Improving Dam Safety Analysis by Using Physically-Based Techniques to Derive Estimates of Atmospherically Maximum Precipitation

by

Allyson K. Bingeman

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Civil Engineering

Waterloo, Ontario, Canada, 2001

© Allyson K. Bingeman, 2001

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-65223-8

Canada

The UNIVERSITY OF WATERLOO requires the signatures of all persons using or photocopying this thesis. Please sign below and give address and date.

Abstract

A high or very high consequence dam is a large dam whose failure would have large consequences to life and/or property downstream. Knowledge of the magnitudes of extreme floods and their associated annual exceedence probabilities (AEP) are needed to determine the risk that such a dam might fail.

Traditionally, the largest "physically possible" precipitation event (the Probable Maximum Precipitation, PMP) and its associated flood event (the Probable Maximum Flood, PMF) have been calculated with a combination of statistical and meteorological techniques developed by the World Meteorological Organization (WMO). These techniques work reasonably well in flatter terrain, but may occasionally produce unrealistic results in mountainous terrain. This research focuses on improving safety studies for hydrologic structures such as dams, by using physically-based techniques to estimate the PMP and PMF and to calculate the associated AEP. This research contributes in three areas. The first area is in using an atmospheric model to estimate maximum precipitation. Secondly, the research demonstrated that simulated streamflow may be used to generate frequency curves and their associated confidence limits. The final contribution was in demonstrating that the frequency statistics indicated that the traditional PMP overestimates the PMF, while the atmospheric model estimates were more in line with accepted AEPs for a PMF.

This research was performed on the upper Columbia River Basin in southwestern British Columbia. The basin is an alpine basin, with annual precipitation varying from 2500 mm on the west to 500 mm on the east. Severe precipitation events generally begin over the Pacific Ocean, but are somewhat moderated by the intervening mountain ranges. There are several hydroelectric and flood-control dams operated by BCHydro on the Canadian portion of the river. One of these, Mica Dam, was used as the focus of this research.

The Mesoscale Compressible Community (MC2) model (Recherche en Prévision Numérique) is an atmospheric model designed to forecast weather at a fine resolution. The

MC2 model was used to calculate a physically-based estimate of the maximum, atmospherically possible, precipitation (referred to as the Probable Maximum Storm, PMS). The numerical experiments with this model suggested that an atmospheric maximum precipitation does in fact exist, and it can be calculated with the model. The method is less subjective than the traditional WMO method and not subject to the same issues of data quality. Also, the model accounts implicitly for topography in its calculation of precipitation. The model determined a maximum 24-hour precipitation of 73.4 mm as an average over the Mica Dam basin (this number is preliminary, further research into the atmospheric model may result in a larger number). This PMS produced by the atmospheric model was larger than any previously observed precipitation event, but lower than the PMP produced with the WMO method, indicating that the WMO method may overestimate the PMP in mountainous terrain. The MC2 model is recommended for developing the maximum atmospherically possible precipitation, but further meteorological research is recommended to ensure that all of the assumptions used in MC2 and the PMS module are suitable for this purpose. The PMS and the PMP were both used as input to the physically-based hydrological model WATFLOOD/SPL.

A flood frequency curve was developed to assess the AEP to the floods generated by the PMS and the PMP. The AEP were used to compare the relative magnitudes of the floods caused by the PMS and PMP, and determine if they were within the presumed probability range of the PMF (10^{-4} to 10^{-6}).

The derivation of a frequency curve is dependent upon the time series length of the data, which is often too short for meaningful extrapolation to the return intervals for a PMF. In this research, historical meteorological data were available and used in a hydrological model to develop a long, deterministically simulated streamflow time series of 95 years. The use of the simulated data decreased the sampling uncertainty due to a short time series. The

simulated data generated frequency curves that were similar to frequency curves derived with observed data.

However, the simulated streamflow data are based on uncertain atmospheric variables that are transformed by an atmospheric model and by WATFLOOD/SPL. This thesis addresses the consequence of the parameter uncertainty in WATFLOOD/SPL. The 95% confidence limits for the frequency curves were derived through a Monte Carlo analysis of the parameter variation. An investigation into the behavior of the model showed that the parameter set within WATFLOOD/SPL was robust and there was only one optimum parameter set within the limits of the parameter space. Due to time constraints, a method to use the variation in a five-year time series as an analogue for the variation in the full time series was developed. The confidence limits grew wider as the return period increased, although further research into the behavior of the parameters may help reduce the width of the confidence limits.

The frequency curve and its confidence limits were used to estimate the range of return interval of the floods produced by the PMS and the PMP. These return intervals of the floods were used to determine if the floods were consistent with the PMF. The 100-year snowpack, the 100-year melting temperature sequence and the PMP together generated a flood with an AEP that was much smaller than the probability range for a PMF ($<<10^{-7}$). The 100-year snowpack, the 100-year melting temperature sequence and the PMS together generated a flood with an AEP that was also smaller than the probability range for a PMF ($<10^{-7}$), however, it was closer to the presumed probability range than the combination with the PMP. This suggests that the PMS may be a more realistic estimate of maximum precipitation for PMF estimation, and is still somewhat conservative.

This research has improved the current methods for safety analysis of hydrological structures, and recommends the use of physically-based methods to derive PMPs and PMFs.

Comparing the PMF estimates with frequency curves should help validate results and provide a higher level of confidence in extreme rain produced flooding.

Acknowledgments

I would first like to thank my supervisors, Dr. Nicholas Kouwen and Dr. Eric D. Soulis, for their support and advice. Their insight was invaluable to help me understand the issues in my research. I would also like to thank all of my officemates, past and present, for their help and interesting discussions.

This research was funded by graduate scholarships from the Natural Sciences and Engineering Research Council (NSERC), and by operating grants from my supervisors.

BCHydro (Vancouver, B.C.) provided the majority of the data used in this research. I would like to acknowledge BCHydro for giving me permission to use the data and for providing additional insight to help me to improve my research.

The atmospheric model MC2-PMS was programmed by Recherche en Prévision Numérique (Montreal, PQ). I would like to thank Pierre Pellerin for taking the time to train me and answer my (many) questions about how to use the model. Pierre was also invaluable in helping me to interpret my results and solve problems with the atmospheric model.

I would like to acknowledge my friends from Waterloo Christian Fellowship and from The Father's House Church. They have prayed for me and taught me many things, and I appreciate their constant love and support.

I would also like to acknowledge the love and encouragement of my family members. I would like to thank my parents and my brothers for listening to me and then confidently stating that you were sure I would be able to figure it out soon. To my husband Mark, I would like to say thanks for sitting through technical discussions about floods and statistics and being my "sounding board." Your encouragement kept me going when technical issues seemed insurmountable. I love you all.

Finally, I would like to praise my Lord and Savior, Jesus Christ, who died on the cross to save me from my sins and give me eternal life. He has transformed my life, and helped me in so many ways.

Table of Contents

ABSTRACT ..... **IV**

ACKNOWLEDGMENTS ..... **VIII**

1 INTRODUCTION **1**
1.1 OBJECTIVES 8

2 STUDY AREA **10**

3 BACKGROUND AND LITERATURE REVIEW **17**
3.1 PROBABLE MAXIMUM PRECIPITATION (PMP) ESTIMATION 17
3.1.1 World Meteorological Organization (WMO) Method 17
3.1.2 Physically-based Estimation Techniques 21
3.2 NUMERICAL WEATHER MODELING 23
3.2.1 Use of atmospheric models in the literature 24
3.2.2 Mesoscale Compressible Community – Probable Maximum Storm (MC2-PMS) Model 25
3.2.2.1 MC2 model description 26
3.2.2.2 PMS module description 27
3.2.3 High-Resolution Boundary Layer Model 36
3.2.3.1 Maximum and Minimum Temperature Calculation 37
3.2.3.2 Precipitation Calculation 39
3.2.3.3 Objective Analysis 40
3.3 HYDROLOGIC MODELING 42
3.3.1 WATFLOOD/SPL subroutines 44
3.3.2 Approaches to Calibration and Validation 49
3.3.3 Uncertainty in Model Estimates 57
3.4 REGIONAL FREQUENCY ANALYSIS AND THE METHOD OF L-MOMENTS 67
3.4.1 Index-Flood Procedure 68
3.4.2 L-moments 70
3.4.2.1 Introduction 70
3.4.2.2 Regional L-moments 73
3.4.3 Steps Involved in Regional Frequency Analysis 75
3.4.3.1 Data screening 76
3.4.3.2 Identification of regions 76
3.4.3.3 Choice of a frequency distribution 81
3.4.3.4 Estimation of the at-site frequency distribution parameters 82
3.4.4 Description of Data used for Frequency Analysis 82
3.5 CHAPTER SUMMARY 87

4 CALCULATING THE PROBABLE MAXIMUM FLOOD WITH THE PROBABLE MAXIMUM STORM MODEL **88**
4.1 CALCULATING THE PROBABLE MAXIMUM STORM 88

4.1.1 Setup Information for the MC2-PMS Model 89
4.1.2 Adjusting the location of the storm 95
4.1.3 Parameter Variation to determine the Probable Maximum Storm 100
4.1.3.1 Parameter Variation for MC2-PMS perturbations 101
4.1.3.2 The relationship between the parameters and storm direction 102
4.1.3.3 The relationship between the parameters and storm magnitude 103
4.1.4 Comparing the PMS to Other Significant Precipitation Events 108
4.1.5 Summary of the MC2-PMS method to develop precipitation events 111
4.2 GENERATING THE FLOODS CAUSED BY THE PROBABLE MAXIMUM STORM 112
4.2.1 Comparison of the PMS flood to the historical 1983 storm 113
4.2.2 Comparison of the PMS flood to the traditional PMF estimate 114
4.3 SUMMARY OF THE MC2-PMS/WATFLOOD METHOD TO DEVELOP STORMS AND FLOODS 117

5 GENERATING IMPROVED FLOOD FREQUENCY CURVES 118
5.1 VALIDATION OF THE USE OF SIMULATED STREAMFLOW DATA 119
5.2 DERIVATION OF THE FLOOD FREQUENCY CURVE 136
5.2.1 Identification of Regions and Choosing Appropriate Distributions 137
5.2.2 Calculating the Frequency Distributions for Mica Dam 141
5.2.3 Discussion of Flood Frequency Curves 147
5.3 DERIVATION OF THE CONFIDENCE LIMITS FOR THE FLOOD FREQUENCY CURVE 148
5.3.1 Investigation of the Objective Function 150
5.3.2 Choice of the Parameter Distributions 162
5.3.3 Calculating the 95-year Confidence Limits 166
5.3.4 Calculating the Regional Confidence Limits 174
5.4 CHAPTER SUMMARY 180

6 COMPARING THE PROBABILITIES OF THE FLOODS 182

7 CONCLUSIONS 186
7.1 CONTRIBUTIONS 187

8 RECOMMENDATIONS FOR FURTHER RESEARCH 189

9 REFERENCES 191

APPENDIX A 199

APPENDIX B 217

APPENDIX C 227

APPENDIX D 236

APPENDIX E 244

List of Figures

Figure 2-1 – Digital Elevation Model for the Columbia River Basin (GTOPO30).....11
Figure 2-2 – Location of Streamflow Stations and Major Dams in the Columbia River basin (Fisheries and Environment Canada, British Columbia Active Hydrometric Stations, December 1977).....13
Figure 2-3 – Landsat MSS image for the Columbia River Basin.....16
Figure 3-1 – Atmospheric Perturbation: Pressure Wave (Domain: 18,000 by 18,000 km).....34
Figure 3-2 – Atmospheric Perturbation: Temperature Wave (Domain: 18,000 by 18,000 km)35
Figure 3-3 – Atmospheric Perturbation: Wind Speed and Direction (Domain: 18,000 by 18,000 km).....36
Figure 3-4 – Weight calculation to make the gridpoints match an observation station.....40
Figure 3-5 – Matching the second derivative in objective analysis.....42
Figure 3-6 – Calibration tool for WATFLOOD/SPL.....56
Figure 3-7 – Comparison of 96-Years of Streamflow – observed vs simulated by WATFLOOD/SPL with Meteorological Forcing Data (Daily Precipitation and Temperature) from HRBL for the Columbia River at Nicholson location.....59
Figure 3-8 – Comparison of 23-Years of Reservoir Inflow – observed vs simulated by WATFLOOD/SPL with Meteorological Forcing Data (Daily Precipitation and Temperature) from HRBL for the Mica Dam location.....61
Figure 3-9 – Contour Plot of Annual Streamflow Volume Differences in % (Rmodel-Rmeasured), 1971-1994.....63
Figure 3-10 – Definition of terms in calculation of Probability Weighted Moments.....71
Figure 3-11 – L-moment ratio diagram (from Hosking and Wallis, 1997). Two and three parameter distributions are shown as points and lines, respectively. E=Exponential, G=Gumbel, L=Logistic, N=Normal, U=Uniform, GLO=Generalized Logistic, GEV=Generalized Extreme Value, GPA=Generalized Pareto, LN3=Lognormal, PE3=Pearson Type III, OLB = Overall Lower Bound.....73
Figure 3-12 – Definition of Physiographic Parameters.....84
Figure 4-1 – Precipitation accumulation (mm) from a storm that traveled through Alaska....90
Figure 4-2 – The perturbation embedded in the July 13, 1983 general circulation.....91
Figure 4-3 – Relative Humidity Profile in the a) Warm Front and b) Cold Front (Vertical axis is Pressure in mb, Horizontal axis is Relative Humidity in %).....92
Figure 4-4 – Calculation Domains for MC2: 150km domain showing location of perturbation93
Figure 4-5 – Calculation Domains for MC2: 150 km domain showing 50 km domain.....94
Figure 4-6 – Calculation Domains for MC2: 50 km domain showing 10 km domain.....95
Figure 4-7 – Precipitation (mm) from Perturbation located at: a) 2 degrees west and 1 degree north, b) 2 degrees west and 2 degrees north.....98
Figure 4-8 – Comparison of Precipitation versus Elevation using 10km2 grids.....100
Figure 4-9 – a) Low Precipitation Storm (#25), b) High Precipitation Storm (#1).....104
Figure 4-10 – Average Mica Precipitation versus Ax and Bx, using the nine original simulations.....106
Figure 4-11 – Average Mica Precipitation versus Ax and Bx, including the extra simulations and showing the maximum precipitation.....108
Figure 4-12 – 24-hour Precipitation from the Probable Maximum Storm derived by MC2-PMS.....109

Figure 4-13 – a) Precipitation (in mm) from July 11, 1983 b) Difference (in mm) between the precipitation for the Probable Maximum Storm and July 11, 1983 110
Figure 4-14 – Comparison of Mica Dam inflow with the Historical Storm and the PMS 114
Figure 4-15 – Comparison of Mica Dam inflow with various PMF scenarios 116
Figure 5-1 - Comparison plot of simulated and observed peak flows for Columbia River at Nicholson station 121
Figure 5-2 – Histogram of Residuals of Peak Flows for Columbia River at Nicholson (90 data points) 121
Figure 5-3 – Frequency Curves for Columbia River near Fairmont Hot Springs 125
Figure 5-4 – Frequency Curves for Illecillewaet River at Greeley 126
Figure 5-5 – Frequency Curve for St. Mary's near Marysville 127
Figure 5-6 – Frequency Curves for Columbia River at Nicholson 128
Figure 5-7 – Frequency Curves for Mica Dam 129
Figure 5-8 – Comparison of 100-year return interval flows versus the number of years of record 132
Figure 5-9 – Ratio of Q100 for simulated series compared to variability in observed Q100 ratio: the grey boxes are the variability in observed data (Q100 for a short data series/Q100 for all years), the black boxes are the Q100 for the simulated short data compared to the Q100 observed estimate. 134
Figure 5-10 – Comparison of Stationarity: Q100 estimates from 10-year sequences of observed and simulated data, divided by Q100 observed 135
Figure 5-11 – Final Clustering of Streamflow Stations 139
Figure 5-12 – Comparison of Frequency Curves for Mica Dam calculated using Regional Frequency Analysis (regular Hosking and Wallis 1997 method): the observed curve used the observed data records, the simulated curve used the 95-year simulated streamflow data. 143
Figure 5-13 – Predicted L-CV versus Observed L-CV, showing results of regression 145
Figure 5-14 – Predicted L-CV versus Simulated L-CV, showing results of regression 145
Figure 5-15 – Comparison of Frequency Curves for Mica Dam calculated using Regional Frequency Analysis (Schaefer 1990 method): the observed curve used the observed data records, the simulated curve used the 95-year simulated streamflow data. 147
Figure 5-16 – Acceptable Shapes for the Objective function (error increases as distance from the calibrated value increases) 155
Figure 5-17 – Average difference between 5-year peaks (as percent of peak) for Mica Dam for the Barren classes of the Melt Factor and Base Temperature parameters 160
Figure 5-18 – Examples of histograms of the 10,000 values for BASE temperature, generated from the Beta-1 distribution (the mode is indicated with a red square). The three parameters are: a) BASE, Barren Class (mode = 3); b) BASE, High Elevation Dense Forest Class (mode = 1); BASE, Low Elevation Dense Forest Class (mode = -2) 166
Figure 5-19 – Histograms of the 10,000 flood frequency curves for Mica Dam 168
Figure 5-20 – Percent of 10,000 simulations whose flood frequency curves cross lines of the histogram for Mica Dam 170
Figure 5-21 – Comparison of Lower Confidence Limits for Mica Dam calculated with the Monte Carlo and as the average of the 95-year simulations 172
Figure 5-22 – Comparison of Upper Confidence Limits for Mica Dam calculated with the Monte Carlo and as the average and maximum of the 95-year simulations 173

Figure 5-23 – Regional Frequency Curve (Schaefer, 1990, method) for Mica Dam Combined with the Regional Confidence Limits Determined from the Monte Carlo Analysis 177
Figure 5-24 – Regional Frequency Curve for Mica Dam, showing Confidence Limits and amount of variation due to snowmelt parameters (MF and BASE).....180

List of Tables

Table 2-1 – List of Streamflow Stations and B.C. Hydro Dams with Number of Years of Record 14
Table 3-1 - Parameters for the Atmospheric Perturbation 33
Table 3-2 - Parameter Values and Limits for WATFLOOD/SPL (Columbia River Domain) 53
Table 3-3 – Definitions of Physiographic Parameters 83
Table 3-4 – Cross Correlation Table for Hydrologically Significant Variables 86
Table 4-1 – Comparison of the 10-km precipitation with the center of the perturbation placed in various locations 97
Table 4-2 – Values of parameter variables for each storm 101
Table 4-3 – Comparison of Storm Magnitudes for 27 simulations (based on 10-km precipitation estimates) 105
Table 4-4 – Extra simulations to refine the maximum storm 107
Table 5-1 – List of Which Simulated Curve Best Approximated the Observed Curve 130
Table 5-2 – Results of Homogeneity Test 140
Table 5-3 – Acceptable Distributions for Each Region 141
Table 5-4 - Heterogeneity statistics for the entire network of 36 stations and dams for the simulated streamflow (95 years) data. 142
Table 5-5 – Regression Parameters for Observed L-moments 144
Table 5-6 – Regression Parameters for Simulated L-moments 144
Table 5-7 – Summary of Relative Sensitivities for each parameter 152
Table 5-8 – Summary of Objective Function Investigation 155
Table 5-9 – Calculation of RMS Error with several combinations of MF and BASE located on the optimal ridge (values for Mica Dam) 161
Table 5-10 – Regression Results for Lower Confidence Limit at Mica Dam 176
Table 5-11 – Regression Results for Upper Confidence Limit at Mica Dam 176
Table 6-1 – Comparing the Probabilities of the Different Floods (Streamflow on June 1 for each simulation) 183

1 Introduction

A high-consequence dam is a large dam whose failure would have large consequences to life and/or property downstream. The magnitudes of extreme floods and their associated annual exceedance probabilities (AEP) are necessary to ensure that such a dam is safe. Traditionally, the largest "physically possible" precipitation event (the Probable Maximum Precipitation, PMP) and its associated flood event (the Probable Maximum Flood, PMF) have been calculated with a combination of statistical and meteorological techniques. These techniques have occasionally produced unrealistic results (Jarrett and Tomlinson, 2000). Recently, however, numerical weather prediction (NWP) models and hydrological models have become more robust and can now be used together to compute these events in a physically-based manner. The probability of exceedance of this flood, which can be estimated by various means, can be used to help evaluate the assumptions and the magnitude of the PMP and PMF. The length of the time series of the observed data limits these methods. NWP and hydrological models can be used in conjunction to increase the time series, and so improve the frequency curve. This research focuses on improving safety analysis for hydrologic structures by using physically-based techniques to estimate the PMP and PMF and calculate the associated AEP.

Many dams are built for flood control purposes, and they must be able to store extraordinary floods (usually the PMF). For instance, should a dam fail above a large city, there would be extreme damage to property, and potentially substantial loss of life. Such high-risk dams are designed to withstand the Probable Maximum Flood (PMF). The PMF, as defined by the U.S. Federal Energy Regulatory Commission (1993), is "the flood that may be

expected from the most severe combination of critical meteorological and hydrologic conditions that are reasonably possible in the drainage basin under study."

The largest theoretical flood varies with season, watershed size and location, and watershed topography. Typically, the PMF has not included climate change factors, although these may become significant in the future. In Canada, a spring flood is often associated with snowmelt. A thunderstorm or a frontal passage may cause a summer or fall flood. In the Canadian Rocky Mountains, snow remains well into the summer and the largest flood may result from a combination of snowmelt and a summer storm occurring together. When multiple flood scenarios exist, they all must be simulated. The PMF is defined as the largest of all of these flood events. The dam is tested to ensure that the PMF may be passed safely in all seasons.

The estimation of the probable maximum flood involves several tasks. The first task involves identifying the possible risk scenarios for the dam. For instance, one risk scenario may include a large amount of snowmelt runoff and a large precipitation event occurring as the reservoir fills. Therefore, a "maximum" snow accumulation, a "fastest" melting temperature sequence, and a "maximum" precipitation are all required for this dam. The second task involves estimation of each of these quantities. The maximum precipitation, referred to as the Probable Maximum Precipitation (PMP), is the most difficult of these three quantities to estimate. The estimation techniques are partly statistical and partly meteorological. It is difficult to ensure that the estimated PMP is physically possible (i.e. that the true atmospheric limit has been found, but not surpassed). The third task involves hydrological simulation of each risk scenario. The hydrological conditions that generate the

largest possible flood are used. Finally, simulations are used to pass the floods through the dam to test the safety of the dam.

The estimation of the PMF uses historical data. Therefore, it is re-estimated periodically, as more data are collected. Occasionally, the revised PMF will be significantly higher. For example, Jarrett and Tomlinson (2000) described a situation where the revised PMF for Olympus Dam in Colorado was almost four times larger than the original estimate. When this occurs, the dams may fail the safety check, leading to expensive spillway re-design and re-construction. There is, therefore, a considerable amount of concern about the techniques for PMF estimation. According to the National Research Council (NRC, 1988), there is continuing interest in extreme or rare floods (probability of exceedance of 10^{-3} to 10^{-7}) in the hydrologic and engineering communities for the purposes of planning and design of structures such as dam. The concerns are often focused on the uncertainty of the PMP (Jarrett and Tomlinson, 2000).

A World Meteorological Organization manual (WMO, 1986) describes the techniques to estimate the PMP. The manual explains that the method for estimating the PMP cannot be standardized and may need to be modified for a particular region (WMO, 1986, p. 4). The techniques depend on the size and location of the basin of interest, the amount and quality of data available at the site, and meteorological conditions that produce severe precipitation events. These problems are particularly severe in orographic regions, such as the Rocky Mountains. As such, the manual states that the PMP must be considered an estimate and that its accuracy cannot be assessed in an objective manner (WMO, 1986, p. 3).

Due to these concerns, BCHydro invited the University of Waterloo and several other groups to work in a collaborative project to develop physically-based estimation techniques

for the PMF. The use of physically-based estimation techniques alleviates these problems by using realistic atmospheric and hydrologic models to estimate severe storms and floods. The Columbia River basin was used as the research basin. There are four large dams operated by BCHydro on this river. This research focused on Mica Dam in particular, which has a drainage area of approximately 20,000 km^2. It is the most upstream dam, and the river is unregulated above Mica. The basin is described in greater detail in Chapter 2.

The Meso-scale Compressible Community (MC2) model (Benoit, *et al.*, 1997a) was used in conjunction with the Probable Maximum Storm (PMS) module (Benoit, *et al.*, 1997b) to develop a physically-based estimate of the largest physically possible storm. The models were developed at Recherche en Prévision Numérique (RPN) in Montreal, Quebec. The PMS module was developed for the Columbia River basin as part of the collaborative agreement with BCHydro. (Herein, the MC2-PMS estimate of the largest possible storm is termed the Probable Maximum Storm, PMS, to avoid confusion with the PMP estimate calculated with the WMO (1986) method.) The storms are calculated by creating a theoretical perturbation (pressure and temperature waves in the atmosphere) in the Pacific Ocean. The perturbations are controlled by parameters that can assume a range of realistic values. To derive the largest atmospherically possible storm, the parameters were varied to change the characteristics of the pressure and temperature waves. The atmospheric physics as calculated by the MC2 model require that the storm be physically possible. This analysis was a preliminary investigation of the PMS module, and so the analysis was limited to the characteristics of the pressure and temperature waves. Further analysis is required to determine the effect of changes in other parameters.

The WATFLOOD/SPL distributed, physically-based hydrological model (Kouwen, *et al.*, 2000a) was used to generate the flood that resulted from the PMS. WATFLOOD/SPL computes streamflow on a catchment basis but calculates runoff on a grid basis, allowing it to use NWP model data as input. The model simulates the hydrological budget, and so reduces the uncertainty associated with storm to flood transformation. The flood that resulted from the PMS was compared to the flood that resulted from the PMP (at Mica Dam). In this way, a physically-based estimate of the PMF was derived.

Risk analysis also requires an estimate of the frequency of a flood. Smith (1998) argued that the PMF has a frequency, which can be used to make informed policy decisions. The estimation of flood frequency curves, especially where measurements are limited, has been a subject of extensive research. The accuracy of flood frequency curves is generally low when time series are short and observation networks are sparse. Regionalization methods (e.g. Hosking and Wallis, 1997) have been developed to improve the estimation of frequency curves when time series are short. These methods are of limited use when the observation network is sparse, and therefore regional frequency analysis is not a solution for many regions in Canada and the world. The problem of sparse observation networks can be solved with the use of simulated streamflow data. Data from atmospheric models and distributed hydrological models could be used to augment and/or replace the observed data. The data may be replaced when there is reason to doubt the accuracy of the observed data.

This is the approach taken in this thesis. This concept is not a new concept (see, for example, Lamb, 1999), however, this application uses an atmospheric model applied for an unusually long period, longer than the observed streamflow record. When a long time series of simulated streamflow data are used in regional frequency analysis the result is a

potentially more accurate frequency curve in spite of uncertainty in the modeling steps. This research explored this possibility.

There are 32 Water Survey of Canada (WSC) streamflow observation stations in the Columbia River basin and four dams operated by BCHydro, with an average observed time series of 34 years (with a range from 5 to 91 years). In contrast, background meteorological data were available for a period of 96 years. The High-Resolution Boundary Layer (HRBL) model (Danard, 1996b) used this meteorological data to calculate gridded temperature and precipitation data over the Columbia River basin for the years 1899 to 1994, inclusive (96 years). These data were used in the WATFLOOD/SPL model to calculate a 96-year simulated streamflow time series. These data were used to improve the estimate of the regional frequency curve for the Mica Dam. First, however, the use of the simulated data to estimate frequency curves was validated. Secondly, the regional flood frequency analysis method (Hosking and Wallis, 1997, as modified by Schaefer, 1990) was used to generate a flood frequency curve. Thirdly, a method was developed to determine the confidence limits for the frequency curve developed from simulated streamflow. For the second and third steps, the regional analysis was performed for the Mica Dam reservoir, so that the PMF flows for the PMS and PMP precipitation could be compared. In this way, the use of a long time series of simulated streamflow to derive frequency curves was investigated.

To validate the use of simulated streamflow data to estimate frequency curves, individual frequency curves for each WSC station and BCHydro dam were calculated for each data set. Therefore, observed and simulated streamflow frequency curves were available for comparison at 36 locations within the domain. Frequency curves derived from the simulated streamfiow were similar to the frequency curves derived from observed

streamflow, and the longer time series improved the frequency curve estimation for high-flow, low-probability floods.

The flood frequency curve was generated with the regional flood frequency analysis method (Hosking and Wallis, 1997). The method increases accuracy of frequency curves by "trading space for time." That is, data are pooled from a number of locations and used to calculate the frequency curves. Therefore, the effective time series length increases. The method involves establishing a set of homogeneous regions, finding a suitable frequency distribution, and then fitting the frequency distribution and calculating the frequency curve at the site of interest. This method results in discontinuity between regions, leading to difficulties with establishing the frequency curves at an ungauged location. Schaefer (1990) developed a method to remove the discontinuity between regions and improve the L-moment estimates for precipitation data. Daviau, *et al.* (2000) used a similar method for streamflow: relationships between geo-statistical data and the L-moment estimates to derive smooth spatially varying L-moments estimates. This research also developed relationships to describe the spatial variation in L-moments: physiographic data were used to describe the spatial variation in L-moment statistics for streamflow. The use of the Schaefer (1990) method improved the estimate of the simulated frequency curve, as it matched the observed frequency curve more closely than the Hosking and Wallis simulated frequency curve.

Simulated streamflow data contain modeling uncertainty, and are therefore less certain than observed streamflow data. The Hosking and Wallis (1997) method includes a method to develop confidence limits for the frequency curve, however, the method assumes that the data contain only statistical uncertainty (e.g. due to sample size), and it was therefore unusable for modeling uncertainty. A method to develop the confidence limits for simulated

streamflow from the WATFLOOD/SPL model was developed. A Monte Carlo analysis was performed on the WATFLOOD/SPL model to develop the modeling uncertainty due to the model parameters. The Monte Carlo allowed the parameters to vary within their physically-possible ranges, and the resulting range of the peak flow was developed. The range of peak flow for each station was used to develop confidence limits for each station, which were then merged to create regional confidence limits. The regional confidence limits were applied to the regional frequency curve for Mica Dam.

The regional flood frequency curve based on the simulated streamflow was used to assign annual exceedance probabilities to the floods generated by the PMS and the PMP. The confidence limits for the frequency curve were used to compare the relative magnitudes of the PMS and PMP, and determine their suitability for use in calculating the PMF.

In this way, the techniques for estimating the PMP and PMF were improved through the use of physically-based atmospheric and hydrologic models, and the flood frequency curves were improved through the use of a long simulated streamflow time series. These tools allow for improved risk analysis for dams.

1.1 Objectives

In summary, the main objective of this research was to develop physically-based techniques for estimating the PMP and the PMF, and to improve the flood frequency curves so that annual exceedance probabilities for the PMF could be defined. The contributions of this research are:

- Investigated the use of an atmospheric model (MC2-PMS) to derive an extreme precipitation estimate, and derived the Probable Maximum Storm (subject to verification by meteorologists).

- Used the Probable Maximum Storm in a distributed hydrological model (WATFLOOD/SPL) to calculate the corresponding flood.
- Improved the flood frequency curve through the use of a long simulated time series of streamflow, and through the application of the Schaefer (1990) method.
- Examined the effect of model parameters on the range of output for extreme events from a hydrological model (WATFLOOD/SPL), and developed a method to derive confidence limits for flood frequency curves calculated from simulated streamflow data.

The remainder of this thesis is organized as follows: the study area is described in Chapter 2, the background information is presented in the literature review in Chapter 3, the Probable Maximum Storm and Probable Maximum Flood are derived in Chapter 4, and the improvements to the flood frequency curve are described in Chapter 5. Chapter 6 uses the flood frequency curve to compare the relative magnitudes of the PMS and the PMP and their effect on the PMF. Finally, Chapters 7 and 8 present the Conclusions and Recommendations from this research.

2 Study Area

This research was performed for the Columbia River Basin within southeastern British Columbia. This location was chosen for several reasons. There were sufficient data to allow the atmospheric and hydrologic models to be developed and calibrated for this basin. The HRBL model (Danard, *et al.*, 1996b) was available to calculate precipitation and temperature for a 96-year time period over the entire basin. The MC2-PMS (Benoit, *et al.*, 1997a and 1997b) model was developed and calibrated for this basin so that extreme storms could be calculated. Finally, there were sufficient observed streamflow data in and near the basin to allow the hydrological model WATFLOOD/SPL to be calibrated.

The basin is located in the Rocky Mountain Range, and has a drainage area of approximately 50,000 km^2. The river begins at Columbia Lake and flows north-west for approximately 330 km through the Rocky Mountain trench before veering south to go through the states of Washington and Oregon, as can be seen in the 30-arcsecond digital elevation model (DEM) of the basin (shown in Figure 2-1). The locations of Golden, Mica Dam, Revelstoke Dam, and Castlegar are shown on the image. The rivers are also superimposed on the DEM. The lighter shades indicate higher elevations; the highest peaks in the basin are approximately 3000 m above sea level, while the lowest elevations are approximately 400 m. The large elevation range of this basin complicates the normal PMP estimation procedures.

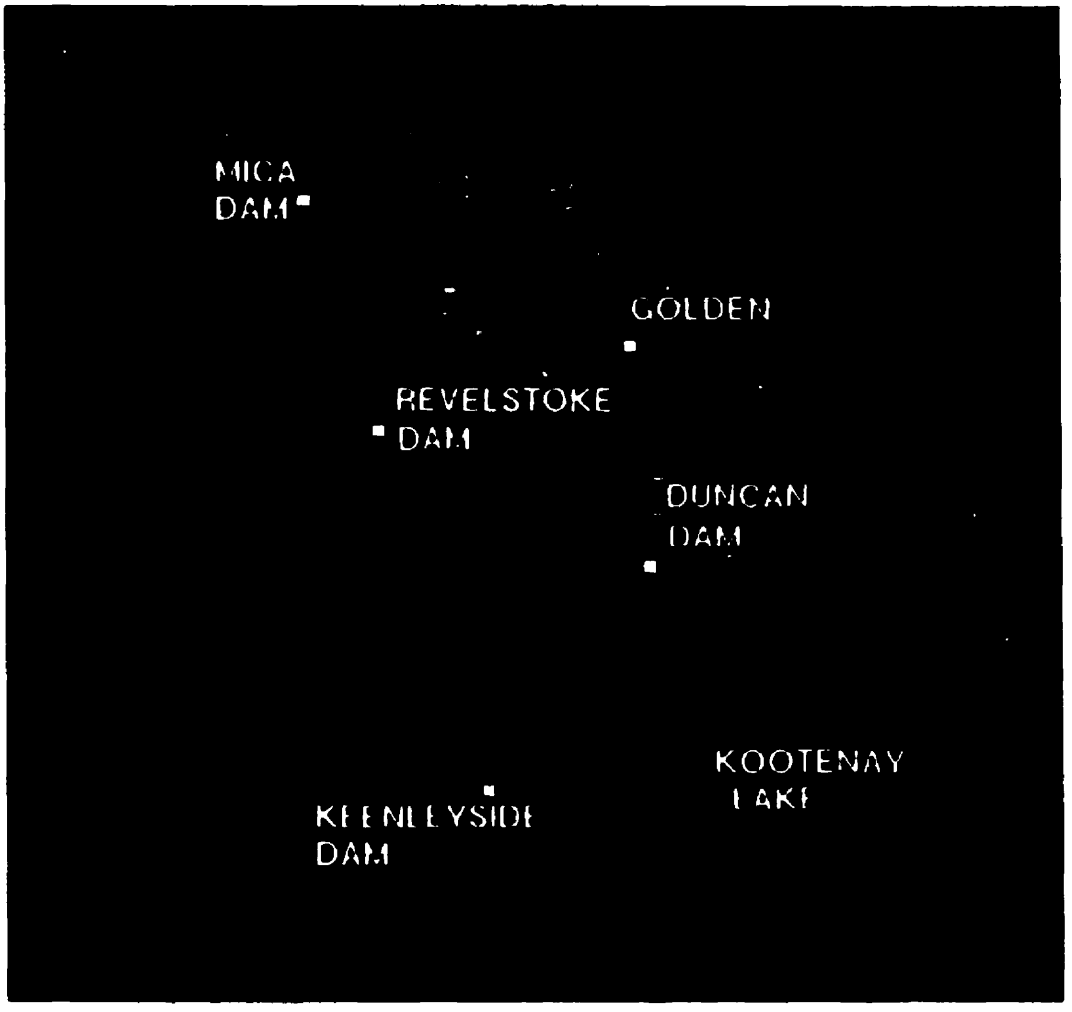


Figure 2-1 – Digital Elevation Model for the Columbia River Basin (GTOPO30)

Most of the storms that affect British Columbia develop in the Pacific Ocean. However, the Columbia River Basin is sheltered from the ocean by several mountain ranges, and is therefore comparatively drier than the coastal areas of British Columbia. The PMP is the maximum summer or fall storm (which would occur when the dam reservoirs are at or near full). A winter storm would fall as snow (and therefore affect the PMF through the depth of the snowpack). A spring storm would occur before snowmelt has filled the dam reservoirs, and could be controlled operationally. Summer storms that begin in the Pacific

tend to travel northward (along the mountain ranges) as opposed to eastward (across the mountain ranges). Storms can only travel eastward if a high pressure system already exists in the north (Pellerin, 2000, personal communication). These storms can be very severe in the Columbia River Basin. One such storm occurred on July 11-13, 1983. This storm began in the Pacific Ocean, and affected mainly the northern portion of the basin.

The mean annual precipitation in the basin ranges from 500 to 2500 mm. There is more precipitation in the western half of the basin, and lower precipitation in the eastern portion. This variation is due mainly to the orographic effects of the mountains. The July mean temperature for the area is 10 to 20 °C, while the January mean temperature is -20 to -10 °C (The Cartographic Department of the Clarendon Press, 1977). Approximately half of the precipitation falls during the winter as snow.

There are 32 streamflow stations operated by Water Survey of Canada on the Canadian portion of this river. Figure 2-2 shows a map of the stations, where the streamflow stations are indicated by black dots. The map also shows the locations of the four major BCHydro dams. The Mica Dam has a drainage area of approximately 20,000 km^2 and is farthest upstream. Above this point, the river is unregulated. The Mica Dam is situated just south of where the river exits the Rocky Mountain Trench. Revelstoke Dam is approximately 130 km south of Mica Dam and has a local drainage area of approximately 4000 km^2. Keenleyside Dam is located approximately 200 km further south, with a local drainage area of approximately 8000 km^2. Keenleyside Dam controls most of the river discharge that enters the United States. Duncan Dam is located on a tributary of the Columbia River (Kootenay River, which joins the Columbia River south of Keenleyside Dam), and has a drainage area of approximately 2000 km^2.

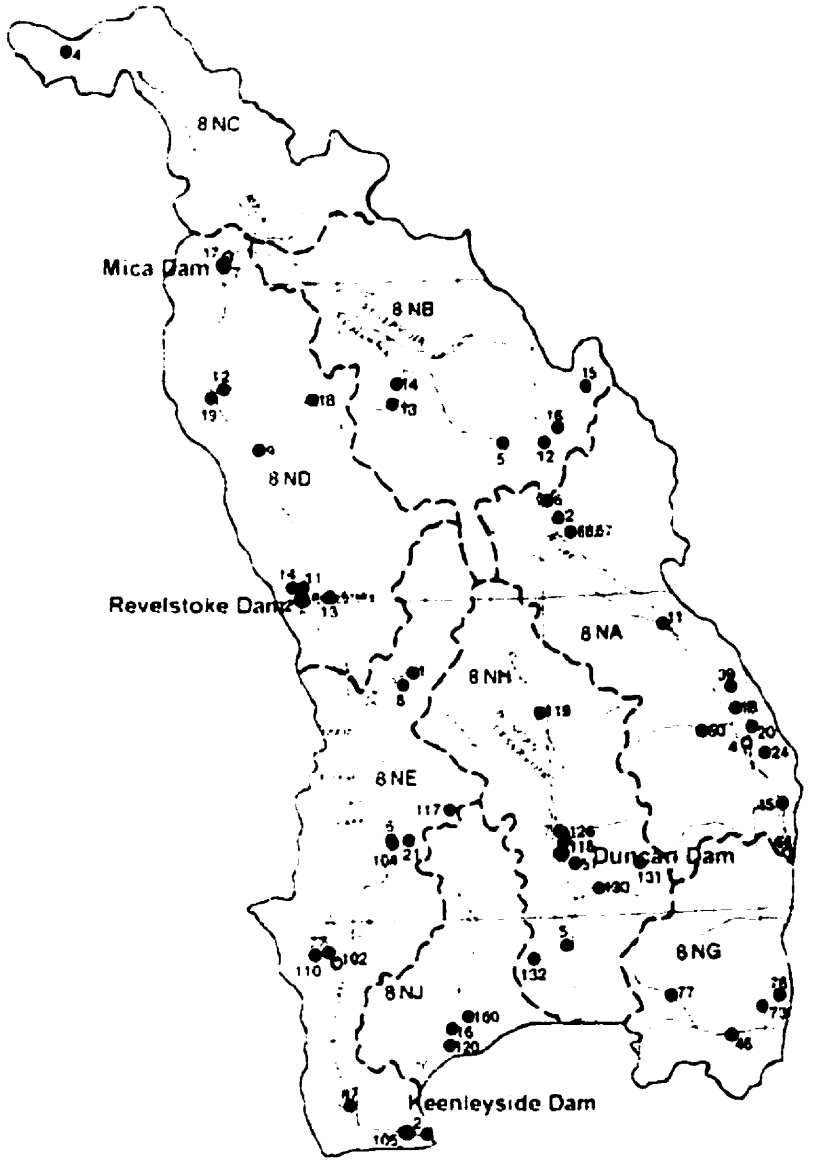


Figure 2-2 – Location of Streamflow Stations and Major Dams in the Columbia River basin (Fisheries and Environment Canada, *British Columbia Active Hydrometric Stations*, December 1977)

Table 2-1 lists the 32 streamflow stations with their WSC number, name, and years of operation, and the four BCHydro dams with the numbers of years of reservoir inflow data used in this research. The time series length of each station varied considerably, from 5 years to 91 years, with an average of 34 years. A portion of these data were used to calibrate the distributed hydrological model WATFLOOD/SPL in the separate collaborative research

project with BCHydro (Kouwen, *et al.*, 2000a). The largest streamflows ever recorded at several of the streamflow gauging stations were due to the July 11-13, 1983 storm. This storm also caused the largest recorded storm inflow (inflow that was caused by a storm) into Kinbasket Lake formed by the Mica Dam. Mica Dam is used as the example basin for the PMS developed in this research.

