an annual low flow value. A regional approach was developed to regress the flow quantiles of flow duration curves to the physiographic parameters of their corresponding watershed.

In the current study several alternative environmental instream flow requirements which are deemed as a minimum to maintain the river ecosystem were estimated for rivers in Newfoundland and Labrador. Flow spell analysis was conducted based on a regional regression approach for rivers in the study area to have an estimate of how long streamflow will be below a certain amount (instream requirement), and how large the deficit volume is.

The next section provides a summary of the results and conclusion obtained from the current study.

7.2 Conclusions: Regional Low Flow Frequency Analysis

1) The method of L-moments allows one to objectively test the homogeneity of the regions under study. The discordancy measures based on L-moment ratios of the observed sample data screen out the data and facilitate the homogeneity test by taking out the discordant sites in the region. No discordant sites were found in the regions of Newfoundland and Labrador. The homogeneity test resulted in two acceptably homogeneous regions of Newfoundland and Labrador using both 1-day and 7-day minimum annual flows.
2) The conventional goodness-of-fit test indicated that the three-parameter lognormal distribution has the best fit among other frequency distributions for both homogeneous regions of Newfoundland and Labrador, 1-day and 7-day minimum annual flow.
3) The regional estimation using the index flow method based on L-moments produced reliable results using three-parameter lognormal distribution.
4) An index flow was estimation at the ungauged (or gauged with short records) locations in Newfoundland and Labrador were obtained using the drainage areas of watersheds.

5) The performance of the regional models for Newfoundland and Labrador, both 1-day and 7-day minimum annual flows were analyzed, using a new subset of the data with short records. The results were promising. Therefore, these regional models can be used for future predictions of low flows with different return periods for any location in Newfoundland and Labrador.

7.3 Conclusions: Regional Flow Duration Analysis

1) The period of record (POR) and annual based flow duration curves (FDC) were constructed for the rivers under study for Newfoundland and Labrador region. Sixteen flow quantiles of interest, representing the high, median and low flows of flow duration curves were determined for POR and annual FDCs.
2) The physiographic parameters of watersheds under study were extracted from the maps. The possible significant site characteristics for river basins in Newfoundland and Labrador include: drainage area, fraction of lake area, fraction of forest area, fraction of swamp area, fraction of barren area, fraction of lake and swamp area, fraction of area controlled by lakes and swamps, lake and swamp factor, lake attenuation factor, length of main channel, elevation difference of main channel, slope of main channel, drainage density, and shape factor.
3) Quantiles of flow duration curves of the rivers were regressed against their corresponding physiographic parameters. The most important predictor variables of different flow quantiles were drainage area and drainage density of watersheds. All the regression equations were statistically significant and had a good fit.

4) Using the sets of prediction equations for quantiles of the flow duration curve, one can construct the complete flow duration curve for any ungauged location by having an estimate of the physiographic parameters of its watershed area.

5) The performance of the regional flow duration prediction models for both period of record and annual based FDCs were examined by using a new sets of data with short record. The results showed acceptable agreement between observed and predicted flow duration. Therefore, these regional models can be used for future prediction of flow duration curve for rivers in Newfoundland and Labrador.

7.4 Conclusions: Regional Flow Spell Analysis

1) Several methods were selected to represent the environmental instream flow requirements as threshold values for daily streamflow sequence of the rivers in Newfoundland and Labrador. The methods used for estimating the threshold flow values were: the Tennant's method; 25% MAF; 85^{th} and 95^{th} percentiles of period of record and annual FDCs; and 7Q10.

2) Flow spells were determined in terms of their duration, volume and intensity for all the different thresholds for each river in Newfoundland and Labrador. Among all the spells the annual maximum flow spell is the most concerned spells for each river.

3) In a regional approach, the threshold values obtained based on the different methodologies for rivers in the study area were regressed against their corresponding drainage area. Strong relationships were found for both Newfoundland and Labrador rivers which can be used for predicting the minimum environmental instream flow required for any ungauged location in these areas.

4) To regionalize the flow spells, the mean of annual maximum flow spell variables (volume and intensity) were regressed against the threshold values which these spells were constructed based on, for rivers in Newfoundland and Labrador. The relationships were satisfactory and one can implement them for future predictions.

7.5 Recommendations

1) For the regional estimation of the T-year return period of low flows (1-day and 7-day) at the gauged or ungauged locations in Newfoundland and Labrador, the provided equations 4-6 to 4-9 in Section 4.5 are recommended.
2) The equations provided in Table 5-2 and 5-3 are recommended to use for estimating the flow quantiles and creating the regional period of record flow duration curves and annual based flow duration curve respectively.
3) The equations provided in Table 6-5 and 6-6 are recommended for the purpose of estimating the instream flow requirements based on different methodologies for rivers in Newfoundland and Labrador. In addition, Table 6-7 to 6-10 provide the regional equations to estimate the mean annual maximum flow spell volume and intensity in Newfoundland and Labrador.

Bibliography

Acreman, M. C., & Sinclair, C. D. (1986). Classification of drainage basins according to their physical characteristics; an application for flood frequency analysis in Scotland. *Journal of Hydrology, 84*(3–4), 365-380.

Alaouze, C. M. (1989). Reservoir releases to uses with different reliability requirements. *JAWRA Journal of the American Water Resources Association, 25*(6), 1163-1168.

Alaouze, C. M. (1991). Transferable water entitlements which satisfy heterogeneous risk preferences. *Australian Journal of Agricultural and Resource Economics, 35*(2), 197-208.

Archfield, S.A., Vogel, R. M., & Brandt, S. B. (2007). Estimation of flow-duration curves at ungauged sites in southern New England. Proceedings of the American Society of Civil Engineers World Water and Environmental Resources Congress, Tampa, FL, 2007.

Armentrout, G.W., Wilson, J.F. (1987). *Assessment of low flows in streams in northeastern Wyoming*. USGS Water Resources Investigations Report, 85-4246, 30 pp.

Barnett, V, & Lewis, T. (1994). *Outliers in Statistical data.* 3rd edition, Willey, Chichester, Toronto, 584 pp.

Beable, M. E., McKerchar, A. I. (1982). *Regional flood estimation in New Zealand,* Published for the National Water and Soil Conservation Organization by the Water and Soil Division, Ministry of Works and Development Wellington, New Zealand.

Bovee, K.D. (1974). The determination, assessment, and design of instream value studies for the Northern Great Plains Region. Northern Great Plains Resource Program.

Burn, D. (1989). Cluster analysis as applied to regional flood frequency. *Journal of Water Resources Planning and Management, 115*(5), 567-582.

Burn, D. H., & DeWit, W. J. (1996). Spatial characterization of drought events using synthetic hydrology. *Canadian Journal of Civil Engineering, 23*(6), 1231-1240.

Caissie, D., & El-Jabi, N. (1995). Comparison and regionalization of hydrologically based instream flow techniques in Atlantic Canada. *Canadian Journal of Civil Engineering, 22*(2), 235-246.