Table 2-1 – List of Streamflow Stations and B.C. Hydro Dams with Number of Years of Record

Name of Station	WSC #	Yrs	Name of Station	WSC #	Yrs
Columbia River at Nicholson	08NA002	91	Mather Creek below Houle Creek	08NG076	23
Kicking Horse River at Golden	08NA006	32	Gold River above Palmer Creek	08NB014	23
Spillimacheen R. near Spillimacheen	08NA011	47	Stitt Creek at the Mouth	08ND018	23
Incomappleux River near Beaton	08NE001	46	Kirbyville Creek near the Mouth	08ND019	23
Kuskanax Creek near Nakusp	08NE006	33	Kuskanax Creek at 1048m Contour	08NE117	22
Kaslo River below Kemp Creek	08NH005	38	St. Mary River below Morris Cr.	08NG077	23
Lardeau River at Marblehead	08NH007	53	Fry Creek below Carney Creek	08NH130	23
Columbia R. near Fairmont Hot Springs	08NA045	50	Cranberry Cr. Above B.C. Hydro Intake	08NE123	5
Columbia River at Donald	08NB005	51	Keen Creek below Kyawats Creek	08NH132	22
St. Mary River near Marysville	08NG046	48	Lemon Cr. Above South Lemon Cr.	08NJ160	23
Barnes Creek near Needle	08NE077	45	Gold River above Bachelor Creek	08NB013	21
Beaton Creek near Beaton	08NE008	43	Blaeberry River below Ensign Cr.	08NB015	22
Goldstream R. below Old Camp Cr.	08ND012	33	Split Creek at the Mouth	08NB016	22
Duncan River below B.B. Creek	08NH119	33	Carney Creek below Pambrun Cr.	08NH131	23
Illecillewaet River at Greeley	08ND013	32	Arrow Dam		13
Jordan River above Kirkup Creek	08ND014	25	Duncan Dam		13

Blaeberry R. above Willowbank Cr.	08NB012	26	Mica Dam	23
Canoe River below Kimmel Creek	08NC004	23	Revelstoke Dam	13

Figure 2-3 is an image of the basin generated from Landsat MSS (Multi-Spectral Scanner) imagery, taken in 1982. The lighter shades correspond to barren areas and glaciers, while the darker shades correspond to agricultural land uses such as crops and forests. The predominant landcover in the region is sub-alpine forest (particularly in the Rocky Mountain Trench, above Mica Dam). There is relatively little farming or commercial development in the region; the development is mainly in the low areas near the Columbia River and the Kootenay River. However, the forests are commercially exploited.

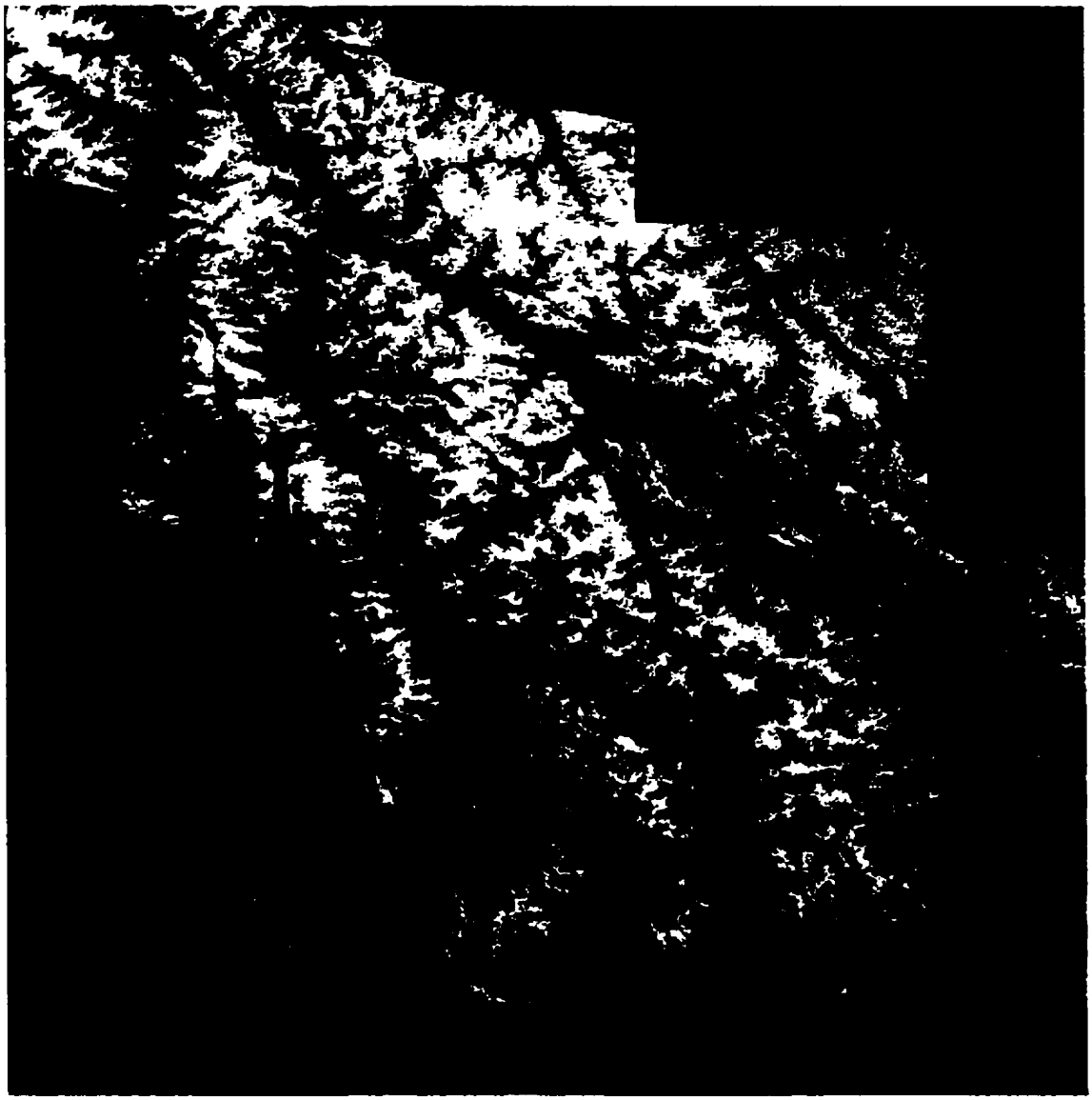


Figure 2-3 – Landsat MSS image for the Columbia River Basin

3 Background and Literature Review

This Chapter introduces the background and relevant literature for this research. The Chapter is organized into four sections, which introduce the method to estimate Probable Maximum Precipitation, Numerical Weather Models, Hydrological Models, and Regional Frequency Analysis, respectively. The relevant literature is cited within each section. In addition, the methods and models used in this research are presented, and any preliminary work (such as model calibration and validation) is described.

3.1 Probable Maximum Precipitation (PMP) Estimation

3.1.1 World Meteorological Organization (WMO) Method

The Probable Maximum Precipitation (PMP) is defined as "the greatest depth of precipitation for a given duration meteorologically possible for a given size storm area at a particular location at a particular time of year, with no allowance made for long-term climatic trends" (WMO, 1986). The World Meteorological Organization (WMO) has established a manual of estimation techniques for the Probable Maximum Precipitation (PMP). This section gives a very brief description of the method; further detail can be found in the WMO manual (WMO, 1986).

Often, a PMP analysis is performed for a large area, and the spatial variation in PMP estimates is necessary. For instance, Miller (1993) prepared PMP estimates for the entire Columbia River Basin above the Lower Border Dam site (an area of 155,700 km^2). To calculate the PMP estimate for a particular subwatershed, it is necessary to know the spatial variation in PMP estimates. One common method is to choose an area that represents point precipitation and thereby to subdivide the watershed into a number of points or grids (e.g. Miller, 1993, used 10 km by 10 km grids). The PMP analysis is then carried out for each

gridpoint, and an index map of the PMP estimates is prepared for a particular duration. It is frequently necessary to smooth the PMP estimates prior to drawing the isohyets on the index maps. These index maps are difficult to make; generally only one index map is prepared. Curves are prepared to adjust the index map values for different sized areas and storm durations. Therefore, the steps to calculate the PMP for a particular subwatershed are: create an index map from the PMP analyses for each gridpoint; calculate the average PMP estimate for the subwatershed from the index map; adjust the PMP estimates for the desired duration from the depth-duration curve; and adjust the PMP estimates for the desired area from the depth-area curve (the subwatershed has a different area than the area of each gridpoint).

The following describes the PMP method for one point. However, as described above, it is often applied over the entire area of interest, and an index map is prepared to show spatial variation. The method for flat terrain will be presented first, and the adjustments for mountainous terrain will follow.

The analysis of precipitation data to determine the PMP for flat terrain involves several tasks. The first task involves establishing the storm database, and determining which storms can be transposed into the area. The storms are evaluated to determine their depth-area-duration characteristics, and the storms are maximized using the precipitable water. Finally, envelope curves are drawn around the maximized storms to develop the PMP estimates. These tasks will be explained in greater detail below.

The reliability of PMP analyses depends on the adequacy of the storm samples. A small number of storms are not likely to yield an accurate estimate of maximum precipitation. However, the historical data may show that only a few, or even no, storms have passed directly over the basin (only storms of a sufficient size are included in the

analysis). Therefore, WMO recommends transposing storms to the area. When this is done, it is necessary to determine transposition limits, based on the meteorology and topography of the surrounding area. In essence, the transposition limits define the limits of storms that "could have" occurred over the area. They can be transposed using only minor modifications to their rainfall amounts. The storm database consists of all storms that either occurred over the area or occurred within the transposition limits.

A depth-area-duration analysis must be performed for each storm in the storm database. Depth-area-duration analysis is described in another WMO manual (WMO No. 237, TP 129). The analysis generates a table for each storm, detailing the depth of precipitation for each combination of area and duration.

The storms are maximized according to the maximum precipitable water available in the area. The precipitable water in the atmosphere is calculated from the dew point and the elevation. The first step is to find the seasonal variation in the maximum persisting 12-hour 1000-mb dew point for the area. The historical database is examined to find the maximum persisting 12-hour 1000-mb dew point data for each day of the year. The maximum dew point data are plotted, and an envelope curve showing a smooth annual variation is drawn. For a particular storm, the maximum dew point is read from the envelope curve according to the date the storm occurred. The method also requires the 12-hour 1000-mb dew point for the storm. These two dew points are used to calculate the precipitable water at the time of the storm and the maximum precipitable water available for that date. The storm precipitation is scaled up according to the ratio of the precipitable water observed during the storm and the maximum precipitable water. For a storm that requires transposition, the storm precipitation is scaled up according to the ratio of the precipitable water available at the time

and original location of the storm, and the maximum precipitable water available at the transposed location.

Envelope curves of the maximized storms are used for the PMP. The storms are plotted as depth vs. duration and depth vs. area, and a smooth line that connects all of the maximum values is drawn. It is not necessary that one storm provides the maximum depth at all durations and areas, and in fact, it is very unlikely. A maximum may be "undercut" if there is sufficient reason to be suspicious of the value (such as, it is close to the transposition limits and may not be fully transposable). These curves become the PMP estimate, for each combination of area and duration.

The PMP estimates in mountainous regions are complicated by orographic influence on precipitation. There are a large number of methods to deal with orographic precipitation, and each region requires a different approach. The most common method to estimate the PMP is to remove the orographic influence from the storm data, perform the PMP analysis on the non-orographic (or convergence) precipitation, maximize the orographic influence, and then add the maximized convergence precipitation and the maximized orographic influence to form the PMP. The orographic influence may be removed in several ways. The recommended method is to model the storm in an atmospheric model with topography, and then to re-model the storm without the topography. This method can only be used if the atmospheric model has been validated with several storms. For instance, Miller (1993) tried to use a boundary layer model to calculate the orographic precipitation, but was unsuccessful in validation. This method, despite its problems, is recommended by the WMO because it holds the greatest promise for reliable orographic precipitation estimates. Another method is to evaluate the storms meteorologically and to separate the convergence and orographic

components of the storms. For instance, the valley precipitation can be used to estimate the convergence precipitation. If there is no valley precipitation available, the free-atmospheric forced precipitation must be calculated (the precipitation that would have occurred if there were no mountains). The PMP analysis is then performed on the convergence precipitation from each storm. The maximum orographic precipitation may be calculated by an orographic model, but this model may only be applied in certain "simple" mountain areas (e.g. a single continuous ridge). Another method for calculating the orographic precipitation is to calculate an orographic intensification factor. This factor combines the storm intensification effects and the orographic effects. The total PMP precipitation is then the convergence precipitation multiplied by the orographic intensification factor (e.g. Miller, 1993, used this method). The addition of topography makes the estimation procedure very complicated.

3.1.2 Physically-based Estimation Techniques

This research examined a method to determine the PMF using physically-based atmospheric and hydrologic models. The emphasis in the research was to ensure that the atmospheric and hydrologic processes were modeled accurately and that the estimated PMF was physically possible.

The WMO method for calculating the PMP for a watershed is based partly on meteorological processes, and partly on statistics. However, there are some problems with the method, such as: differences in availability and quality of data; site specific issues such as topography; and the simplifying assumptions about atmospheric processes. For instance, the method assumes linearity in the storm maximization and transposition procedure (i.e. that precipitation increases as moisture increases) and that topographic and convergence

precipitation can be separated. Abbs (1999) used a numerical weather model to test these (and other) assumptions, and found that they are generally not valid. Jarrett and Costa (1988) showed that storm transposition from a low elevation to a higher elevation in Colorado is unfounded by meteorological, hydrological and paleoflood information. The use of envelope curves to define the maximum curves assumes that the data set contains at least some of the true maximum points. There have been instances where a PMP estimate was exceeded, with the result that the PMP was revised upwards (e.g. Klemeš, 2000). However, there have also been instances where the PMP magnitude depended on data that were later determined to be erroneously high (e.g. Jarrett and Crow, 1988) or the magnitude was unsupported by paleohydrologic research (e.g. Pruess, *et al.*, 1998). The most important assumption is that the PMP involves a type of precipitation system that has been observed in the past. It is not known whether a storm system will behave similarly to observed systems at the levels of precipitation that must be predicted. These problems may reduce the accuracy of the PMP estimates and so the WMO recommended using caution when applying its PMP calculation method (WMO, 1986). It is recommended (and frequently necessary) to examine alternate methods for a particular site. In the particular case of mountainous regions, the WMO recommended using atmospheric models to calculate orographic effects (WMO, 1986). It was therefore reasonable to investigate the use of an atmospheric model to derive the PMP.

This research examined the possibility of estimating the PMP with an atmospheric model, and using it to calculate the PMF. Many atmospheric models are available in the literature. For this research, two atmospheric models were used. The first model, called the High-Resolution Boundary Layer model, estimated a 96-year historical time series of precipitation and temperature (Danard and Galbraith, 1996b) (see Section 3.2.3), in addition

to the 100-year melting temperature sequence and the 100-year snowpack depth (Danard and Galbraith, 1995). These data were used to calibrate the hydrological model and as antecedent conditions for the PMP, respectively. The second model, called MC2-PMS (Benoit, *et al.*, 1997a and 1997b), is a forecast model used in simulation mode to develop a large storm. An atmospheric perturbation was used as the initial conditions for the model. The model controlled the development of the storm and calculated the precipitation. The model physics ensured that the storm was physically possible. These two models are described in Section 3.2.

This research also required a hydrological model to develop the PMF from the PMP and the antecedent conditions. Since the hydrological model would be accepting input from the atmospheric models, a grid-based hydrological model was advantageous. In addition, it was required that the hydrological model be physically based or at least behaviorally correct. This model is described in Section 3.3.

Thirdly, this research required flood frequency curves, to estimate the severity of the PMF estimates. However, the observed streamflow database was very short, and therefore the 96-year simulated streamflow time series from the hydrological model was used. The regional frequency analysis method by Hosking and Wallis (1997) was used, and is described in Section 3.4.

3.2 Numerical Weather Modeling

The knowledge of atmospheric processes has improved over time and computers have become more powerful and it has become possible to numerically model the atmosphere at a useful resolution (e.g. 10 km or better). This section introduces numerical weather modeling,

and in particular, the two atmospheric models used in this research. Atmospheric modeling is placed in context, and then the two models are described in turn.

3.2.1 Use of atmospheric models in the literature

There are two main categories of atmospheric models, which have arisen simultaneously to cover different needs. Climate models are in the first category. They are used mainly for long-term climate change predictions. These are often calculated with General Circulation Models (GCMs) using low resolution grids over a large portion of the globe (see, for example, Mimikou, *et al.*, 2000). The emphasis is usually on modeling the atmospheric physics. These models are not used in this research and will not be discussed further.

Various weather models make up the second category of atmospheric models. These models operate on a high-resolution grid to be able to model local weather phenomena. The emphasis in these models is on the atmospheric dynamics and data assimilation. Some weather models are used for short-term weather modeling, and others are used to fill-in historical data.

There are several Canadian atmospheric models produced by the Canadian Meteorological Center that can be used for short-term weather modeling. The models include: the Regional Finite Element (RFE) model (J. Mailhot, *et al.*, 1997), the Global Environmental Multiscale (GEM) model (Côté, *et al.*, 1998), and the Mesoscale Compressible Community (MC2) model (Benoit, *et al.*, 1997a). These models generally operate in a forecast mode, where the initial atmospheric conditions are specified and the model physics are used to predict future weather conditions (currently, the GEM model is used for operational weather forecasting in Canada).

Another application for weather models is the simulation of past historical events. Often the atmospheric models are linked with hydrological models; the atmospheric models are used to re-create historical weather scenarios and the hydrological models estimate the streamflows. Several authors have used simulated historical weather data in hydrological models (for instance, Kouwen, *et al.* (2000), Lamb (1999), Z. Yu, *et al.* (1999), Kite, *et al.* (1995)). Some atmospheric models are limited by their need for extensive initial conditions and boundary conditions, and so only brief periods of time may be generated (e.g. Z. Yu, *et al.*, 1999, used a model to generate data during three storms, the earliest of which was in 1984). Other simpler atmospheric models (e.g. Kouwen, *et al.*, 2000 used data from a simple atmospheric model by Danard and Galbraith, 1996b) can generate data for longer periods of time.

3.2.2 Mesoscale Compressible Community – Probable Maximum Storm (MC2-PMS) Model

Two different atmospheric models were used in this research. The first model was the Mesoscale Compressible Community (MC2) model, developed by Recherche en Prévision Numérique (RPN) (Benoit, *et al.*, 1997a), which was used to develop extreme rainfall estimates. The model requires initial atmospheric conditions, and calculates the forecasted atmospheric conditions. RPN has developed a module called the Probable Maximum Storm (PMS) module (Benoit, *et al.*, 1997b), which was added to the MC2 model. The PMS module calculates the initial conditions for an extreme storm, and the MC2 model (in forecast mode) develops the storm. The MC2 model physics ensure that the storm complies with the physical laws governing atmospheric processes. This section describes the MC2 model, and then the PMS module is described.

3.2.2.1 MC2 model description

The Meso-Scale Compressible Community Model (MC2) developed by Recherche en Prévision Numérique (RPN) is an atmospheric model that can run at multiple resolutions and is suitable for fine-resolution weather forecasting and simulation. The model is described by Benoit, *et al.* (1997a). MC2 is a non-hydrostatic, finite difference, semi-Lagrangian, limited area model. The horizontal variables are distributed in a polar stereographic map projection, and the vertical variables are distributed according to a modified Gal-Chen height coordinate (Gal-Chen and Sommerville, 1975). The semi-implicit, semi-Lagrangian integration method for time was found to be more stable than other integration methods for mountainous terrain (Pinty, *et al.*, 1995). This research used version 4.7 of the MC2 model.

The physics package (described in J. Mailhot, *et al.*, 1998) was programmed using a modular philosophy. Depending on the resolution of the simulation, the modules used for each physical process can be changed. For instance, a low resolution simulation may use one module for a particular process. A higher resolution simulation would use a different module for the same process (one that models the process better at a fine scale). The user may specify the modules to be used for each simulation. In this way, the physics package can be used at multiple resolutions. The atmospheric physics package includes modules for the following atmospheric processes:

- Turbulent Vertical Diffusion
- Gravity Wave Drag
- Cloud Processes
- Condensation Processes
- Deep and Shallow Convective Processes
- Land Surface Processes
- Water Surface Processes
- Infrared Radiation
- Solar Radiation

The MC2 model has been validated by a number of authors in a number of regions. A few examples are listed here. W. Yu, *et al.* (1998) compared the MC2 precipitation estimates to Doppler Radar precipitation in Quebec, and found reasonable agreement. Lackman, *et al.* (1998) validated MC2 in the Mackenzie River Basin; the model was found to be able to reproduce precipitation events. Desjardins, *et al.* (1998) validated the sea surface temperatures near Nova Scotia. Kouwen and Innes (2000, 2001) compared the ability of MC2 data and radar data to produce accurate flood forecasts for the RAPHAEL (*Runoff and Atmospheric Processes for Flood Hazard Forecasting and Control* Program in Italy) and MAP (*Mesoscale Alpine Project* in Switzerland and Germany) projects. The MC2 data was found to produce plausible streamflow forecasts. Benoit, et al. (1997b), as part of a collaborative research project for BCHydro, validated the model for the Columbia River Basin. In all of these papers, it was found that MC2 tended to have a timing problem: predicted storms were offset in time from when they were actually observed. However, apart from the time offset, the precipitation distributions in time and space were reasonable and sometimes very good. In this research, the time offset did not affect the results, since the goal was to generate a large storm for insertion into WATFLOOD/SPL – the start and end times were not important.

3.2.2.2 PMS module description

The Probable Maximum Storm module was developed as part of a collaborative research project for BCHydro, and is described in a report by Benoit, *et al.* (1997b). The Probable Maximum Storm (PMS) module creates an atmospheric perturbation, and embeds the perturbation in a real day. The MC2 model develops the perturbation and calculates the

precipitation distribution. When the PMS module is used, the model is referred to as MC2-PMS.

The PMS module calculates an atmospheric perturbation that consists of a temperature wave and a pressure wave. The atmospheric perturbation tends to develop into a cyclonic storm when it is superimposed on an unstable zonal current (such as the jet stream) (Benoit, *et al.*, 1997b). The jet stream is located at different latitudes during the summer and winter seasons, and conditions for developing extreme storms are therefore different for these two seasons. The PMS module used the equations from Nuss and Anthes (1987) to calculate the atmospheric perturbation. These equations combine meteorological constraints with sinusoidal equations to describe the waves. The temperature and pressure waves are out of phase with one another to allow cyclogenesis to occur. The waves slope with height to account for wind effects at higher altitudes. Various parameters modify the equations, and manipulate the perturbation. Benoit, *et al.* (1997b) developed the PMS module but did not determine the appropriate settings for the development of an extreme storm. This research continued the work of Benoit, *et al.* (1997b) and determined appropriate settings for an extreme storm occurring in the warm season. The PMS module settings are described below.

The perturbation is embedded into data for a real day. When the MC2 model is in forecast mode, it runs as a hemispheric model for the first simulation (approximately 18,000 km by 18,000 km). The atmospheric perturbation derived by the PMS module was only 9000 km (east-west) by 5000 km (north-south). (The size of the perturbation may be adjusted in the PMS module.) Therefore, the perturbation was embedded in a real day so that initial atmospheric values would exist at all the points within the hemispheric domain. The RPN database contains several years of data that can be used to initialize the model. This

research used the general circulation for July 13, 1983 as the initial and boundary conditions for the model. This date was used because a large historical storm occurred on this date (July 11-13, 1983); it caused the largest recorded reservoir inflows at Mica Dam (due to a storm); and affected the northern part of the Columbia River Basin.

The July 11-13, 1983 storm developed in the Pacific Ocean. Therefore, the perturbation was originally placed in the Pacific Ocean in the location of the historical storm. The location of the perturbation was modified to find the optimal location for the storm.

The atmospheric perturbation is calculated as a 5-step procedure (Nuss and Anthes, 1987). Each step will be described below.

The first step is to calculate a pressure wave for a particular reference height. The pressure wave is initialized at the 5 km level, and is calculated as the total of: a reference pressure, an east-west variation, and a north-south variation:

$$P = P_o + \Delta P_x + \Delta P_y \qquad (3\text{-}1)$$

where: P_o is the reference pressure, and ΔP_x and ΔP_y represent the east-west and north-south variations respectively.

The east-west variation consists of an asymmetric sinusoidal pressure perturbation that is forced to zero on the northern and southern boundaries:

$$\begin{aligned} \Delta P_x &= a_x c_p(x) G_p(y) \sin\left[\frac{2\pi x}{L_x} + \Phi_2(y)\right] \\ c_p(x) &= d_1\left[d_2 + \sin\left(\frac{2\pi x}{L_x}\right)\right] \\ G_p(y) &= \sin\left(\frac{\pi y}{L_y}\right) \end{aligned} \qquad (3\text{-}2)$$

where a_x, L_x, $\Phi_2(y)$, d_1, d_2, and L_y are constants. The function $c_p(x)$ controls the asymmetry in the wave, and the $G_p(y)$ function forces the equation to zero at the northern and southern boundaries. The variable a_x refers to the amplitude of the east-west pressure wave. L_x and L_y are the lengths of the wave in the east-west and north-south directions respectively. $\Phi_2(y)$ is a phase relation in the north-south direction (assumed constant). Finally, d_1 and d_2 control the amplitude of the asymmetry in the east-west wave.

The north-south variation creates a pressure differential across the perturbation, and includes a streak caused by the jet stream along the flow:

$$\Delta P_y = -a_{y1} \tanh\left[\frac{y - y_c}{p_1 F_j(x) \cdot dy}\right] - a_{y2} \tanh\left(\frac{y - y_c}{p_2 \cdot dy}\right) \quad \text{.....(3-3)}$$
$$F_j(x) = 1 - b \sin\left[\frac{2\pi x}{L_x} + \Phi_1(y)\right]$$

where a_{y1}, a_{y2}, y_c, p_1, p_2, dy, b, and $\Phi_1(y)$ are constants. The $F_j(x)$ function creates the jet streak by varying the pressure differential along the east-west direction. The variables a_{y1} and a_{y2} are amplitude coefficients for the pressure differential. The variable y_c is the center of the domain in the north-south direction. The coefficients p_1 and p_2 affect the density of the pressure differential, and dy is the grid spacing. The variable b affects the amplitude of the jet streak, and $\Phi_1(y)$ is a phase relation in the north south direction (assumed constant).

The second step is to calculate the temperature wave for all levels. The temperature wave is similar to the pressure wave, except that a vertical variation is included. The vertical variation allows the temperature wave to be calculated on all levels. The temperature wave, therefore, is a total of: the surface temperature, an east-west variation, a north-south variation, and a vertically varying lapse rate.

$$T = T_o + \Delta T_x + \Delta T_y + \gamma(z)\Delta z \quad \text{.....(3-4)}$$

The east west variation is calculated as an asymmetric sinusoidal function that is forced to zero at the northern and southern boundaries. However, it also includes vertical variation in the wave amplitude and phase.

$$
\begin{aligned}
\Delta T_x &= b_x D(z) c_T(x) G_T(y) \sin\left[\frac{2\pi x}{L_x} + \Phi_T(y,z)\right] \\
D(z) &= \frac{1}{2} - \frac{1}{2}\tanh\left(\frac{z - z_o level}{dz}\right) \\
c_T(x) &= d_{t1}\left[d_{t2} + \sin\left(\frac{2\pi x}{L_x}\right)\right] \qquad \text{.....(3-5)} \\
G_T(y) &= \sin\left(\frac{2\pi y}{L_y}\right) \\
\Phi_T(y,z) &= \frac{\pi}{2}\left[\frac{z_o^2 - (z - z_o)^2}{z_o^2 - (z_R - z_o)^2}\right] + \Phi_2(y)
\end{aligned}
$$

where b_x, L_x, dz, d_{t1}, d_{t2}, L_y, z_R, and $\Phi_2(y)$ are constants, and z_o is the maximum phase difference. The functions $c_T(x)$ and $G_T(y)$ are analogous to $c_p(x)$ and $G_p(y)$ for the pressure wave, and their constants have analogous meanings. D(z) and $\Phi_T(y,z)$ introduce vertical variation in the amplitude of the wave and in the phase of the wave, respectively. The variable b_x refers to the amplitude of the temperature wave in the east-west direction (analogous to a_x), d_{t1} and d_{t2} are analogous to d_1 and d_2, and L_x and L_y are identical to L_x and L_y for the pressure wave. The variable dz is the vertical grid spacing. Z_R is the reference pressure level (5 km). $\Phi_2(y)$ is the phase relation in the north-south direction (assumed constant).

The north-south variation in temperature is similar to the north-south variation in pressure, but includes a term to introduce an intense low-level temperature front near the surface:

$$\Delta T_y = -b_{y1}\tanh\left(\frac{y-y_c}{p_{b1}\cdot dy}\right) - b_{y2}\tanh\left(\frac{y-y_c}{p_{b2}\cdot dy}\right) + F_R(x,y,z)$$

$$F_R(x,y,z) = f(z)\left[\sin\left(\frac{2\pi x}{L_x}+\pi\right)+\sin\left(\frac{2\pi y}{L_y}\right)\right]\sin^2\left(\frac{\pi x}{L_x}\right)\sin^2\left(\frac{\pi y}{L_y}\right) \quad \text{.....(3-6)}$$

$$f(z) = a_f\left[\frac{1}{2}-\frac{1}{2}\tanh\left(\frac{2z-3}{dz}\right)\right]$$

where b_{y1}, b_{y2}, y_c, p_{b1}, p_{b2}, and dy are constants that are analogous to a_{y1}, a_{y2}, y_c, p_1, p_2, and dy for the pressure wave (there is no jet streak term for the temperature wave). The $F_R(x,y,z)$ function produces the low-level temperature front near the surface, and it decreases as height increases. The variable a_f is a constant to describe the rate of decrease with height.

The vertical variation in lapse rate is calculated as a parabolic function:

$$\gamma(z)\Delta z = 2sk - 2s[k(z+k)]^{1/2}$$

$$k = \frac{\left(\frac{\Delta T}{2s}\right)^2}{z_o - 2\left(\frac{\Delta T}{2s}\right)} \quad \text{.....(3-7)}$$

where s and ΔT are constants to indicate the lapse rate at the surface and the temperature difference between the surface and 10 km.

The third step is to specify the three-dimensional moisture structure for the relative humidity. The moisture field is arbitrary since there are no atmospheric constraints for a particular wave. The humidity fields used in this research are described in Section 4.1.1.

The last two steps are to calculate the pressure on all levels and to calculate the winds on all levels. The hydrostatic equation describes the relationship between pressure and temperature in the atmosphere, and so it is used to integrate pressure on all levels. Finally, the winds are calculated from the temperature and pressure, using the nonlinear balance equation.

Table 3-1 summarizes the parameters for the pressure and temperature waves, along with typical values for those parameters. These may all be adjusted to develop different storms. Preliminary research at RPN (Pellerin, 2000, personal communication) indicated that the perturbation was most sensitive to a_x, b_x, and s. In this research, these parameters were varied within the ranges given to generate extreme storms, as described in Section 4.1.

Table 3-1 - Parameters for the Atmospheric Perturbation

Pressure Wave Variables		**Temperature Wave Variables**	
Po (reference pressure)	500 mb	To (reference temperature)	289 K
Lx (length in x)	4000 km	Lx (length in x)	4000 km
Ly (length in y)	5000 km	Ly (length in y)	5000 km
Amplitude factors		Amplitude factors	
a_x	3-15 mb	b_x	5-15 °C
a_{y1}	10 mb	b_{y1}	12 °C
a_{y2}	18 mb	b_{y2}	7 °C
		a_f	1 °C
Packing and Structure		Packing and Structure	
p_1	9	p_{b1}	11
p_2	12	p_{b2}	8
d_1	0.25	d_{t1}	0.25
d_2	3	d_{t2}	3
B	0	ΔT	60 °C
		s	6.5-10 °C/km
Phase relations		Phase relations	
$\phi 1$	$3\pi/8$	$\phi 2(y)$	π
$\phi 2(y)$	π		
From: *Recherche en Prévision Numérique (1997)*			

The following figures show the atmospheric perturbation created with the above parameters. The atmospheric variables (temperature, wind speed, etc) are available at a number of pressure levels, here they are shown at surface level. The pressure and temperature waves in Figure 3-1 and Figure 3-2 are orthogonal to each other. The pressure wave is low (-18 mb), high (+26 mb), low (-23 mb). The temperature wave is high, low,

high. Figure 3-3 shows the wind speed and direction. Note that the waves have created two circular wind patterns. MC2 will develop the cyclone.

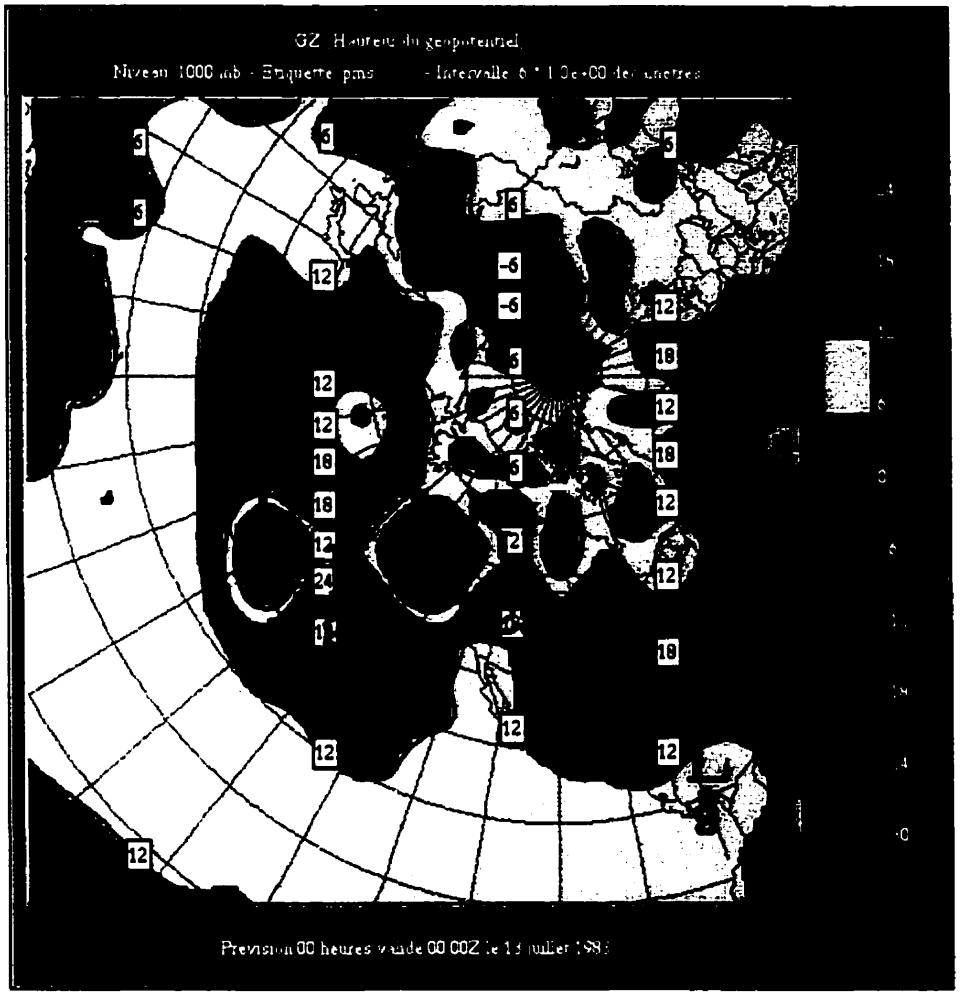


Figure 3-1 – Atmospheric Perturbation: Pressure Wave (Domain: 18,000 by 18,000 km)

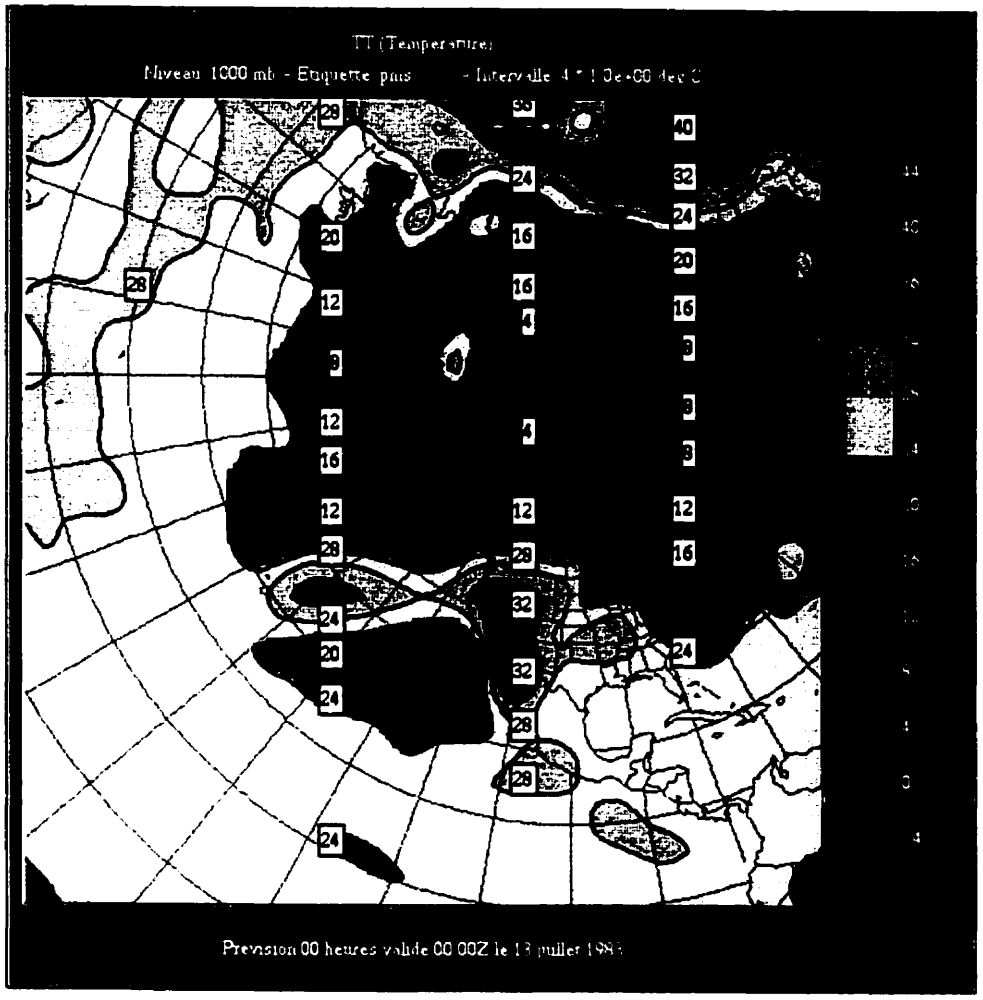


Figure 3-2 – Atmospheric Perturbation: Temperature Wave (Domain: 18,000 by 18,000 km)

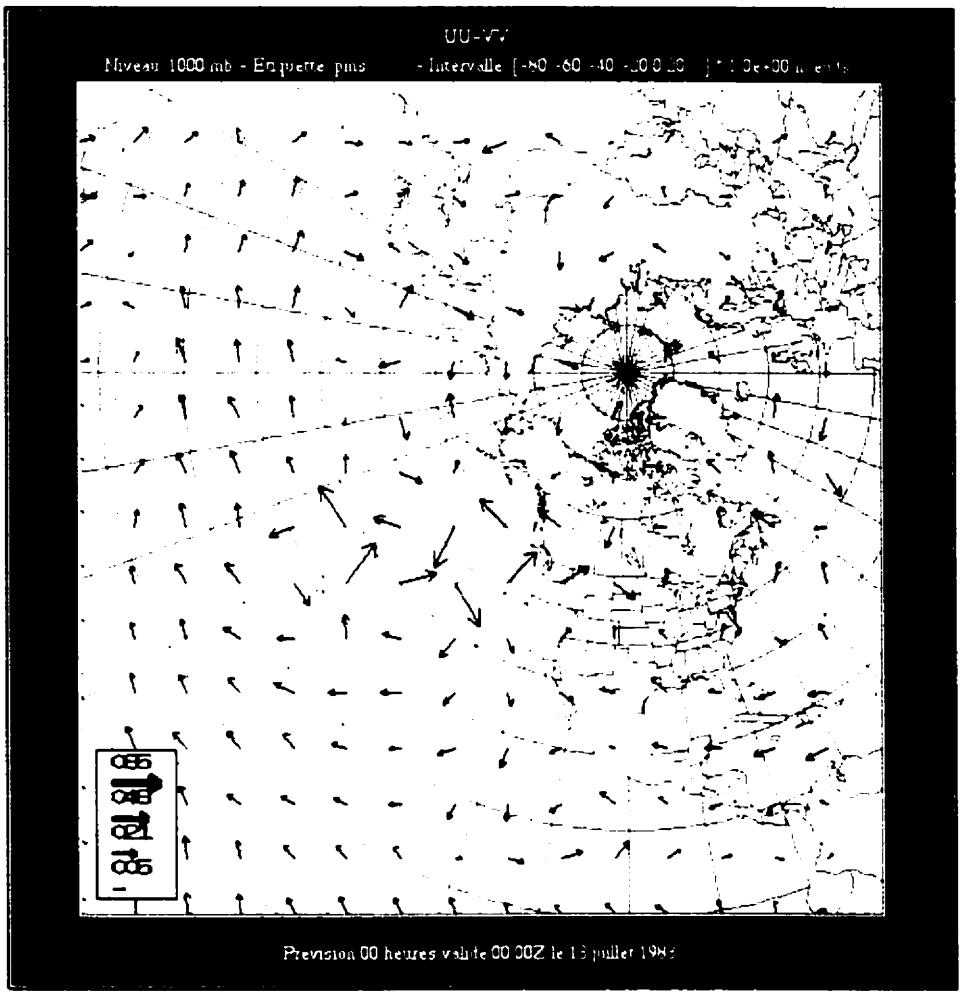


Figure 3-3 – Atmospheric Perturbation: Wind Speed and Direction (Domain: 18,000 by 18,000 km)

3.2.3 High-Resolution Boundary Layer Model

The second atmospheric model used in this research was the High-Resolution Boundary Layer (HRBL) Model by Atmospheric Dynamics Corporation. It was used to model temperature and precipitation for the years 1899-1994 (inclusive). Danard and Galbraith also used this model to determine the 100-year snowpack and 100-year melting temperature sequence. The model is described in a series of reports by Danard and Galbraith (1994, 1995, 1996a (with Davies), 1996b, and 1997), and is described briefly in this Section.

The model calculations were performed on a 2' (latitude) by 4' (longitude) grid (approximately 3.7 km by 4.7 km). The outermost latitudes and longitudes were 48°4'N, 52°2'N, 115°8'W, and 119°4'W (120 rows by 60 columns). A border region of five grid squares was added on all sides to account for boundary effects, so the calculation grid was 130 rows by 70 columns. The model was used to calculate gridded temperature and precipitation for the time period from 1899 to 1994.

There were two steps in the model: a trial field was calculated from the methods described below, and then objective analysis was used to fit the trial field to the observed meteorological station values. The physical equations used in the trial field calculations and the objective analysis method are presented in the following sections.

The input data for the trial field were generated from the 190 km LFM grid (Limited area, Finite Mesh grid by U.S. National Centers for Environmental Prediction). The LFM data were interpolated to the 2' by 4' grid with bi-cubic splines. However, the LFM data were only available from 1971 to 1994. Prior to this time, other data sources were interpolated to the 190 km grid. When no other data were available, an analog method was used to fill in missing data.

3.2.3.1 Maximum and Minimum Temperature Calculation

This Section describes the calculation of the trial field for the daily maximum and minimum temperature.

The LFM data at 0000 UTC were used to calculate the maximum temperatures, and the data at 1200 UTC were used for minimum temperatures. The following data were used to calculate temperatures: 850 mb temperature, 700 mb temperature, 850 mb height, grid point height, and surface pressure.

If the surface pressure was below 850 mb (i.e. the ground was above the 850 mb level), the temperature was found using the 700 mb and 850 mb temperatures. The 700 mb and 850 mb temperatures (from the LFM data) were interpolated to the grid points using bi-cubic spline interpolation. The temperature was then interpolated vertically at each grid point by assuming a linear variation of temperature with height.

If the surface pressure was above the 850 mb level, a sea-level temperature was calculated by assuming the hydrostatic equation applies:

$$T = \frac{2gZ_{85}}{R\ln\left(p/p_{85}\right)} - T_{85} - 546.4 \qquad (3\text{-}8)$$

where g is the acceleration of gravity, R is the gas constant for 1 kg of dry air, Z_{85} is the 850 mb height, p and p_{85} are the sea-level and 850 mb pressures, and T_{85} is the 850 mb temperature. (The value of 546.4 (2×273.2) is used to convert °K to °C.) Bi-cubic spline interpolation was used to interpolate the sea-level and 850 mb temperatures to the grid points, and then the temperature was interpolated vertically, as before.

After the temperatures were calculated at all the grid points and meteorological stations as described above, the average bias was calculated. The average difference between the model estimates and the observations at the meteorological stations was calculated, and subtracted from the model estimates for the entire grid. The bias was generally negative for minimum temperatures (the model underestimated minimum temperature), and positive for maximum temperatures (the model overestimated maximum temperature). The model estimates with the bias removed constituted the trial field.

The trial field was then modified with the objective analysis described in Section 3.2.3.3.

3.2.3.2 Precipitation Calculation

The model also calculated daily precipitation using the horizontal moisture convergence as a predictor. Danard (1971) found that this was a useful predictor for precipitation in a mountainous terrain. The horizontal moisture convergence was found with:

$$C = -\Delta z \nabla \cdot q\rho\bar{V} \qquad (3\text{-}9)$$

where Δz = 1000m (the assumed thickness of the layer), and q, ρ, and V are the specific humidity, density, and wind velocity respectively. The calculations were performed on the surface ς=1500m ($\varsigma = Z - Z_s$) where Z is the height above sea-level and Z_s is smoothed terrain elevation. This surface follows the terrain. The q, ρ, and V were all smoothed to avoid noise. Downwind displacement of precipitation was accounted for by advecting C with the 700 mb wind.

Since the LFM data were available at 1200 UTC and 0000 UTC, an average was used for the climate day. The climate day was calculated by combining C at 1200 UTC, at 0000 UTC, and 1200 UTC the following day with weights of 0.25, 0.5, and 0.25 respectively.

Once the climate day C was calculated, the precipitation was found using the following equation:

$$P_m = a_0 + a_1 C + a_2 C^2 + a_3(x - \bar{x}) + a_4(y - \bar{y}) + a_5 h \qquad (3\text{-}10)$$

where P_m is the model precipitation, a_0 to a_5 are coefficients determined for the day by minimizing $\Sigma(P_{obs}-P_m)^2$ for all stations, C is the climate day horizontal moisture convergence interpolated to the grid point, x and y are distances to the east and north in grid units, $\bar{x}$ and $\bar{y}$ are the coordinates of the center of the domain, and h is the grid point height. This regression was performed for each day. The coefficients were modified if the equation gave negative precipitation. The equation was prone to error, as it depended on the amount and

quality of observed precipitation on each historical day. However, the overall pattern was considered to be appropriate for the topography, and corrections based on streamflow volume were used to further improve the results.

Various corrections were performed when observed precipitation was sparse. For instance, the calculated precipitation was restrained to be within 10 mm of the maximum observed precipitation. If less than 14 stations were available, a simpler equation or the average precipitation was used instead. For details, see Danard and Galbraith (1996b).

The values of P_m were then combined with observed precipitation in the objective analysis.

3.2.3.3 Objective Analysis

The objective analysis method used in the HRBL model is described in Danard, *et al.* (1993). Objective analysis was used to make the gridded values match the observed values at the observation stations, but yet retain the shape and structure of the gridded data. This method was used to modify both the precipitation and the temperature trial fields.

The first step in objective analysis was to make the trial field match the observations at the stations. The terms used in calculating the weights to make the trial field match the observation stations are illustrated in Figure 3-4.

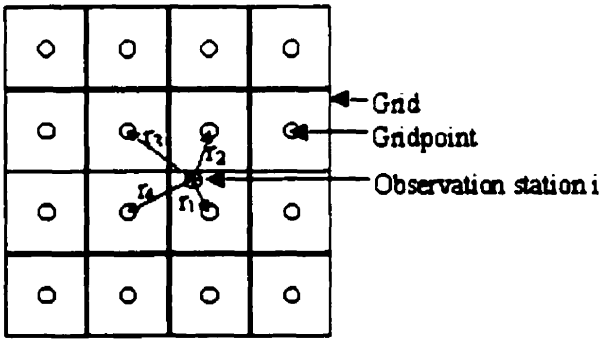


Figure 3-4 – Weight calculation to make the gridpoints match an observation station

Weights were calculated according to the distance of the station (i) to the four surrounding grid points (m):

$$W_{mi} = \frac{1 - r_m^2}{\sum_{n=1}^{4}\left(1 - r_n^2\right)} \quad m = 1,2,3,4 \qquad (3\text{-}11)$$

where r is the distance from the station to the grid point. The differences between an observation (O_i) and the trial field value (T_i), interpolated to the station location, were weighted by the above W_{mi} factors. These weighted differences were used to adjust the trial field in the following equation:

$$G^* = G + \frac{\sum_i W_{mi} \times (O_i - T_i)}{\sum_i W_{mi}} \qquad (3\text{-}12)$$

where G* is the modified field, G is the trial field, the summation is performed for all the stations i within the four squares surrounding the grid point, and m takes on the value 1,2,3,4 according to the location of the station with respect to the grid point. Only grid points that had an observation station in one (or more) squares around them were modified, but the other grid points were not yet modified. The second step of objective analysis was to adjust the unaffected grid points to reduce discontinuities caused by the first step. The unaffected grid points were adjusted so that the second derivative of the new field matched the second derivative of the trial field at all points:

$$\nabla^2 G^{**} = \nabla^2 G \qquad (3\text{-}13)$$

where G** is the new field, and G is the trial field.

The new field, G** matched the observed data at station locations, and retained the shape of the trial field, as indicated in Figure 3-5.

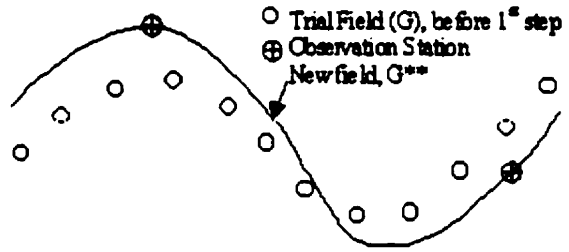


Figure 3-5 – Matching the second derivative in objective analysis

The G** field was the "best guess" for the temperature and precipitation. However, there were two large sources of error in the G** field. The first error was caused by errors in the observed temperature and precipitation. The second error was due to the shape of the trial field. For precipitation, the trial field was calculated by a regression equation (Equation 3-10) that was prone to error. The errors were minimized during hydrologic modeling with the use of a Precipitation Adjustment Factor (PAF) field (Section 3.3.3).

3.3 Hydrologic Modeling

There are many hydrological models currently in use. Refsgaard and Knudsen (1996) have listed three general groupings of hydrological models: empirical, lumped conceptual, and distributed physically-based models. An empirical model is one that is based on empirically derived equations, which have little or no physical basis. A lumped conceptual model is one that "lumps" the watershed into a single element, and uses representative descriptions of hydrologic processes. These models use parameterizations of hydrologic processes that are conceptually sound but do not model the detail of the processes in a watershed. A distributed physically-based model is one that subdivides a domain (watershed) into hydrologically significant subgroups and uses detailed physically-based descriptions of the hydrologic processes. There are, of course, models which fall between

these categories, as some hydrologic processes are not understood very well, and cannot be modeled in a physically-based manner at a suitable grid scale (Beven, 1989).

Of these three categories of hydrological models, the most sophisticated and philosophically attractive model is the distributed physically-based model. It is not, however, the ideal model for all purposes. Refsgaard and Knudsen (1996) found that lumped conceptual models and distributed physically-based models worked equally well, if there were sufficient data for calibration. The parameters for lumped conceptual models are watershed-specific, while the parameters for physically-based models are based on landcover or soil information, and can be transferred with little or no modification to another watershed. Therefore, the physically-based models were superior for cases when there were insufficient data, or there were changes in the basin (deforestation, urbanization, etc.).

This application used a hydrological model to calculate the floods produced by large theoretical storms. These storms have never occurred, and there are therefore no observed streamflow data available for calibration and validation of a lumped conceptual model. In addition, the large amounts of precipitation would render invalid any calibration based on average streamflows. Therefore, this application required a distributed physically-based model.

This research used the distributed hydrological model WATFLOOD/SPL (Kouwen, 2000). This model is an integrated data management and largely physically based streamflow simulation and forecasting model package. The model accepts rain and temperature as input, and simulates hydrologic processes to determine streamflow. The model works on a grid basis enabling it to use distributed meteorological data as input, but integrates runoff to calculate streamflow on a catchment basis.

The remainder of this section will describe the WATFLOOD/SPL model, the calibration and validation of this type of hydrological model, and the approaches for determining the level of model uncertainty.

3.3.1 WATFLOOD/SPL subroutines

The WATFLOOD/SPL modeling system consists of two parts. WATFLOOD is the data management system that includes a number of data pre-processing programs, and SPL is the hydrological simulation model.

To account for the spatial variability of the hydrological variables, WATFLOOD/SPL uses the Grouped Response Unit (GRU) method to group hydrologically similar response units (Tao and Kouwen, 1989; Kouwen, *et al.*, 1990; Kouwen, *et al.*, 1993). A GRU is a hydrologic computational unit that consists of a grouping of areas that can be expected to react similarly to meteorological conditions. The Columbia River Basin has been set up with eight different landcover types: barren area, high elevation dense forest, low elevation dense forest, high elevation light forest, low elevation light forest, glaciers, water, and impervious. LANDSAT MSS imagery from 1989 to 1991 was used to determine the landcover types. In the GRU method, all similarly vegetated areas (not necessarily contiguous) within a grid element are grouped into one aggregate response unit and called a GRU. Experience to date has shown that five to eight classes are usually sufficient to represent the variability of land cover. The hydrologic response of each class is computed according to the geometry of the grid, and the response (e.g. streamflow) is weighted according to its percent cover of that grid element or sub-watershed.

The meteorological forcing data are assumed to be uniform over the grid cell, and it is assumed that all pixels belonging to a land cover group respond in a similar way with respect

to infiltration, surface runoff, interflow, evaporation, snowmelt and drainage to ground water. Therefore, model parameters are associated with land cover class and are invariant over the modeled domain. In this way, there are very few "watershed specific parameters," only parameters pertaining to land cover, which are readily transferred to or from other watersheds. There are four parameters that are associated with the type of rivers in the modeled area and the underlying geology. These are channel roughness (for both channel and floodplain) and two groundwater flow parameters. These parameters are watershed specific, although related to the physiography of the area, and do not vary greatly.

The vertical water balance component of the WATFLOOD/SPL model is a conventional hydrological model. Where it differs is in the method that watersheds and regions are subdivided to preserve the hydrological responses of greatly differing surface areas, namely by employing the GRU or pixel grouping approach. Details of the hydrological abstractions in WATFLOOD/SPL are available in previous publications (Kouwen, *et al.*, 2000, Donald et al., 1995; Kouwen, *et al.*, 1993; Tao and Kouwen, 1989). Brief descriptions of the algorithms for snowmelt, surface storage, infiltration, soil moisture, evaporation, interception, overland flow and base flow are given below.

The snowmelt algorithm is based on the Anderson method (1976). It differs by using hourly time steps instead of days and using separate calculations for each land cover in each grid instead of basin-wide calculations. Snowcover depletion curves (SDC) are used to summarize the relationship between snowcover distribution and an average snowcover property, such as depth of water equivalent, for a given area (Donald, *et al.*, 1995). More specifically, these curves provide the amount of snow covered area for a given depth of water equivalent for each land cover class.

Surface storage is modeled according to the ASCE Manual of Engineering Practice No. 37 for the design of sanitary and storm sewers (ASCE, 1969), which gives typical values of retention for various surface types. It is assumed that the limiting value of depression storage (S_d) is reached exponentially (Linsley, *et al.*, 1949)

The Philip formula (Philip, 1954) was chosen to represent the important physical aspects of the infiltration process. A three-zone scheme is used to manage soil moisture and therefore control infiltration. The zones are:

UZ Upper zone storage (saturated)

IZ Intermediate zone storage (unsaturated)

LZ Lower zone storage (saturated)

Infiltrated water initially is accumulated in the Upper Zone Storage (UZS). Water within this layer percolates downward or is exfiltrated to nearby streams as interflow. A simple storage-discharge relation represents interflow. Upper zone to lower zone drainage is the same simple linear function as for interflow. Interflow and drainage occur simultaneously and are prorated if the amount calculated cannot be supplied from UZS. The moisture content of the intermediate zone (IZ), through the Philip formula, affects the infiltration rate of rain and melt water. It is used only as an index to provide a method of calculating the progress of the wetting front. When the temperature is less than 0 °C the soil moisture is not changed.

The specific retention of the soil in the upper zone is an analogue for the field capacity and is used to limit the amount of water that can become interflow or drain to the saturated zone. Retained water can be evaporated but not drained.

In WATFLOOD/SPL, the Priestley-Taylor method (1972), Hargreaves and Samani method (1982), and pan evaporation data can be used to calculate potential evaporation. Comparisons between these methods within the WATFLOOD/SPL system have shown that one method is not greatly superior to any other in long term simulation. The choice of the method is mostly based on the availability of the data. For the Columbia River Basin, radiation data were not available and therefore the Hargreaves equation (Hargreaves and Samani, 1982) was used to estimate the potential evapotranspiration (PET). The PET is reduced to the actual evapotranspiration (AET) with the use of three coefficients, which are functions of soil moisture availability, degree-days and vegetation type respectively. Evaporation of intercepted water is assumed to occur preferentially to soil water evaporation.

The procedure used for tracking interception storage and interception evaporation follows the model developed by Linsley, *et al.* (1949).

When the infiltration capacity is exceeded by the water supply, and the depression storage has been satisfied, water is discharged to the channel drainage system. The relationship employed is based on Manning's formula. The internal slopes (i.e., the slope of the local relief, not the average slope) of the GRUs are explicitly incorporated in this calculation while the roughness value is a parameter for each GRU (land cover class).

An exponential ground water depletion function is used to gradually deplete the lower zone storage. Ground water is replenished by recharge from the UZS. GRUs from all land covers within one grid element contribute to a single lower zone reservoir for the grid.

The total inflow to the river system is found by adding the surface runoff components, the interflow, and the base flow. These flows, along with flows produced by

upstream grids, are all added to the upstream end of the channel traversing the runoff producing grid.

SPL is a storage model and as such, all water storage quantities need to be initialized. The storages are surface (snow and/or water), upper zone, lower zone, and channel storage. The water storage on the surface is always assumed to be zero when a modeling run is initiated. This is a reasonable startup condition when a run is initiated during a dry period. Snow storage is usually initiated using snowcourse data although it is preferable to start a run when no snow is present in the watershed. Upper zone storage is initialized using the Antecedent Precipitation Index. Channel and lower zone storages are initialized using measured streamflow at a downstream gauging station. Prorated flows, based on drainage area, are used to determine the initial channel storage using the storage-discharge functions in reverse.

The WATFLOOD/SPL model was calibrated and validated for the Columbia River basin as part of the collaborative research project for B.C. Hydro (Kouwen, *et al.*, 2000). A grid size of 10 km by 10 km was chosen for this basin. This resolution was sufficient to represent the topography of the basin. The HRBL model data were available at a resolution of 3.7 km by 4.7 km for the time period of 1899 to 1994, and these were aggregated to form 10 km by 10 km estimates. The observed streamflow were obtained from Water Survey Canada and B.C. Hydro (36 stations in total). Most of these stations recorded observed streamflow during the years 1972 to 1994. Therefore, calibration was performed for the years 1981 to 1985, and validation was performed for other years. To avoid possible errors in the startup conditions for the WATFLOOD/SPL model, however, it was recommended that a two-year spin-up period be used. The simulated data from these two years were

discarded. The parameters were calibrated for five of the landclasses (barren, high elevation dense forest, low elevation dense forest, high elevation light forest, low elevation light forest). The parameters for glaciers, water, and impervious were set to textbook values. The variation in river type for this basin was represented by three river classes: valley, high elevation mountain, and low elevation mountain. The model required approximately 3-4 hours to simulate the streamflow for the entire 96-year time period for the Columbia River basin (using one CPU on an Origin 200 180 MHz computer).

The HRBL model gave daily estimates of maximum and minimum temperatures, and daily precipitation totals. When used in WATFLOOD/SPL, the maximum and minimum temperatures were assumed to occur at noon and midnight, respectively. This resulted in a slight offset from the true maximum and minimum times (which often occur in the early afternoon hours and pre-dawn hours, respectively), but this did not greatly affect daily streamflow estimates. Temperature was assumed to follow a sinusoidal function between these two time periods, and 3-hourly temperature was calculated for input into the WATFLOOD/SPL model. The precipitation was assumed to occur evenly throughout the day (the precipitation was divided by 24 and an equal amount of precipitation was applied in each hour).

3.3.2 Approaches to Calibration and Validation

Physically based hydrological models are superior to lumped models in their ability to model hydrological processes because they can be transferred between watersheds without extensive re-calibration. However, the initial calibration is more difficult for physically based models since they have a larger number of parameters. The literature reveals three approaches to this problem.

The first acknowledges that multiple "local optima" exist in the parameter space of physically-based hydrological models. The models are calibrated so that an error function is minimized (e.g. root mean squared, RMS, error). Two or more parameter sets may give similar values of the error function. Several authors contend that, while the models may produce good results with a particular set of parameters, there are hidden errors in the model (e.g. Beven, 1989, Grayson, *et al.*, 1992, and Beven, 1993). The errors in one physical process can be cancelled by errors in other physical processes. They contend that it is difficult to determine the "best" set of parameters, if it exists. They recommend either using simple models whenever possible or using multiple sets of parameters. For instance, during calibration Vertessy and Elsenbeer (1999) used a distributed, physically based model and found that the "best" parameters varied depending on the event used for calibration, and therefore they used multiple sets of parameters in their analysis. These authors recommend great caution with physically-based hydrological models. However, this approach does not solve the problem of modeling non-gauged watersheds or changing conditions.

The second approach is more optimistic. Several authors have developed algorithms to search for the "global optimum" parameters, and not simply a local optimum (e.g. Thyer, *et al.*, 1999, Gupta, *et al.*, 1999). The models are calibrated to minimize an error function, but the algorithms are able to jump between different local minima of the error function. These are essentially automatic calibration algorithms, which are programmed to search for several optima.

The third approach takes advantage of the physical modelling in the hydrological model. The use of physically-based parameters (i.e. parameters that can be measured in the field) and the use of multi-response data (i.e. calibration using outputs other than streamflow)

are recommended in this approach. This approach solves the problem of modelling non-gauged watersheds and changing conditions.

The amount of calibration required by a model may be reduced with measurable parameters. Refsgaard (1997) recommends selecting parameterization schemes that allow the user to measure the parameter values in the field. Refsgaard (1997) also recommends that physically-realistic intervals be determined for each parameter. Some very simple models do not require any calibration (e.g. Lange, *et al.*, 1999), as all of their parameters may be measured in the field. However, these simple models are limited in the hydrological conditions they can simulate and can only be applied to certain small watersheds. More complex models will require some calibration, as not all of the parameters can be measured in the field.

In terms of calibration, distributed physically-based hydrological models have an advantage over lumped conceptual models as they are more suited to model domains that encompass widely varying hydrological conditions and processes. Therefore, multiple-site data and multiple-response data, where they exist, can be used to help calibrate and validate a model. Klemeš (1986) proposed a series of hierarchical validation tests, which include multi-site validation. The model must be validated at each hierarchical level in order to be considered valid. Therefore, the model must pass a split-sample test (single location), followed by a proxy-basin test (multiple locations), followed by other tests when the end-purpose of the model requires further validation. Mroczkowski, *et al.* (1997) extended the hierarchical tests to include interior data such as soil moisture, snowdepth, evaporation, internal streamflows, etc. This will allow the user to test the internal physics of the model to ensure that the model simulates the hydrological processes in a reasonable manner.

Refsgaard (1997) recommended that a model be validated on all of the outputs that are required for the research. He stated that "comprehensive validation procedures specifically adapted for each particular application of a distributed model should be used" (pg. 95, Refsgaard, 1997). Therefore, it would seem that a model may be considered validated if each of the major hydrologic processes modelled have been validated.

The calibration and validation of WATFLOOD/SPL has followed this third approach. The parameterization schemes in WATFLOOD/SPL have been chosen so that all the parameters have physical meanings. Although not all of the parameters may be measured in the field, all of the parameters have physically definable limits. These limits have been established using textbook values and through experience with the model. The parameters with their limits are listed in Table 3-2. When parameters remain within these limits, the hydrological processes within WATFLOOD/SPL operate correctly.

A calibration of WATFLOOD/SPL proceeds in the following manner:

- the model is calibrated by hand until the internal variables (such as evaporation, snowmelt, etc.) show that the model physics are operating realistically in all grids;
- the parameters are "fine-tuned" with an automatic calibration procedure, to match the calculated streamflow to the measured streamflow (all internal and external streamflow observations are used); and
- the internal variables are checked again to ensure that the model physics are realistic.

In the first two of these steps, the parameter limits from Table 3-2 are used as a guideline for calibration. The method cannot be automated, as it is necessary to verify the internal parameters.

Table 3-2 - Parameter Values and Limits for WATFLOOD/SPL (Columbia River Domain)

Parameter	Name	Number of Classes	Lower Bound	Upper Bound
Unsaturated Zone Moisture Coefficient	A5	1	0.98	0.999
Surface Permeability	AK	5	1	100
Surface Permeability under snow	AKfs	5	1	100
Interflow storage-discharge coefficient	REC	5	0	0.1
Overland flow conveyance parameter	R3	5	5	90
Overland flow conveyance parameter under snow	R3fs	5	5	90
Soil retention coefficient	RETN	5	0	50
Upper to lower zone drainage coefficient	AK2	5	0.01	1
Upper to lower zone drainage coefficient under snow	AK2fs	5	0.01	1
Lower zone drainage function	LZF	3 river classes	10^{-7}	10^{-5}
Lower zone drainage function exponent	PWR	3 river classes	0.3	3
River roughness coefficient	R2	3 river classes	0.1	4
Melt factor	MF	5	0.05	0.25
Base temperature	BASE	5	-5	5
Potential Evapotranspiration Factor	FPET	5	1	3
Evapotranspiration Factor for Tall Vegetation	FTALL	5	0.5	1.2
Number of parameters: 70				

The initial manual calibration is accomplished by adjusting parameters to match various components of observed hydrographs. For instance, the base temperature for snowmelt is adjusted so the initial rise of the computed spring melt hydrograph occurs at the proper time. If computed peaks are consistently late and low throughout the domain, the river roughness is reduced. If the peaks are late and low only in the smaller watersheds, it is

likely that the interflow discharge coefficient is set too low. Recession curves can be matched using the lower zone function parameters and the parameter governing recharge, using log plots of flow versus time. Evaporation rates are adjusted to ensure annual volumes of runoff are correctly computed. Boyle, *et al.* (2000) reported on a similar approach to match the various segments of the hydrograph that they labelled as "driven", "nondriven quick", and "nondriven slow" (corresponding to flow driven by rainfall, the "fast" portion of the recession curve, and the "slow" portion of the recession curve respectively).

Once the parameters are given these initial values, an automatic scheme can be used. For the "fine-tuning" step, the model employs the Hooke and Jeeves (1961) automatic pattern search optimization algorithm taken from Monro (1971). The parameters for optimization are recharge and interflow coefficients, soil permeability, overland flow roughness, channel roughness, melt factors, base temperatures, soil retention, lower zone drainage coefficients, and an unsaturated zone soil moisture coefficient. The optimization is based on minimizing the root-mean-square error of streamflow estimates for all of the streamflow stations.

Finally, the internal variables are checked again to ensure that the model physics are realistic. The main tool for checking the internal model physics is shown in Figure 3-6. This figure plots the time variation of various state variables for the high elevation dense forest class in one grid cell. The major storage locations (lower zone storage, upper zone storage and depression storage, with and without snow cover) are plotted in the upper portion of the figure. The middle portion shows the fraction of the area covered with snow, the snow water equivalent and heat deficit. Finally, the lower portion shows the cumulative precipitation, runoff and evaporation for the high dense forest class. In figures such as Figure 3-6, all state variables can be tracked for all land cover classes on any grid. Snowmelt and snow

accumulation, infiltration, evaporation and other processes may all be examined in this figure. The internal parameters of WATFLOOD/SPL can also be viewed with a program called EnSim Hydrologic (Calder, 1999). EnSim Hydrologic is capable of displaying watershed data in 2-D format, or as a time variation plot for a single grid square. An examination of several such figures can be used to diagnose any parameter problems in WATFLOOD/SPL, and plots can be compared to field data when available.

The calibration of the WATFLOOD/SPL model for the Columbia River basin is described in Kouwen, *et al.* (2000).

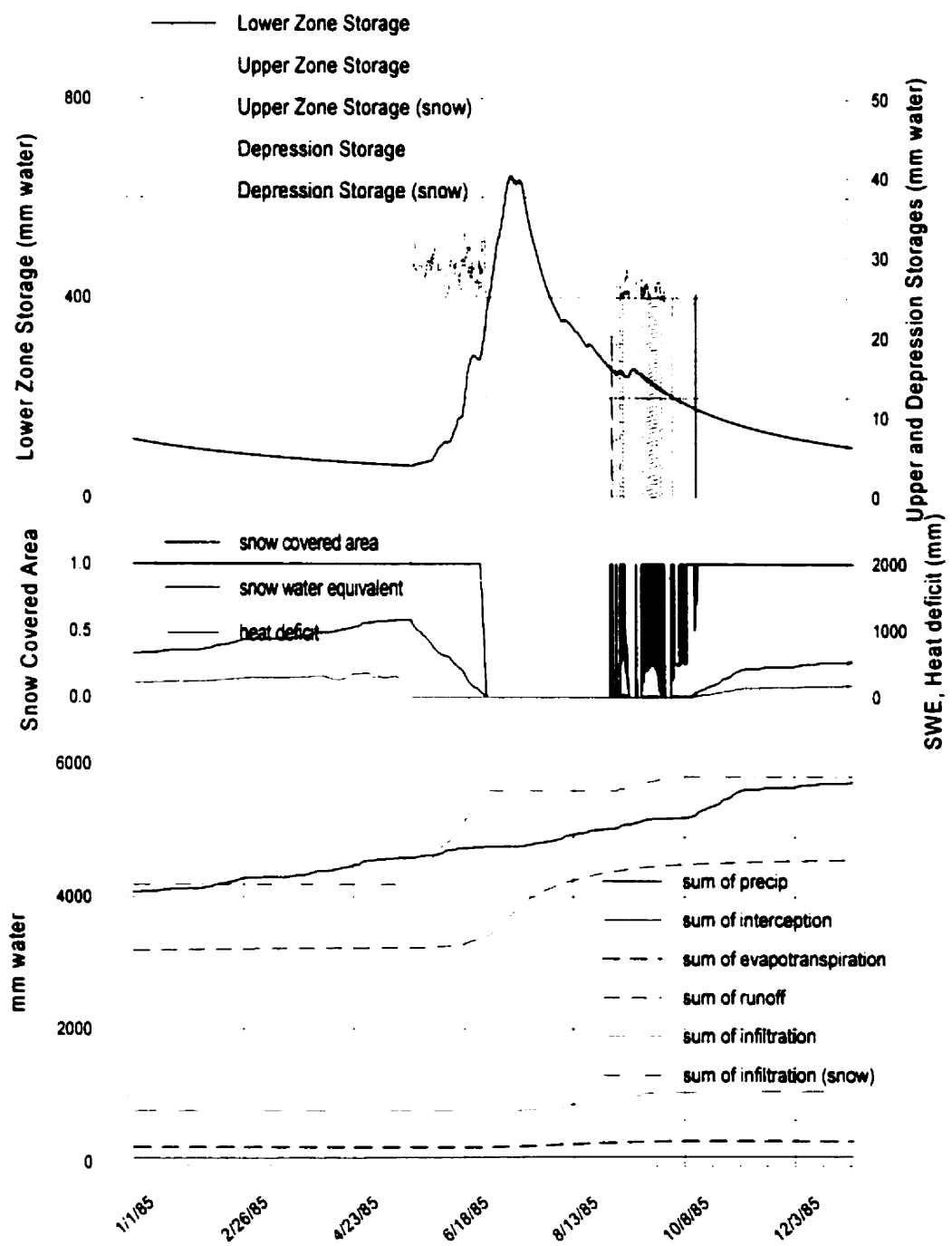


Figure 3-6 – Calibration tool for WATFLOOD/SPL

In addition to the implicit calibration and validation of internal streamflow and internal state variables in the model, a series of studies of multi-site and multi-response calibration and validation has been performed with WATFLOOD/SPL (Bingeman, *et al.*,

2001). The validations included comparisons of measured and computed soil moisture, evaporation, snowmelt and accumulation, groundwater flow, and peak flows. A set of parameters has been found that generates physically-realistic streamflow for multiple basins: farmland in Southern Ontario (Kouwen, *et al.*, 1993), the Rocky Mountains in British Columbia (Kouwen, *et al.*, 2000), the BOREAS study areas in Northern Saskatchewan and Manitoba (Neff, 1996) and the Mackenzie river basin in north-western Canada (Seglenieks, *et al.*, 1998). These parameters are called the "global parameter set."

3.3.3 Uncertainty in Model Estimates

Although each of the major hydrological processes within WATFLOOD/SPL have been calibrated and validated, the output data are still subject to modeling uncertainty because of the site-specific nature of these tests. Three types of errors lead to uncertainty in the estimates of the model. They are caused by model formulation errors, input data errors, and calibration errors (Lei and Schilling, 1996). Each of these errors will be discussed in turn.

The first type of error, model formulation error, refers to an error in the model itself, and is difficult to evaluate. Two examples of this kind of error are: an important hydrologic process may be missing from the model; or the process is included but modeled inaccurately. Lei and Schilling (1996) recommended that the model structure and equations be examined prior to calibration. The hydrological model WATFLOOD/SPL has been extensively examined with multiple validation studies (Bingeman, *et al.*, 2001). In addition, the model has been successfully applied on many watersheds in Canada, as described in Section 3.3.2, indicating a great degree of model stability. These tests indicate that the WATFLOOD/SPL

hydrological model has low model formulation error, although it is acknowledged that this applies only at the scale that it is usually applied, namely a grid size of 1 to 25 km.

The second error, input data error, refers to inaccurate model forcing data. This research used the HRBL model estimates of precipitation and temperature data as input for the hydrological model WATFLOOD/SPL. Therefore, the precipitation and temperature data are also subject to modeling uncertainty, and were examined. Typically, precipitation data are more difficult to both measure and predict than temperature data. There are several reasons for this, including problems with the observation network and problems with interpolation schemes. Most observation stations are located in valleys, and very few data exist for higher elevations. Observation stations may not record a rain event if the observation network is too sparse. The gauge may undercatch because of wind effects, causing the measurements to be inaccurate. The interpolation scheme used in the HRBL model was a regression equation based on the horizontal moisture convergence, and the regression equation could not model the precipitation perfectly. Therefore, the examination focused mainly on the precipitation data.

To test the precipitation data, the data were used to calculate streamflow, and a comparison of streamflow volume was performed (Kouwen, *et al.*, 2000). One streamflow station, Columbia River at Nicholson, had 91 years of observed streamflow, and could be used to validate almost all of the 96 available years of data. The streamflow estimates closely matched the observed streamflow for all 91 years (Figure 3-7).

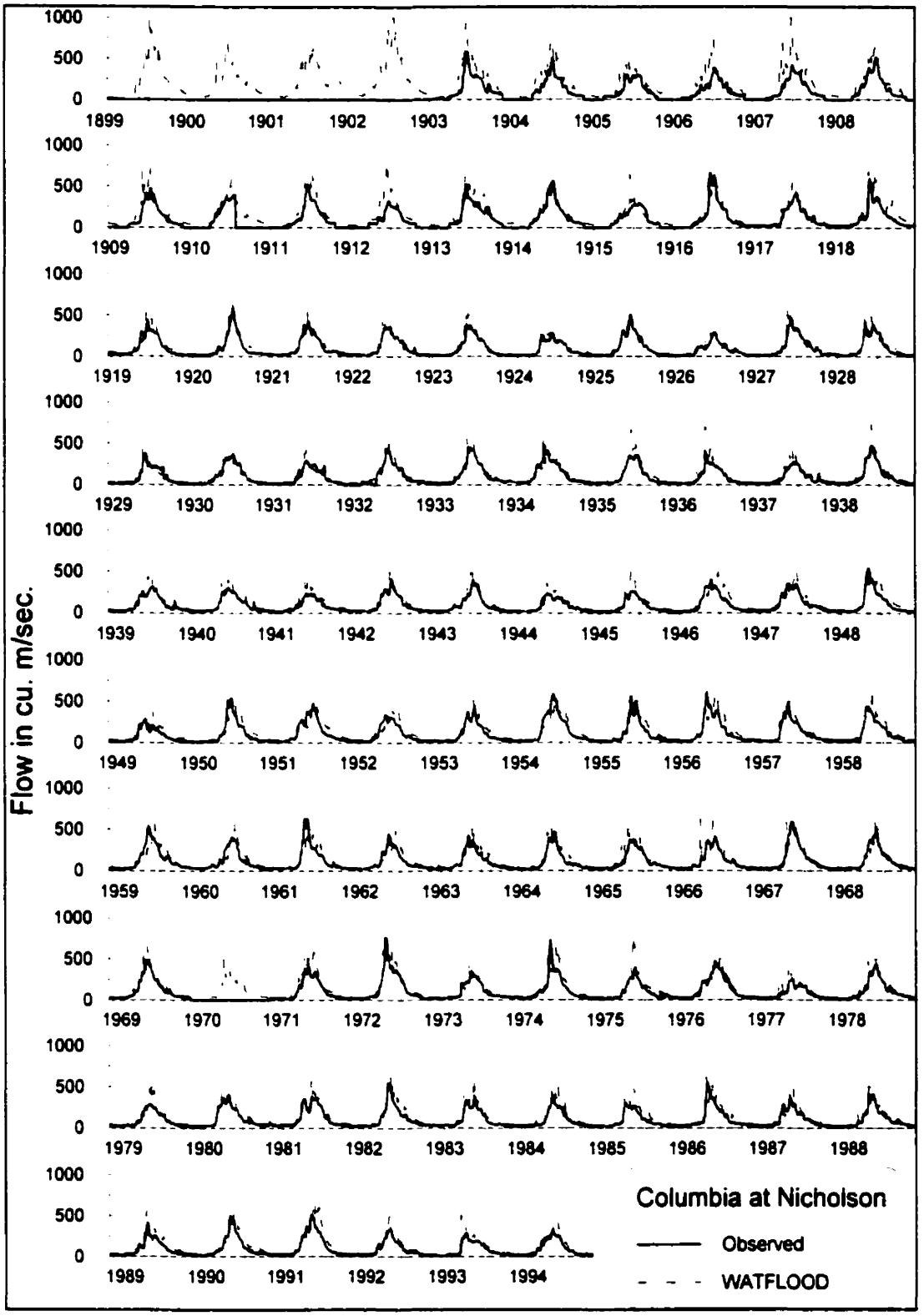


Figure 3-7 – Comparison of 96-Years of Streamflow – observed vs simulated by WATFLOOD/SPL with Meteorological Forcing Data (Daily Precipitation and Temperature) from HRBL for the Columbia River at Nicholson location

The streamflow estimates also matched closely for the inflow to Mica Dam (Figure 3-8). There were 23 years of observed inflow available for comparison (1972 to 1994). For both of these locations, the observed and simulated hydrographs matched very well. The timing of the snowmelt was accurate for both locations, and the volume of snowmelt also matched. The summer precipitation peak flows were generally well represented. Finally, the recession curves and low winter flows were also in agreement between the observed and simulated hydrographs. Therefore, it was reasonable to state that the temperature and precipitation data from the HRBL model were suitable for the prediction of long-term streamflow.