Castellarin, A., Burn, D. H., & Brath, A. (2001). Assessing the effectiveness of hydrological similarity measures for flood frequency analysis. *Journal of Hydrology, 241*(3–4), 270-285.

Center for Ecology & Hydrology (1999). *Flood estimation handbook: 1-5*. London, U.K.: Natural Environment Research Council.

Chang, T. J., & Stenson, J. R. (1990). Is it realistic to define a 100-year drought for water management? *Water Resources Bulletin, 26*(5), 823-829.

Chen, Y. D., Huang, G., Shao, Q., & Xu, C. -. (2006). Regional analysis of low flow using L-moments for dongjiang basin, south china. *Hydrological Sciences Journal, 51*(6), 1051-1064.

Chiang, S.L., & Johnson, F.W. (1976). Low flow criteria for diversions and impoundments. *Journal of the Water Resources Planning and Management Division, 102*: 227-238.

Chowdhury, J. U., Stedinger, J. R., & Lu, L. (1991). Goodness-of-fit tests for regional generalized extreme value flood distributions. *Water Resources Research, 27*(7), 1765-1776.

Clausen, B., & Pearson, C. P. (1995). Regional frequency analysis of annual maximum streamflow drought. *Journal of Hydrology, 173*(1–4), 111-130.

Collings, M.R. (1972), Methodology for determining instream flow requirements for fish, Proceedings of Instream flow methodology workshop, Washington State Department of Ecology, Olympia, Washington, 1972.

Cunnane, C. (1978). Unbiased plotting positions — A review. *Journal of Hydrology, 37*(3-4), 205-222.

Dalrymple, T. (1960). *Flood-frequency analyses*. Washington, D.C., U.S. G.P.O. Geological Survey (U.S.).

De Coursey, D. G. (1973). Objective regionalization of peak flow rates in floods and droughts, Proceedings of the Second International Symposium in Hydrology, Fort Collins, Colorado, pp. 395-405. Water Resources Publications, Fort Collins, Colo.

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

Duan, N. (1983). Smearing estimate: A nonparametric retransformation method. *Journal of the American Statistical Association, 78*(383), 605-610

El-Jabi, N., Ashkar, F., & Issa, M. (1997). Application stochastique du phénomène d'étiage. *Canadian Journal of Civil Engineering, 24*(2), 191-200.

Environment Canada, (1993) Consolidated Frequency Analysis, version 3.1.

Fennessey, N., & Vogel, R. (1990). Regional Flow - Duration curves for ungauged sites in Massachusetts. *Journal of Water Resources Planning and Management, 116*(4), 530-549.

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

Gingras, D., Adamowski, K., & Pilon, P. J. (1994). Regional flood equations for the provinces of Ontario and Quebec. *JAWRA Journal of the American Water Resources Association, 30*(1), 55-67.

Gordon, N. D., McMahon, T. A., Finlayson, B. L., Gippel, C. J., & Nathan, R. J. (2004). *Stream hydrology-an introduction for ecologists (2nd edition)* John Wiley & Sons.

Government of Newfoundland and Labrador (1991), Characteristics and estimation of minimum streamflows for the Island of Newfoundland., Government of Newfoundland and Labrador, Department of Environment and Conservation, Water Resources Division.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., & Wallis, J. R. (1979). Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research, 15*(5), 1049-1054.

Guttman, N. B. (1993). The use of L-moments in the determination of regional precipitation climates. *Journal of Climate, 6*(12), 2309-2325.

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological), 52*(1), 105-124.

Hosking, J. R. M., & Wallis, J. R. (1987). Parameter and quantile estimation for the generalized pareto distribution. *Technometrics, 29*(3), 339-349.

Hosking, J. R. M., & Wallis, J. R. (1993). Some statistics useful in regional frequency analysis. *Water Resources Research, 29*(2), 21-281.

Hosking, J. R. M., & Wallis, J. R. (1997). *Regional frequency analysis : An approach based on L-moments*. Cambridge; New York: Cambridge University Press.

Hughes, D.A., O'Keeffe, J.H., Smakhtin, V.Y., & King, J. (1997). Development of an operating rule model to simulate time series of reservoir releases for instream flow requirements. *Water SA, 23(1)*, 21-30.

Institute of Hydrology (1980), Low flow estimation in studies (1-4), Wallingford, UK.

Jowett, I.G. (1997). Instream flow methods: A comparison of approaches, *Regulated Rivers: Research and Management*, *13*, 115–127.

Kachroo, R. K., Mkhandi, S. H., & Parida, B. P. (2000). Flood frequency analysis of southern Africa: I. delineation of homogeneous regions. *Hydrological Sciences Journal, 45*(3), 437-447.

Karim, K., Gubbels, M. E., & Goulter, I. C. (1995). Review of determination of instream flow requirements with special application to Australia. *JAWRA Journal of the American Water Resources Association, 31*(6), 1063-1077.

Kendal, M. G., & Gibbons, J.D. (1990), *Rank Correlation methods* 5th edition, Oxford University Press, New York, NY, USA, 260 pp.

Kroll, C. N., & Vogel, R. M. (2002). Probability distribution of low streamflow series in the United States. *Journal of Hydrologic Engineering, 7*, 137-146.

Landwehr, J. M., Matalas, N. C., & Wallis, J. R. (1979a). Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research*, *15*, 1055-64.

Landwehr, J. M., Matalas, N. C., & Wallis, J. R.(1979b). Estimation of parameters and quantiles of Wakeby distributions. *Water Resources Research*, *15*, 1361-79.

Lane, E.W., Lei, K. (1949). Stream flow variability, Proceedings of the American Society of Civil Engineers Conference 75, 935-994.

Lim, Y. H., & Lye, L. M. (2003). Regional flood estimation for ungauged basins in Sarawak, Malaysia. *Hydrological Sciences Journal, 48*, 79-94.

Loaiciga, H., & Marino, M. (1988). Fitting minima of flows via maximum likelihood. *Journal of Water Resources Planning and Management, 114*(1), 78-90.

Mallory, S. J. L., McKenzie, R. S. (1993). Water resources modeling of flow diversions, Proceedings of the Sixth South African National Hydrology Symposium, Pietermaritzburg, South Africa, vol. 1, pp 429-436.

Matalas, N. C., & Wallis, J. R. (1973). Eureka! it fits a Pearson type: 3 distribution. *Water Resources Research, 9(2)*, 281-289.

Matalas, N. C., Slack, J. R., & Wallis, J. R. (1975). Regional skew in search of a parent. *Water Resources Research, 11*(6), 815-826.

McCuen, R. H. (1985). *Statistical methods for engineers*. Prentice-Hall, Englewood Cliffs, N. J.

McMahon, T.A. (1976). Low flow analyses of streams: details of computational procedure and annotated bibliography, Monash University, Department of Civil Engineering, Clayton (Australia), Research Report No. 5/1976, 60 pp.

McMahon, T. A., & Mein, R. G. (1986). *River and reservoir yield.* Water Resources Publication. Littleton, Colo., U.S.A.