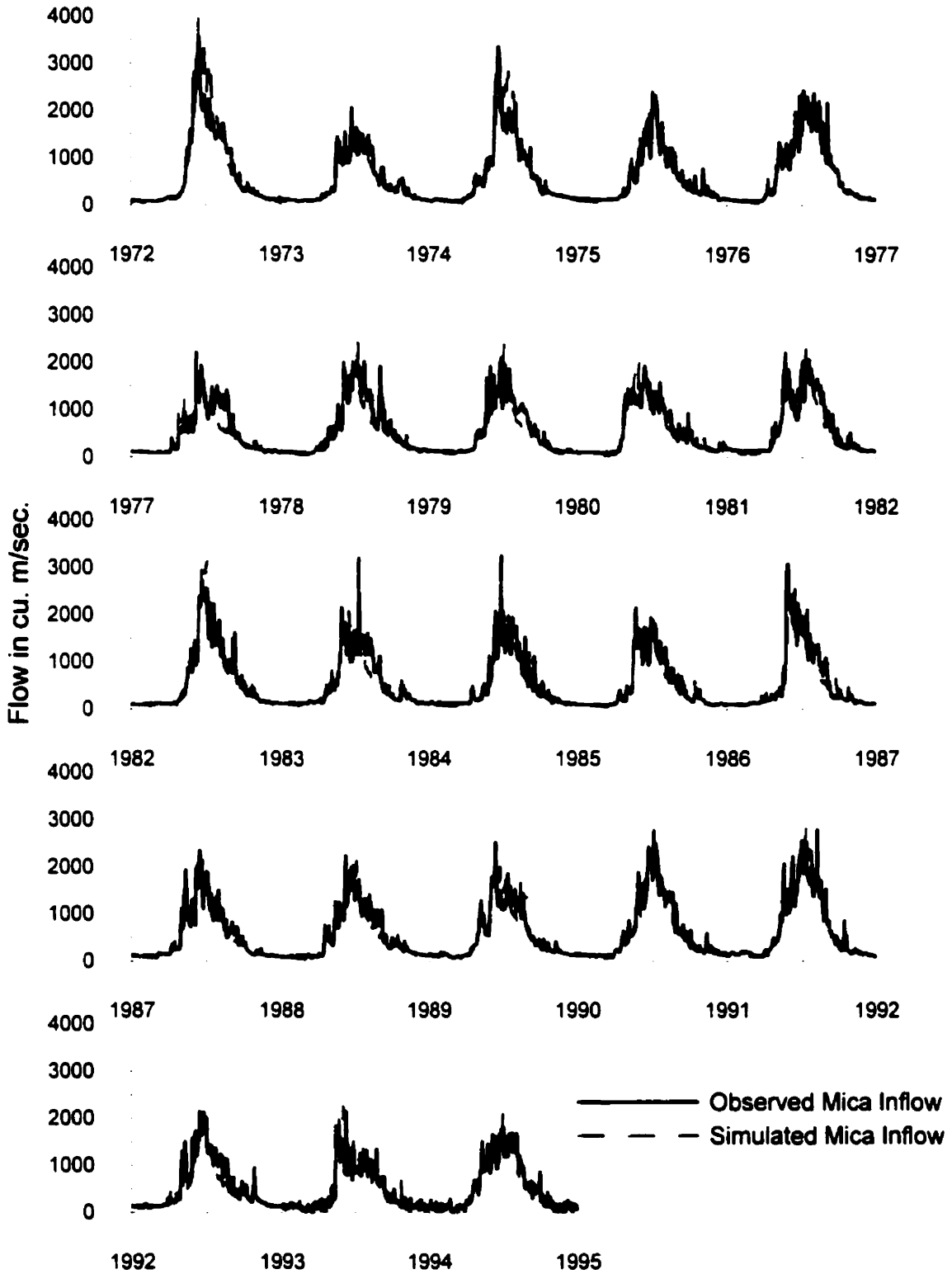


Figure 3-8 – Comparison of 23-Years of Reservoir Inflow – observed vs simulated by WATFLOOD/SPL with Meteorological Forcing Data (Daily Precipitation and Temperature) from HRBL for the Mica Dam location

The hydrographs shown above indicated good agreement between the observed and simulated streamflow data, indicating that the temperatures for predicting freezing and thawing were acceptable. Similarly, the hydrographs indicated that the timing of the precipitation was suitable for predicting streamflow. However, some stations were consistently over- or under-estimated. This consistent over- or under-estimation indicated an error in the precipitation estimates from the HRBL model. The average errors for the 32 streamflow stations for the period 1972 to 1994 were calculated for each streamflow station, and a contour plot of error was created (Figure 3-9). These years were chosen because LFM data were available to create the precipitation data during this period (and therefore the precipitation data were most accurate), and because most of the streamflow stations were active during this period. This error plot indicates the amount of error in the most accurate portion of the time period modeled (1972 to 1994). This figure shows that the streamflow volume differences at a station can be very significant, up to 80% error. In general, the model underestimates streamflow in the northwest corner, and overestimates in the southeast corner. This pattern further indicated consistent errors in the precipitation data. Therefore, a precipitation adjustment factor (PAF) field was created based on this error field. The PAF field was used to adjust the precipitation estimates. The streamflow estimates improved, as stations that were generally under-estimated received a greater amount of precipitation (and vice versa). In this way, the input data errors were minimized.

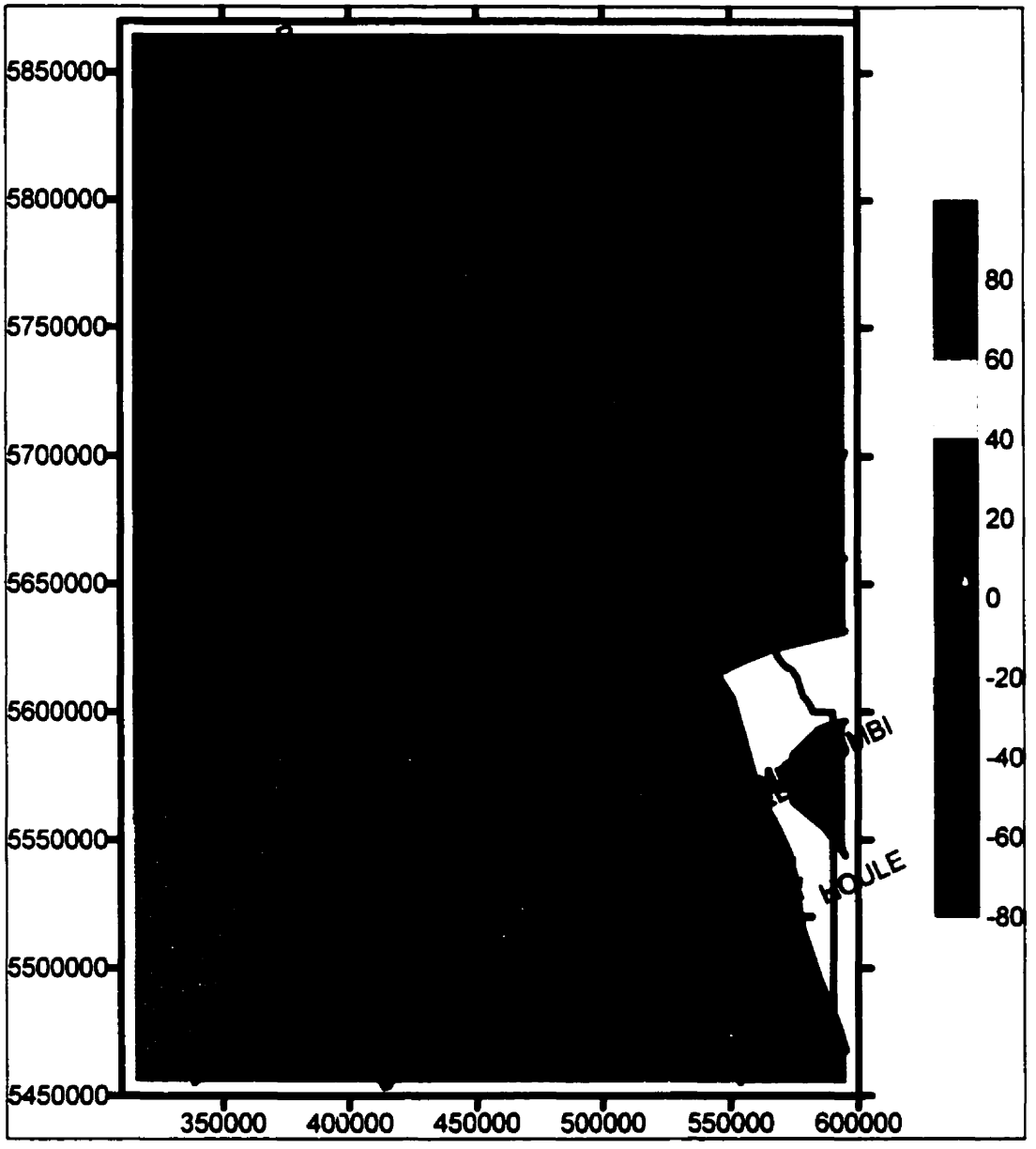


Figure 3-9 – Contour Plot of Annual Streamflow Volume Differences in % (R_{model}-$R_{measured}$), 1971-1994

The temperature data were verified briefly by comparing the timing of the simulated snowmelt and snow accumulation to observed snow measurements. Wong (2000) compared the simulated snow water equivalent (SWE) to observed snowcourse measurements and observed snow pillow data in the Columbia River basin. There were some significant differences in the simulated SWE and the snowcourse data, but these were attributed mainly

to differences in the elevation of the snowcourse and the 100 km^2 grid. That is, the snowcourse was above or below the grid elevation (an average over 100 km^2), and would therefore record more or less snow, respectively. The snow pillow data provided a time series of snow water equivalent, and these data indicated that the timing of the snowmelt and snow accumulation was correct. These results indicated that the temperature data derived by Danard, *et al.* (1996b) were adequate for long-term hydrological modeling.

The third type of error in a hydrological model is uncertainty in the parameter calibration. In comparison to the number of hydrological models available, there has been relatively little focus on this type of error. In general, hydrological models are highly non-linear and analytical techniques to convert parameter uncertainty into output uncertainty are difficult to derive. Monte Carlo analysis can be used, but many hydrological models require a large amount of computer time for a single simulation. A Monte Carlo analysis allows a distribution to be defined for each parameter. These distributions are randomly sampled, and a model simulation is performed for each sample to derive the output distribution. Therefore, a Monte Carlo analysis can be very time consuming for a hydrological model. However, the confidence of model predictions is improved through the estimation of parameter uncertainty.

Where the uncertainty due to errors in the parameters have been estimated, most authors recommend a Monte Carlo analysis. A variation such as importance sampling may also be used. Importance sampling can be used when there is sufficient information about the behaviour of the model and parameters to be able to confine the Monte Carlo to a small range in the parameters. However, these authors accept the "equifinality" concept (Beven, 1993), where multiple sets of parameters can deliver similar or identical quality of hydrological predictions. The parameter distributions for the Monte Carlo are chosen

according to this assumption. For instance, several authors (e.g. Binley, *et al.*, 1991) recommend that the model be calibrated for numerous calibration events. Each calibration event will have different optimal parameters, and so the mean and standard deviation of each parameter can be defined from the different calibration events. A normal distribution is used to describe the parameter uncertainty, using the calculated mean and standard deviation. Once the distribution has been chosen, the Monte Carlo analysis is performed (a variation, such as importance sampling, may also be used instead), and confidence limits are derived.

There were two problems with performing this type of Monte Carlo analysis for WATFLOOD/SPL. First, the uncertainty analyses performed in the literature use the "equifinality" approach to parameter calibration, which implies multiple sets of parameters are equally valid. However, the calibration process of WATFLOOD/SPL ensures that the "true" optimum parameter values are found. The calibration process checks all of the internal variables within the WATFLOOD/SPL model, and therefore ensures that the hydrological processes are operating in a realistic manner. There is still error in the parameter values (due to possible input errors), but the calibrated parameter values represent a most likely set of parameters, or a "mode." The parameter distributions were chosen based on the mode and the boundaries of the parameters. Secondly, a single 96-year simulation takes approximately 4 hours on a single processor of the SGI Origin 200 computer, and therefore computer time was a limitation for the analysis. These two issues were addressed in this research.

The calibration philosophy of WATFLOOD/SPL assumes that the "optimum" parameter values can be determined through the calibration process. The parameter set can be used for multiple basins, and the internal variables of the model have been validated (as

described in Section 3.3.2). Therefore, if the calibrated parameters are not truly the optimum parameter values (due to uncontrollable uncertainties in input data), then they are "close" to optimum; that is, they are on the same "hill" of the objective function. The physically possible boundaries for the parameters were checked to ensure that they remain on a single hill of the objective function. (In this research, the term "objective function" is a generic term to represent the degree of fit between the observed streamflow and the calculated streamflow. A "hill" on the objective function denotes an area of parameter space where the function rises to a "good" fit and then falls away again.) There are multiple hills in the objective function, as other combinations of parameters may lead to reasonable streamflow estimates. However, these other hills would result from two (or more) errors in modeling the hydrological processes that cancel each other. Should this condition occur, a check on each of the process plots (Figure 3-6) would reveal some sort of non-plausible condition. Furthermore, unrealistic parameters affecting river routing would become evident using the hierarchical validation approach with a large number of streamflow stations simultaneously. A set of parameters from another "hill" of the objective function would therefore be rejected during the validation process of WATFLOOD/SPL. Therefore, in keeping with the calibration philosophy of WATFLOOD/SPL, it was also necessary to choose the parameter distributions so that the Monte Carlo simulations remained on the same "hill" of the objective function.

The second difficulty with the Monte Carlo analysis was simulation time. Since all parameters were allowed to vary simultaneously in the Monte Carlo analysis, approximately 100 simulations per variable (on average) were necessary to achieve realistic 95% confidence limits (Crosetto, *et al.*, 2000). For 70 variables, this was a minimum of 7000 simulations!

Therefore, it was not feasible to perform the Monte Carlo analysis for the entire 95-year simulation. Instead, the full 95-year simulation was represented by a five-year sequence that included a variety of wet and dry years. It was assumed that the confidence limits from the five-year analysis could be used in the 95-year analysis. This assumption was checked as part of the analysis and shown to be accurate.

The detailed methodology to resolve these conflicts is described in detail in Chapter 5.

3.4 Regional Frequency Analysis and the Method of L-Moments

The atmospheric model and the hydrological model were used to derive a physically-based PMF in this research. One method to check the magnitude of the PMF is to calculate the probability of the PMF on a frequency curve (Smith, 1998). The return period of the PMF may range from 10,000 years to 1,000,000 years (Smith, 1998). In addition, a frequency curve would allow risk-based analysis to be performed for the dam. This research investigated the use of a simulated streamflow time series to derive flood frequency curves. The simulated streamflow were obtained by using the 96-year HRBL meteorological data as forcing data for the WATFLOOD/SPL model. The regional frequency analysis method using L-moments (Hosking and Wallis, 1997) was chosen for this research because it estimates extreme flood quantiles with less uncertainty than conventional methods (Pilon and Adamowski, 1992, Hosking and Wallis, 1987). This section describes the regional frequency analysis method.

Frequency analysis is a standard statistical method (see, for example, Yevjevich, 1972) used to estimate extreme events. A probability distribution is fitted to the observed

data, and then the extreme values are estimated using the probability distribution. Two problems with this method are:

- Frequency analysis is often limited by a short time series.
- It is difficult to move from a gauged site to a nearby ungauged site, since the probability distribution depends on the statistics of the gauged site, which are not available for the ungauged site.

Regional frequency analysis is a method to deal with these issues.

Regional frequency analysis has been in use for many years; an early example is Dalrymple (1960). Regional frequency analysis "trades space for time". Gauges that are similar are grouped together, thus increasing the effective amount of data and reducing the uncertainty in the frequency distribution parameters. The frequency distribution at ungauged locations may then be estimated by using the probability distribution of the group.

Two concepts must be presented before regional frequency analysis can be discussed. Accordingly, the next two sections present the index flood procedure and L-moments, both of which are included in regional frequency analysis. The following section describes the regional frequency analysis method of Hosking and Wallis (1997).

3.4.1 Index-Flood Procedure

The index-flood procedure is a simple way to pool data from different locations. The name comes from its original application to flood data in hydrology (e.g. Dalrymple, 1960). The index-flood procedure assumes that the probability distributions for the sites are simply a scale factor multiplied by the regional probability distribution, such as:

$$Q_i(F) = \mu_i q(F) \quad i = 1 \ldots N \qquad (3\text{-}14)$$

where $Q_i(F)$ is the quantile function of the frequency distribution of each site, μ_i is the index-flood of each site, and $q(F)$ is the quantile function of the frequency distribution of the group of sites. This method assumes that each site has the same underlying probability distribution. The index-flood (μ_i) is often chosen as the at-site mean, but a median, trimmed mean, or another percentile may also be used.

The index-flood procedure makes the following assumptions about the data.

- Observations at a site are identically distributed.
- Observations at a site are temporally independent.
- Observations at different sites are spatially independent.
- Frequency distributions are identical at all sites, and follow a regional frequency distribution, except for a scale factor.
- The regional frequency distribution is correctly specified.

These assumptions are never fully met with environmental data. There may be some serial dependence in any environmental data and correlation between sites is common, since neighbouring sites have similar weather patterns. The fourth assumption can only be approximately satisfied, by careful grouping of the sites. The frequency distribution is chosen as the best fit of the data in the region, and so the last assumption is also only approximately satisfied. However, research (e.g. Pilon and Adamowski, 1992) has shown that it is possible to use the index-flood procedure to yield suitable extreme quantile estimates.

3.4.2 L-moments

3.4.2.1 Introduction

Frequency distributions have traditionally been described by their moments. The moments are:

$$\begin{aligned} &\mu = E(X) \quad (mean) \\ &\mu_r = E(X-\mu)^r \quad r = 2,3,\ldots \quad (higher\ central\ moments) \end{aligned} \qquad \text{.....(3-15)}$$

where X is any random variable (such as precipitation or streamflow).

The mean describes the location of the center of the distribution. The second moment describes the dispersion of the distribution. Dimensionless versions of the higher moments are usually used, such as:

$$\begin{aligned} &\gamma = \mu_3 / \mu_2^{3/2} \quad (skewness) \\ &\kappa = \mu_4 / \mu_2^2 \quad (kurtosis) \end{aligned} \qquad \text{.....(3-16)}$$

These moments can be estimated with a data sample. However, these estimates are known to be biased and bounded by the sample size. Skewness and kurtosis are also sensitive to outliers in the data, since they use the difference between the value and the mean to the third or fourth power. Therefore, they are unreliable as measures of distribution shape.

L-moments are alternative measures of distribution shape. They are modifications of the "probability weighted moments" of Greenwood, *et al.* (1979). The two most useful probability weighted moments are

$$\alpha_r = \int_0^1 x(u)(1-u)^r du, \quad \beta_r = \int_0^1 x(u)u^r du \qquad \text{.....(3-17)}$$

where x(u) is the inverse of the cumulative density function, and the integration is performed over the range of u, the probability of non-exceedance (Figure 3-10).

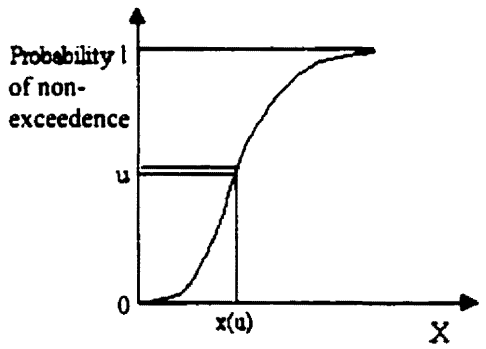


Figure 3-10 – Definition of terms in calculation of Probability Weighted Moments

Probability weighted moments involve powers of the functions u and 1-u, while conventional moments involve powers of the data, x(u). Various authors have related α_r and β_r to the conventional measures of distribution shape (e.g. Hosking and Wallis, 1985), and these relationships are termed the L-moments. The term "L-moments" is used because the L-moments are calculated as linear combinations of probability weighted moments.

To estimate the L-moments for a data sample, first arrange the data values in ascending order: $x_{1:n} \leq x_{2:n} \leq \ldots \leq x_{n:n}$. The first four sample β_r's can be found with the following formulas,

$$
\begin{aligned}
b_0 &= n^{-1}\sum_{j=1}^{n} x_{j:n} \\
b_1 &= n^{-1}\sum_{j=2}^{n} \frac{(j-1)}{(n-1)} x_{j:n} \\
b_2 &= n^{-1}\sum_{j=3}^{n} \frac{(j-1)(j-2)}{(n-1)(n-2)} x_{j:n} \\
b_3 &= n^{-1}\sum_{j=4}^{n} \frac{(j-1)(j-2)(j-3)}{(n-1)(n-2)(n-3)} x_{j:n}
\end{aligned}
\qquad \text{.....(3-18)}
$$

The sample L-moments are then found using,

$$
\begin{aligned}
l_1 &= b_0 \\
l_2 &= 2b_1 - b_0 \\
l_3 &= 6b_2 - 6b_1 + b_0 \\
l_4 &= 20b_3 - 30b_2 + 12b_1 - b_0
\end{aligned}
\qquad \text{.....(3-19)}
$$

The first L-moment, l_1, is analogous (and equal) to the conventional mean. The second L-moment, l_2, is analogous to the conventional standard deviation. The conventional coefficient of variation (CV) is calculated as the standard deviation divided by the mean. Therefore, the L-moment equivalent, the L-CV, is found as $t=l_2/l_1$ (no subscript is used with t by convention). The conventional skewness and kurtosis are dimensionless, and so the L-moment equivalents are also dimensionless. The L-skewness is found as $t_3=l_3/l_2$. Similarly, the L-kurtosis is found as $t_4=l_4/l_2$.

L-moments have been found to be better estimates of distribution shape than conventional moments. The estimates of L-skewness and L-kurtosis have been found to be much less biased than conventional estimates. The L-moments are not bounded algebraically by the sample size. L-moments have a natural bound $|\tau_r| \leq 1$ (where τ_r represents the population statistic), making interpretation of the magnitude of the moment easier. L-moments are less affected by outliers in the distribution, since they give less weight to the tail of the distribution (u^r is used instead of $(x(u)-\mu)^r$, and in the tail $u^r \rightarrow 1$, whereas $(x(u)-\mu)^r \rightarrow \infty$). L-moments have also been shown to discriminate between different distributions better. Therefore, L-moments are more reliable than ordinary moments for estimating the distribution shape.

Figure 3-11 shows the ranges of several common distributions in terms of L-skewness and L-kurtosis. This figure can be used to visually fit distributions, or the equations supplied by Hosking and Wallis (1997) can also be used to fit distributions. The parameters for each

distribution can be calculated from the L-moments in a similar manner that parameters can be calculated from regular moments. For example, the parameters for a normal distribution are the mean and the standard deviation. Using L-moments, the parameters become the mean (λ_1) and $\pi^{0.5}\lambda_2$ (λ_2 is the L-standard deviation).

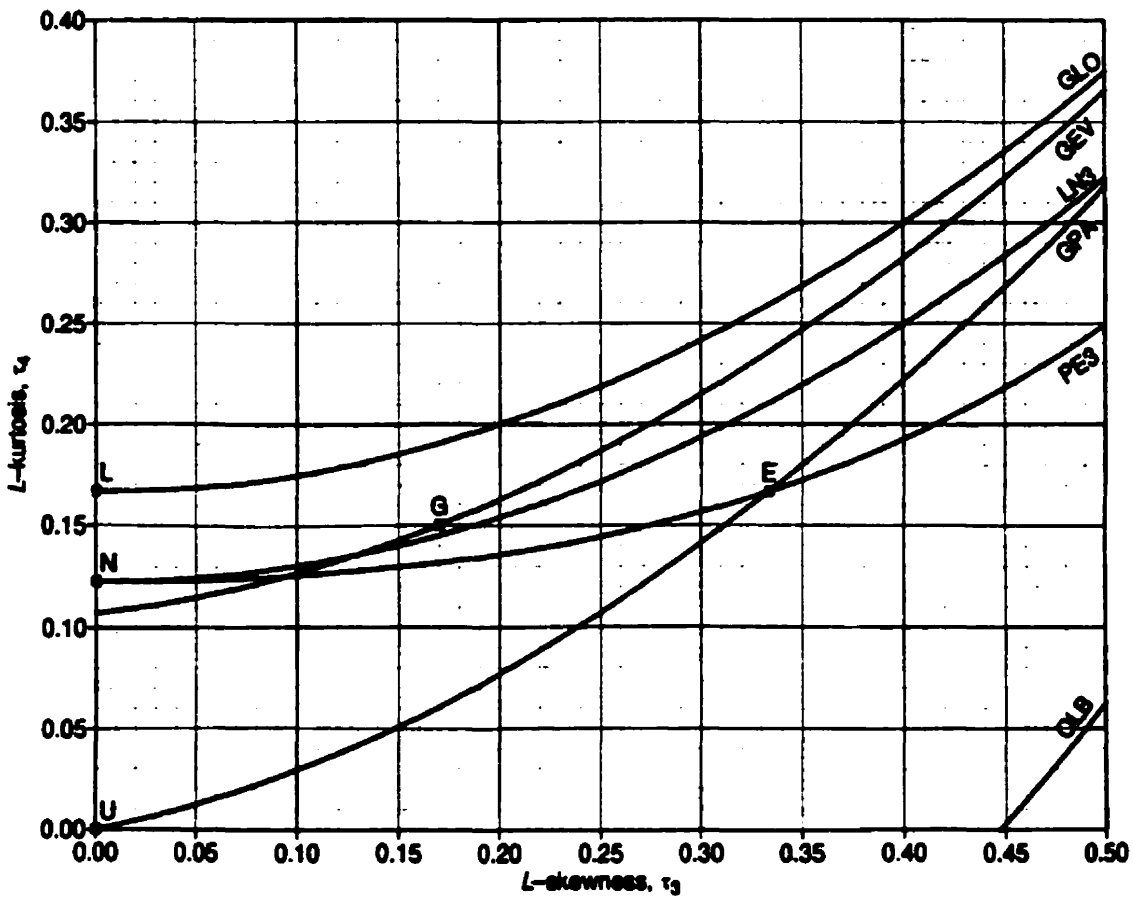


Figure 3-11 – L-moment ratio diagram (from Hosking and Wallis, 1997). Two and three parameter distributions are shown as points and lines, respectively. E=Exponential, G=Gumbel, L=Logistic, N=Normal, U=Uniform, GLO=Generalized Logistic, GEV=Generalized Extreme Value, GPA=Generalized Pareto, LN3=Lognormal, PE3=Pearson Type III, OLB = Overall Lower Bound

3.4.2.2 Regional L-moments

This section discusses the link between L-moments and regional frequency analysis.

In regional frequency analysis, several locations are grouped together into a region, and a distribution is found for the region. It is necessary to develop a regional estimate of the L-moments in order to fit a distribution to the region. Hosking and Wallis (1997) recommend the use of a weighted average of the at-site estimates,

$$\theta_k^r = \sum_{i=1}^{N} n_i \theta_k^{(i)} \Big/ \sum_{i=1}^{N} n_i \quad \text{.....(3-20)}$$

where N is the number of stations, θ_k is the L-moment of interest, and n_i is the number of data values at each station. Therefore, the at-site L-moments are found from the equations in Section 3.4.2.1, and the weighted average is found according to the length of record at each site. This method results in discontinuities in the frequency domain between regions.

A second method for calculating the regional estimate of the L-moments is to derive a relationship between the L-moments and other watershed variables. Runoff varies smoothly between regions in a similar way that precipitation varies smoothly, although there may be local differences due to topography and/or soils. These can be accounted for by using topographical and/or soils data in the relationship to allow the regional L-moments to vary smoothly between regions). This method was described by Schaefer (1990), and later used in Canada by Adamowski, *et al.* (1996). These authors used precipitation data, and derived relationships across a super-region between the L-moments for homogeneous sub-regions, mean annual precipitation for the sub-regions, and other sub-region variables. This research also used this concept: a relationship between the L-moments and basin-averaged topographic and physiographic characteristics was found and used to estimate the regional L-moments. However, the relationships were formed from the network of individual stations, not from homogeneous sub-regions. A similar method has been used for streamflow data; an example is the work of Daviau, *et al.* (2000), who used a GIS system to develop relationships

between L-moments and geostatistical data for streamflow gauges in Eastern Canada. This method will be referred to as the Schaefer (1990) method, although it has been modified (e.g. Daviau, *et al.*, 2000) to be performed for the network of stations.

The Schaefer (1990) method has some advantages over the Hosking and Wallis (1997) method. One advantage is that it does not assume that the at-site L-moments are perfect. By allowing the L-moments to be represented by regressions over all sites available, anomalous data (high or low) may be smoothed over. The Hosking and Wallis (1997) method uses the average of all of the L-moments within the region, and one anomalous station could strongly affect the results. A second advantage is that the regionalized L-moments are formed from a larger data pool (all sites available instead of the sites within the region) and can therefore be expected to be more accurate. Therefore, the Schaefer (1990) method lowers data uncertainty.

When the index-flood procedure is used with μ_i equal to the at-site mean for each station, the L-location (the first L-moment) will always be equal to 1. Therefore, distributions are fitted to the following regional L-moments: $1, t^R, t_3^R, t_4^R$.

Since the regional estimates of L-moments are found from at-site estimates, the necessity of using L-moments is clear. L-moments are more reliable than ordinary moments for short data records. It is important that the at-site estimates are as accurate as possible, so that the regional estimate will also be accurate.

3.4.3 Steps Involved in Regional Frequency Analysis

There are four main steps to the Hosking and Wallis (1997) regional frequency analysis method. These are: 1) Data screening, 2) Identification of regions, 3) Choice of a

frequency distribution, and 4) Estimation of the at-site frequency distribution parameters. These steps will be discussed below.

3.4.3.1 Data screening

Due to errors in recording or transcribing of data (e.g. incorrect logger setup, or changes in the rating curve due to vegetation growth), the data must be screened to remove invalid data. For instance, occasionally negative streamflow data are recorded, or an incorrect rating curve causes errors in the streamflow estimation procedure.

Hosking and Wallis (1997) presented a discordancy measure to help identify discordant data. The discordancy measure is

$$D_i = \frac{1}{3} N(u_i - \bar{u})^T \left[(u_i - \bar{u})(u_i - \bar{u})^T \right]^{-1} (u_i - \bar{u}) \quad \text{.....(3-21)}$$

where u is the vector of L-moments, and N is the number of stations. The discordancy measure is essentially the Mahalanobis distance (e.g. Pao, 1989) of the station in the 5-space of the L-moments from the mean of the stations. For "large" values of D_i, the data in that station should be examined for errors. (For $N \geq 15$, D_i should be less than or equal to 3.) It is possible, however, that the data are correct, even for large D_i.

Regional frequency analysis was performed with streamflow station data. These data (obtained from Water Survey Canada and B.C. Hydro) are generally valid. The analysis was also performed with the output of a hydrological model, and the calibration process ensured that the output was realistic. Therefore, these data were screened before this research.

3.4.3.2 Identification of regions

This step is perhaps the key step in regional frequency analysis. Hosking and Wallis (1997) assumed that all stations within the region have identical frequency distributions.

This research incorporated the variation in L-moments across the region by forming regressions between L-moments and watershed characteristics (Schaefer, 1990). This method relaxed the requirement for completely homogeneous regions; however, homogeneous regions were still required for identification of a suitable frequency distribution. This section describes the general method for forming homogeneous regions according to Hosking and Wallis (1997), and a specific clustering algorithm by Burn, *et al.* (1997).

Hosking and Wallis (1997) called their regions homogeneous, since the frequency distribution was assumed identical at all stations. They recommend that the stations be assigned to regions according to non-statistical parameters. This will allow an ungauged location to be classified into a region, although it is not possible to calculate statistical parameters for the location. In addition, the criterion to test for region homogeneity is based on the L-moments of the stations, so it would be inappropriate to use the L-moments for calculating both region definition and region homogeneity. Therefore, climatic and physiographic variables may be used to define a region, but the L-CV may not be used.

Hosking and Wallis (1997) presented a criterion for homogeneity based on the L-moments of the stations. The criterion for homogeneity of a region is based on

$$H = \frac{V - \mu_V}{\sigma_V}, \quad where \quad V = \left\{ \sum_{i=1}^{N} n_i \left(t^{(i)} - t^R \right)^2 \Big/ \sum_{i=1}^{N} n_i \right\}^{1/2} \quad \text{.....(3-22)}$$

where V is the weighted standard deviation of the at-site sample L-CVs (t), and μ_V and σ_V are the mean and standard deviation of V, found through simulation. The simulation is performed by fitting a Kappa distribution to the regional average L-moment ratios, 1, t^R, t_3^R, t_4^R. A Kappa distribution is used because it is a four-parameter distribution and therefore

makes fewer assumptions about the shape of the distribution than the more common three-parameter distributions. The region is assumed homogeneous (since a single Kappa distribution is used), and data are simulated for each station, according to the number of years of record at each station. For each set of data simulated, V is calculated, and the mean and standard deviation of V are calculated at the end of the simulation. If H is sufficiently large, the region is declared heterogeneous. Hosking and Wallis (1997) suggest that the region is "acceptably homogeneous" if H<1 and "possibly homogeneous" if H≤2. Schaefer (1997), however, points out that, due to variability in meteorological data from local site changes through time (station location, growth of trees, etc.), these limits should be changed to 2 and 3 respectively.

There are two other methods of calculating the heterogeneity measure, V. The first measure, V2, measures the at-site dispersion of sample L-moments based on L-CV and L-skewness. The second measure, V3, measures the at-site dispersion of sample L-moments based on L-skewness and L-kurtosis.

$$V_2 = \left\{ \sum_{i=1}^{N} n_i \left\{ \left(t^{(i)} - t^R \right)^2 + \left(t_3^{(i)} - t_3^R \right)^2 \right\}^{1/2} \Big/ \sum_{i=1}^{N} n_i \right\}^{1/2} \quad \text{.....(3-23)}$$

$$V_3 = \left\{ \sum_{i=1}^{N} n_i \left\{ \left(t_3^{(i)} - t_3^R \right)^2 + \left(t_4^{(i)} - t_4^R \right)^2 \right\}^{1/2} \Big/ \sum_{i=1}^{N} n_i \right\}^{1/2} \quad \text{.....(3-24)}$$

The identification of homogeneous regions is therefore an iterative process:

- Identify regions from non-statistical parameters
- Calculate H for each region
- If H is too large for some regions, go back and redefine the regions

One clustering algorithm that follows the above recommendations is the clustering algorithm developed by Burn, *et al.* (1997). This algorithm used a combination of geographic position and other hydrologic data to form regions. Geographically contiguous regions can be identified.

Many clustering algorithms operate by determining a "distance" between each set of two watersheds. The two watersheds with the shortest distance between them are clustered together, and then the distances are re-calculated between the new cluster and all of the other watersheds. Again, the two watersheds with the shortest distance between them are clustered, and the process repeats until the desired number of clusters are obtained.

The Burn, *et al.* (1997) algorithm used the following distance measure:

$$D_{ij}^{d} = \frac{D_{ij} + \frac{d_{ij}}{d_{\max}} w}{1 + w} \qquad (3\text{-}25)$$

where: D_{ij} is the dissimilarity between two watersheds (i and j) in terms of one or more hydrologic data (defined below), d_{ij} is the geographic distance between two watersheds (i and j), d_{max} is the largest geographic distance between any two watersheds, and w is a weighting coefficient. A low weighting coefficient implies that the geographic location has limited effect on the formation of clusters. As the weighting coefficient increases, a greater emphasis is placed on distance, to preserve geographic continuity in the clusters. Burn, *et al.* (1997) used a weighting coefficient of 0.3 with a set of streamflow stations in the Saskatchewan-Nelson River basin, covering Saskatchewan, Manitoba, and western Ontario. The appropriate value of w was determined by visual inspection of the clusters formed by the algorithm.

The dissimilarity between two watersheds in terms of the hydrologic data was found with the following:

$$D_{ij} = \frac{1}{p}\sum_{k=1}^{p}\frac{|X_{ik} - X_{jk}|}{X_{ik} + X_{jk}} \quad \text{.....(3-26)}$$

where: p is the number of hydrologic properties, and X_{ik} is the value of the kth hydrologic response property for catchment i. This measure, called the Canberra dissimilarity metric, was developed by Lance and Williams (1966).

These distance measures were used in the standard clustering algorithm, to develop the desired number of clusters. The clusters were first tested for homogeneity, using the Hosking and Wallis (1997) homogeneity test. If the clusters were homogeneous, they were accepted. If they were not, the clustering algorithm was re-applied (using only the watersheds not already in clusters) and the number of desired clusters increased by one to develop different clusters. After several levels of this, a number of homogeneous, or possibly homogeneous, clusters would be formed, and several "residue" watersheds (watersheds that could not be placed in a cluster) may have been left behind. Since the clustering algorithm was applied several times, it was termed "multi-level clustering."

After the desired number of clusters was formed, the Burn, *et al.* (1997) algorithm entered a second stage, termed the enhancement stage. The residue watersheds were tested to determine if they could be added to the existing clusters. Discordant stations within a cluster were tested to determine if they should be moved to another cluster, and thereby improve both clusters. In this way, the "best" clusters were formed.

3.4.3.3 Choice of a frequency distribution

There are several frequency distributions that can be used to describe a particular set of data. The distributions have different upper and lower bounds, and different tail weights. It is important to ensure that the best fit to a frequency distribution is made, otherwise the extreme estimates may be in considerable error.

A goodness-of-fit statistic was developed by Hosking and Wallis (1997) to help determine which three-parameter distributions can be used to describe the data. The statistic calculates how well the candidate distribution simulates the fourth L-moment, since this L-moment is not used to fit the distribution. The first step is to assemble a series of candidate three-parameter distributions. Next, data are simulated at each station (according to the number of years of record) by assuming a Kappa distribution for the region. For the mth simulation, the regional average L-kurtosis must be calculated ($t_4^{(m)}$). After all the simulations, the bias (B_4) and standard deviation (σ_4) of t_4^R are found:

$$B_4 = N_{sim}^{-1} \sum_{m=1}^{N_{sim}} \left(t_4^{(m)} - t_4^R \right)$$

$$\sigma_4 = \left[\left(N_{sim} - 1 \right)^{-1} \left\{ \sum_{m=1}^{N_{sim}} \left(t_4^{(m)} - t_4^R \right)^2 - N_{sim} B_4^2 \right\} \right]^{1/2} \quad \text{(3-27)}$$

Then, for each distribution, the goodness-of-fit measure is found as:

$$Z^{DIST} = \left(\tau_4^{DIST} - t_4^R + B_4 \right) / \sigma_4 \quad \text{(3-28)}$$

where τ_4^{DIST} is the L-kurtosis for the distribution. A reasonable criterion for an acceptable fit is $|Z^{DIST}| \leq 1.64$. If no distributions are found to be acceptable, a four (or more) parameter distribution such as the Kappa or the Wakeby distributions should be used. If several distributions are acceptable, the most robust distribution to misspecification in parameters should be used.

The Schaefer (1990) algorithm required that one distribution be chosen for all regions, to avoid the problem of discontinuity at the edges of the regions. After the distribution was chosen, the regions were discarded, as they were no longer necessary.

3.4.3.4 Estimation of the at-site frequency distribution parameters

The final step to regional frequency analysis is the estimation of the frequency distribution parameters. The at-site mean was used as the index-flood, to scale the regional frequency distribution up or down. The parameters of the regional frequency distribution were estimated according to the method of regionalization. In the Hosking and Wallis (1997) method, the weighted averages of the L-moments for all stations within a region were used as the regional average L-moments. In the Schaefer (1990) method, the interpolation functions based on basin characteristics were used to estimate the regional average L-moments. The regional L-moments were used to calculate the distribution parameters. After the distribution parameters and the index-flood were calculated, the distribution was used to calculate the flood frequency curves.

3.4.4 Description of Data used for Frequency Analysis

As described in Section 3.4.3, non-statistical hydrologic data are required for regional frequency analysis. This research used two types of data: climatological data and physiographic parameters (Solomon, *et al.*, 1968). These data were used to form homogeneous regions with the Burn, *et al.* (1997) algorithm (Section 3.4.3.2). In addition, the Schaefer (1990) algorithm required relationships to describe the variation in L-moments, and the non-statistical hydrologic data were used to determine these relationships. This section describes these data.

Two climatological values were derived from the streamflow data. The first value was the average Julian date of the largest flood. The second value was the average of the ratio of the peak flood to the mean streamflow. These two values, while derived from the streamflow data, are still independent from the L-moments of the data and can therefore be used for clustering.

The physiographic parameters were derived as defined by Solomon, *et al.* (1968). The parameters were calculated from a digital elevation model (DEM), called GTOPO30 (U.S. Geological Survey National Mapping Division), that is composed of 30 arc-second grid squares for all of Canada. The following parameters were calculated from the DTM: elevation, slope, azimuth, distance to ocean (8 compass directions), shield effect (8 compass directions), and barrier height (8 compass directions). These parameters are defined in Table 3-3. Solomon, *et al.* (1968) found these physiographic parameters to be useful for estimating monthly precipitation, temperature and runoff. The watersheds for each of the 36 streamflow gauges and dams were delineated in the digital terrain model, and the average value of each parameter over each of the basins was found. The drainage area was also found and used as a parameter.