Mimikou, M., & Kaemaki, S. (1985). Regionalization of flow duration characteristics. *Journal of Hydrology, 82*(1–2), 77-91.

Modarres, R. (2008). Regional frequency distribution type of low flow in north of Iran by L-moments. *Water Resources Management, 22*(7), 823-841

Modarres, R. (2010). Regional dry spells frequency analysis by L-moment and multivariate analysis. *Water Resources Management, 24*(10), 2365-2380.

Modarres, R., & Sarhadi, A. (2010). Frequency distribution of extreme hydrologic drought of southeastern semiarid region, Iran. *Journal of Hydrologic Engineering, 15*(4), 255-264.

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

Mostofi Zadeh, S., Lye, L. M., & Khan, A. A. (2012). Regional frequency analysis of low flow using L-moments for Labrador, Canada, Proceeding of 125th CSCE annual conference, Edmonton, Canada.

Nash, J., & Sutcliffe, J. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology, 10*(3), 282-290.

Nathan, R. J., & McMahon, T. A. (1990). Evaluation of automated techniques for base flow and recession analyses. *Water Resources Research, 26*(7), 1465-1473.

Niadas, I. A. (2005). Regional flow duration curve estimation in small ungauged catchments using instantaneous flow measurements and a censored data approach. *Journal of Hydrology, 314*(1–4), 48-66.

Pearson, C. P. (1991a). New Zealand regional flood frequency analysis using L-moments. *Journal of Hydrology New Zealand, 30*, pp53-64.

Pearson, C. P. (1991b). Regional flood frequency for small New Zealand basins, 2: flood frequency group. *Journal of Hydrology New Zealand, 30,* 77-92.

Pearson C. P. (1995). Regional frequency analysis of low flows in New Zealand, *Journal of Hydrology New Zealand, 33(2),* 94-122.

Pitman, W.V. (1993). Simulation of run-of-river schemes using monthly data, Proceedings of the Sixth South African National Hydrology Symposium, Pietermaritzburg, South Africa, vol. 1, 445-452.

Prudhomme, C., Gilles, G. (1997). The typology of regional QdF models. Acta Hydrotechnica 15/19 — Proceedings of Poster Presentations, FRIEND'97,. 75–78.

Quimpo, R., Alejandrino, A., & McNally, T. (1983). Regionalized flow duration for Philippines. *Journal of Water Resources Planning and Management, 109*(4), 320-330.

Richter, S. H., & Lye, L. M. (1995). Relationships of flow and basin variables on the Island of Newfoundland, Canada. Proceeding of 48th CWRA Annual Conference, Fredericton, New Brunswick, Canada, 1995.

Robson, A., & Reed, D.,. (1999). Flood estimation handbook, volume 3 : Statistical procedures for flood frequency estimation. Wallingford: Institute of Hydrology.

Rocky Durrans, S., & Tomic, S. (1996). Regionalization of low-flow frequency estimates: An Alabama case study. *Water Resources Bulletin-American Water recourses Association, 32*(1), 23-38.

Searcy, J. C. (1959). Flow duration curves, United States Geological Survey, Washington DC, Water Supply Paper 1542A.

Shaughnessy, B. (1997). Duration, volumes and intensities of flow spells from hydrological methods of estimating environmental instream flows in Newfoundland, Canada. Master of Science thesis, University of Glasgow, Glasgow, United Kingdom.

Shi P., Chen X., Qu S.-M., Zhang Z.-C., & Ma J.-L. (2010). Regional frequency analysis of low flow based on L moments: Case study in Karst area, southwest China. *Journal of Hydrologic Engineering, 15*(5), 370-377.

Singh, K. P. (1971). Model flow duration and streamflow variability. *Water Resources Research, 7*(4), 1031-1036.

Singh, R., Mishra, S., & Chowdhary, H. (2001). Regional flow-duration models for large number of ungauged Himalayan catchments for planning microhydro projects. *Journal of Hydrologic Engineering, 6*(4), 310-316.

Smakhtin, V. U. (2001). Low flow hydrology: A review. *Journal of Hydrology, 240*(3–4), 147-186.

Smakhtin, V. Y., Hughes, D. A., & Creuse-Naudin, E. (1997). Regionalization of daily flow characteristics in part of the eastern cape, south Africa. *Hydrological Sciences Journal, 42*(6), 919-936.

Stahl, K., & Demuth, S. (1999). Linking streamflow drought to the occurrence of atmospheric circulation patterns. *Hydrological Sciences Journal, 44*(3), 467-482.

Tallaksen, L. M., Madsen, H., & Clausen, B. (1997). On the definition and modelling of streamflow drought duration and deficit volume. *Hydrological Sciences Journal, 42*(1), 15-33.

Tasker, G. D. (1987). A comparison of methods for estimating low flow characteristics of streams. *JAWRA Journal of the American Water Resources Association, 23*(6), 1077-1083.

Tate, E. L., & Freeman, S. N. (2000). Three modelling approaches for seasonal streamflow droughts in southern Africa: The use of censored data. *Hydrological Sciences Journal, 45*(1), 27-42.

Tate, E. L., Meigh, J. R., Prudhomme, C. P., McCartney, M. P. (2000). *Drought assessment in southern africa using river flow data,* Department for International Development.

Tennant, D. L. (1976). Instream flow regimens for fish, wildlife, recreation and related environmental resources. *Fisheries, 1(4),* 6-10.

Thomas, W. O. (1987). Techniques used by U.S. Geological Survey in estimating the magnitude and frequency of floods., Proceeding of 18th Binghamton Geomorpho. Sympos., Published by Unwin and Hyman, London, pp. 267-288.

Tlalka, A., & Tlalka, J. (1987). Low discharge periods of rivers in the upper Vistula River basin, Poland. *JAWRA Journal of the American Water Resources Association, 23*(2), 227-231.

Tucci, C., Silveria, A., Sanchez, J., & Albuquerque, F. (1995). Flow regionalization in the upper Paraguay basin, Brazil. *Hydrological Sciences Journal, 40*(4), 485-497.

Ullah, W. (1992). *Water resources atlas of Newfoundland.* St. John's, NL., Canada: Water Resources Division, Dept. of Environment and Lands, Government of Newfoundland and Labrador.

Viola, F., Noto, L. V., Cannarozzo, M., & La Loggia, G. (2011). Regional flow duration curves for ungauged sites in sicily. *Hydrology and Earth System Sciences, 15*(1), 323-331.

Vogel, R. M., & Fennessey, N. M. (1993). L moment diagrams should replace product moment diagrams. *Water Resources Research, 29*(6), 1745.

Vogel, R. M., & Fennessey, N. (1994). Flow - Duration curves. I: New interpretation and confidence intervals. *Journal of Water Resources Planning and Management, 120*(4), 485-504.

Vogel, R. M., & Fennessey, N. M. (1995). Flow duration curves II: A review of applications in water resources planning. *JAWRA Journal of the American Water Resources Association, 31*(6), 1029-1039.