Table 3-3 – Definitions of Physiographic Parameters

Elevation	Average elevation in the grid square
Slope	Local slope of the grid square
Azimuth	Direction of slope (in degrees with North defined as zero)
Distance to Ocean (DTO)	The distance to the ocean in the given compass direction (Figure 3-12a)
Shield Effect (SHE)	The sum of the elevation differential of all ascending stretches of terrain encountered when travelling from the ocean shore in the given compass direction to the corresponding point (Figure 3-12b)
Barrier Height (BH)	The difference between the elevation of the square and the highest elevation encountered in the given compass direction between the square and the ocean (Figure 3-12c)

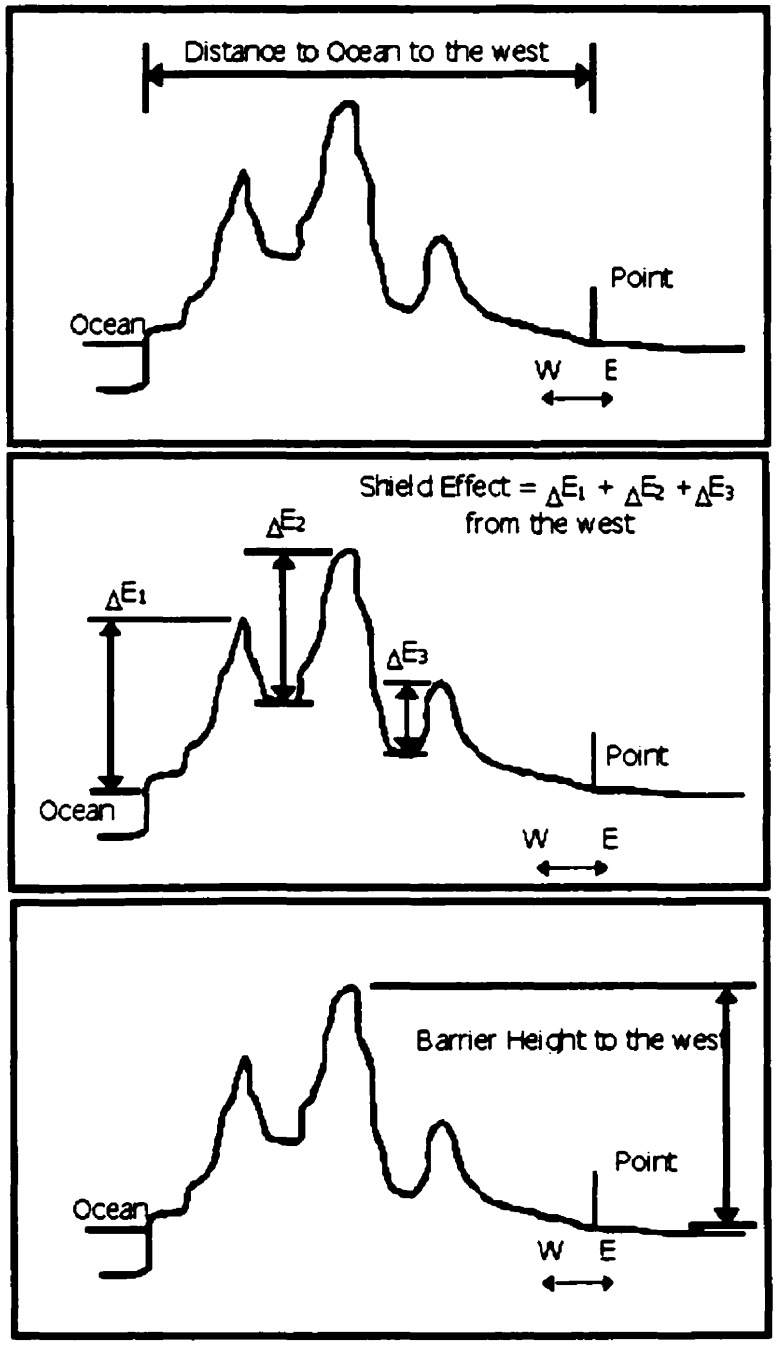


Figure 3-12 – Definition of Physiographic Parameters

The shield effect, which is the sum of all differential elevation increases, was calculated using a signed 2-byte integer (which has a maximum value of 32767). (Due to disk space considerations, the 2-byte integer was used.) The mountainous terrain in British

Columbia cause the maximum value to be reached for the majority of the pixels in the watershed in three of the eight compass directions (northwest, south, and southeast). Therefore, these three directions could not be used, as there was little or no variation in these parameters between watersheds and therefore they could not be used for forming regressions.

Due to significant cross-correlation, not all of the remaining physiographic parameters were used. To lower cross-correlation, only four of the eight compass directions were used. There were two choices for compass directions: north, east, south and west, or northeast, southeast, southwest, and northwest. Either set of compass directions would retain the majority of the information in the parameters. When linear regressions were calculated between the parameters and the L-moments, the regressions were better with the northeast, southeast, southwest, and northwest directions than with the north, east, south and west directions. Most summer storm systems in British Columbia travel in the northeast and southeast directions, because they are controlled by the jet stream. The jet stream is an unstable wind current located just below the tropopause that blows from west to east across North America. The jet stream, due to the locations of the polar and tropical air masses, the Coriolis effect of the rotation of the earth, the influence of the mountain ranges, and other factors, often blows northeastwards or southeastwards over British Columbia instead of directly east (Benoit, *et al.*, 1997b). In addition, variations in topography affect precipitation (e.g. less precipitation falls on the lee side of a mountain range, while more precipitation falls on the facing side of a mountain range). Many of the watersheds for the streamflow observation stations used in this research are oriented in a northwest, northeast, southwest, or southeast direction. Therefore, it was appropriate that these directions should give better

regressions. Therefore, the northeast, southeast, southwest, and northwest directions for distance to ocean, shield effect, and barrier height were used to help define the regions.

The following fifteen variables were used in the analysis: drainage area, slope, azimuth, average Julian date of peak flow, ratio of peak flow to mean flow, distance to ocean (northeast, northwest, southeast, southwest), barrier height (northeast, northwest, southeast, southwest), and shield effect (northeast, southwest). The cross-correlation table is included in Table 3-4. There is still some high cross-correlation, most notably in the distance to ocean variables, but these variables were kept since they do contain distinct and potentially important information.

Table 3-4 – Cross Correlation Table for Hydrologically Significant Variables

	Area	Ratio	Jday	Azi.	Slope	bh-ne	bh-nw	bh-se	bh-sw	dto-ne	dto-nw	dto-se	dto-sw	she-ne	she-sw
Area	1.000														
Ratio	0.125	1.000													
Jday	-0.470	-0.568	1.000												
Azi.	-0.175	0.152	0.201	1.000											
Slope	-0.087	-0.080	0.185	0.329	1.000										
bh-ne	0.057	-0.437	0.043	-0.341	-0.318	1.000									
bh-nw	0.280	-0.177	-0.391	-0.523	-0.511	0.482	1.000								
bh-se	0.013	-0.463	0.155	-0.444	-0.205	0.885	0.454	1.000							
bh-sw	0.481	-0.209	-0.436	-0.533	-0.388	0.562	0.784	0.465	1.000						
dto-ne	-0.122	-0.730	0.584	-0.189	0.274	0.163	0.057	0.350	0.092	1.000					
dto-nw	-0.028	-0.662	0.467	-0.219	0.146	0.074	0.165	0.190	0.193	0.944	1.000				
dto-se	0.008	0.626	-0.382	0.189	-0.159	0.098	-0.174	0.000	-0.168	-0.871	-0.962	1.000			
dto-sw	0.063	-0.371	0.170	-0.172	0.203	-0.318	0.163	-0.267	0.183	0.658	0.812	-0.910	1.000		
she-ne	-0.214	-0.566	0.535	0.066	0.319	0.557	-0.260	0.624	-0.095	0.563	0.348	-0.153	-0.169	1.000	
she-sw	0.179	0.095	-0.367	-0.037	0.170	-0.367	0.016	-0.513	0.253	-0.083	0.119	-0.266	0.570	-0.493	1.000

Legend:
Area – Drainage area of the basin
Ratio – Ratio of average peak flow to mean flow
Jday – Average julian date of peak flow
Azi. – Azimuth angle of the basin
Slope – Slope of the basin
BH-ne – Average barrier height of the basin in the northeast direction
BH-nw – Average barrier height of the basin in the northwest direction
BH-se – Average barrier height of the basin in the southeast direction
BH-sw – Average barrier height of the basin in the southwest direction
DTO-ne – Average distance to the ocean of the basin in the northeast direction
DTO-nw – Average distance to the ocean of the basin in the northwest direction
DTO-se – Average distance to the ocean of the basin in the southeast direction
DTO-sw – Average distance to the ocean of the basin in the southwest direction
SHE-ne – Average shield effect of the basin in the northeast direction
SHE-sw – Average shield effect of the basin in the southwest direction

These fifteen variables were used to generate relationships to describe the variation in the statistical parameters.

3.5 Chapter Summary

This chapter has been used to describe the relevant literature and background information for this research. In the following chapters, the MC2-PMS model will be used to develop a maximum precipitation event and it will be converted to a flood with the WATFLOOD/SPL model. The return intervals of the two floods will be used to compare the flood calculated from the PMS and the traditional PMF. The return intervals will be calculated from a frequency curve developed with the regional frequency analysis method.

4 Calculating the Probable Maximum Flood with the Probable Maximum Storm Model

The Probable Maximum Flood (PMF) is the theoretically largest possible flood in a particular area. For the Columbia River Basin, this may involve a combination of events, such as a large snowpack, a quick melting temperature sequence, and a large storm occurring together. One aim of this research was to improve the estimation of the PMF through improved estimation of the large storm, termed the Probable Maximum Precipitation (PMP). Previously, these estimates were based on a combination of meteorological and statistical methods. In this research, a physically-based atmospheric model was used as a possible alternative.

The MC2-PMS model was used to generate a "Probable Maximum Storm" (PMS). This was the largest storm that the model could generate, and was therefore the largest storm theoretically possible, given the model setup. The PMS was used in place of the Probable Maximum Precipitation to generate a flood. This section will describe the process used to develop the PMS and calculate the "new" PMF. The first section will describe the procedure to develop the PMS with the MC2-PMS model, and the second section will describe the conversion of the storm into a flood with the WATFLOOD/SPL model.

4.1 Calculating the Probable Maximum Storm

This section describes the derivation of the Probable Maximum Storm (PMS). First, the setup information for the MC2-PMS model is presented. The modifications of the perturbation to derive the PMS are described: the storm location and the wave parameters were modified. Finally, the PMS is compared to other significant precipitation events.

4.1.1 Setup Information for the MC2-PMS Model

The first step of this research was setting up the MC2-PMS model so that the calculated storm would affect the basin of interest, the Columbia River Basin. This includes date and location choices, and nesting information. Various historical storms that have occurred over the Columbia River basin were examined to determine common meteorological characteristics. Large winter storms tend to begin in the middle of the Pacific Ocean as a small perturbation. If conditions are appropriate, the perturbation in the Ocean develops and collects humidity from the ocean as it travels to the coast. Therefore, the model was setup to create a perturbation in the Pacific Ocean. The perturbation was described in detail in the literature review in Section 3.2.3.

The perturbation was placed in the mid-Pacific, in the approximate location of the winter storms. The perturbation had a two-day travel time to the coast (during which the storm could collect humidity from the ocean). It was necessary to embed the perturbation inside the general circulation (meteorological conditions) for a particular day, in order to establish initial and boundary conditions for the model. Since a summer storm that would occur when the reservoirs were at or near full was desired, the perturbation was embedded into the general circulation data for September 20, 1983 (the reservoirs are filled by snowmelt in the spring and early summer, and are full by late summer). However, it was found that the storm traveled north towards Alaska, and very little rain fell on southeastern British Columbia. Figure 4-1 shows the 6-day precipitation from one of these storms at the 150-km grid size. Most of the rain fell in the North Pacific and very little fell over the Columbia River Basin (red box).

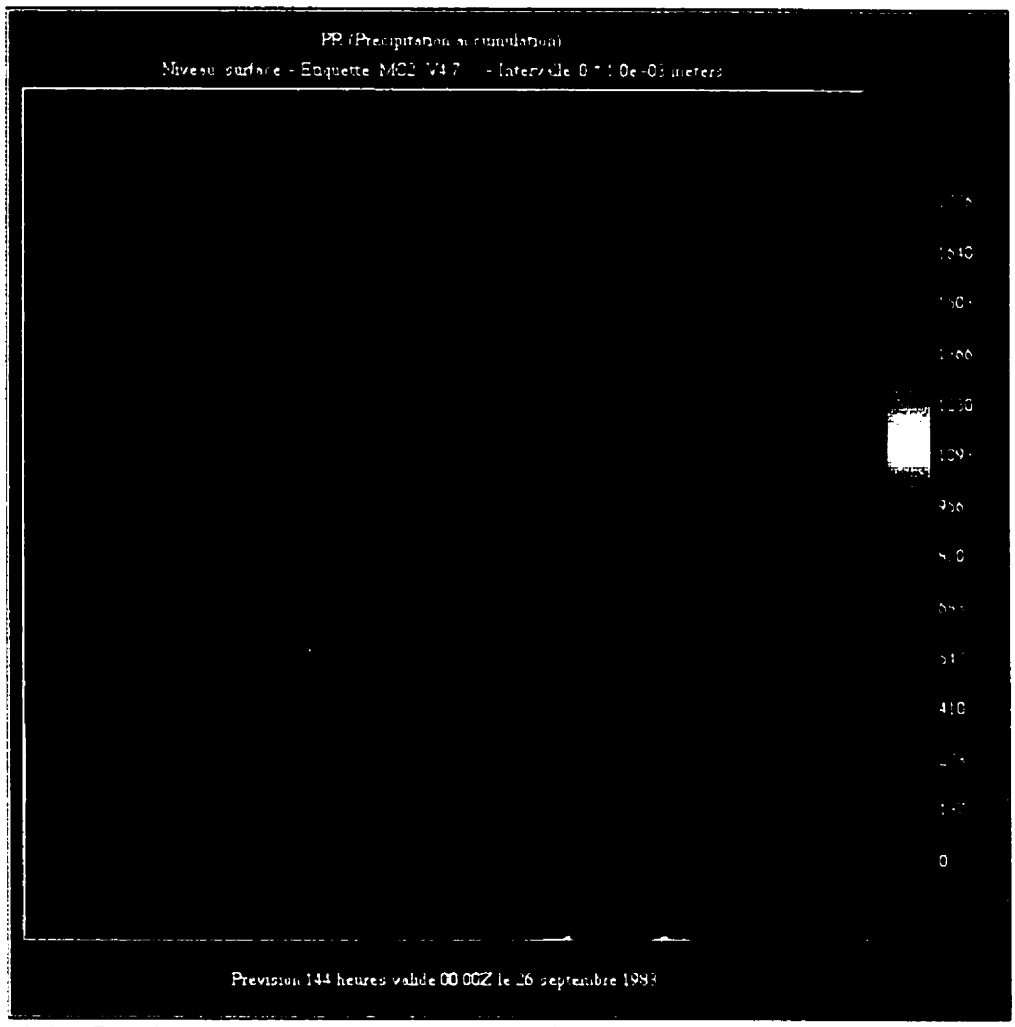


Figure 4-1 – Precipitation accumulation (mm) from a storm that traveled through Alaska

Pierre Pellerin at RPN (personal communication, 2000) examined two large historical summer storms for the Columbia River Basin (July 11-13, 1983 and August 25-27, 1984). He found that, unlike winter storms, large summer storms for the Columbia River Basin do not originate in the mid-Pacific. The storms begin near the coast and will tend to travel northwards unless a high-pressure system already exists in the north to force the storm inland. Therefore, it was decided to move the perturbation closer to the coast, and to embed the perturbation inside the July 13, 1983 general circulation data. This date was chosen because a large historical storm occurred on this date, and a high-pressure system existed in

the north. Figure 4-2 shows the geopotential height plot of the perturbation embedded in the July 13, 1983 general circulation. Geopotential height approximates the actual height (in this Figure, in decameters) of a pressure surface (in this Figure, 1000 mb) above mean sea-level. The pressure wave is visible in this Figure as a low-high-low pattern just west of the continental United States. Note the existence of the high-pressure system to the north-east of the perturbation in northern British Columbia and Alberta. With these changes, the storm affected the Columbia River Basin.

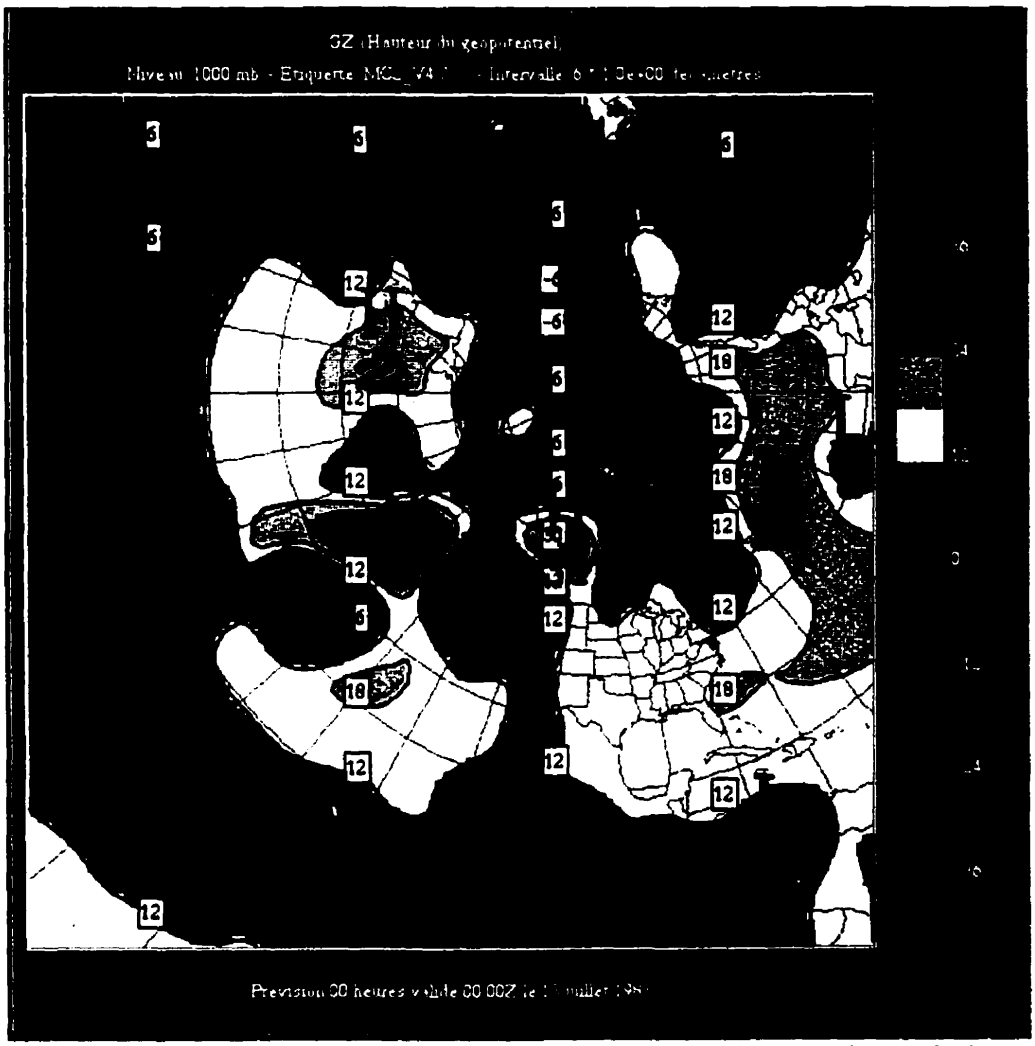


Figure 4-2 – The perturbation embedded in the July 13, 1983 general circulation

However, the new location for the storm was less than a one-day travel time to the coast. Therefore the storms would not have enough time to collect humidity from the ocean before reaching the coast. Wet and dry humidity profiles were created to address this problem. The curves in Figure 4-3 were used (Pellerin, personal communication, 2000). These curves were based on observed summer storms. They have not yet been validated to determine if they produce the largest possible storm.

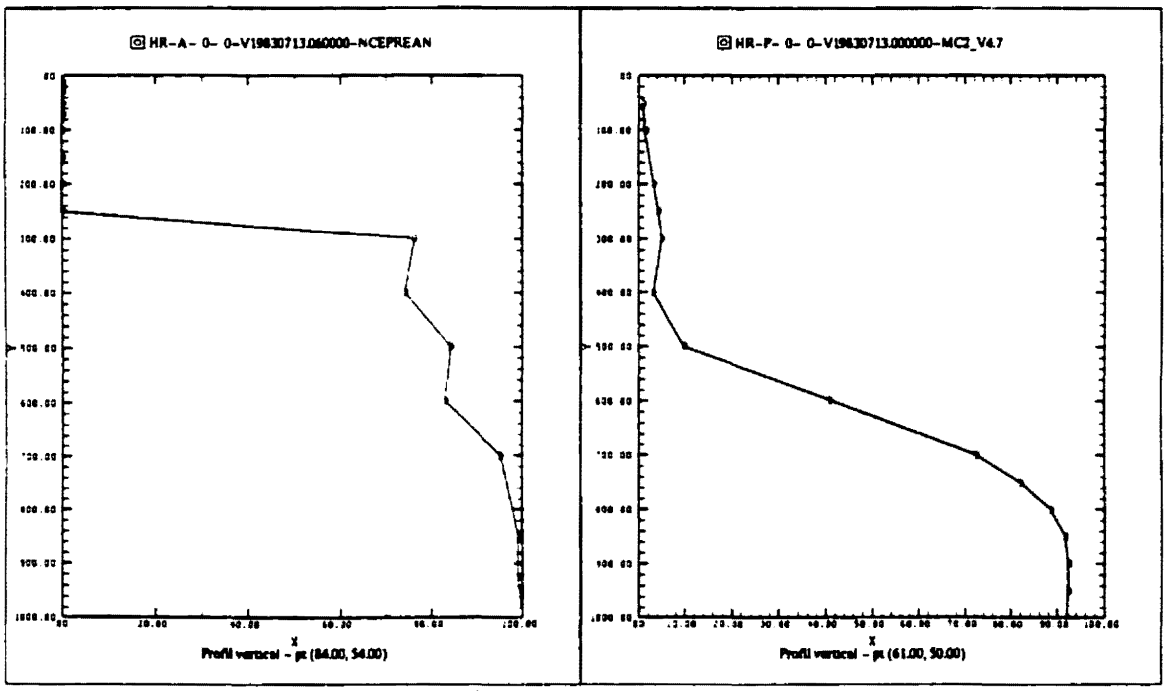


Figure 4-3 – Relative Humidity Profile in the a) Warm Front and b) Cold Front (Vertical axis is Pressure in mb, Horizontal axis is Relative Humidity in %)

Once the model was setup so that storms were occurring over the Columbia River Basin, it was necessary to "cascade" the model down to a resolution appropriate for the hydrological model. In this research, data at a resolution of 10 km were required for the WATFLOOD/SPL model. A three-part simulation was performed: 150 km, 50 km, and 10 km. The simulation domain for the 150 km resolution is shown in Figure 4-4, where the highlighted section is the location of the theoretical perturbation for the Probable Maximum

Storm. The 150 km simulation was a hemispheric model, to allow for simulation of global atmospheric phenomena. The highlighted section in Figure 4-5 shows the portion of the hemispheric simulation that was simulated in the 50 km resolution. The highlighted section in Figure 4-6 shows the portion of the 50 km simulation that was modeled in the 10 km resolution. Note that the Columbia River Basin is included as part of the domain. The lower mainland watersheds are also included, but are not examined as part of this research.

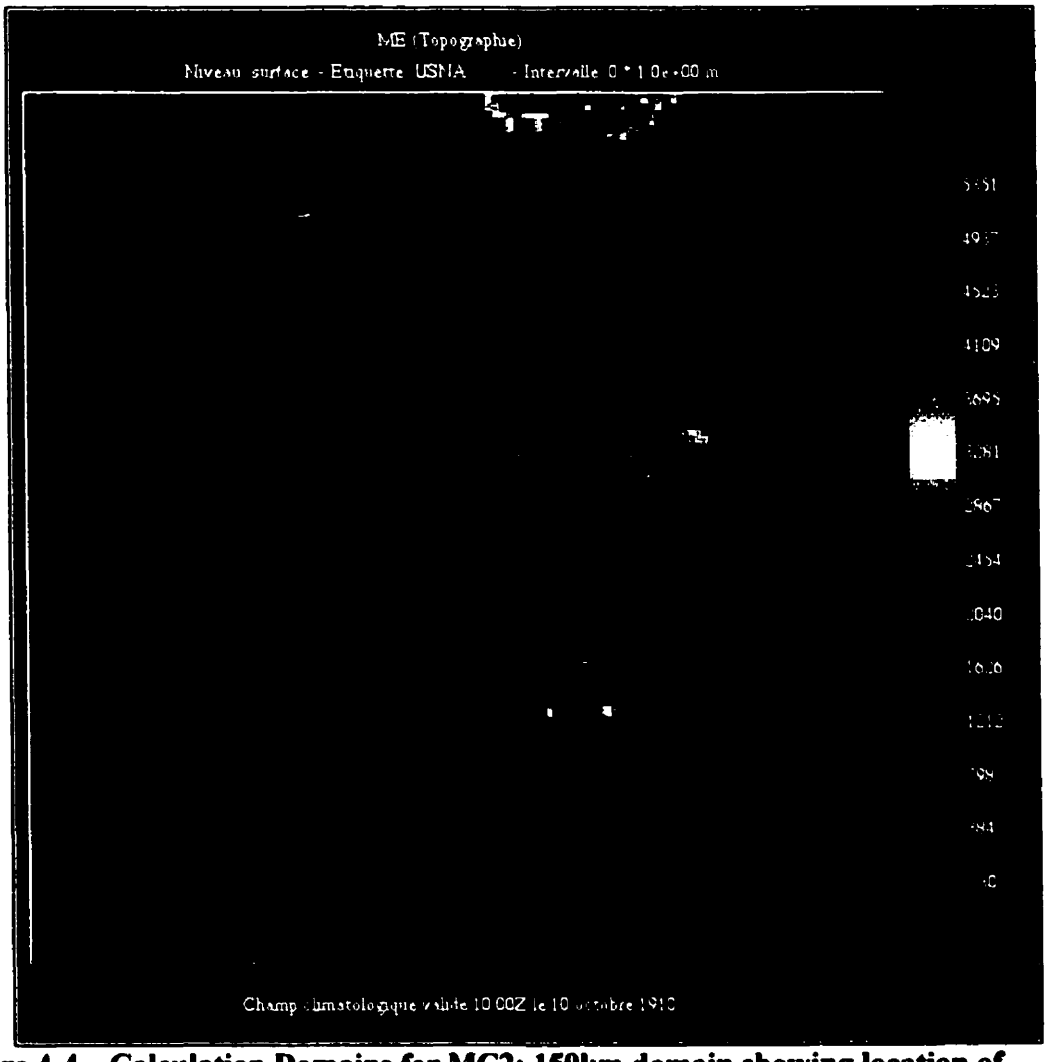


Figure 4-4 – Calculation Domains for MC2: 150km domain showing location of perturbation

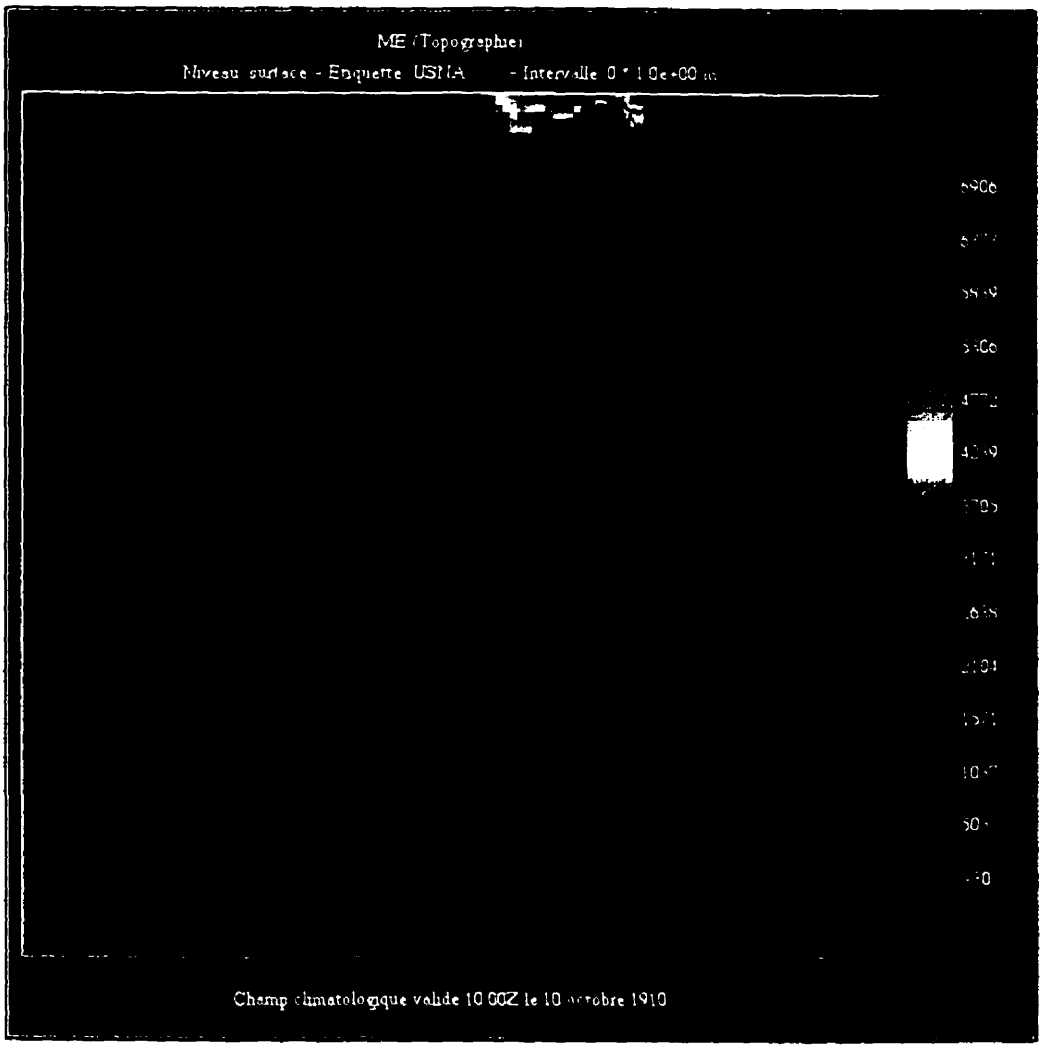


Figure 4-5 – Calculation Domains for MC2: 150 km domain showing 50 km domain

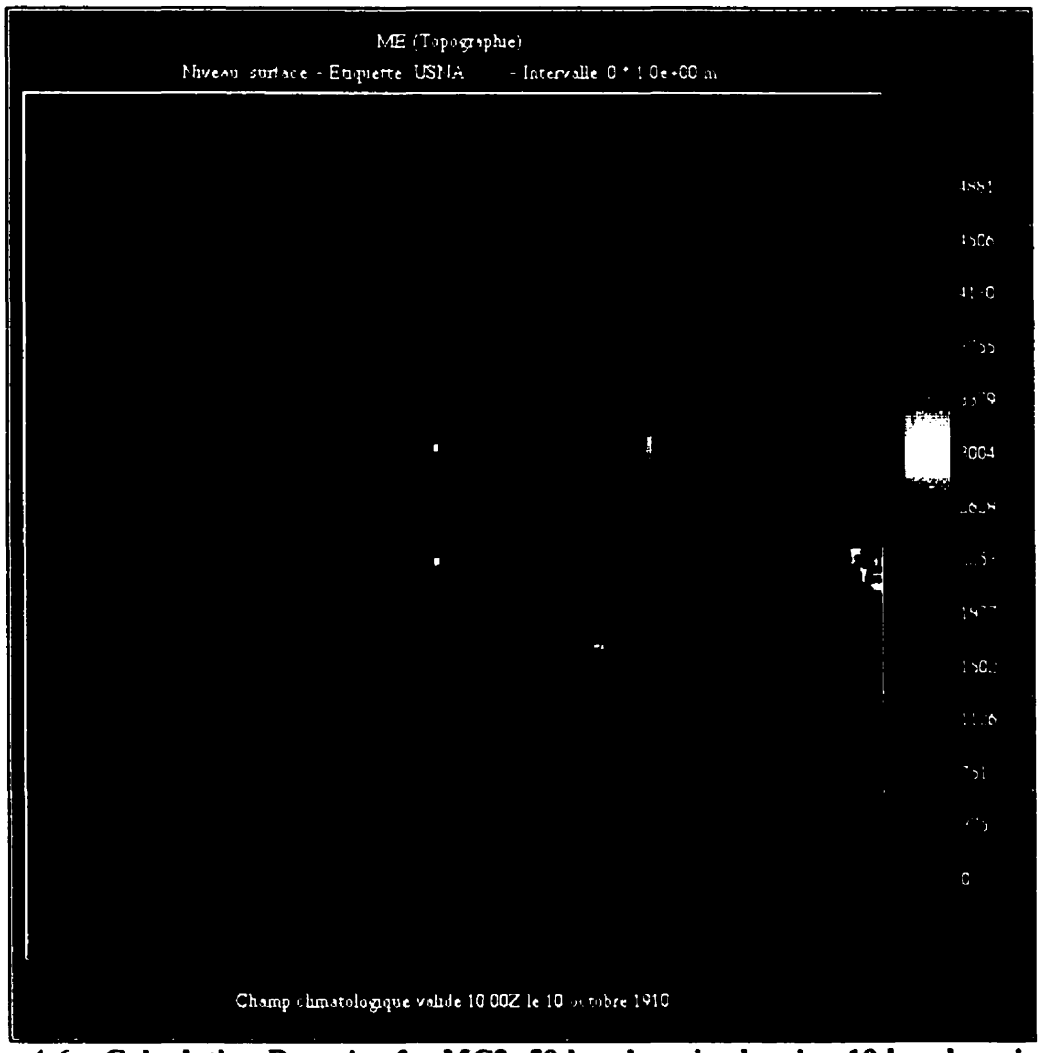


Figure 4-6 – Calculation Domains for MC2: 50 km domain showing 10 km domain

4.1.2 Adjusting the location of the storm

The location that Pellerin (personal communication, 2000) used for the perturbation generated a short, intense storm over the Columbia River Basin. After 12 hours of simulation, a cyclone with a depression of 32 mb was visible just to the west of Vancouver Island. For the rest of the simulation, the low pressure point remained west of Vancouver Island, but an "arm" of the cyclone passed from south to north over the basin between hours 6 and 30. (Note that at the 10-km resolution, the "arm" passed over the basin between hours

0 and 24, since three hours are lost with each cascade.) After 48 hours, the cyclone had dissipated, and a new storm was developing in northwestern U.S.A. A series of figures showing the development of the storm every three hours is available in Appendix A (Figures A-1 to A-17). For the Columbia River Basin, the precipitation occurred as a result of the "arm" passing over the basin (between hours 0 and 24). Therefore, the 24-hour precipitation was used in this research, so that the precipitation from the passage of the arm was obtained.

The perturbation location was based on the true location of the storm in July 1983. However, it was thought that another location might increase the storm's intensity and/or duration. Probable Maximum Flood analysis is often performed with a 3-day Probable Maximum Precipitation, so a longer storm would be desired.

The storm could not be moved to the east, since the storm must begin over the ocean. Therefore, various locations to the west of the original location were used to generate storms. As the storm moved west, it had more time to develop before it reached the coast. The storm was moved by degrees west and then north or south, as shown in Table 4-1. The precipitation was calculated at the 150-km, 50-km, and 10-km resolutions, as shown in Figure 4-4 to Figure 4-6. The average precipitation for the 24 hours over all of the Columbia River Basin grid squares was calculated and shown in Table 4-1. The Table also shows the average precipitation over the Mica Dam basin, because the July 1983 storm affected mainly the northern part of the basin near Mica Dam. For the entire Columbia River Basin, the largest precipitation occurred during the storm that was 2 degrees west and 2 degrees north of the original location. For the Mica Dam Basin, the largest precipitation occurred during the storm that was 2 degrees west and 1 degree north of the original location.

Table 4-1 – Comparison of the 10-km precipitation with the center of the perturbation placed in various locations

Location	Avg. Columbia River Basin Precip.	Avg. Mica Dam Basin Precip.
Original, historical July 1983 storm location	52.3 mm	55.2 mm
1 degree west, 1 degree north	51.1 mm	49.4 mm
1 degree west	57.4 mm	53.1 mm
1 degree west, 1 degree south	51.3 mm	49.2 mm
2 degrees west, 2 degrees north	59.1 mm	56.1 mm
2 degrees west, 1 degree north	57.5 mm	56.9 mm
2 degrees west	54.9 mm	54.1 mm
2 degrees west, 1 degree south	43.8 mm	39.3 mm
3 degrees west	48.6 mm	46.2 mm
4 degrees west	48.2 mm	45.4 mm

It should be noted that the higher average precipitation for the Columbia River Basin is due to the location of the precipitation event with relation to the Mica Dam basin. A large amount of precipitation occurred in the valley below Mica Dam, and these data were not included in the average for Mica Dam. For instance, the Illecillewaet River at Greeley station (WSC #08ND013, drainage area of 1170 km^2) is located in the valley below Mica Dam, and received an average of 106.5 mm for the storm that was 2 degrees west and 1 degree north of the original location.

Figure 4-7 shows the calculated 24-hour precipitation for the two storms highlighted above. The approximate limits of the Columbia River basin are shown with red rectangles. The storm that was 2 degrees west and 1 degree north consisted of one intense band of precipitation near Mica Dam. The storm that was 2 degrees west and 2 degrees north was a broader, less intense band of precipitation that was not focused on Mica Dam. This visual comparison shows that the storm two degrees west and 1 degree north of the original location had a better rainfall distribution over Mica Dam. Therefore this storm location was used for further analysis, since the PMS estimate for Mica Dam was desired.

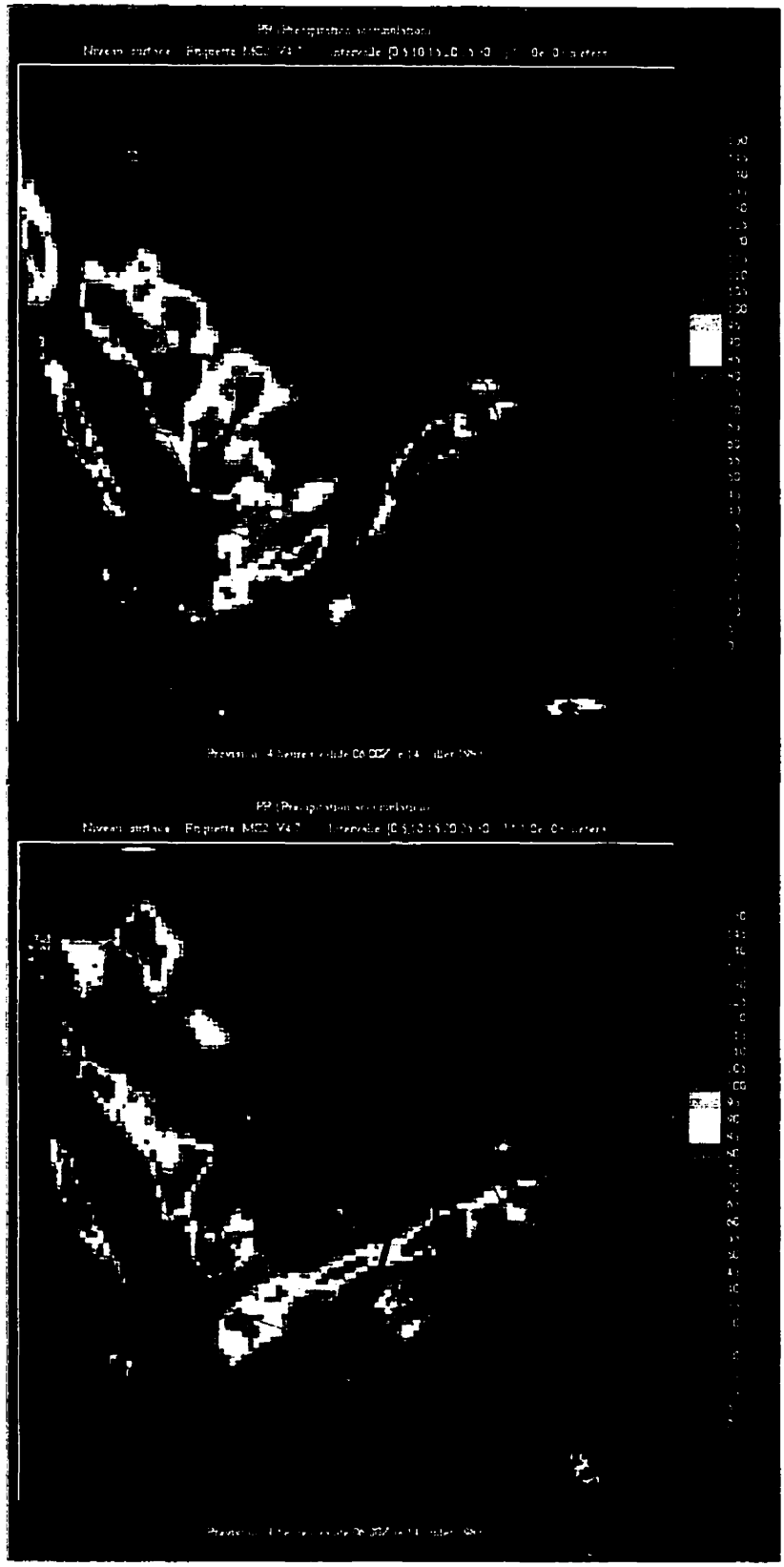


Figure 4-7 – Precipitation (mm) from Perturbation located at: a) 2 degrees west and 1 degree north, b) 2 degrees west and 2 degrees north

It was found that the location of the perturbation did not affect the duration of the storm. Each location generated a storm that lasted approximately 24 hours. In addition, the location of the perturbation did not greatly affect the direction of the storm. The central pressure low for the storms stayed just to the west of Vancouver Island. As the location of the perturbation moved west, the strength of the precipitation band near Mica Dam varied. The strongest precipitation band near Mica Dam occurred for the storm that was 2 degrees west and 1 degree north of the original location.