Vogel, R. M., & Kroll, C. N. (1992). Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. *Water Resources Research, 28*(9), 2451-2458.

Vogel, R. M., & Wilson, I. (1996). Probability distribution of annual maximum, mean, and minimum streamflows in the United States. *Journal of Hydrologic Engineering, 1*(2), 69-76.

Wallis, J. R., Matalas, N. C., & Slack, J. R. (1977). Just a moment. *Water Resources Research, 10*, 211-219.

Wijayaratne, L. H., & Golub, E. (1991). MULTIYEAR DROUGHT SIMULATION1. *JAWRA Journal of the American Water Resources Association, 27*(3), 387-395.

Wilcock, D.N., & Hanna, J.E. (1987). Derivation of flow duration curves in Northern Ireland, Proceeding of Institution of Civil Engineers, Vol. 83(2), 381-396.

Wiltshire, S. E. (1985). Grouping basins for regional flood frequency analysis. *Hydrological Sciences Journal, 30*(1), 151-159.

Wiltshire, S. E. (1986a). Regional flood frequency analysis I: Homogeneity statistics. *Hydrological Sciences Journal, 31*(3), 321-333.

Wiltshire, S. E. (1986b). Regional flood frequency analysis II: Multivariate classification of drainage basins in Britain. *Hydrological Sciences Journal, 31*(3), 335-346.

Wiltshire, S. E. (1986c). Identification of homogeneous regions for flood frequency analysis. *Journal of Hydrology, 84*(3–4), 287-302.

Woo, M., & Tariiule, A. (1994). Streamflow droughts of northern Nigerian rivers. *Hydrological Sciences Journal, 39*(1), 19-34.

Yevjevich, V. (1967). An objective approach to definitions and investigations of continental hydrologic droughts, Colorado State University, Fort Collins, 18 pp.

Yevjevich, V., (1972). *Probability and statistics in Hydrology*, Water Resources Publications, Fort Collins, CO, USA, 302 pp.

Appendix

A1-Discordancy measure

```
clear all;
close all;
clc;


format long

%open excel file with L-moment ratios (L-CV, L-Sk, L-ku) in it (it
should be in
%the same folder as this M-file-sheet1 of this excel file contains the
%values

[type, sheets] = xlsfinfo('Discordancy.xlsx');
U = xlsread('Discordancy.xlsx', 'Sheet1');

iter=input(' Enter number of sites in the region: ');

%in this loop each row of the big U matrix will be divided into separate
%matrices, and then they would transposed to vertical matrices.
for i=1:iter,

    eval(['U_' num2str(i) '=U(' num2str(i) ',:);'])
    eval(['U_' num2str(i) '=transpose(U_' num2str(i) ');'])

 end

%this loop will add all this separate U(i) matrices together, and yield
sum
%to calculate the average of them. U_mean

sum1=0;
for i=1:iter

  sum1=sum1+eval(['U_' num2str(i)]);

end
U_mean=sum1/iter


%this loop calculates the Ai variable
%A_i=(U_i-U_mean)*transpose(U_i-U_mean)
for i=1:iter

    eval(['A_' num2str(i) '=(U_' num2str(i) '-U_mean)*transpose(U_'
num2str(i) '-U_mean);']);
```

```
end


%this loop calculates the sum of all A_i, and put it in matrix A.
sum2=0;
for i=1:iter

    sum2=sum2+ eval(['A_' num2str(i)]);
end

A=sum2


%this loop works on calculating D_is
%D_i=N/3*transpose(U_i-U_mean)*inv(A)*(U_i-U_mean)

for i=1:iter

eval(['D_' num2str(i) '=(iter/3)*transpose(U_' num2str(i) '-
U_mean)*inv(A)*(U_' num2str(i) '-U_mean)'])

end


%this loop will combine all D_is together and make one matrix named D
%this matrix will then be put on the sheet2 of the excel file.
%
D=[0];
for i=1:iter

    D=vertcat(D,eval(['D_' num2str(i)]));

end


%writing the result matrix in the sheet2 of Discordancy file
xlswrite('Discordancy.xlsx', D, 'Sheet2', 'B3');
```