The majority of the precipitation (that fell over the Columbia River Basin) fell in a precipitation band near Mica Dam. The topographic effect of the mountains caused the precipitation band to be narrow and confined to the mountain pass. Figure 4-8 illustrates that the majority of precipitation fell over the lower elevations of the domain. The figure was made from the storm that was 2 degrees west and 1 degree north of the original location, and the elevation was the average elevation for the 10 km^2 grid. The use of the average values for a 10 km^2 grid is not as accurate as point precipitation and point elevation data, but it suffices for illustrating the topographic effect. Above approximately 1600 m, the 24-hour precipitation decreased. These results are in agreement with Jarrett (1990a), who found no meteorological, hydrological, or paleohydrological evidence of significant precipitation at high elevations in the American Rocky Mountains (2300 m in Colorado, but only 1600 m in Montana – Jarrett, 1990b). The WMO method for PMP calculation does not ensure that precipitation occurs mainly in the valleys and mountain passes. Therefore, the MC2-PMS method is superior to the WMO method in its calculation of the topographic effect on precipitation.

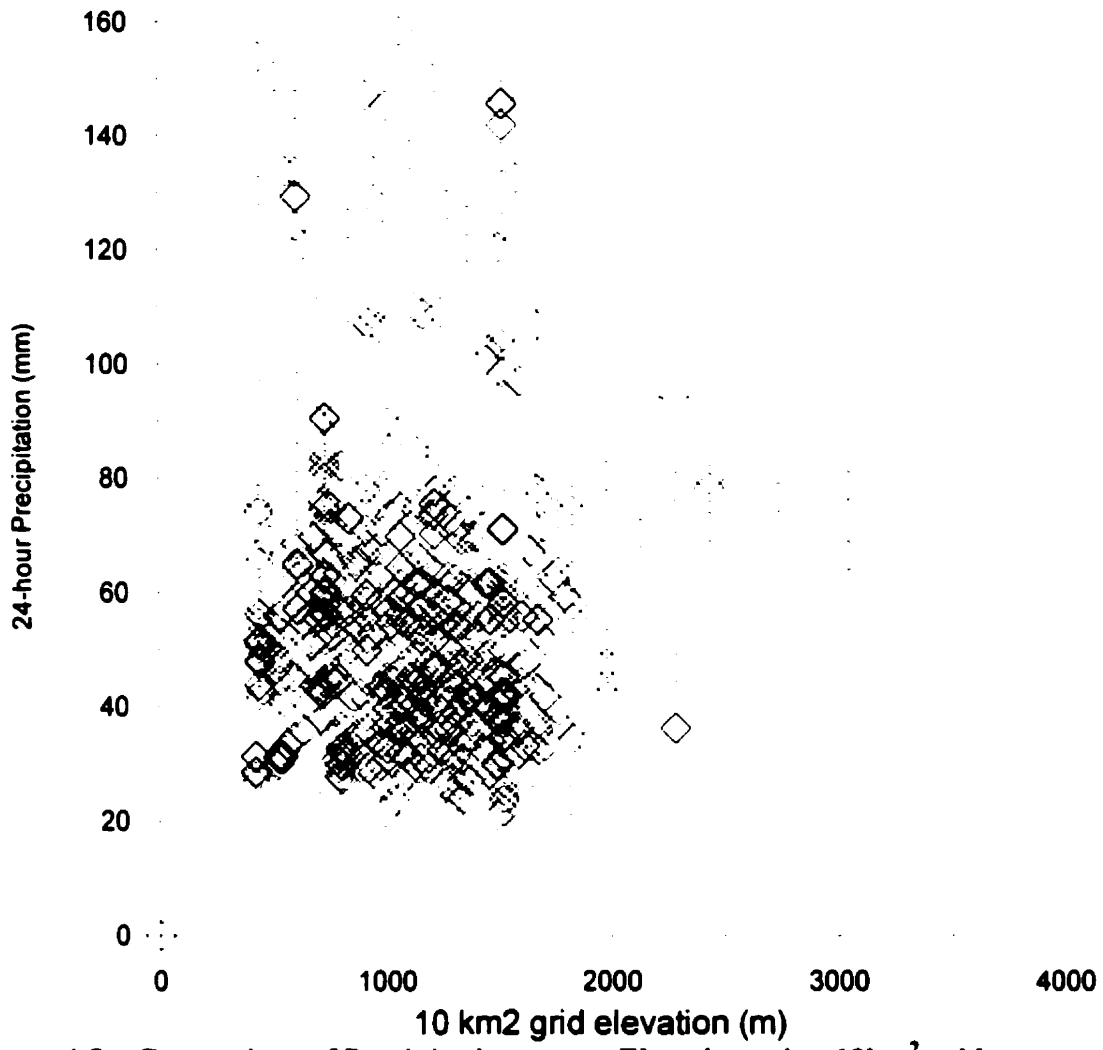


Figure 4-8 – Comparison of Precipitation versus Elevation using 10km^2 grids

4.1.3 Parameter Variation to determine the Probable Maximum Storm

Once the storm location was chosen, the pressure and temperature wave parameters were varied to determine the maximum storm. The MC2-PMS model is a combination of the MC2 model and an added PMS module. The PMS module creates an atmospheric perturbation, and the MC2 model develops the perturbation. In this research, the search for a maximum storm consisted of modifying the PMS module parameters only. The internal meteorological parameters of MC2 were not altered because this was outside the scope of this preliminary hydrological investigation. This section develops the relationships between

the wave characteristics and the storm magnitudes and directions. First, the methodology for the parameter variation is presented. Next, the results of the parameter variation are examined to determine the effect of the parameters on the storm location and the effect of the parameters on the storm magnitude.

4.1.3.1 Parameter Variation for MC2-PMS perturbations

Several PMS module parameters controlled the characteristics of the pressure and temperature waves (described in Section 3.2.2.2). Preliminary studies at RPN showed that three parameters in particular, the amplitude of the temperature wave, the amplitude of the pressure wave, and the lapse rate were the most important parameters (Pellerin, personal communication, 2000). These parameters were varied to determine how they would affect the characteristics of the storm. In particular, two relationships were desired: between the parameter values and the direction of the storms; and between the parameter values and the magnitude of the storm.

Accordingly, the three variables were allowed to vary across their "normal" range. The normal range for the amplitude of the temperature wave is 5°C to 15 °C. For the amplitude of the pressure wave, the normal range is 3 mb to 15 mb. The normal range for the lapse rate is 6.5 to 10 °C/km. A high, medium, and low value of each variable was chosen, and every combination was used to generate a storm (27 storms). All storms were generated with the perturbation located two degrees west and one degree north of the original location, as determined from Table 4-1. The variables are listed in the table below.

Table 4-2 – Values of parameter variables for each storm

Storm #	Pressure wave amplitude – ax (mb)	Temperature wave amplitude – bx (°C)	Lapse Rate – s (°C/km)
1	15	15	6.5
2	7	15	6.5
3	3	15	6.5

4	15	10	6.5
5	7	10	6.5
6	3	10	6.5
7	15	5	6.5
8	7	5	6.5
9	3	5	6.5
10	15	15	8
11	7	15	8
12	3	15	8
13	15	10	8
14	7	10	8
15	3	10	8
16	15	5	8
17	7	5	8
18	3	5	8
19	15	15	10
20	7	15	10
21	3	15	10
22	15	10	10
23	7	10	10
24	3	10	10
25	15	5	10
26	7	5	10
27	3	5	10

These 27 storms were generated, and then cascaded from 150 km grid squares to 50 km grid squares, and finally down to 10 km grid squares (Figure 4-4 to Figure 4-6). Each cascade refined the precipitation estimate, as the mountain topography improved.

4.1.3.2 The relationship between the parameters and storm direction

The movement of the low pressure system was used to determine the travel direction of the storm. However, the 27 storms all behaved in the same way. In each case, the central low formed and deepened, and traveled north to Vancouver Island. The low pressure system stayed west of Vancouver Island for the entire length of the storm. A low pressure "arm" of the storm, which passed from south to north over the basin between hours 6 and 18, caused the precipitation over the Columbia River Basin. Each storm produced similar pressure plots

at each time step (the time variation in pressure was similar to that of the series of figures (Figures A-1 to A-17) shown in Appendix A for the storm described earlier). The storms differed mainly in the depth of the low pressure system. These data showed that the storm travel direction was not sensitive to characteristics of the perturbation.

Similarly, the travel direction for the storm was also not affected by changing the location of the perturbation (Section 4.1.2). In each case, the low traveled north to Vancouver Island, and a low pressure "arm" passed over the Columbia River Basin.

Further experiments were performed with the perturbation embedded inside September 20, 1995. These storms traveled in a different direction from storms that were embedded inside July 13, 1983. However, varying the parameters showed that all storms embedded in September 20, 1995, traveled in the same direction.

Generally, the storm travel direction seemed insensitive to the characteristics of the perturbation; however, it was sensitive to the conditions surrounding the perturbation. Therefore, it was assumed that the travel direction is dependent only upon the atmospheric characteristics of the day chosen for embedding the storm. If a storm over another basin were desired, a different day would have to be chosen as the embedding day. The historical record would need to be examined to find a day with appropriate conditions (such as the day a significant historical storm occurred).

4.1.3.3 The relationship between the parameters and storm magnitude

The magnitude of the storms varied widely. For instance, Figure 4-9 shows a small storm with very little precipitation, and a large storm with significant precipitation. The red rectangles show the approximate limits of the Columbia River Basin.

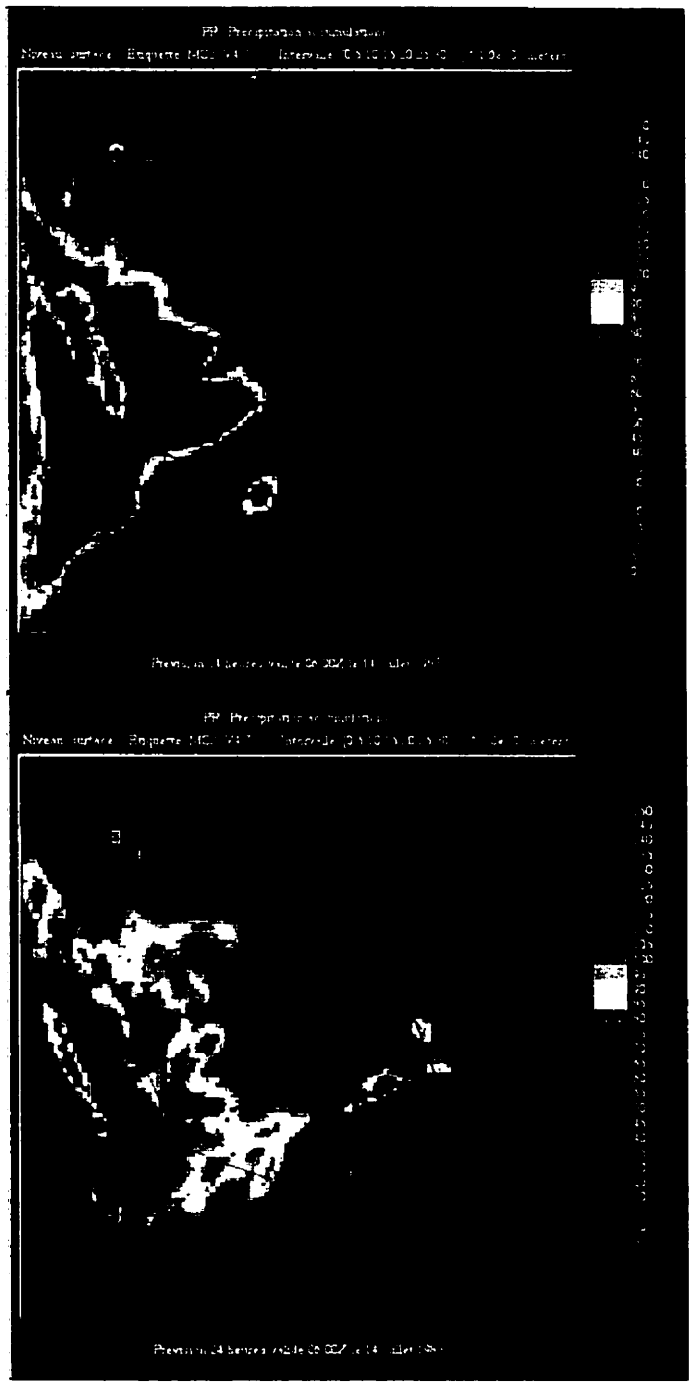


Figure 4-9 – a) Low Precipitation Storm (#25), b) High Precipitation Storm (#1)

Due to this wide variation, an examination of the relationship between the parameters and the storm magnitude required a definition of storm magnitude. There were many possible indicators, such as: intensity, total depth, and areal coverage. This research was

concerned with calculating the Probable Maximum Flood. As such, a statistic that could be used to evaluate the effect of the storm on runoff was used.

The runoff is related to the total depth of precipitation that falls over a particular watershed. Therefore an appropriate statistic would be the average (over all grid squares in a particular basin) of the total depth of precipitation. Since the precipitation for all 27 of the storms occurred during the first 24 hours of the 10-km simulation, the total depth of precipitation was taken as the precipitation after 24 hours. The averaging was performed for the basin upstream of Mica Dam, since the storm was focused on the area near Mica Dam.

This statistic was calculated for all 27 storms, and displayed in Table 4-3. The highest precipitation (55.0 mm) occurred in storm 1, and the lowest precipitation (8.4 mm) occurred in storm 25. The table has been arranged so that the effect of the lapse rate (s) may be clearly seen. In all cases, when the wave amplitudes were kept constant, the effect of the lapse rate on precipitation was small. The average precipitation varied by less than 7% when the lapse rate was changed. For instance, storms 1, 10 and 19 had constant values for the wave amplitudes (ax=15, bx=15), and three different lapse rates. However, the precipitation was not greatly affected (51.1 mm to 55.0 mm). The lapse rate was therefore considered a minor variable. The storms with the lapse rate set equal to 10 °C/km generally had the highest precipitation.

Table 4-3 – Comparison of Storm Magnitudes for 27 simulations (based on 10-km precipitation estimates)

Ax, Bx	Storm (s=6.5)	Avg. Mica Precip. (mm)	Storm (s=8)	Avg. Mica Precip. (mm)	Storm (s=10)	Avg. Mica Precip. (mm)
15, 15	1	55.0	10	51.1	19	54.9
7, 15	2	39.9	11	39.9	20	42.3
3, 15	3	25.5	12	26.2	21	26.2
15, 10	4	42.0	13	42.0	22	43.0
7, 10	5	36.1	14	38.3	23	38.7
3, 10	6	29.5	15	29.5	24	30.0

15, 5	7	8.9	16	8.9	25	8.4
7, 5	8	17.3	17	17.3	26	17.3
3, 5	9	18.8	18	19.6	27	19.6

The table shows that the two wave amplitudes had a significant effect on precipitation. In addition, the data showed a high degree of interaction between the variables. Therefore, the plot in Figure 4-10 was created to show the relationship between the wave amplitudes and the average precipitation. When one variable was high and the other was low, very little precipitation resulted: the conditions were insufficient to generate significant precipitation. If both amplitudes were low, a small storm developed. However, significant precipitation resulted when both amplitudes are large. The maximum storm occurred in the vicinity of ax=15 and bx=15.

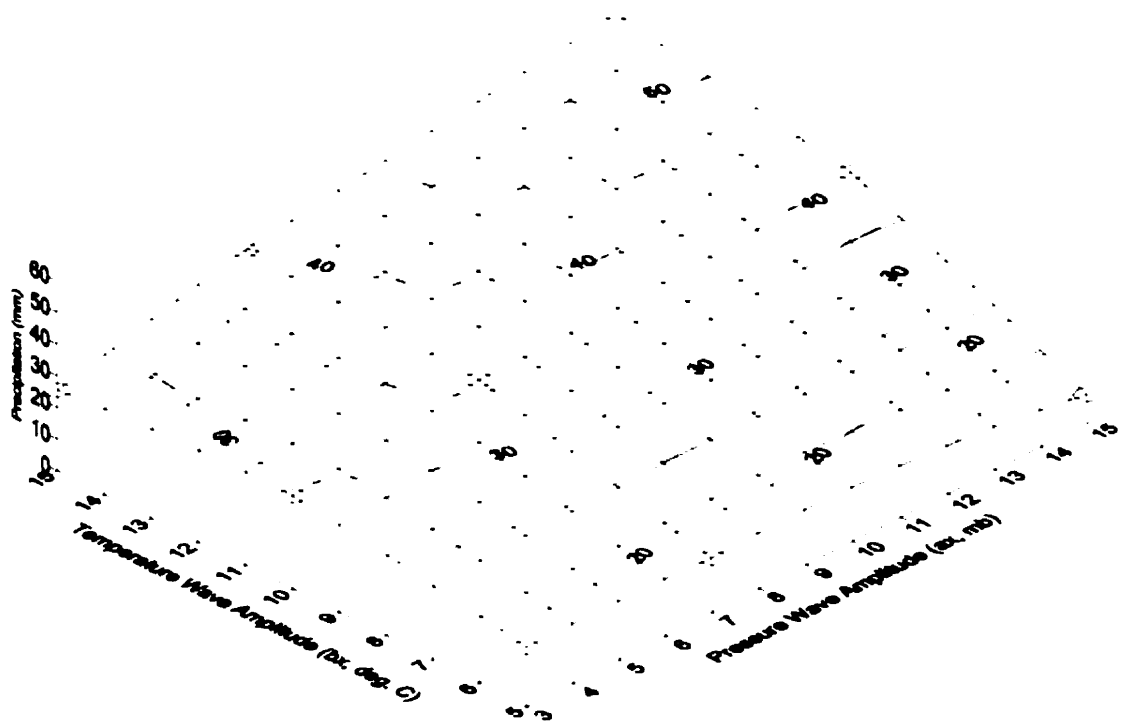


Figure 4-10 – Average Mica Precipitation versus Ax and Bx, using the nine original simulations

These data showed that large precipitation occurred when the temperature and pressure wave amplitudes were both large. Therefore, several simulations in the vicinity near ax=15 and bx=15 were performed in order to refine the estimation of the maximum storm. All of the storms were calculated with the lapse rate, s, equal to 10 °C/km. The extra storms, along with the values of the two wave amplitudes, are listed in Table 4-4.

Table 4-4 – Extra simulations to refine the maximum storm

Storm #	Ax	Bx	Storm #	Ax	Bx
28	13	13	36	12	13
29	11	11	37	11	14
30	12	12	38	10	13
31	13	15	39	9	13
32	15	13	40	10	12
33	11	13	41	10	14
34	13	11	42	9	12
35	11	12	43	9	14

These storms were generated with MC2-PMS. The average for Mica Dam of the 24-hour precipitation estimates was calculated for each storm. These data were used to generate another surface plot of precipitation versus the two amplitudes (Figure 4-11). This Figure shows that a maximum in precipitation was found (at ax equal to 10 mb and bx equal to 13 °C). This storm was the Probable Maximum Storm.

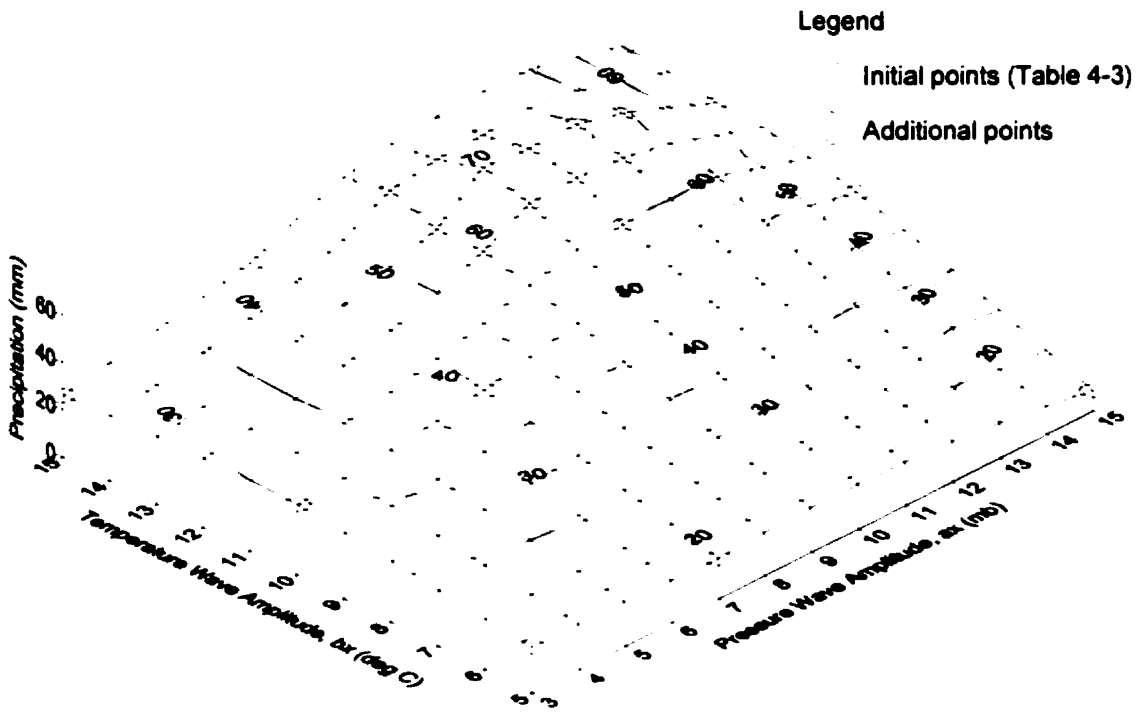


Figure 4-11 – Average Mica Precipitation versus Ax and Bx, including the extra simulations and showing the maximum precipitation

4.1.4 Comparing the PMS to Other Significant Precipitation Events

The Probable Maximum Storm is shown in Figure 4-12. This was the largest precipitation that could be generated by MC2 with this particular perturbation embedded in this particular day. Figure 4-12a shows the 24-hour precipitation for the entire domain of the 10-km simulation, while Figure 4-12b shows the 24-hour precipitation over the Columbia River modeling area only. The vertical lines in Figure 4-12b indicate the area above Mica Dam. The total precipitation ranged from approximately 19 mm to 111 mm, with an average of 73.4 mm over the Mica Dam basin. The topographic effect of the mountains on the precipitation is visible in Figure 4-12. The use of new amplitude parameters caused the precipitation to move slightly; it occurred mainly in two areas: in the valley above Mica Dam, and in the valley below Mica Dam. The Illecillewaet River at Greeley station (WSC #08ND013) is located near Rogers Pass (the horizontal lines in Figure 4-12b correspond to

the Illecillewaet River basin). The average basin precipitation over the Illecillewaet River basin for the 1983 storm was 89 mm, and corresponded to a 500-year return interval flow. The average basin precipitation for the PMS storm was 96 mm, the decrease from 106 mm caused by the precipitation moving out of the Rogers Pass and into the valley. (It should be noted that this storm was created in order to maximize precipitation for the Mica Dam basin, and it therefore does not necessarily contain the maximum precipitation for any other basin.) The topographic effect of the mountains was well represented in the precipitation plots from the MC2-PMS model. The WMO method does not tend to concentrate the precipitation in valleys and mountain passes, and therefore this method was an improvement to the calculation of maximum precipitation.

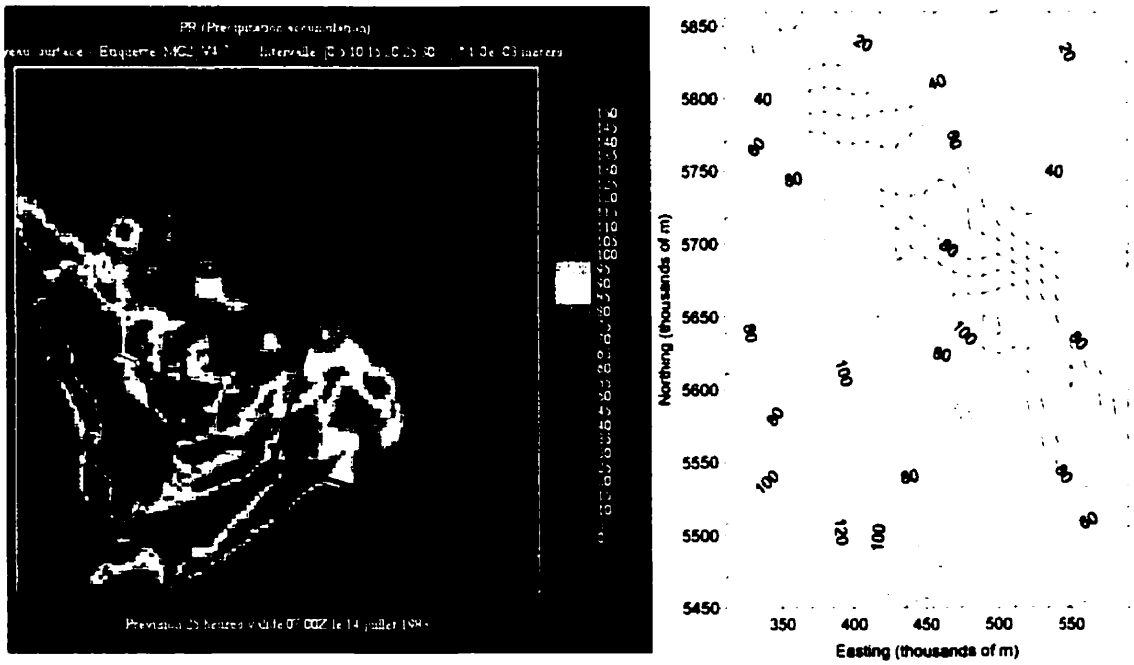


Figure 4-12 – 24-hour Precipitation from the Probable Maximum Storm derived by MC2-PMS

The Probable Maximum Storm precipitation was compared to the precipitation from the July 11-13, 1983 storm. The data for the 1983 storm were taken from the HRBL model

precipitation estimates by Danard (1996b). In the three-day period from July 11 to 13, 1983, the largest amounts of precipitation fell on July 11, 1983. Figure 4-13a shows a contour plot of the historical precipitation for this date (adjusted with the precipitation adjustment factor field described in Section 3.3.3). An average of 32.4 mm of precipitation fell over the Mica Dam basin on this date (range: 0 to 76 mm). The difference between the PMS and the July 11, 1983 data was calculated and plotted on a contour plot in Figure 4-13b. For most of the Columbia River Basin domain, there was a greater amount of precipitation for the PMS storm. There were small sections where the PMS storm was 0 to 20 mm less than the 1983 storm, but these were mainly outside of the Mica Dam basin. Therefore, the Mica Dam basin experienced greater precipitation during the PMS than it experienced during the 1983 storm.

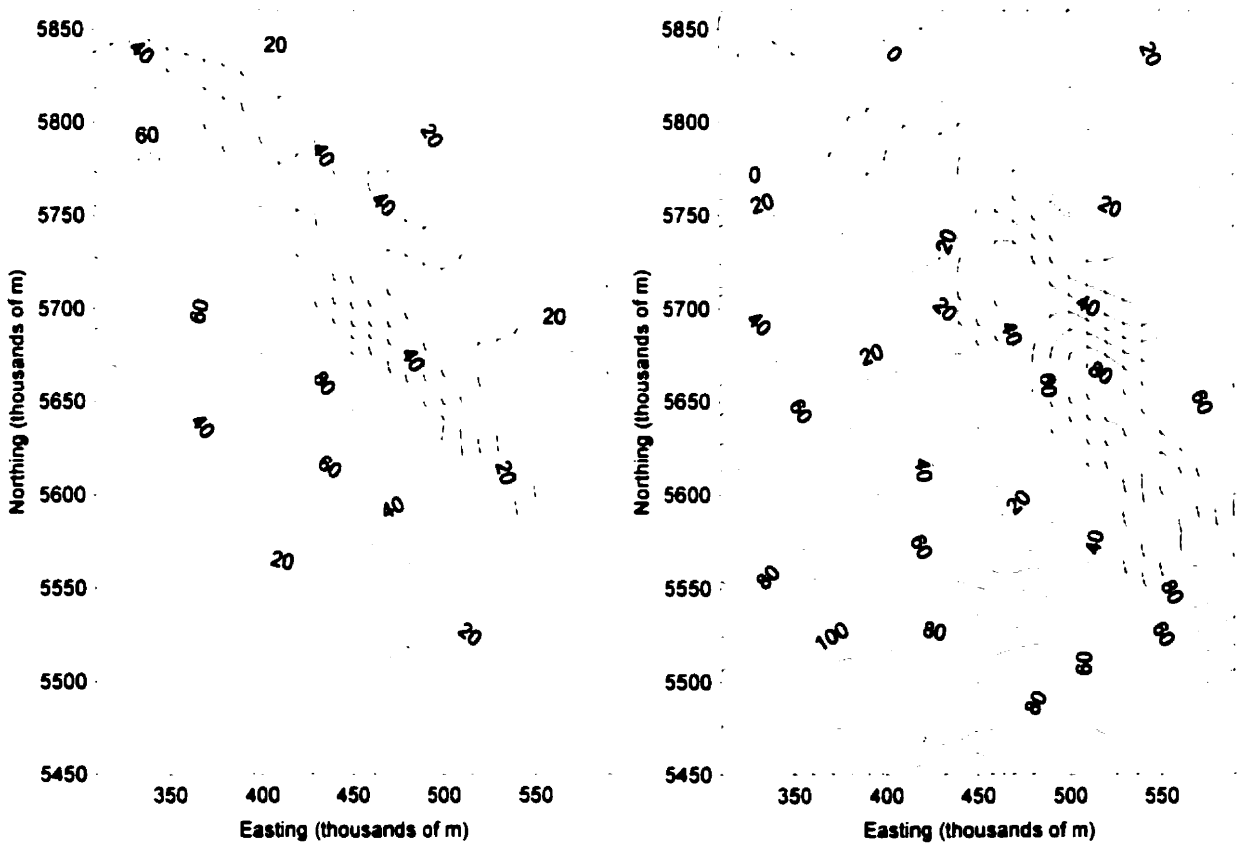


Figure 4-13 – a) Precipitation (in mm) from July 11, 1983 b) Difference (in mm) between the precipitation for the Probable Maximum Storm and July 11, 1983

Miller (1993) estimated the Probable Maximum Precipitation (PMP) for the Columbia River Basin. The PMP estimate has not been officially adopted by BCHydro and is therefore confidential at this time, so a full comparison between the PMP and the PMS cannot be made. However, an average 24-hour precipitation of 73.4 mm over the Mica Dam basin for the PMS was on the order of 50% of Miller's PMP estimate for the Mica Dam basin.

These comparisons show that the MC2-PMS estimate of the Probable Maximum Storm is larger than the 1983 historical storm (and hence, is larger than all observed historical storms), but lower than the PMP estimate.

4.1.5 Summary of the MC2-PMS method to develop precipitation events

The PMS was calculated with the MC2-PMS model using physically-plausible techniques, and was therefore a physically-based maximum storm within the range of perturbations allowed in this study. The mountain topography was evident in the precipitation of the PMS: the majority of the precipitation occurred within the valleys and in the mountain passes (e.g. Figure 4-8). This result matched with research by Jarrett (1990a), who found that significant precipitation events do not occur at high elevation (in Colorado, no significant precipitation occurred above 2300 m, but this decreases at latitude increases to become 1600m in Montana – Jarrett, 1990b). Other research by Jarrett and Costa (1988) showed that storm transposition from a region of low elevation to a region of higher elevation is not supported by hydrological, meteorological, or paleohydrologic data. The mountain topography was not as easily discerned in the PMP precipitation plots (Miller, 1993), and significant precipitation occurred at high elevation. The PMP estimate for the Columbia River Basin depends on storm transposition, and may be inaccurate. The WMO

method cannot recommend a standard method for topographic regions because of the great variation in the effects of topography upon precipitation (WMO, 1986, p. 139) and site-specific analyses are often approximations of the topographic effect. In this research, the magnitude of the PMS suggested that the WMO method may overestimate the true atmospheric maximum storm. These results are very important for determining safety at hydrologic structures (e.g. dams) in mountainous terrain.

These data suggest that the MC2-PMS model is suitable for estimating severe precipitation events. However, the internal parameter settings within MC2 have not yet been verified. A meteorological analysis should be performed to determine the appropriate settings for severe precipitation events. The PMS estimate is subject to change, but the preliminary results presented in this research indicate that the model is suitable for this use.

4.2 Generating the Floods caused by the Probable Maximum Storm

Once the Probable Maximum Storm was found, it was necessary to convert the precipitation into streamflow. In order to calculate the flood caused by the PMS, the storm must be entered into a hydrological model. There were several reasons for using a distributed, physically-based hydrological model, such as WATFLOOD/SPL. Firstly, WATFLOOD/SPL was designed to accept the gridded output of an atmospheric model (such as MC2), and the gridded precipitation could be used directly in the hydrological model. Secondly, the PMS and PMP represent larger magnitude storms than any historical storm. Physically based models are required when estimating flows that have not been observed in the past (Refsgaard and Knudsen, 1996). Thirdly, the hydrological processes behave non-linearly, and simple prediction techniques do not allow for accurate forecasting of runoff, as shown in the research below. Finally, the PMS must be combined with a large snowpack and

a fast melting temperature sequence to form one scenario of the PMF. The snowpack and the temperature sequence have not been historically observed. The combination must be simulated together in a physically based manner, due to the interactions between the antecedent conditions and the runoff from the storm. Therefore, the WATFLOOD/SPL model was used to convert the PMS into a flood.

The flood derived from the PMS is compared to a historical flood and to the theoretical PMF.

4.2.1 Comparison of the PMS flood to the historical 1983 storm

The first comparison was between the historical 1983 storm and the flood from the PMS. In Figure 4-14, there are three hydrographs. The grey trace shows the observed inflow hydrograph for Mica Dam during the July 11-13 storm. The solid black trace is a calculated hydrograph, based on using the 1983 data (Danard, 1996b) as forcing data for WATFLOOD/SPL. There are some differences between these traces; in particular, the WATFLOOD/SPL model overestimates the streamflow before the storm. However, during the July 11-13 storm event, the traces were very similar and the WATFLOOD/SPL model accurately predicted the volume of runoff. The dashed black trace used the HRBL model data for 1983 as forcing data for WATFLOOD/SPL, except that the data for July 11 were replaced by the PMS precipitation (the data for July 12 and July 13 were unchanged). A comparison of the solid black and dashed black traces shows the volume of streamflow was approximately doubled with the PMS on July 11 compared to the historical 1983 storm.

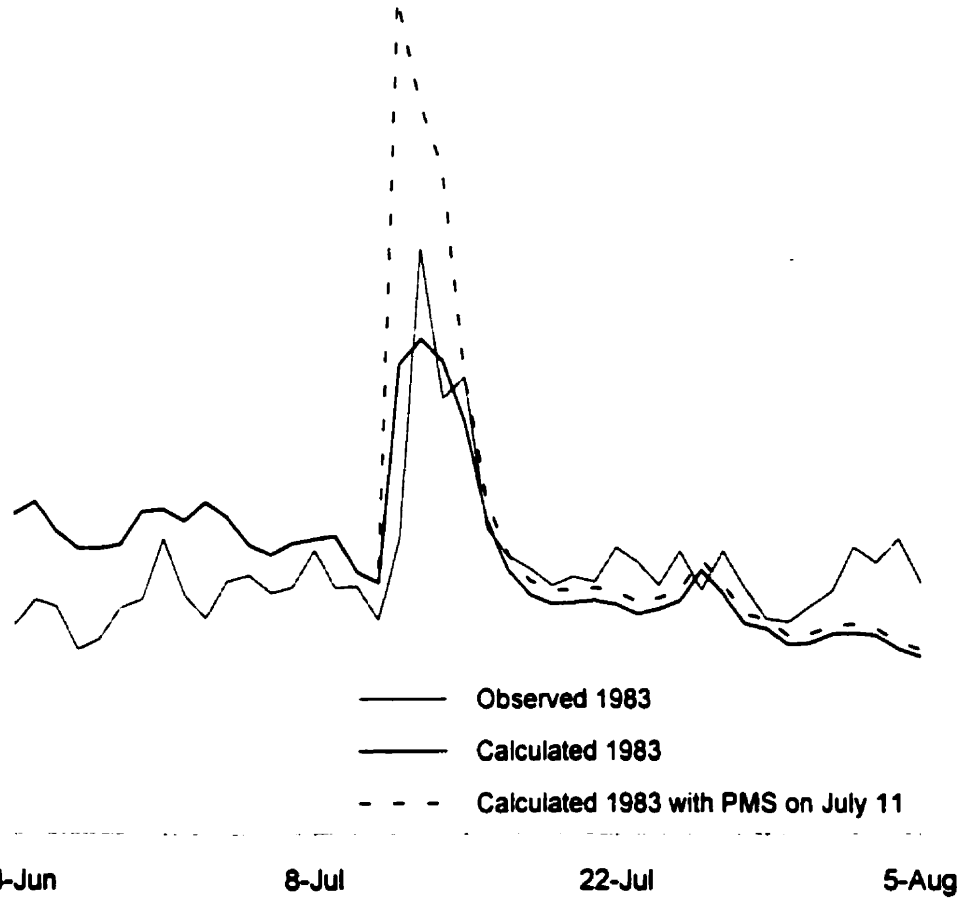


Figure 4-14 – Comparison of Mica Dam inflow with the Historical Storm and the PMS

4.2.2 Comparison of the PMS flood to the traditional PMF estimate

The flood from the PMS was also compared to the PMF estimates. However, the comparison was qualitative instead of quantitative because the PMF estimate has not been officially adopted by BCHydro and is therefore confidential at this time.

One particular scenario for the PMF consists of a maximized snowpack and a maximized melting temperature sequence, followed by the 3-day Probable Maximum Precipitation (PMP). The 100-year snowpack for April 1 was used as the maximized snowpack. The 100-year melting temperature sequence was derived for May 15 to June 4,

and used in place of the maximized melting temperature sequence. Danard (1995) calculated these variables. In addition, Danard (1995) also derived a distributed PMP, centered on the Mica Dam basin, based on the Miller (1993) PMP estimates. The PMP was placed on one of three different dates: May 15, June 1, and June 15. In addition, a "pre-storm" was also calculated to ensure that the watershed would be wet prior to the PMP (i.e. soil moisture equal to the porosity). The pre-storm was to begin two days before the PMP. The worst possible date of these three dates was chosen for use in calculating the PMF.

In order to simulate the PMF from the Danard (1995) data, the WATFLOOD/SPL model required continuous meteorological forcing data from April 1 (the date of the maximized snowpack) until after the PMP ended. Therefore, the data were embedded inside data for another year (available from the Danard, 1996b, model). The year 1972 was chosen because it was the wettest year since the Mica Dam was constructed, and sufficient data existed to be able to create 6-hour estimates of precipitation and temperature.

Figure 4-15 shows four inflow hydrographs for Mica Dam. The solid black line is the unaltered 1972 data from the HRBL model, while the dashed black hydrograph is the 1972 data with the PMS storm on June 1. The dashed grey line is the PMF calculation with the 100-year snowpack on April 1, the 100-year melting temperature sequence from May 15 to June 4, and the 1-day PMP on June 1 (the PMP was shortened to match with the length of the PMS). The solid grey hydrograph is the same snowpack and melting temperature sequence, but with the PMS storm on June 1 instead of the PMP. The effect of the maximized snowpack and melting temperature sequences can be observed as the difference between the two black lines (regular 1972) and the two grey lines (with snowpack and melting temperatures). A comparison of the grey lines shows that the volume of runoff from the

PMP (after subtracting base flow) was approximately eight times the volume of runoff from the PMS (after subtracting base flow), while the peak flow for the PMP was twice the peak flow of the PMS. These differences between the PMP and PMS are very large, which could have significant implications for many dams.

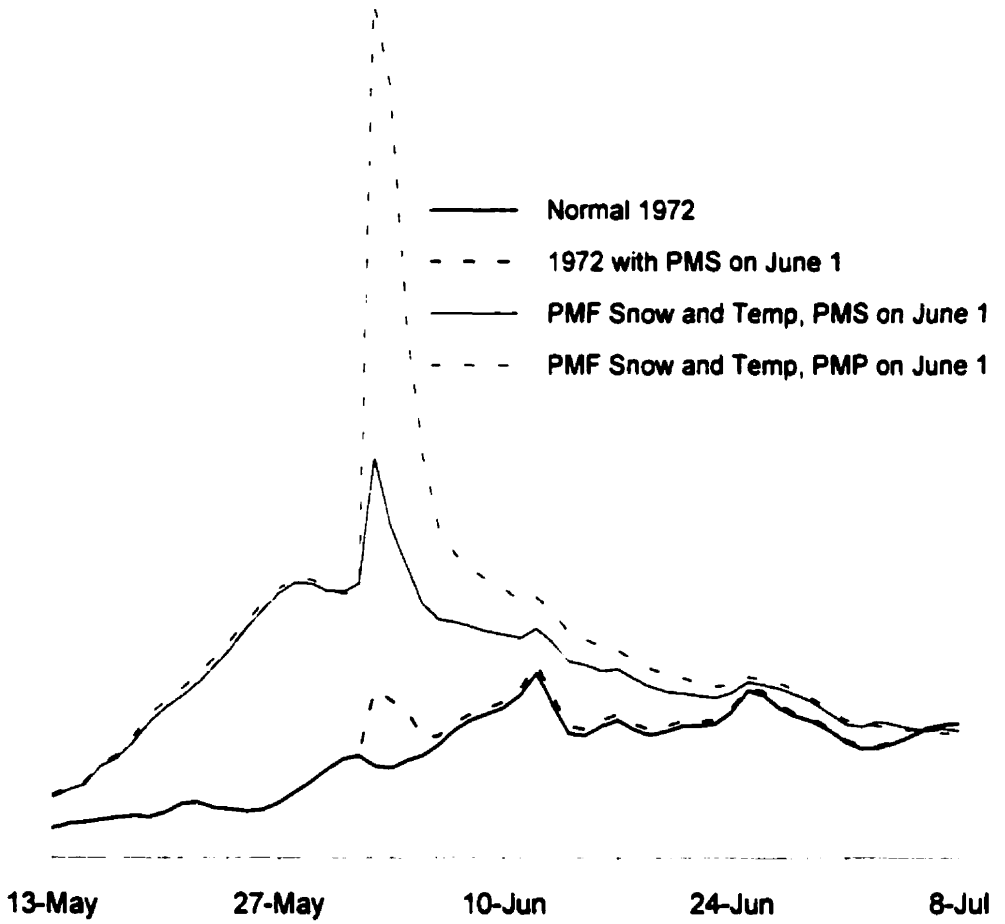


Figure 4-15 – Comparison of Mica Dam inflow with various PMF scenarios

Therefore, the flood caused by the PMS was larger than observed flood events, but smaller than the PMF. These results were expected, since the PMS was larger than the 1983 historical storm, and smaller than the PMP.

4.3 Summary of the MC2-PMS/WATFLOOD method to develop storms and floods

The combined use of the MC2-PMS and WATFLOOD/SPL models appears to be a viable method to calculate the PMF.

The flood that was calculated from the PMS was larger than observed flood events, but smaller than the PMF (calculated with the PMP). The WMO (1986) method does not ensure that the PMP is physically possible, and the use of an overestimated PMP would cause an overestimated PMF. In contrast, the PMS was developed in a physically-based manner, and the use of the WATFLOOD/SPL model ensured that the resulting flood was also physically-based. These results indicated that the PMF may have been overestimated.

However, before the MC2-PMS method can be used in practice, meteorologists must carry out an in-depth analysis of the physical constraints set on the internal workings of the numerical weather model MC2. For example, the following variables and processes will need to be assessed with regard to the impact on the PMS (Pellerin, 2000):

- Atmospheric energy
- Sea surface temperature
- Precipitation scheme
- Grid resolution

The effects of these and other variables and processes should be examined to determine the effect on the maximum precipitation. Nevertheless, the MC2-PMS approach promises to account directly for geographical features that can not be accounted for by the traditional approach of transposing storms.

5 Generating Improved Flood Frequency Curves

The flood frequency curve for the Mica Dam basin was estimated in order to allow the return intervals of the floods from the PMS and PMP to be compared. The regional frequency analysis using L-moments (Hosking and Wallis, 1997) method was used to calculate the frequency curves due to its robustness. This research used two sources of streamflow data: observed streamflow and calculated streamflow from a hydrological model. All of the analysis was performed with the annual maximum daily flows. The use of a long continuous time series of deterministically simulated streamflow for flood frequency curve estimation is a relatively new concept that is not yet fully established (e.g. Lamb, 1999, and Cameron, *et al.*, 1999). This application of the concept is unique in that the deterministically simulated streamflow time series was *longer* than the observed streamflow time series, and therefore it was less prone to sampling errors. The time series for simulated streamflow was 96 years long (1899-1994), whereas the time series for the observed streamflow was an average of 34 years long (with a range of 5 years to 91 years). Therefore, it was possible to compare the frequency curves generated by each source, and so validate the concept. After the concept was validated for this hydrological model and data set, the flood frequency curve and its confidence limits were derived from the simulated streamflow. This chapter will describe the validation of the use of simulated streamflow data, the derivation of the Mica Dam basin flood frequency curve and the derivation of the confidence limits for the flood frequency curve. Chapter 6 will compare the peak flows calculated in Chapter 4 to the flood frequency curve computed in this chapter.

5.1 Validation of the use of simulated streamflow data

The use of hydrological models to deterministically generate continuous streamflow for a long time series is relatively new. Relatively few authors have been able to generate a long time series of streamflow for flood frequency analysis, because of short observed meteorological records. Lamb (1999) was able to generate continuous streamflow for 10 years, while Cameron, *et al.* (1999) were able to generate continuous streamflow for 21 years.

In this research, a 96-year continuous time series of distributed meteorological forcing data were available. There were, however, several differences between the Lamb (1999) and Cameron, *et al.* (1999) studies and this study. The other studies compared the frequency curves derived from observed and simulated data and found that the calculated data produced reasonable flood frequency curves. However, both used multiple parameter sets, and parameter sets were accepted or rejected based on their ability to reproduce peak streamflow estimates for a particular basin (in addition to a suitable hydrograph). In contrast, the WATFLOOD/SPL model is a distributed physically-based hydrological model with a single optimal parameter set that applies to all sub-basins, and the calibration process focuses on the generation of correct hydrological processes (not solely on the hydrograph and/or peak flows). The High Resolution Boundary Layer (HRBL) model data (Danard, 1996b) was available for the time period 1899 to 1994 (96 years), which is longer than the average of 34 years for the streamflow observation stations on the Columbia River (range from 5 to 91 years). A longer time record is beneficial for flood frequency analysis purposes and this was one of the reasons for simulating the 96-year record. Although the modeled meteorological data were subject to modeling error, the model was based on observations, and there is

evidence that the synthesized streamflow was close to reality. However, the use of simulated streamflow to generate a frequency curve must be validated with this data set and with this model. In addition, although the WATFLOOD/SPL has been validated for many of its hydrological processes (e.g. Bingeman, *et al.*, 2001, Carlaw, 2000, Cranmer, *et al.*, 2001, Mousavi and Kouwen, 2000, Wong, 2000, and others), it has not been validated for the purpose of frequency curve estimation.

Hydrographs of the observed and simulated streamflow were presented in Figure 3-7 (Columbia River at Nicholson station) and Figure 3-8 (Mica Dam inflow). These showed that the meteorological data could be used to generate reasonable streamflow time series. However, the frequency curves were based on the annual maxima for each year. Therefore, to further test the simulated streamflow data, the residuals of the peak flows for the station with the longest time series (Columbia River at Nicholson, 90 years) were examined to determine their characteristics. Figure 5-1 shows the observed and simulated peak flows plotted against one another. There is a slight bias in this Figure; the lower peak flows tended to be overestimated, while the higher peak flows tended to be underestimated. However, the peak flows were generally close to the 45° line. Figure 5-2 shows the histogram of the differences between the simulated streamflow and observed streamflow peak flows. The mean of the residuals was slightly negative (-3 m^3/s), and they had a slight negative skew. However, the plot shows that the residuals were distributed around zero, indicating that there was very little systematic error. The residuals were not correlated in time (the correlation coefficient for a 1-year lag was 0.08), which also indicated that there was very little systematic error.

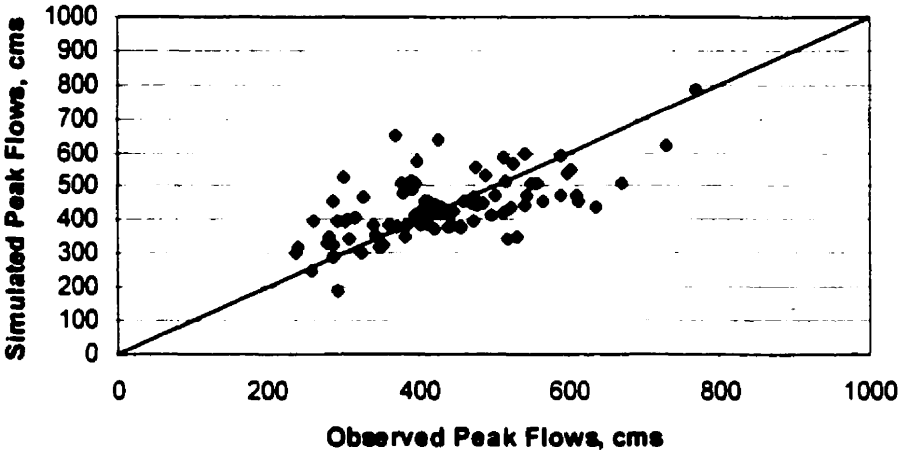


Figure 5-1 - Comparison plot of simulated and observed peak flows for Columbia River at Nicholson station

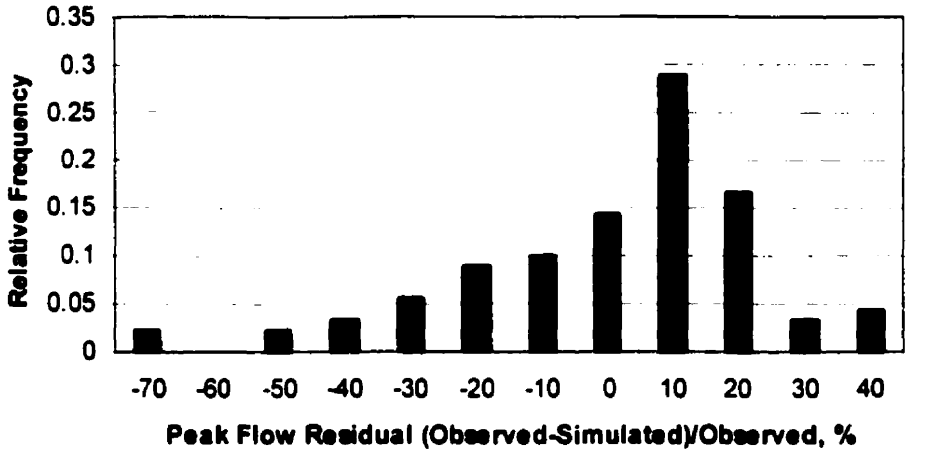


Figure 5-2 – Histogram of Residuals of Peak Flows for Columbia River at Nicholson (90 data points)

Due to model spin-up errors in WATFLOOD/SPL, it was not possible to use the entire 96-year simulated streamflow time series. The model spin-up errors are caused by starting the model with incorrect storage values (e.g. depression storage equals zero, soil moisture equals the antecedent precipitation index (Viessman, *et al.*, 1989), etc.). To determine the length of time that spin-up errors affect the streamflow, the model is started in

two different years, and the results compared to determine when the streamflow estimates match. For the Columbia River domain, the spin-up period is approximately two years. However, the following method may be used to decrease the model spin-up time. A typical year may be used as a lead-in year (the year 1971 was used). The lead-in year was used to begin the 96-year simulation (1899-1994), and the first year (1899) of output was discarded. In this way, instead of starting with zero storage on January 1, 1899, the model starts with a typical watershed condition. Only one year of data was discarded, so 95 years of simulation were available for analysis (1900-1994). A brief sensitivity study was performed to determine the effect of using a lead-in year on the streamflow. When the streamflow data calculated with a lead-in year were compared to streamflow data calculated without a lead-in year, the simulated annual peak streamflow was only slightly affected for the years 1900 and 1901 (1899 was discarded). After these years, the peak streamflow with and without the lead-in year were identical. Consequently, the frequency curves were not significantly affected by the use of the lead-in year; however, the extra year of data would tend to increase the accuracy of the curves.

In the HRBL model, there were non-random errors in the regressions between precipitation and horizontal convergence (see Section 3.3.3). There were very few precipitation stations located at high elevation, and therefore the regressions did not represent the precipitation at high elevation very well. Also, the use of horizontal convergence to predict precipitation may also cause non-random error in the precipitation estimate. Therefore, two simulations were carried out with the WATFLOOD/SPL model. The first pass used the unmodified HRBL model data, while the second pass used a precipitation adjustment factor (PAF) field based on the errors from the first simulation (Kouwen, *et al.*,

2000). The PAF field was generated from the errors in streamflow volume to correct precipitation errors in the HRBL model, and it had a greater effect on flow volumes than on flow peaks. This analysis compared the output from both runs to the observed frequency curves. In this analysis, the PAF field was generated with a single-pass correction of the streamflow volume (see Section 3.3.3). In this way, the bias was removed but the dispersion was left unchanged. It is likely that additional passes to correct the streamflow volume would improve the frequency curves.

In order to perform the validation of the model's ability to reproduce frequency curves, frequency curves were generated from the observed data and from the simulated data for each station. The observed streamflow database included 32 streamflow stations from Water Survey of Canada, and four B.C. Hydro dams, as described in Chapter 2. The 95-year simulated streamflow series were "shortened" to match with the observed streamflow database. The shortened data series were used to eliminate the possibility of introducing a bias due to the longer historical time series. The 1910's and 1920's were a fairly dry period, while the 1960's and 1970's were wetter. The comparison of a short time series (which may only include wetter (or drier) periods) to the full time series (which would include all periods) would be inappropriate. For instance, the Incomappleux River near Beaton had observed data in the years 1914-1915 and 1952-1995 (46 years) and so the data for these same years (1914-1915, 1952-1994) were extracted from the 95-year simulated streamflow data. The shortened data series are referred to as the "short series" in this research. Two short series were extracted: the first series was the simulation that used the unmodified HRBL model data, and the second series was the simulation that used a PAF field.

The L-moments (Hosking and Wallis, 1997) were used to generate the frequency curves for the observed and simulated streamflow series. Individual frequency curves were created for each station (regional analysis was not used at this point). The L-moments for each station were calculated for the observed data, the short series without PAF field, and the short series with PAF field.

Finally, a Wakeby distribution was fitted to each set of L-moments. The Wakeby distribution was chosen because it is a five-parameter distribution and is therefore very flexible and fits most data. Higher-order L-moments are much more robust than higher-order conventional moments (Hosking and Wallis, 1997), and therefore it is possible to use a greater number of the moments for distribution fitting than would typically be used in a conventional analysis with central moments. In addition, it is easier to compare frequency curves if they are generated with the same probability distribution, and a typical three-parameter distribution may not fit all three sets of L-moments.

There was considerable variation in the frequency curves. For brevity, only five of the larger sub-basins are discussed in detail. However, the frequency curves for all 36 stations are included in Appendix B.

The Columbia River near Fairmont Hot Springs station is an 891 km^2 basin in the northwest arm of the Columbia River. The curves are plotted in Figure 5-3. The WATFLOOD/SPL model consistently overestimated the streamflow at this station; the suspicion of bias in the HRBL model data was partly based on the results of this station. The frequency curves supported these results. The simulated curve overestimated the observed frequency curve by a factor of about two at all probabilities. The simulated-with-PAF field curve showed that the PAF field has over-corrected for this error somewhat. The PAF field

improved the results; the simulated-with-PAF field frequency curve corresponded to the observed frequency curve for high probabilities of exceedance. The average peak flows were modeled correctly, but the shape of the peak flows was incorrect, and therefore the use of simulated data to create a frequency curve was partially valid at this station.

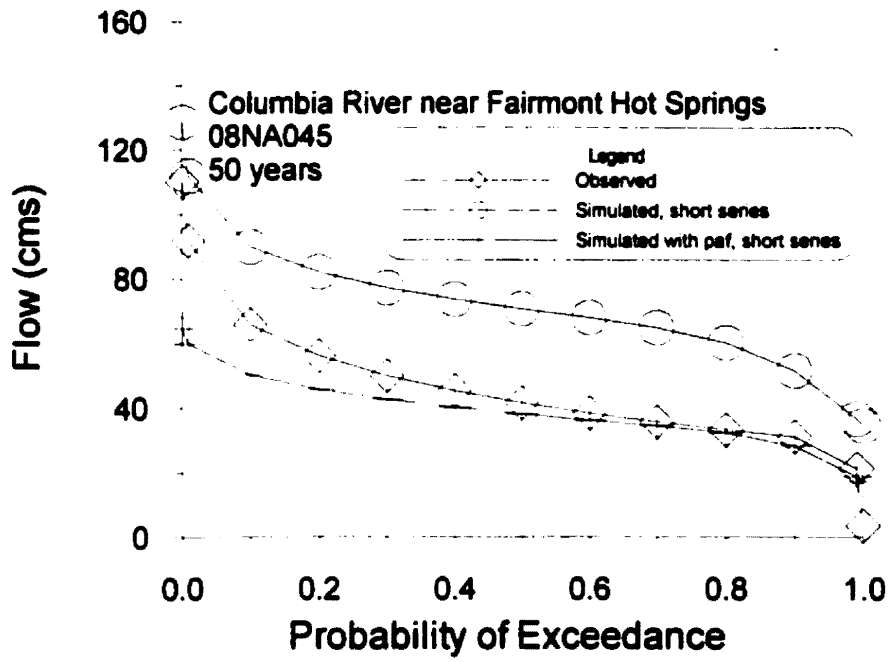


Figure 5-3 – Frequency Curves for Columbia River near Fairmont Hot Springs

The Illecillewaet River at Greeley station is an 1170 km^2 basin in the western part of the basin, below Mica. Its headwaters are located near Rogers Pass. The curves are plotted in Figure 5-4. The simulated-with-PAF field curve followed the observed curve closely; they were coincident at most probabilities. This station experienced its 500-year flow from the 1983 storm. The simulated streamflow data were able to match this value. The PAF field improved the estimate of the frequency curve, and the use of simulated data to create the frequency curve was reasonable.

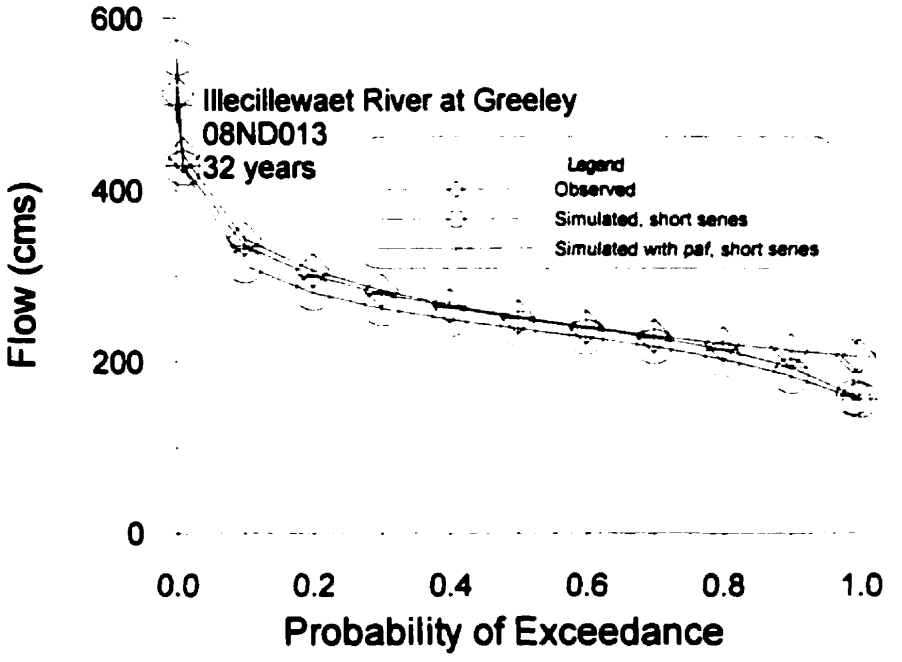


Figure 5-4 – Frequency Curves for Illecillewaet River at Greeley

The St. Mary River near Marysville station is a 1480 km^2 basin that is not in the Columbia River Basin, but is modeled due to its proximity to the Columbia River. The basin is south of the headwaters of the Columbia River. The curves are plotted in Figure 5-5. The curves showed that the observed and simulated frequency curves were similar in shape, however the simulated curves underestimated the observed flood frequency curve. At this station, the PAF field lowered the streamflow values, which increased the difference between the observed and simulated flood frequency curves. There was a 20-25% error for the simulated-with-PAF field frequency curve, and a 10-15% error for the simulated-without-PAF field frequency curve. At this station, the PAF field did not appear reasonable, and since the peaks were underestimated, the simulated streamflow data could not be used to estimate the frequency curve. The reason for this discrepancy is the nearness of the St. Mary River to watersheds above the Fairmont Hot Springs gauge, where the flows were greatly overestimated (Figure 5-4). The overestimation at the nearby gauges caused an over-

correction at the St. Mary gauge. This problem can be traced back to the single pass HRBL model, where possibly too much emphasis was placed on "rubbersheeting" to all the rainfall observations. This problem may be improved with the use of a two- or three-pass PAF field.

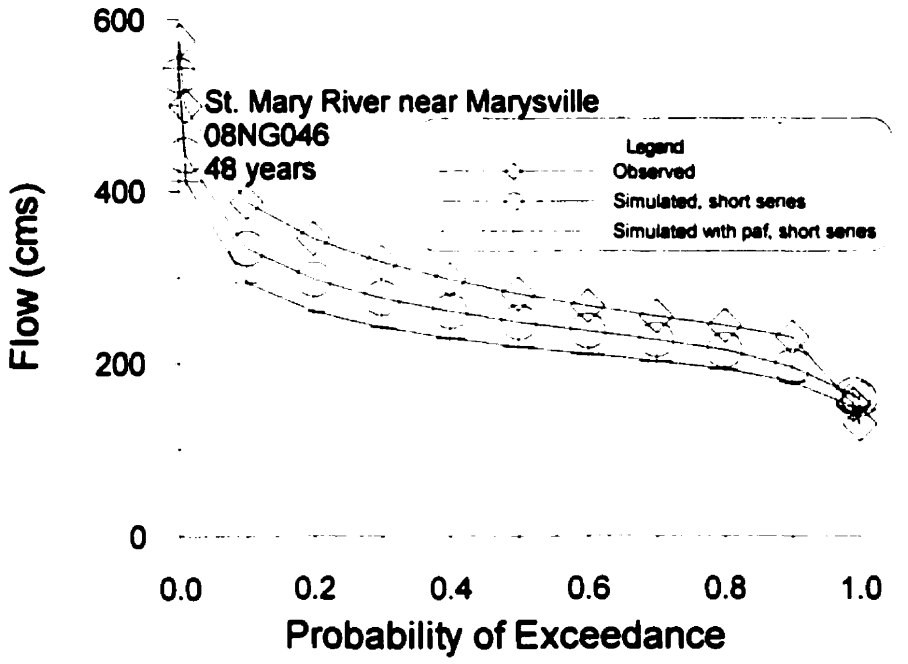


Figure 5-5 – Frequency Curve for St. Mary's near Marysville

The Columbia River at Nicholson station is a 6660 km^2 basin in the northwest arm of the Columbia River, below Fairmont Hot Springs. Since this station is downstream of the Fairmont Hot Springs station, it was also consistently overestimated, in part due to the overestimation at Fairmont Hot Springs. This station is the oldest Columbia River Basin station (91 years), and these data have been relied upon to verify the earlier years of the HRBL model data. The curves are plotted in Figure 5-6. As expected, the simulated-without-PAF field curve showed that the streamflow was overestimated. However, the PAF field corrected the error, and the simulated-with-PAF field curve was almost coincident with the observed frequency curve. This result was very important, as it showed that the WATFLOOD/SPL model was well calibrated, and calculated the streamflow accurately for a

period of 90 years. The simulated streamflow data could be used to develop the frequency curve at this station.

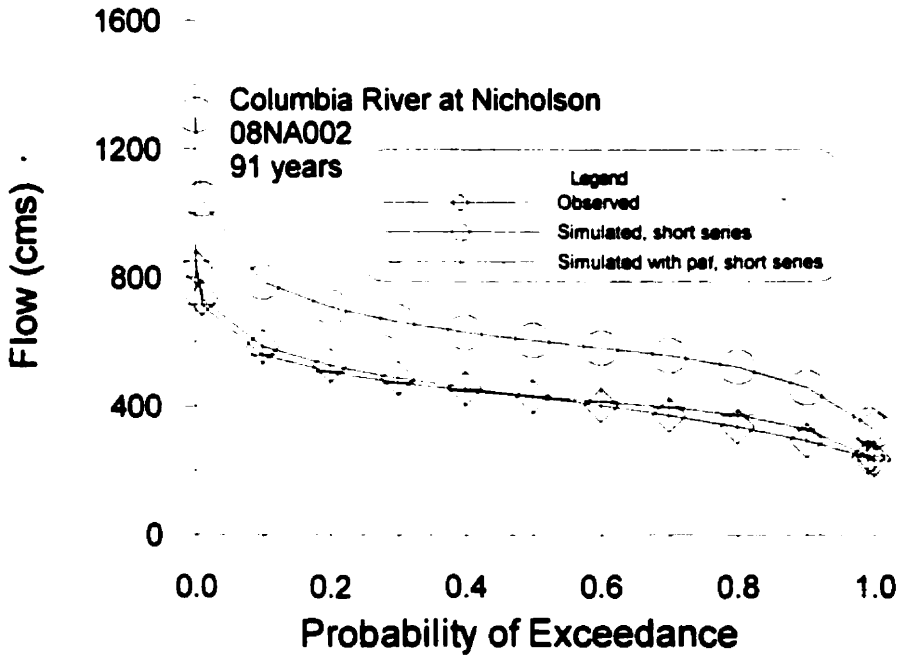


Figure 5-6 – Frequency Curves for Columbia River at Nicholson

Finally, the Mica Dam has a drainage area of approximately 20,000 km^2 and is located at the northern end of the north-south arm of the Columbia River. The reservoir inflow data for Mica Dam were calculated by B.C. Hydro. The curves are plotted in Figure 5-7. The three curves were coincident for most probabilities, and at high probabilities, the simulated curves overestimated the observed curve. There was very little difference between the simulated with and without PAF field curves, mainly because the Mica Dam drainage area contained areas where precipitation was overestimated and other areas where it was underestimated. The use of simulated streamflow to calculate the frequency curve was valid at this station.

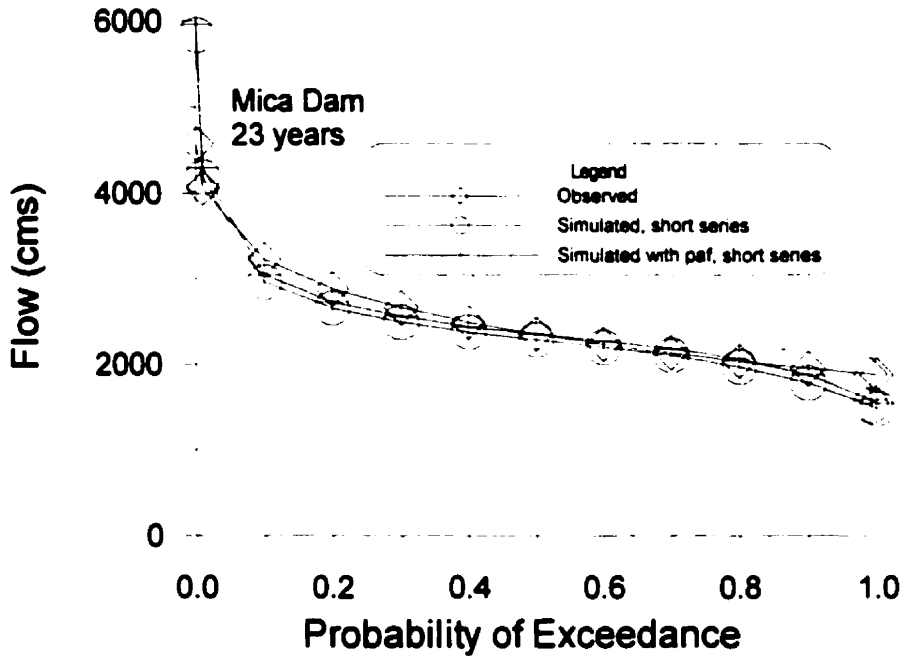


Figure 5-7 – Frequency Curves for Mica Dam

In general, approximately two-thirds of the stations showed good agreement between the observed curve and at least one of the simulated curves, where "good agreement" referred to coincidence for some or all of the main part of the frequency curve. The errors ranged from -50% to +100%, with the larger errors occurring at the stations with low streamflow. This suggested that the frequency curves derived from simulated streamflow data were able to model the observed frequency curves.

The simulated-without-PAF field and simulated-with-PAF field frequency curves were compared to determine which curve was more accurate. The results are listed in Table 5-1. It was found that the simulated-with-PAF field curve was closer to the observed frequency curve for 16 stations (six stations showed only marginal improvement). The simulated-without-PAF field was closer for eight stations (two stations showed only marginal differences). Twelve stations showed no improvement when the PAF field was used (it is likely that these drainage basins contained areas that were overestimated and areas that were

underestimated, and these cancelled). These data suggested that the PAF field was an important adjustment, but further refinements could be made to improve the eight stations where the PAF field worsened the results. A second or third pass to calculate the PAF field may improve the results, or a topographical correction could be applied. In general, the PAF field improved the estimation of peak flows, and consequently the PAF field adjustment was adopted.

Table 5-1 – List of Which Simulated Curve Best Approximated the Observed Curve

Simulated with PAF field was better	Simulated without PAF field was better	Both were about equal in accuracy
Columbia River near Fairmont Hot Springs	Kicking Horse River at Golden	Blaeberry River above Willowbank Creek
Split Creek at the Mouth	Stitt Creek at the Mouth	Gold River above Palmer Creek
Mather Creek below Houle Creek	Beaton Creek near Beaton	Canoe River below Kimmel Creek
Kuskanax Creek at 1040 M Contour	St. Mary River near Marysville	Goldstream River below Old Camp Creek
Lardeau River at Marblehead	Carney Creek below Pambrun Creek	Barnes Creek near Needles
Jordan River above Kirkup Creek	Fry Creek below Carney Creek	Kaslo River below Kemp Creek
Columbia River at Nicholson	Count: 6	Keen Creek below Kyawats Creek
Columbia River at Donald		Gold River above Bachelor Creek
Arrow Dam		Lemon Creek above South Lemon Creek
Revelstoke Dam		Cranberry Creek above BCHydro Intake
Count: 10		Duncan Dam
		Mica Dam
Marginally better:	Marginally better:	Count: 12
Blaeberry River below Ensign Creek	Kirbyville Creek near the Mouth	
Illecillewaet River at Greeley	St. Mary River below Morris Creek	
Incomappleux River near Beaton	Count: 2	
Kuskanax Creek near Nakusp		
Duncan River below BB Creek		
Spillimacheen River near Spillimacheen		
Count: 6		

It was also necessary to evaluate the ability of the hydrological model to reproduce the high-flow end of the frequency curves. Frequency curves are often used to predict the magnitude of large flood events, and the large flows were of interest to this research. In general, the high-flow tails of the observed frequency curves were poorly estimated by the simulated curves, but this was related to the length of observed record. Figure 5-8 shows the simulated 100-year return interval flow divided by the observed 100-year return interval flow, versus the number of years of record for each station. The ratios of the 100-year return interval flow for the simulated without PAF data were plotted with black diamonds, while the ratios of the simulated with PAF data were plotted with grey squares. In general, the simulated with PAF data were closer to the ideal ratio (1.0) than the simulated without PAF. The average error for the simulated without PAF data was 0.29, while the average error for the simulated with PAF was 0.23. This indicated that the PAF field improved the estimation of extreme peak flows. For the simulated with PAF data, stations with short records (less than 30 years) had ratios that varied from 0.4 to 1.4. Stations with 40 or more years of data showed improvement, with most stations between 0.6 and 1.2. Frequency curves derived from short data records may be inaccurate, but the accuracy improves as the record length increases. In this case, as the data record became longer, the simulated with PAF frequency curve estimated the observed curve more accurately. When stations with short data records were not included in the analysis, the data showed that the simulated with PAF data from WATFLOOD/SPL were accurate for calculating frequency curves, even at the high-flow tail.

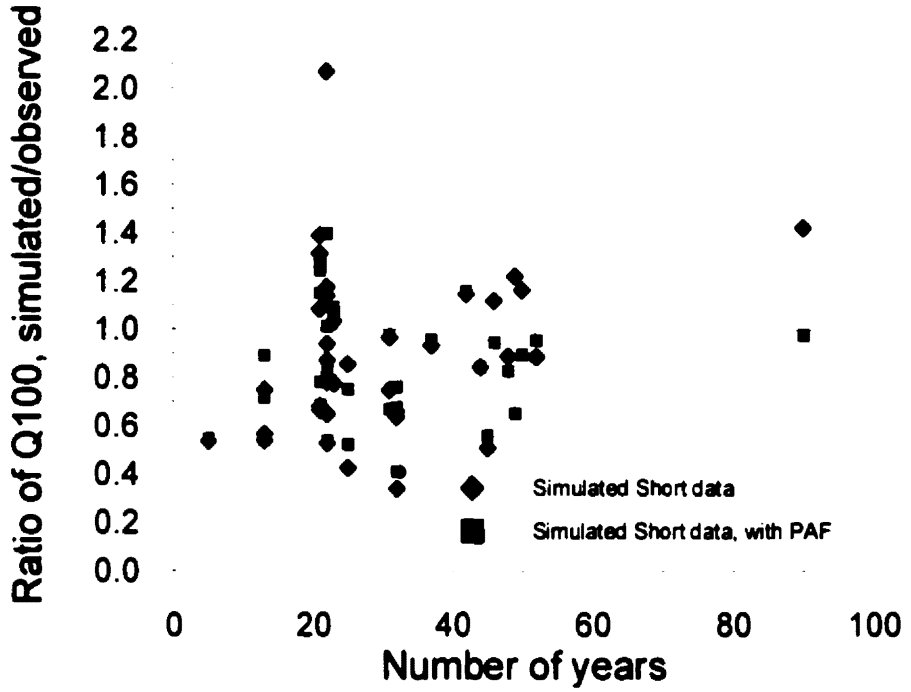


Figure 5-8 – Comparison of 100-year return interval flows versus the number of years of record

The natural variation within the observed data was used to determine whether the errors in the Q100 ratios were reasonable. The comparison was only performed for the simulated with PAF data (the simulated without PAF data were not compared because the previous analysis indicated that the simulated with PAF data were superior for estimating extreme peak flows). Figure 5-9 shows the results. The open grey boxes depict the natural variation within the observed data. The data points were calculated by performing the frequency analysis on continuous subsets of the data for each station. First, 10-year sequences were removed from the observed time series, starting with the first data point, then the second data point, and so on. Then, 11-year sequences were removed from the observed time series, and so on until all data points for each observed time series were used. After each frequency curve was calculated, the 100-year return interval flow was compared to the 100-year return interval flow using all of the observed data. The ratios for the simulated

short data with PAF were superimposed upon the natural variation. It was found that most of the stations could be classified as within the natural variation. Three stations were significantly below the area of natural variation in Q100 estimation: Kuskanax Creek near Nakusp (08NE006), Incomappleux River near Beaton (08NE001), and Columbia River near Fairmont Hot Springs (08NA045). The first two of these stations are located in an area where the HRBL model underestimated precipitation, and the PAF was used to increase precipitation. However, the other stations nearby prevented the PAF from fully correcting the flows at these two stations. The shape of the simulated frequency curve for the Columbia River at Fairmont Hot Springs station did not match the observed frequency curve. This station was over-corrected by the PAF, which lowered the streamflows too much. The PAF field could be improved to decrease the underestimation at these locations. These results have demonstrated that the error in estimation of the high-flow end of the frequency curve is within the natural variation of the observed data. Therefore, the model was able to estimate the extreme flows on the frequency curves.

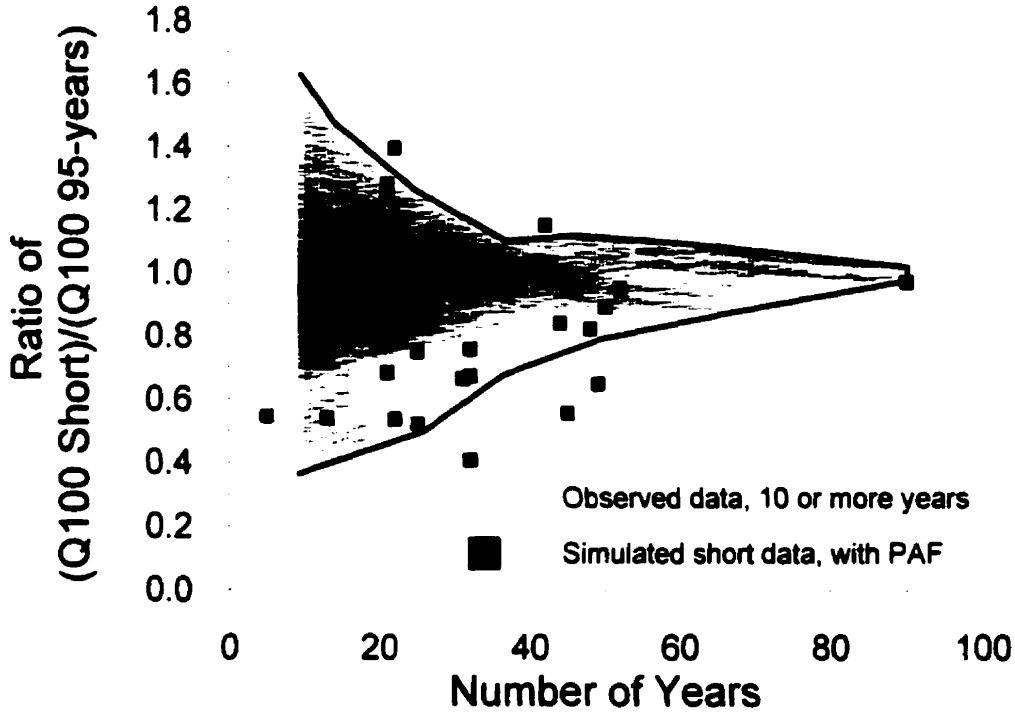


Figure 5-9 – Ratio of Q100 for simulated series compared to variability in observed Q100 ratio: the grey boxes are the variability in observed data (Q100 for a short data series/Q100 for all years), the black boxes are the Q100 for the simulated short data compared to the Q100 observed estimate.

Finally, the data were examined to determine if they were stationary in time. The Columbia River at Nicholson station had 91 years of observed data available for analysis. Ten-year consecutive sequences of annual maxima of both the observed data and the simulated with PAF data were used to create frequency curves. The 100-year return interval flow of each frequency curve was compared to the 100-year return interval flow of the frequency curve calculated with the full time series of observed data. There is a slight downward trend in 100-year return interval flow estimates (a linear trendline goes down 0.3 over 90 years). However, the slight downward trend is mainly due to the large overestimation in the 1920's and 1930's. A 10-year moving average of the observed data reveals that the remainder of the variation is due to local high and low variations that

correspond to the wetness or dryness of each ten-year sequence, with no downward trend. The local high and low variations in the two data series corresponded very well. They compare particularly well for the period when the LFM data were available for calculating the meteorological data (1971-1994). Therefore, these data showed that the frequency curves calculated with simulated data compared very well with the observed frequency curves, and also that the data were mainly stationary over time.