A-2 Kappa Distribution

(Translated FORTRAN code provided by Hosking and Wallis, 1997)

```
clc, clear all


%PARAMETER ESTIMATION VIA L-MOMENTS FOR THE KAPPA DISTRIBUTION
%
%  PARAMETERS OF ROUTINE:
%  XMOM    * INPUT* ARRAY OF LENGTH 4. CONTAINS THE L-MOMENTS LAMBDA-
            1,LAMBDA-2,TAU-3, TAU-4.
%  PARA   *OUTPUT* ARRAY OF LENGTH 4. ON EXIT, CONTAINS THE PARAMETERS
IN THE ORDER XI,
%                ALPHA, K, H.
%  IFAIL  *OUTPUT* FAIL FLAG. ON EXIT, IT IS SET AS FOLLOWS.
%          0  SUCCESSFUL EXIT
%          1  L-MOMENTS INVALID
%          2  (TAU-3, TAU-4) LIES ABOVE THE GENERALIZED-LOGISTIC LINE
%             (SUGGESTS THAT L-MOMENTS ARE NOT CONSISTENT WITH ANY KAPPA
%              DISTRIBUTION WITH H.GT.-1)
%          3  ITERATION FAILED TO CONVERGE
%          4  UNABLE TO MAKE PROGRESS FROM CURRENT POINT IN ITERATION
%          5  ITERATION ENCOUNTERED NUMERICAL DIFFICULTIES - OVERFLOW
%             WOULD HAVE BEEN LIKELY TO OCCUR
%          6  ITERATION FOR H AND K CONVERGED, BUT OVERFLOW WOULD HAVE
%             OCCURRED WHEN CALCULATING XI AND ALPHA
%
%  N.B.  PARAMETERS ARE SOMETIMES NOT UNIQUELY DEFINED BY THE FIRST 4
%  L-MOMENTS. IN SUCH CASES THE ROUTINE RETURNS THE SOLUTION FOR WHICH
%  THE H PARAMETER IS LARGEST.
%
%  OTHER ROUTINES USED: DLGAMA,DIGAMD
%
%  THE SHAPE PARAMETERS K AND H ARE ESTIMATED USING NEWTON-RAPHSON
%  ITERATION ON THE RELATIONSHIP BETWEEN (TAU-3,TAU-4) AND (K,H).
%  THE CONVERGENCE CRITERION IS THAT TAU-3 AND TAU-4 CALCULATED FROM
%  THE ESTIMATED VALUES OF K AND H SHOULD DIFFER BY LESS THAN 'EPS'
%  FROM THE VALUES SUPPLIED IN ARRAY XMOM.

format long

zero=0; half=0.5; one=1; two=2; three=3; four=4;
five=5; six=6; twelve=12; twenty=20; thirty=30;
p725=0.725; p8=0.8;

%         EPS,MAXIT CONTROL THE TEST FOR CONVERGENCE OF N-R ITERATION
%         MAXSR IS THE MAX. NO. OF STEPLENGTH REDUCTIONS PER ITERATION
%         HSTART IS THE STARTING VALUE FOR H
%         BIG IS USED TO INITIALIZE THE CRITERION FUNCTION
%         OFLEXP IS SUCH THAT DEXP(OFLEXP) JUST DOES NOT CAUSE OVERFLOW
%         OFLGAM IS SUCH THAT DEXP(DLGAMA(OFLGAM)) JUST DOES NOT CAUSE
```

```
%           OVERFLOW
%
eps=10^(-6); maxit=20; maxsr=10; hstart=1.001; big=10;
oflexp=170; oflgam=53;

%Enter the input array consist of tHE L-moments LAMBDA-1,LAMBDA-2, TAU-
3,TAU-4.
xmom=input('Enter Lambda-1 Lambda-2 Tau-3 Tau-4:');
t3=xmom(3);
t4=xmom(4);

para=[0,0,0,0];
% test for feasibility
    if(xmom(2)<=zero),ifail=1; end;
    if ((abs(t3)>=one)||(abs(t4)>=one)), ifail=1; end;
    if (t4 <=(five*t3*t3-one)/four), ifail=1; end;
    if (t4 >= (five*t3*t3+one)/six), ifail=2; end;

%   set starting values for n-r iteration:
%   g is chosen to give the correct value of tau-3 on the
%   assumption that h=1 (i.e. a generalized pareto fit) -
%   but h is actually set to 1.001 to avoid numerical
%   difficulties which can sometimes arise when h=1 exactly

g=(one-three*t3)/(one+t3);
h=hstart;
z=g+h*p725;
xdist=big;

%START OF NEWTON-RAPHSON ITERATION

for it=1:maxit
    %reduce steplength until we are nearer to the required
    %values of tau-3 and tau-4 than we were at the previous step
    for i=1:maxsr
          % - calculate current tau-3 and tau-4
      % notation:
      % u. ratios of gamma functions which occur in the pwm's beta-sub-r
      %alam. - l-moments (apart from a location and scale shift)
      %tau.  - l-moment ratios

      if (g > oflgam), ifail=5; end;
      if (h > zero)
          u1=exp(dlgama(one/h)-dlgama(one/h+one+g));
          u2=exp(dlgama(two/h)-dlgama(two/h+one+g));
          u3=exp(dlgama(three/h)-dlgama(three/h+one+g));
          u4=exp(dlgama(four/h)-dlgama(four/h+one+g));
      else
          u1=exp(dlgama(-one/h-g)-dlgama(-one/h+one));
          u2=exp(dlgama(-two/h-g)-dlgama(-two/h+one));
          u3=exp(dlgama(-three/h-g)-dlgama(-three/h+one));
          u4=exp(dlgama(-four/h-g)-dlgama(-four/h+one));
```

```
        end
        alam2=u1-two*u2;
        alam3=-u1+six*u2-six*u3;
        alam4=u1-twelve*u2+thirty*u3-twenty*u4;
        if(alam2 == zero),ifail=5; end;
        tau3=alam3/alam2;
        tau4=alam4/alam2;
        e1=tau3-t3;
        e2=tau4-t4;

        % if nearer than before, exit this loop
        dist=max(abs(e1), abs(e2));
        if (dist < xdist)
            if(dist < eps)
                %Converged
                ifail=0;
                para(4)=h;
                para(3)=g;
                temp=dlgama(one+g);
                if(temp > oflexp), ifail=6; end;
                gam=exp(temp);
                temp=(one+g)*log(abs(h));
                if (temp > oflexp), ifail=6; end;
                hh=exp(temp);
                para(2)=xmom(2)*g*hh/(alam2*gam);
                para(1)=xmom(1)-para(2)/g*(one-gam*u1/hh);
                s_1=num2str(ifail);
                s_2=num2str(para(1));
                s_3=num2str(para(2));
                s_4=num2str(para(3));
                s_5=num2str(para(4));
                disp(['ifail=',s_1,'   para(1)=', s_2, '
para(2)=',s_3,...
                    '   para(3)=',s_4, '   para(4)=', s_5]);

            else
               % not converged: calculate next step
               %notation:
               %u1g  - derivative of u1 w.r.t. g
               %dl2g - derivative of alam2 w.r.t. g
               %d..  - matrix of derivatives of tau-3 and tau-4 w.r.t. g
and h
               %h..  - inverse of derivative matrix
               %del. - steplength
               xg=g;
               xh=h;
               xz=z;
               xdist=dist;
               rhh=one/(h*h);
               if(h > zero)
                   u1g=-u1*digamd(one/h+one+g);
                   u2g=-u2*digamd(two/h+one+g);
                   u3g=-u3*digamd(three/h+one+g);
```

```
            u4g=-u4*digamd(four/h+one+g);
            u1h=rhh*(-u1g-u1*digamd(one/h));
            u2h=two*rhh*(-u2g-u2*digamd(two/h));
            u3h=three*rhh*(-u3g-u3*digamd(three/h));
            u4h=four*rhh*(-u4g-u4*digamd(four/h));
        else
            u1g=-u1*digamd(-one/h-g);
            u2g=-u2*digamd(-two/h-g);
            u3g=-u3*digamd(-three/h-g);
            u4g=-u4*digamd(-four/h-g);
            u1h=rhh*(-u1g-u1*digamd(-one/h+one));
            u2h=two*rhh*(-u2g-u2*digamd(-two/h+one));
            u3h=three*rhh*(-u3g-u3*digamd(-three/h+one));
            u4h=four*rhh*(-u4g-u4*digamd(-four/h+one));
        end
        dl2g=u1g-two*u2g;
        dl2h=u1h-two*u2h;
        dl3g=-u1g+six*u2g-six*u3g;
        dl3h=-u1h+six*u2h-six*u3h;
        dl4g=u1g-twelve*u2g+thirty*u3g-twenty*u4g;
        dl4h=u1h-twelve*u2h+thirty*u3h-twenty*u4h;
        d11=(dl3g-tau3*dl2g)/alam2;
        d12=(dl3h-tau3*dl2h)/alam2;
        d21=(dl4g-tau4*dl2g)/alam2;
        d22=(dl4h-tau4*dl2h)/alam2;
        det=d11*d22-d12*d21;
        h11= d22/det;
        h12=-d12/det;
        h21=-d21/det;
        h22= d11/det;
        del1=e1*h11+e2*h12;
        del2=e1*h21+e2*h22;
        %take next n-r step
        g=xg-del1;
        h=xh-del2;
        z=g+h*p725;
        %reduce step if g and h are outside the parameter space
        factor=one;
        if(g <= -one), factor=p8*(xg+one)/del1; end;
        if(h <= -one), factor=min(factor,p8*(xh+one)/del2); end;
        if(z <= -one), factor=min(factor,p8*(xz+one)/(xz-z)); end;
        if((h <= zero)&& (g*h <=-one)),
factor=min(factor,p8*(xg*xh+one)...
                /(xg*xh-g*h)); end;
        if (factor == one)
           %end of newton-raphson iteration
           break
        else
           del1=del1*factor;
           del2=del2*factor;
           g=xg-del1;
           h=xh-del2;
           z=g+h*p725;
        end
```

```
            end
        else
            %otherwise, halve the steplength and try again
            del1=half*del1;
            del2=half*del2;
            g=xg-del1;
            h=xh-del2;
        end
    end

    %too many steplength reductions ifail=4
     %test for convergence
end
```

```
function [ dlgamafn ] = dlgama( x )
% dlgama calculate the logarithm of gamma function
%base on algorithm acm291, ommun. assoc. comput. mach. (1966)


format long


small=10^(-7); crit=13; big=10^9; toobig=2*10^36;

% c0 is 0.5*log(2*pi)
% c1...c7 are the coeffts of the asymptotic expansion of dlgama

c0=0.918938533204672742; c1=0.833333333333333333e-1;
c2=-0.277777777777777778e-2; c3=0.793650793650793651e-3;
c4=-0.595238095238095238e-3; c5=0.841750841750841751e-3;
c6=-0.191752691752691753e-2; c7=0.641025641025641026e-2;

%s1 is -(euler's constant), s2 is pi**2/12

s1=-0.577215664901532861; s2=0.822467033424113218;
zero=0; half=0.5; one=1; two=2;
dlgamafn=zero;

if (x <= zero)
    x_1=num2str(x);
    disp(['*** error*** routine dlgamma',x_1, 'argument out of range']);
end
if (x > toobig)
    x_1=num2str(x);
```

```
    disp(['*** error*** routine dlgamma',x_1, 'argument out of range']);
end

%use small-x approximation if x is near 0, 1 or 2

if(abs(x-two)> small)
    if (abs(x-one)> small)
        if (x > small)
            sum1=zero;
            y=x;
            if (y >= crit)
                sum1=sum1+(y-half)*log(y)-y+c0;
                sum2=zero;
                if(y >= big)
                    dlgamafn=sum1+sum2;
                    return
                else
                   z=one/(y*y);
                   sum2=((((((c7*z+c6)*z+c5)*z+c4)*z+c3)*z+c2)*z+c1)/y;
                   dlgamafn=sum1+sum2;
                   return
                end
            else
                z=one;
                z=z*y;
                y=y+one;
                while (y < crit)
                    z=z*y;
                    y=y+one;
                end
                sum1=sum1-log(z);
                %use asymptotic expansion if y .ge. crit
                sum1=sum1+(y-half)*log(y)-y+c0;
                sum2=zero;
                if (y >= big)
                    dlgamafn=sum1+sum2;
                    return
                else
                    z=one/(y*y);
                    sum2=((((((c7*z+c6)*z+c5)*z+c4)*z+c3)*z+c2)*z+c1)/y;
                    dlgamafn=sum1+sum2;
                    return
                end

            end
        else
            dlgamafn=-log(x)+s1*x;
            return
        end
    else
       xx=x-one;
       dlgamafn=dlgamafn+xx*(s1+xx*s2);
       return
```

```
    end
else
   dlgamafn=log(x-one);
   xx=x-two;
   dlgamafn=dlgamma+xx*(s1+xx*s2);
   return
end
```

```
function [ digamdfn ] = digamd( x )
%digamma function (euler's psi function)- the first derivative
%of log(gamma(x)) based on algorithm as103, appl. statist. (1976)
%vol. 25 no. 3

format long

 zero=0; half=0.5; one=1;
small=10^(-9); crit=13;

% c1...c7 are the coeffts of the asymptotic expansion of digamd
% d1 is -(euler's constant)

c1=0.833333333333333333e-1; c2=-0.833333333333333333e-2;
c3=0.396825396825396825e-2; c4=-0.416666666666666666e-2;
c5=0.757575757575757575e-2; c6=-0.210927960927960928e-1;
c7=0.833333333333333333e-1; d1=-0.577215664901532861e0;

 digamdfn=zero;

if (x <= zero)
    x_1=num2str(x);
    disp(['*** error*** routine dlgamma',x_1, 'argument out of range']);
end
% use small-x approximation if x. le. small
if (x > small)
    y=x;
    if (y >= crit)
        digamdfn=digamdfn+log(y)-half/y;
        y=one/(y*y);
        sum=((((((c7*y+c6)*y+c5)*y+c4)*y+c3)*y+c2)*y+c1)*y;
        digamdfn=digamdfn-sum;
        return
    else
        digamdfn=digamdfn-one/y;
        y=y+one;
        while (y < crit)
           digamdfn=digamdfn-one/y;
           y=y+one;
        end
          digamdfn=digamdfn+log(y)-half/y;
```

```
            y=one/(y*y);
            sum=((((((c7*y+c6)*y+c5)*y+c4)*y+c3)*y+c2)*y+c1)*y;
            digamdfn=digamdfn-sum;
            return

        end
    else
        digamdfn=d1-one/x;
        return
    end
    end
```

A-3 Heterogeneity Test

(Translated FORTRAN code provided by Hosking and Wallis, 1997)

```
clc;
clear all;


v=input(' Enter the weighted sd of sample L-CVs for the region: ');
ns=input(' Enter the number of sites in this region: ');
nrg=input(' Enter the number of regions to be simulated: ');
eps=input(' Enter the location parameter of kappa distribution: ');
alpha=input(' Enter the scale parameter of kappa distribution: ');
k=input(' Enter the shape parameter of kappa distribution: ');
h=input(' Enter the 4th parameter of kappa distribution: ');


%open excel file with number of records at each site within the region i
% in it (it should be in the same folder as this M-file
%sheet1 of this excel file contains the values
%
[type, sheets] = xlsfinfo('Sites_records.xlsx');
SitesMatrix = xlsread('Sites_records.xlsx', 'Sheet1');



disp ('simulating...please wait');
disp ('  ');

for k1=1:nrg,

    for k2=1:ns,
        nrec=SitesMatrix(k2);
        y=0;
        for i=1:nrec,
            y(i)=eps+alpha/k*(1-((1-(rand)^h)/h)^k);
        end
        y_sort=sort(y);
        x=y_sort/mean(y);
        x1=0;

        for j=1:nrec,
            x1(j)=x(j)*(j-1);
        end


        x2=sum(x1)/(nrec*(nrec-1));  %b1
        x3=2*x2-mean(x);              %l2=2*b1-b0
        x4(k2)=x3/mean(x);            %l-CV=l2/l1
    end
```

```
    for k3=1:ns,
        x5(k3)=x4(k3)*SitesMatrix(k3);
    end

    x6=sum(x5)/sum(SitesMatrix);


    for l=1:ns,
        x7(l)=SitesMatrix(l)*((x4(l)-x6)^2)/sum(SitesMatrix);
    end


    x8(k1)=sqrt(sum(x7));
    k1
end




H=(v-mean(x8))/std(x8);

beep

disp ('Results:');
disp ('================================');
disp ('   ')



if and(lt(H,1), ge(H,0))
    disp ('The region is homogeneous');
    disp ('  ');

elseif H<0
    disp ('The L-moments are correlated');
    disp ('  ');

elseif and (ge(H,1), lt(H,2))
    disp('The region is possibly heterogeneous');
    disp('  ');

else
    disp('The region is definitely heterogeneous: ');
    disp('  ');

end

fprintf ('The heterogeneity measure, H=%6.2f\n', H);
fprintf ('The mean of simulated regions is, mean=%6.4f\n', mean(x8));
fprintf( 'The standard deviation of simulated regions is, std=%6.4f\n',
std(x8));
```

A-4 Goodness-of-fit test

(Translated FORTRAN code provided by Hosking and Wallis, 1997)

```
clear all;
clc;

%this program calculates the goodness of fit measure 'z'
%in the first part it computes the bias and standard deviation of the
%sample regional L-Kurtosis.

%In the next part this program computes one part of calculations needed
in
%goodness of fit test. (Calculating tau-4 for each candidate
distribution)


%the candidate distribution names are as follow:
% GLO=Generalized Logistic Distribution
% GEV=Generalized Exterme Value Distribution
% LN3=Lognormal Distribution
% PE3=Pearson type III Distribtuion
% GPA=Generalized Pareto Distribuion



ns=input(' Enter the number of sites in this region: ');
nrg=input(' Enter the number of regions to be simulated: ');


eps=input(' Enter the location parameter of kappa distribution: ');
alpha=input(' Enter the scale parameter of kappa distribution: ');
k=input(' Enter the shape parameter of kappa distribution: ');
h=input(' Enter the 4th parameter of kappa distribution: ');
%distr=input('Enter the candidate distribution name:', 's');

Tau3=input(' Enter regional average L-Skewness tau3 for this region: ');
t4R=input(' Enter regional average L-Kurtosis for this region: ');


%open excel file with number of records at each site within the region
in
%it (it should be in the same folder as this M-file
%sheet1 of this excel file contains the values
%
[type, sheets] = xlsfinfo('Sites_records.xlsx');
SitesMatrix = xlsread('Sites_records.xlsx', 'Sheet1');


disp ('simulating...please wait');
disp ('  ');
```

```
for k1=1:nrg,

    for k2=1:ns,
        nrec=SitesMatrix(k2);
        y=0;
        for i=1:nrec,
            y(i)=eps+alpha/k*(1-((1-(rand)^h)/h)^k);
        end
        mode='descend';
        y_sort=sort(y,mode);
        x=y_sort/mean(y);

        x1=0;
        x2=0;
        x3=0;

        for j=1:nrec,
            x1(j)=x(j)*(j-1);
            x2(j)=x(j)*(j-1)*(j-2);
            x3(j)=x(j)*(j-1)*(j-2)*(j-3);
        end



        b0=mean(x);
        b1=sum(x1)/(nrec*(nrec-1));
        b2=sum(x2)/(nrec*(nrec-1)*(nrec-2));
        b3=sum(x3)/(nrec*(nrec-1)*(nrec-2)*(nrec-3));

        l1=b0;
        l2=2*b1-b0;
        l3=6*b2-6*b1+b0;
        l4=20*b3-30*b2+12*b1-b0;

        t(k2)=l2/l1;
        t3(k2)=l3/l2;
        t4(k2)=l4/l2;

    end

     for i=1:k2,
         t4r(i)=SitesMatrix(i)*t4(i)/sum(SitesMatrix);
     end


     T4(k1)=sum(t4r);

end
```

```
%calculate the bias of t4R

for k1=1:nrg,


   b4(k1)=(T4(k1)-t4R)/nrg;
   b5(k1)=(T4(k1)-t4R)^2;

end


%bias for t4R
B4=sum(b4);

%standard deviation of t4R
B5=sum(b5);
sigma4=sqrt((B5-nrg*B4^2)/(nrg-1));

beep
disp('======================================');
fprintf ('The Bias of regional L-Kurtosis, B4= %8.4f\n', B4);
fprintf ('The Standard deviation of regional L-Kurtosis, Sigma4=
%8.4f\n', sigma4);




%if distr=='GLO'
    %Tau4distr=0.16667*Tau3^0+0.83333*Tau3^2;
%elseif distr=='GEV'
    %Tau4distr=0.10701*Tau3^0+0.11090*Tau3^1+0.84838*Tau3^2-
0.06669*Tau3^3+0.00567*Tau3^4-0.04208*Tau3^5+0.03763*Tau3^6;
%elseif distr=='LN3'
    %Tau4distr=0.12282*Tau3^0+0.77518*Tau3^2+0.12279*Tau3^4-
0.13638*Tau3^6+0.11368*Tau3^8;
%elseif distr=='PE3'
    %Tau4distr=0.12240*Tau3^0+0.30115*Tau3^2+0.95812*Tau3^4-
0.57488*Tau3^6+0.19383*Tau3^8;
%elseif distr=='GPA'
    %Tau4distr=0.20196*Tau3^1+0.95924*Tau3^2-
0.20096*Tau3^3+0.04061*Tau3^4;
%else
    %disp('wrong name was entered for candidate distribution');
%end


Tau4distr(1)=0.16667*Tau3^0+0.83333*Tau3^2;
Tau4distr(2)=0.10701*Tau3^0+0.11090*Tau3^1+0.84838*Tau3^2-
0.06669*Tau3^3+0.00567*Tau3^4-0.04208*Tau3^5+0.03763*Tau3^6;
```

```
Tau4distr(3)=0.12282*Tau3^0+0.77518*Tau3^2+0.12279*Tau3^4-
0.13638*Tau3^6+0.11368*Tau3^8;
Tau4distr(4)=0.12240*Tau3^0+0.30115*Tau3^2+0.95812*Tau3^4-
0.57488*Tau3^6+0.19383*Tau3^8;
Tau4distr(5)=0.20196*Tau3^1+0.95924*Tau3^2-
0.20096*Tau3^3+0.04061*Tau3^4;


%distr(1)='GLO';
%distr(2)='GEV';
%distr(3)='LN3';
%distr(4)='PE3';
%distr(5)='GPA';

distr=['GLO';'GEV';'LN3';'PE3';'GPA'];



for j=1:5,
Zdist(j)=(Tau4distr(j)-t4R+B4)/sigma4;


 fprintf ('The L-Kurtosis of candidate distribution is: %8.6f\n',
Tau4distr(j));

 fprintf ('The goodness of fit measure, Zdist of candidate distribution
%-5.10s', distr(j)), fprintf(' is: %8.6f\n', Zdist(j));
 %disp('The goodness of fit measure, Zdist of candidate distribution',
distr(j), 'is=', Zdist(j));

 if abs(Zdist(j))<= 1.64
     disp('The candidate distribution has accepted fit to the data');
 else
     disp('The candidate distribution does not give an adequate fit to
the data');
 end
     disp('==============');
end
```

A-5 Parameters of Lognormal Distribution

```
clc; clear all;

% In this M-file, 3 parameters of lognormal distribution
% epsilon, alpha and ka will be calculated.

%this program needs regional average l-moment ratios to performe this
parameter
%estimation.

format long

tau2=input(' Enter the L-CV for the region: ');
tau3=input(' Enter the L-Sk for the region: ');
tau4=input (' Enter the L-Ku for the region: ');

E0=2.0466534;
E1=-3.6544371;
E2=1.8396733;
E3=-0.20360244;
F1=-2.0182173;
F2=1.2420401;
F3=-0.21741801;


 k=-
tau3*(E0+E1*tau3^2+E2*tau3^4+E3*tau3^6)/(1+F1*tau3^2+F2*tau3^4+F3*tau3^6
);

fix=-k/(2^0.5);

fixstring=num2str(fix);
disp(['Find fi(x) from the cumulative standard normal table when x is',
fixstring]);

fi=input('fi(x) is equal to: ');

alpha=(tau2*k*exp(-k^2/2))/(1-2*fi);

epsilon=1-alpha*(1-exp(k^2/2))/k;



kstring=num2str(k);
alphastring=num2str(alpha);
epsilonstring=num2str(epsilon);

disp(['the first parameter of lognormal distribution, k=', kstring]);
```

```
disp(['the second parameter of lognormal distribution, alpha=',
alphastring]);
disp(['the third parameter of lognormal distribution, epsilon=',
epsilonstring]);
```

A-6 Quantile of Lognormal Distribution

(Translated FORTRAN code provided by Hosking and Wallis, 1997)

```
clc; clear all;
format long

para(1)=input(' Enter the location parameter of lognormal distribution,
epsilon: ');
para(2)=input(' Enter the scale parameter of lognormal distribution,
alpha: ');
para(3)=input(' Enter the shape parameter of lognormal distribution, ka:
');


[type, sheets] = xlsfinfo('Lognormal Distribution.xlsx');
F = xlsread('Lognormal Distribution.xlsx', 'Sheet1', 'B3:B120');


zero=0; one=1;


U=para(1);
A=para(2);
G=para(3);


for i=1:118


 j=i+2;
 jstring=num2str(j);
 cstring='c';
 cellstring=strcat(cstring, jstring);

if A <= zero
   disp(['*** error*** routine QUALOGN:  Parameters invalid']);
      QUALOGN=zero;
      xlswrite('Lognormal distribution.xlsx', QUALOGN, 'Sheet1',
cellstring );


elseif F(i) <= zero || F(i) >= one

    if F(i)== zero && G < zero
        QUALOGN=U+A/G;
        xlswrite('Lognormal distribution.xlsx', QUALOGN, 'Sheet1',
cellstring );

    elseif F(i)== one && G > zero
```

```
        QUALOGN=U+A/G;

         xlswrite('Lognormal distribution.xlsx', QUALOGN, 'Sheet1',
cellstring );

    else
        disp(['*** error*** routine QUALOGN:  Argument of function
invalid']);
        QUALOGN=zero;
        xlswrite('Lognormal distribution.xlsx', QUALOGN, 'Sheet1',
cellstring );
    end



else

    QUASTNfn=QUASTN(F(i));
    Y=QUASTNfn
    if G ~= zero
    Y=(one -exp(-G*Y))/G;
    end
    QUALOGN=U+A*Y;

xlswrite('Lognormal distribution.xlsx', QUALOGN, 'Sheet1', cellstring );


end

end




function [ QUASTNfn ] = QUASTN( x )
%QUASTN This functions will calculate the cumulative standard normal
%distribuyion (fi)
%
format long

zero=0; half=0.5; one=1;
split1=0.425; split2=5; const1=0.180625; const2=1.6;


% Coefficients of rational-function approximations
A0=0.338713287279636661e1;
A1=0.133141667891784377e3;
A2=0.197159095030655144e4;
A3=0.137316937655094611e5;
A4=0.459219539315498715e5;
```

```
A5=0.672657709270087009e5;
A6=0.334305755835881281e5;
A7=0.250908092873012267e4;
B1=0.423133307016009113e2;
B2=0.687187007492057908e3;
B3=0.539419602142475111e4;
B4=0.212137943015865959e5;
B5=0.393078958000927106e5;
B6=0.287290857357219427e5;
B7=0.522649527885285456e4;



C0=0.142343711074968358e1;
C1=0.463033784615654530e1;
C2=0.576949722146069141e1;
C3=0.364784832476320461e1;
C4=0.127045825245236838e1;
C5=0.241780725177450612;
C6=0.227238449892691846e-1;
C7=0.774545014278341408e-3;
D1=0.205319162663775882e1;
D2=0.167638483018380385e1;
D3=0.689767334985100005;
D4=0.148103976427480075;
D5=0.151986665636164572e-1;
D6=0.547593808499534495e-3;
D7=0.105075007164441684e-8;



E0=0.665790464350110378e1;
E1=0.546378491116411437e1;
E2=0.178482653991729133e1;
E3=0.296560571828504891;
E4=0.265321895265761230e-1;
E5=0.124266094738807844e-2;
E6=0.271155556874348758e-4;
E7=0.201033439929228813e-6;
F1=0.599832206555887938;
F2=0.136929880922735805;
F3=0.148753612908506149e-1;
F4=0.786869131145613259e-3;
F5=0.184631831751005468e-4;
F6=0.142151175831644589e-6;
F7=0.204426310338993979e-14;


Q=x-half;

if abs(Q)> split1
    R=x;
```

```
    if Q >= zero
        R=one-x;
    end

    if R <= zero
        disp(['*** error*** routine quastn argument of function
invalid']);
        QUASTNfn=zero;
        return

    else

    R=sqrt(-log(R));
    if R > split2
        R=R-split2;

QUASTNfn=((((((((E7*R+E6)*R+E5)*R+E4)*R+E3)*R+E2)*R+E1)*R+E0)/(((((((F7*R
+F6)*R+F5)*R+F4)*R+F3)*R+F2)*R+F1)*R+one);


        if Q < zero
            QUASTNfn=-QUASTNfn;
        end
        return



    else

    R=R-const2;

QUASTNfn=((((((((C7*R+C6)*R+C5)*R+C4)*R+C3)*R+C2)*R+C1)*R+C0)/(((((((D7*R
+D6)*R+D5)*R+D4)*R+D3)*R+D2)*R+D1)*R+one);

    if Q < zero
        QUASTNfn=-QUASTNfn;
    end
        return

    end
    end


else


R=const1-Q*Q;
QUASTNfn=Q*(((((((A7*R+A6)*R+A5)*R+A4)*R+A3)*R+A2)*R+A1)*R+A0)/(((((((B7
*R+B6)*R+B5)*R+B4)*R+B3)*R+B2)*R+B1)*R+one);
return
end
```