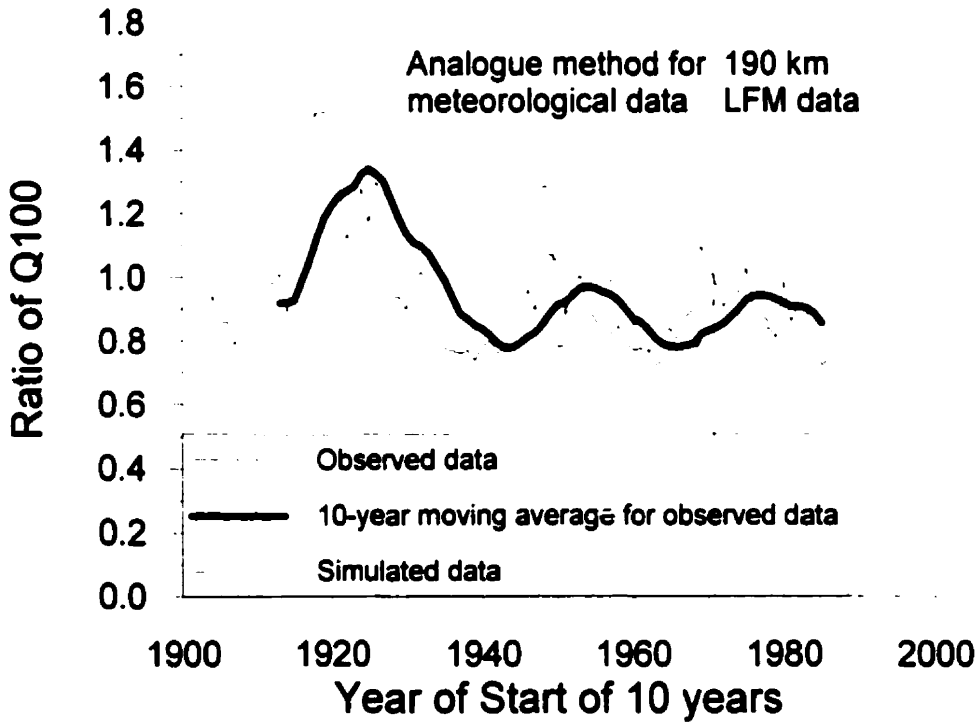


Figure 5-10 – Comparison of Stationarity: Q100 estimates from 10-year sequences of observed and simulated data, divided by Q100 observed

In general, the use of simulated data to calculate frequency curves resulted in good agreement between the simulated and observed curves. However, there was some disagreement between observed and simulated frequency curves for high flow – low probability events. The estimation of the probability of the PMF involves the high-flow – low probability region of the frequency curve, and so the disagreement in this region was

examined. Frequency curves made from short time series may be inaccurate, and so the stations with longer records were examined. There was better agreement for stations with longer time series; in particular, there was excellent agreement for the station with 91 years of data. It appears that frequency curves derived from the long simulated time series (95 years) would improve the frequency curves derived from the observed data series for all stations.

This comparison showed that the use of simulated data to calculate flood frequency curves is valid. The emphasis during calibration of WATFLOOD/SPL was on ensuring that the hydrological processes were reasonable and on obtaining the correct volume of runoff. However, the model was also accurate for the peak flows, as evidenced by the agreement of the frequency curves, although little effort was made to fit the peak flows.

The remainder of this research to derive improved flood frequency curves uses the 95-year simulated data, generated with the PAF field, except where noted. The curves are compared to the observed frequency curves.

5.2 Derivation of the flood frequency curve

This section presents the derivation of the estimate of the frequency curve for Mica Dam. The regional frequency analysis method was described in some detail in the literature review in Section 3.4. The method contains four main steps: data screening, identification of regions, choice of a frequency distribution, and estimation of the at-site frequency distribution parameters. The first of these steps was performed during model calibration and validation.

The frequency curves were derived for two data sets: the observed streamflow data, and the simulated streamflow data. The simulated streamflow was the 95-year time series of

streamflow calculated by using the HRBL meteorological data (adjusted with the PAF field) as forcing data for WATFLOOD/SPL. For the 32 streamflow stations and 4 B.C. Hydro dams, there were a total of 1073 years of observed streamflow, and 3420 years of simulated streamflow.

5.2.1 Identification of Regions and Choosing Appropriate Distributions

The identification of homogeneous regions is an important step in regional analysis. The ability to fit a frequency distribution depends on having homogeneous regions. There were two sets of data (observed streamflow and simulated-with-PAF streamflow) used in this analysis, and it was decided to use the same regions for both sets of streamflow, so that comparisons between the frequency curves could be made. This placed a greater restriction on the choice of regions: they must be homogeneous (or possibly homogeneous) for two different data sets. The 36 streamflow stations and dams were grouped with the Burn, *et al.* (1997) clustering algorithm.

This algorithm (described in Section 3.4.3.2) uses the Canberra dissimilarity metric to calculate the distance between any two stations and form clusters. The Canberra dissimilarity metric was calculated with all fifteen physiographic and climatic variables (described in Section 3.4.4). Several trials with various values of the weighting coefficient between 0.1 and 0.9 were performed. Visual inspection of the clusters was used to select a satisfactory distance weighting coefficient.

The algorithm generated four clusters with a weighting coefficient of 0.3 (observed data) or 0.4 (simulated data). The variation occurred because of the use of different data sets. These results compared well with the weighting coefficient of 0.3 used by Burn, *et al.* (1997). The boundaries of the clusters remained relatively constant for the two data sets.

Some stations near the boundaries shifted between regions for different data sets. The nearly constant division between the regions showed that the data sets were similar, further validating the use of simulated data to derive frequency curves. Where differences between the data sets existed, the two sets of clusters were modified until both data sets had identical homogeneous (or possibly homogeneous) clusters.

The four clusters obtained from the algorithm are shown in Figure 5-11. Two stations were not placed into a cluster (they are marked with a star). Ellipses were placed around the clusters for illustrative purposes (the program does not calculate ellipsoidal regions). The two stations that were not placed in a region were originally placed into regions by the program; however, to allow the clusters to be identical for both data sets, the stations were removed from their regions. One of these stations was outside the Columbia River basin (but is modeled due to its proximity), and the second station was Revelstoke Dam. The clusters showed a high degree of geographic continuity, and they generally corresponded to west and east basins, for high and low elevations. These four regions were significantly different meteorologically (due to the locations of the mountains). Since streamflow is affected by meteorology, it follows that streamflow clusters also show these meteorological regions. Streamflow stations within each region were not independent, since a given storm could occur over several watersheds. A lack of independence does not affect the clustering process, but it lowers the effectiveness of the regional frequency analysis method by reducing the effective amount of data. This is a common problem in regional frequency analysis; however, Hosking and Wallis (1988) found that intersite dependence had little effect on the estimated quantiles of the frequency curves, but the variance of the estimates increased.

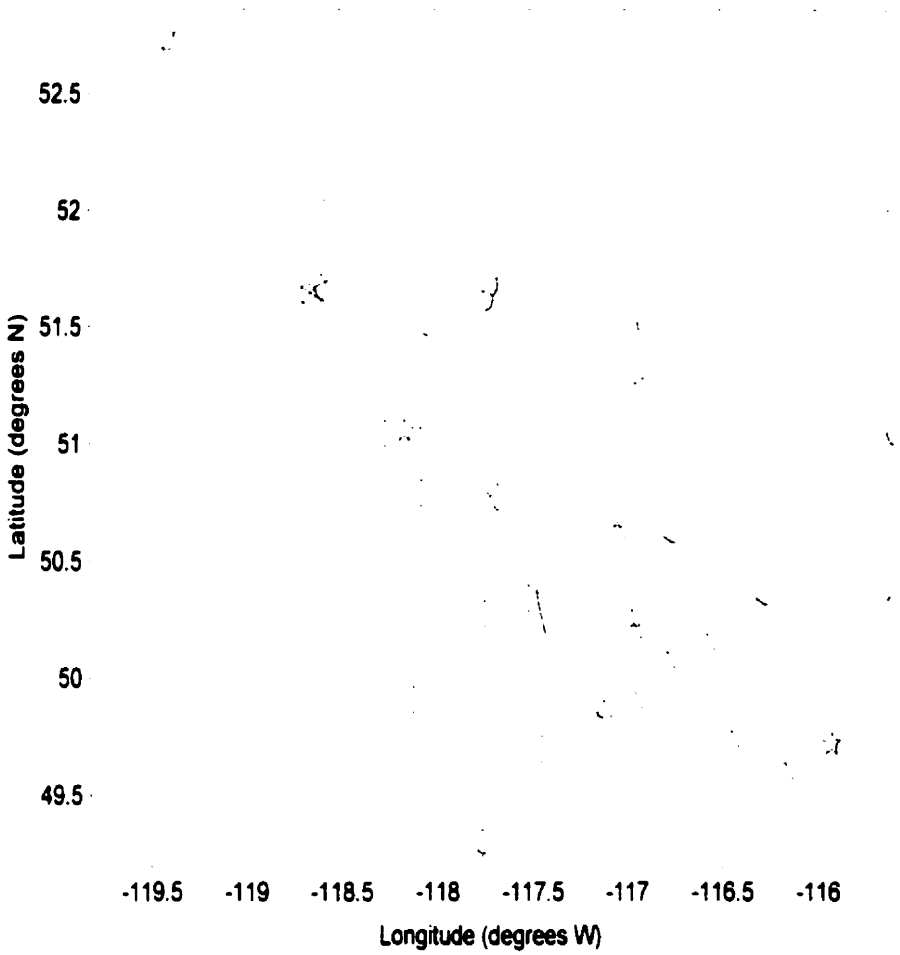


Figure 5-11 – Final Clustering of Streamflow Stations

The four clusters were tested for homogeneity. Hosking and Wallis (1997) provided three tests of homogeneity (Equations 3-22, 3-23, and 3-24). They recommended that all three homogeneity criteria should be less than one, in order for the cluster to qualify as homogeneous. It was found that if the limit of the homogeneity criteria was set to one, then the clustering algorithm was unable to find suitable clusters. However, if values up to two were allowed, then clusters were formed. Schaefer (1997) offers a possible solution to this problem: since variability exists in meteorological data due to local site changes through time (e.g. land use changes), the homogeneity limit should be increased to two. Therefore, the

limit for the homogeneity criteria was increased to two, and the clusters were assumed homogeneous for distribution fitting. Table 5-2 shows the results of the homogeneity test. Mica Dam is located in the first cluster. The clusters for both data sets were homogeneous according to the Schaefer (1997) definition, and possibly heterogeneous according to the Hosking and Wallis (1997) definition. These clusters were accepted for the regional frequency analysis.

Table 5-2 – Results of Homogeneity Test

Cluster	Number of Stations	Location	Total Years of Record	Observed Streamflow Homogeneity Criterio	Simulated (with PAF field) Streamflow: Homogeneity Criterion
1*	6	Northwest	136	0.81 / -1.35 / -0.8	0.54 / -0.09 / -0.74
2	8	Northeast	341	0.73 / 0.68 / 0.3	0.59 / -1.62 / -1.51
3	9	Southeast	276	0.46 / -1.06 / -1.7	-1.44 / 0.13 / -0.24
4	11	Southwest	320	1.68 / -0.52 / -0.5	1.83 / -0.41 / -0.89
* This region contains Mica					

Hosking and Wallis (1997) also provided a test to determine which frequency distributions fit the data (Equation 3-28). This test assumes that the clusters are homogeneous. The results of the test for acceptable distributions for each region and data set are listed in Table 5-3. The distributions are listed in the order of most acceptable to least acceptable. This table shows some differences between the two data sets. To allow comparisons between the frequency curves, the same frequency distribution was used for both data sets. For Region 2, it was not possible to use a simple three parameter distribution for the data (no distributions fit both data sets). For each of the other regions, one or more distributions were acceptable for both data sets. The choice of distribution therefore depended upon the region.

Table 5-3 – Acceptable Distributions for Each Region

Cluster	Location	Observed Streamflow: Valid Distributions	Simulated (with PAF field) Streamflow: Valid Distributions
1	Northwest	GLO, GEV, GNO	P3, GEV, GNO
2	Northeast	P3, GNO, GEV	GLO
3	Southeast	GEV, GNO, P3	GNO
4	Southwest	GEV, GNO, GLO	GNO, GEV, P3

- Cluster 1 contains the inflow to Mica Dam
- GLO = Generalized Logistic, GEV = Generalized Extreme Value, GNO = Generalized Normal, P3 = Pearson Type 3

5.2.2 Calculating the Frequency Distributions for Mica Dam

This research used two methods to develop the frequency curves. The first method (Hosking and Wallis, 1997) calculates the weighted average of the L-moments for a region and uses it to fit the regional frequency curve. This method results in discontinuities in the frequency domain between regions. Runoff varies smoothly between regions in the same way that precipitation varies smoothly. The second method (Schaefer, 1990) derives regression relationships between the L-moments and other non-statistical variables. Schaefer (1990) derived regression relationships for homogeneous sub-regions. In this research, the regression relationships were performed for the network of 36 stations and dams. The network could be used because the entire network (with L-moments calculated from the simulated streamflow data) was only possibly heterogeneous according to the Schaefer (1997) definition (Table 5-4), and therefore, the use of a single region was possible. The regressions were used to calculate regional L-moments that were then used to fit the frequency distribution. The regressions varied smoothly across all regions, avoiding the discontinuity problem, however a single frequency distribution that can fit all of the regions is required. This research used the Wakeby distribution, since there were no three-parameter

distributions that fit all four regions for both data sets. The estimated L-moments from the regressions were used to fit the frequency distributions.

Table 5-4 - Heterogeneity statistics for the entire network of 36 stations and dams for the simulated streamflow (95 years) data.

Heterogeneity measure	Statistic	Hosking and Wallis (1997) Conclusion	Schaefer (1997) Conclusion
H1 (Eq. 3-22)	2.78	Heterogeneous	Possibly Heterogeneous
H2 (Eq. 3-23)	-0.30	Homogeneous	Homogeneous
H3 (Eq. 3-24)	-0.95	Homogeneous	Homogeneous

The first method, Hosking and Wallis (1997), used the data from the six stations in Region 1 to create the frequency curves, a total of 136 years for the observed data and 570 years for the simulated data. Table 5-3 shows that both the Generalized Extreme Value and the Generalized Normal distributions were acceptable for both data sets for this region. However, the Wakeby distribution was used to calculate the frequency curves, in order that the frequency curves could be compared with the curves calculated by the second method.

The observed and simulated regional frequency curves are shown in Figure 5-12. Generally, the simulated frequency curves underestimated the observed frequency curves, but there was reasonable agreement for the probabilities of exceedance of 0.2 to 0.6. However, there were significant differences outside of this region. This research was mainly interested in high-flow, low-probability events (such as the PMF). At a probability of exceedance of 0.001, the simulated curve underestimated the observed curve by more than 2000 m^3/s, or approximately 30% (see Figure 5-12). One possible reason for this result was that the six stations that were included with Mica Dam in region 1 all had short time series (21-23 years), and therefore the accuracy of the observed frequency curve was questionable. The single station analysis indicated that the simulated streamflow data estimated the high-flow, low-

probability streamflow poorly when the time series was short (see Section 5.1). The large discrepancy between the observed and simulated frequency curves was likely due to a combination of errors in the observed curve and errors in the simulated curve. The amount of error in the simulated frequency curve was therefore unknown.

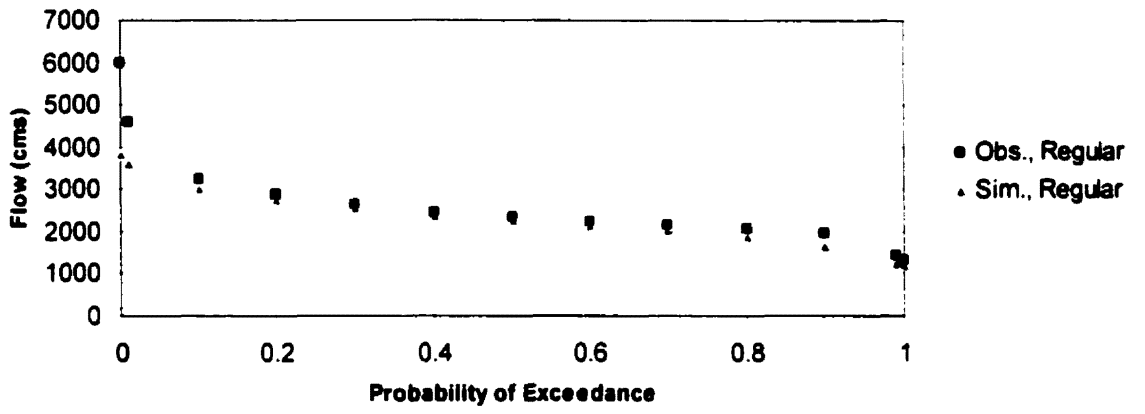


Figure 5-12 – Comparison of Frequency Curves for Mica Dam calculated using Regional Frequency Analysis (regular Hosking and Wallis 1997 method): the observed curve used the observed data records, the simulated curve used the 95-year simulated streamflow data.

The second method (Schaefer, 1990) was also used to estimate the regional L-moments. This method used the data from all 36 stations by calculating the relationships between the L-moments and physiographic parameters. The regressions were created from a total of 1073 years of observed data, or 3420 years of simulated data. Therefore, since this method used a larger data set, it was not as susceptible as the regular Hosking and Wallis (1997) method to errors in the individual station L-moments.

The relationships between the L-moments and the physiographic and climatic variables were examined with linear regressions. It was found that a large number of parameters were required to obtain a suitable fit. However, there were only 36 stations available for the regression, and 15 variables. A subset of these variables, 11 variables, was

required to obtain regressions with an R^2 greater than 50%. An R^2 coefficient of 50% implied that the regression explained 50% of the variation in the dependent variable (the L-moment). The regressions for the observed L-moments, L-CV (t), L-skew (t_3), L-kurtosis (t_4), and the fifth L-moment (t_5), are presented in Table 5-5. Similarly, the regressions for the simulated L-moments are presented in Table 5-6. These two tables list the best regressions with a set of 11 variables. A comparison of these tables shows that the parameter sets that gave the best regressions for the four L-moments and the two data sets were slightly different.

Table 5-5 – Regression Parameters for Observed L-moments

L-moment	Parameters	R^2
L-CV (t)	Constant, area, Julian date of peak, azimuth, slope, DTO-SE, DTO-SW, SHE-NE, SHE-SW, BH-NE, BH-NW, BH-SW	0.584
L-skew (t_3)	Constant, azimuth, slope, DTO-NE, DTO-SE, DTO-SW, SHE-NE, SHE-SW, BH-NE, BH-SE, BH-SW	0.615
L-kurtosis (t_4)	Constant, Julian date of peak, slope, DTO-NE, DTO-SE, DTO-SW, SHE-NE, SHE-SW, BH-NE, BH-NW, BH-SE, BH-SW	0.571
t5	Constant, ratio of peak flow to mean flow, azimuth, slope, DTO-NE, DTO-SE, DTO-SW, SHE-NE, SHE-SW, BH-NE, BH-SE, BH-SW	0.643

Table 5-6 – Regression Parameters for Simulated L-moments

L-moment	Parameters	R^2
L-CV (t)	Constant, Julian date of peak, ratio of peak to mean flow, azimuth, slope, DTO-NE, DTO-NW, DTO-SE, SHE-NE, SHE-SW, BH-SE, BH-SW	0.582
L-skew (t_3)	Constant, area, Julian date of peak, ratio of peak to mean flow, DTO-NE, DTO-NW, DTO-SW, SHE-NE, BH-NE, BH-NW, BH-SE, BH-SW	0.604
L-kurtosis (t_4)	Constant, area, Julian date of peak, ratio of peak to mean flow, slope, DTO-NE, DTO-NW, DTO-SE, DTO-SW, SHE-SW, BH-NW, BH-SE	0.642
t5	Constant, area, ratio of peak to mean flow, azimuth, slope, DTO-NE, DTO-NW, DTO-SE, SHE-SW, BH-NE, BH-NW, BH-SW	0.608

The results of the regressions for the observed and simulated L-CV regressions are shown in Figure 5-13 and Figure 5-14. These Figures show that the regressions captured the variation in observed and simulated L-CV very well.

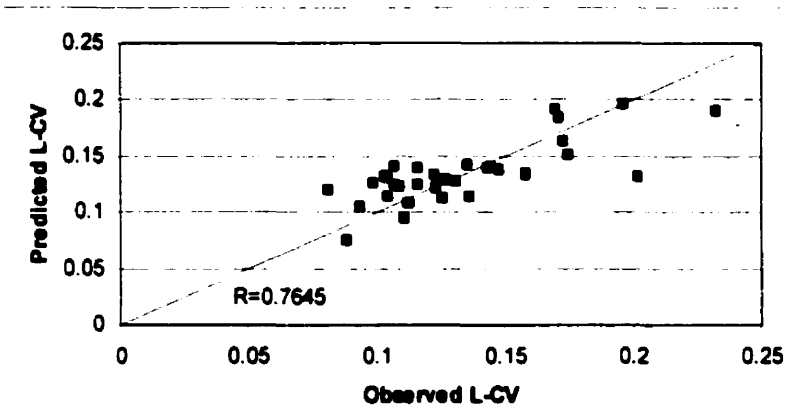


Figure 5-13 – Predicted L-CV versus Observed L-CV, showing results of regression

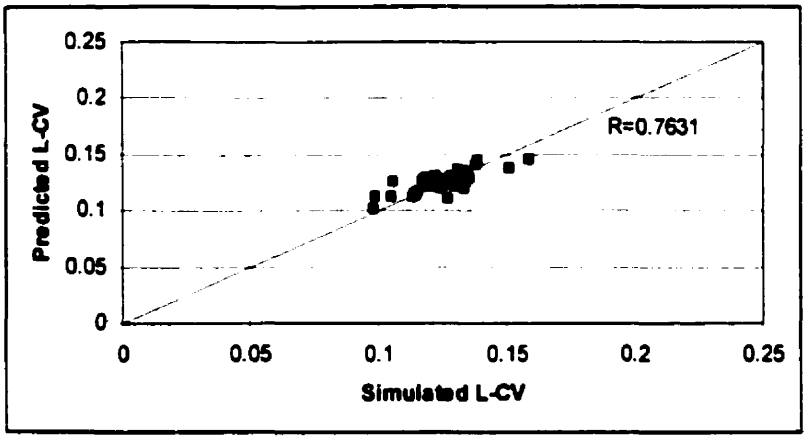


Figure 5-14 – Predicted L-CV versus Simulated L-CV, showing results of regression

In general, the regressions for the simulated L-moments had similar regression coefficients to the regressions for the observed L-moments. The Schaefer (1990) algorithm depends on suitable relationships between the L-moments and the other non-statistical variables. The algorithm allows for the at-site L-moments to contain error, and the error may be different for each data set. Therefore, the best regressions for each parameter were used,

even though they were different for both sets of data. In this way, the regressions that were most indicative of the regional pattern of L-moments in the data were found. The use of regression allows the regional estimate of the L-moment to be set according to regional variations in the at-site L-moments. The regional estimates will be "close" to the at-site L-moments, but "corrected" with the regional variation.

The observed and simulated flood frequency curves were calculated and compared in Figure 5-15. The difference between the frequency curves was significantly smaller with the Schaefer (1990) method. As before, the simulated frequency curve underestimated the observed frequency curve slightly. In the normal range of probabilities (approximately 0.4 to 0.8), the two frequency curves agreed well with each other and with the Hosking and Wallis (1997) curves (compare Figure 5-12 and Figure 5-15). The main difference between the Figures occurred for high-flow, low probability events. For a probability of exceedance of 0.001, the simulated frequency curve underestimated the observed frequency curve by approximately 100 m^3/s (compared to a 2000 m^3/s difference for the Hosking and Wallis, 1997, method). The use of the regressions allowed the estimation of the regional L-moments to take advantage of longer time series stations. Errors in the observed L-moments for the six stations in Region 1 were minimized when a regression was performed with other stations that had longer time series, and there was less overestimation. Similarly, the underestimation in the simulated data was minimized when regressions were formed with stations that were not underestimated. Thus the simulated streamflow data may be used with the Schaefer (1990) method to generate flood frequency curves.

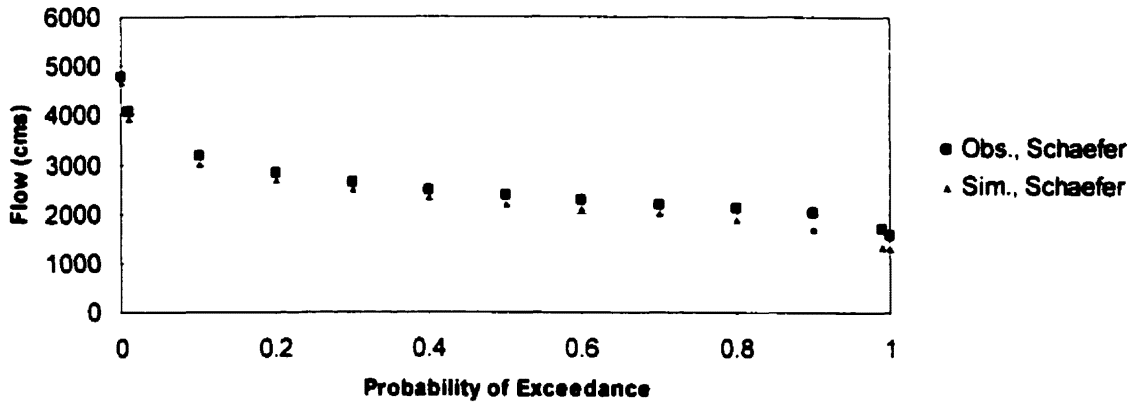


Figure 5-15 – Comparison of Frequency Curves for Mica Dam calculated using Regional Frequency Analysis (Schaefer 1990 method): the observed curve used the observed data records, the simulated curve used the 95-year simulated streamflow data.

5.2.3 Discussion of Flood Frequency Curves

This section derived an improved regional flood frequency curve. The regional frequency curve was based on the 95-year simulated streamflow time series calculated by the WATFLOOD/SPL model. The Hosking and Wallis (1997) regionalization method generated a frequency curve that underestimated the observed frequency curve; however, this problem was alleviated with the use of the Schaefer (1990) algorithm. The 95-year simulated regional frequency curve was compared to the observed regional frequency curve, and reasonable agreement was found.

The final issue in the use of simulated data to derive frequency curves was the derivation of confidence limits. A regional flood frequency curve based on observed data contains uncertainty due to observational error and due to the regionalization process. Hosking and Wallis (1997) presented a Monte Carlo method to derive confidence limits for frequency curves. The Monte Carlo simulation calculated multiple realizations of streamflow data for each station from a distribution fitted to the original data. Using the same regions, the regional frequency curves were re-calculated for each realization, and the

95% confidence limits were found. However, this method was unsuitable for use with simulated data, since simulated data also contain uncertainty due to the model(s). The model uncertainty was likely to be greater than the uncertainty due to observational error and regionalization. Model uncertainty would not be included in the Monte Carlo simulation described above. The next section describes the analysis to calculate confidence limits for the frequency curve derived from simulated data.

5.3 Derivation of the confidence limits for the flood frequency curve

One goal of this research was to determine accurate flood frequency curves. In general, the accuracy of a flood frequency curve improves with the use of a longer time series. This research used a 95-year simulated time series of streamflow to derive the frequency curve. However, although these data increased the accuracy of the frequency curve, they also caused the confidence limits to grow wider because of model uncertainty in the simulated data. This section focuses on defining the confidence limits for the flood frequency curve.

The Hosking and Wallis (1997) method to develop confidence limits for regional flood frequency curves was inappropriate due to the parameter uncertainty in the model. This method accounts only for uncertainty due to measurement error or errors in region definition. However, simulated streamflow data contain parameter uncertainty, and therefore, a Monte Carlo analysis was performed to develop the confidence limits.

There were two problems with performing a classical Monte Carlo analysis. First, the calibration philosophy of WATFLOOD/SPL differs from that of other authors who have performed Monte Carlo analyses. Other authors (e.g. Binley, *et al.*, 1991) have used multiple

parameter sets to develop a mean and standard deviation for each parameter, which are then used to describe the parameter distribution. This method of developing the parameter distributions would not represent the WATFLOOD/SPL parameter space appropriately, since WATFLOOD/SPL does not use multiple parameter sets. The calibration philosophy of WATFLOOD/SPL (Kouwen, *et al.*, 2000) assumes that the calibrated parameter values are close to optimum (i.e. at or near the peak of the optimum multi-dimensional hill on the objective function). The limits of the parameters should be set small enough that all of the simulations occur on a single multi-dimensional hill and do not "jump" onto another hill, since this would represent an invalid parameter set for WATFLOOD/SPL. Thus, the parameter distributions were set according to the calibration philosophy of WATFLOOD/SPL, and not according to the methods established in the literature. Secondly, a Monte Carlo based on the full 95-year time series of simulated streamflow would require too much computer simulation time. Therefore, the Monte Carlo was based on a five-year time series (1981-1985), and variation in the five-year time series was used as an analogue for the 95-year time series variation. This assumption was tested after the analysis was complete; the 95% confidence limits from the five-year runs and the 95-year runs were compared.

Therefore, this analysis presents a method for developing confidence limits for the flood frequency curves that agrees with the calibration philosophy of WATFLOOD/SPL. The following subsections present: the examination of the shape of the objective function, the parameter distributions for the Monte Carlo analysis, the conversion of the five-year Monte Carlo confidence limits to 95-year confidence limits, and the regional estimates of the confidence limits.

5.3.1 Investigation of the Objective Function

The multi-dimensional "hill" of the objective function of WATFLOOD/SPL near the optimum parameters was difficult to define. However, the limits of the "hill" were required to help define the parameter distributions. There were 70 different parameters involved in the Monte Carlo analysis (Table 3-2). Very little was known about the physically possible parameter space of WATFLOOD/SPL, except for the calibrated parameters and the physically possible ranges for the parameters (Table 3-2). It was known that there were inter-relationships between certain parameters (for instance, the melt factor and base temperature parameters are both used in the snowmelt algorithm), however, these inter-relationships were complex and poorly defined. In addition, the limits that are used represent the limits that are physically possible for each parameter. It is possible that the "physically probable" limits (likely parameter values) are smaller than the "physically possible" limits (determined from textbook values).

To help determine the limits of the "hill," the parameters were varied one at a time. This would give an approximate description of the "smoothness" of the objective function, and the approximate extent of the multi-dimensional "hill" of the objective function. Secondly, when the one-at-a-time analysis indicated that the objective function was not smooth, two parameters were varied simultaneously to define interactions between variables.

The first step of investigating the objective function was to modify one parameter at a time. The physically possible range was available for each parameter. Each variable varied between its minimum and maximum, in eleven steps: the first used the minimum value, the eleventh used the maximum value, and the other simulations used values ranging between the minimum and maximum.

The results of the analysis were compared in two different ways. Firstly, the one variable analysis was used to calculate the relative sensitivity of the model to each parameter, to determine which parameters affect peak flows. Secondly, the analysis was used to evaluate the smoothness of the objective function.

The automatic "fine tuning" step of the calibration process minimized the root-mean-square error between observed and simulated streamflow. However, this statistic was not the ideal statistic for analyzing the objective function shape for two reasons:

- Flood frequency analysis is based on the peak flows each year. The peak flows should also be used in evaluating the objective function shape.
- The root-mean-square statistic is always positive, and information regarding over- or under-estimation is lost. This information is useful for developing confidence limits.

The statistic was therefore based on the average difference between the peak flows for the simulation and the peak flows for the calibrated parameters (reference). This statistic was positive when the model overestimated the peak flows, and was negative when the model underestimated the peak flows. The 5-year time period of 1981-1985 (with the years of 1979 and 1980 used as a spin-up time) was used. The statistic was therefore:

$$statistic = Average\ Peak_{simulation} - Average\ Peak_{reference} \quad \text{.....(5-1)}$$

The reference peak flows were calculated using the calibrated parameter values. The calibrated parameter values were accepted as the best possible parameters, and therefore the flows resulting from these parameters were accepted as "perfect." Some authors use a Monte Carlo analysis to assist with calibration by using the observed flows as the reference (e.g. A. Mailhot, *et al.*, 1997). However, this analysis was not used to alter the calibrated parameter values, because doing so would alter the calculation of the water balance inside the

hydrological model. Therefore, it was acceptable to use the simulated peak flows as the reference (instead of the observed peak flows).

The statistic in Equation 5-1 was used to calculate the relative sensitivity of the model to each parameter, according to the equation (Filho, 1995):

$$Sensitivity = \frac{statistic}{\Delta parameter} \times \frac{parameter_{reference}}{average\ peak_{reference}} \quad \text{.....(5-2)}$$

The relative sensitivity is dimensionless, and invariant to the magnitude of the average peak flow or to the parameter value. Therefore, it was used to evaluate the relative importance of the various parameters to the generation of peak flows. Table 5-7 presents the range of relative sensitivity for each parameter.

Table 5-7 – Summary of Relative Sensitivities for each parameter

Parameter	Parameter Name	Relative Sensitivity Range	Rank of importance
A5	Unsaturated Zone Moisture Coefficient	-0.01 to 0.05	13
AK	Surface Permeability	-0.08 to 0.00	11
AKfs	Surface Permeability under snow	-0.14 to 0.00	9
REC	Interflow storage-discharge coefficient	-0.04 to 0.32	7
R3	Overland flow conveyance parameter	0.00	15
R3fs	Overland flow conveyance parameter under snow	0.00	15
RETN	Soil retention coefficient	-0.03 to 0.02	14
AK2	Upper to lower zone drainage coefficient	-0.12 to 0.01	10
AK2fs	Upper to lower zone drainage coefficient under snow	-0.68 to 0.30	3
LZF	Lower zone drainage function	-0.13 to 0.39	5
PWR	Lower zone drainage function exponent	-0.19 to 3.64	1
R2	River roughness coefficient	-0.10 to 0.50	4
MF	Melt factor	-0.55 to 1.10	2
BASE	Base temperature	-0.38 to 0.15	6
FPET	Potential Evapotranspiration Factor	-0.05 to 0.05	12
FTALL	Evapotranspiration Factor for Tall Vegetation	-0.23 to 0.20	8

The most important parameters for peak flow generation were the baseflow parameters (LZF, PWR), the snowmelt parameters (MF, BASE), river roughness (R2), the

interflow coefficients (REC), and the drainage from upper to lower zone under snow (AK2fs). These parameters are hydrologically signficant for peak flows. Base flow can be a large portion of a hydrograph, and as the baseflow rate increases, the peak flow will tend to decrease. The Columbia River basin is an alpine river basin, and therefore the peak flows are often associated with snowmelt events. Increasing the river roughness will tend to decrease peak flows. Increasing interflow will tend to increase peak flows since the flow can reach the river more quickly, while increasing the drainage from interflow to baseflow will lower peak flows since the flow will take longer to reach the river.

Some parameters were not important for the generation of peak flows. In forested areas, the overland flow parameters (R3, R3fs) have little effect because overland flow rarely occurs. These parameters describe the roughness of the ground and affect mainly the rate of runoff. However, a large precipitation event generates enough runoff to overcome depression storage. The unsaturated zone moisture coefficient (A5), the soil retention coefficient (RETN), and the potential evaporation (FPET) generally apply during dry weather processes (periods of low flow) and affect only the initial rise of the hydrograph. Therefore, they only moderately affect the peak flows. Therefore, the lack of sensitivity of the simulated peak flows to these parameters made hydrological sense.

The relative sensitivity could not be used to examine the shape of the objective function, since it was in relation to the amount of change in the parameter value. Therefore, it would tend to remain constant as distance from the calibrated value increases, since the greater change in average peak flow would be matched by a greater change in the parameter. Therefore, the difference between the averages of the peak flows for the simulation and the

reference was used directly to examine the shape of the objective function. The statistic was converted to a percent by calculating:

$$\% \, statistic = \frac{statistic}{average \ peak \ flow} \times 100 \quad \text{(5-3)}$$

The percent statistic was plotted against the parameter value for each parameter. A monotonically increasing (or decreasing) curve indicated that the objective function shifted smoothly from under- to over-estimation (or vice-versa). A convex or concave curve indicated that the objective function came to a maximum or minimum near the optimized parameter value. In both cases, the absolute value of the statistic would be small near the calibrated value of the parameter, and larger on both sides of the calibrated value. Therefore, the objective function would be optimal near the parameter value, and error would increase away from the calibrated value. Acceptable objective function shapes are shown in Figure 5-16. An acceptable objective function shape indicates two important characteristics of the variable:

- The variable can be optimized, as only one optimum exists in the physically possible range. (A second optimal parameter value does not exist within the range.)
- The entire physically possible range is located on a single "hill" of the objective function. (The objective function does not approach a second optimal parameter value within the range.)

Any other shaped curve was examined closely to determine how far it deviated from an ideal curve. Such deviation may affect the parameter distributions. Small deviations were allowed, where the deviations were less than 1% of the average streamflow.

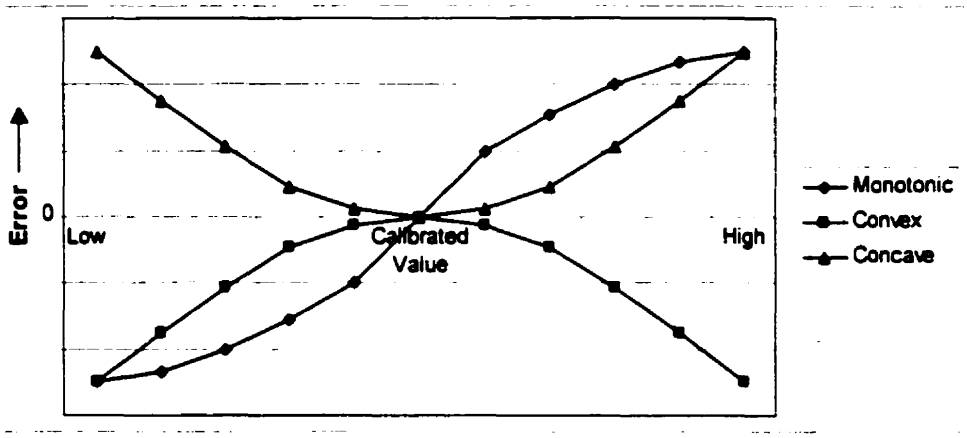


Figure 5-16 – Acceptable Shapes for the Objective function (error increases as distance from the calibrated value increases)

The variation in the statistic around the calibrated values of each variable is included in Table 5-8. The table also includes the range in the objective function statistic. The plots of the objective function statistics at Mica Dam for each of the 70 variables are included in Appendix C. A more detailed description of the objective function shape for each parameter is also included in Appendix C.

Table 5-8 – Summary of Objective Function Investigation

Variable	# of Classes	Range of Flow (% of average)	Shape	Conclusion
A5	1	-0.01% – 0.06%		Not a significant parameter for peak flows
AK	5	-0.26% – 10%	Monotonic	Smooth objective function
AKfs	5	-0.01% – 14%	Monotonic	Smooth objective function
REC	5	-54% – 34%	Monotonic (deviations <1%)	Smooth objective function
R3	5	-0.16% – 0.14%		Not a significant parameter for peak flows
R3fs	5	0%		Not a significant parameter for peak flows
RETN	5	-3.1% – 5.5%	Monotonic (deviations <1%)	Not significant for most land classes and stations. Smooth objective function for the rest.
AK2	5	-19% – 19%	Monotonic (deviations <1%)	Smooth objective function

AK2fs	5	-39% – 107%	Monotonic (deviations up to 3%)	Generally smooth objective function. Variations from desired curves in land class #5.
LZF	3	-53% – 87%	Monotonic (deviations <1%)	Smooth objective function
PWR	3	-78% – 177%	Monotonic (deviations <1%)	Smooth objective function
R2	3	-40% – 12%	Monotonic or Concave/Convex (deviations up to 5%)	Generally smooth objective function. Some variations from desired curves (all river classes).
MF	5	-100% – 54%	Monotonic or Concave/Convex (most deviations <10%)	Mainly smooth objective function. Some large variations from desired results.
BASE	5	-56% – 57%	Monotonic or Concave/Convex (large deviations)	Second mode visible on some traces. Several large variations from desired curves (land classes #1, #3, #5 in particular).
FPET	5	-0.60% – 0.62%		Not a significant parameter for peak flows
FTALL	5	-2% – 3%	Monotonic (deviations <1%)	Smooth objective function

The range of several parameters (A5, R3, R3fs, and FPET) was very small, between -1% and 1% of the average flow for all basins. These parameters also ranked low in relative sensitivity. Therefore, these parameters did not affect the generation of the peak flows.

For each of the remaining parameters, some combinations of land classes and streamflow stations were not significantly affected by the variation. The range of flow for these land class and streamflow station combinations remained between –1% and 1% over the entire physically possible range of the variable. These combinations of land class and streamflow station were not included in the description of the shape of the objective function.

The simulations showed that the objective function was mainly smooth near the calibrated values of the parameters. Most of the remaining parameters in Table 5-8 had traces where the absolute value of the statistic increased as the distance from the calibrated

value of the parameter increased (an ideal objective function shape). Deviations from the ideal objective function shapes that were less than 1% of the average peak flow for the particular station were assumed insignificant. For instance, occasionally a curve trended generally monotonically upwards from negative to positive, yet at one parameter value the statistic went down by 0.5% before continuing the upwards trend. Such a deviation (less than 1%) was considered insignificant. A deviation larger than 1% was recorded in the table.

Several parameters showed a region of sensitivity to the parameter, followed by a region where the effect of the parameter change was insignificant. This would indicate that the peak flows were no longer affected by the value of the parameter above or below a certain value (i.e. the hydrological process no longer affected peak flows). It indicated that there was a "ridge" on the objective function. From a physical perspective, these ridges may be caused by parameters such as saturated conductivity that reach values beyond which water either all infiltrates or ponds. This behaviour was considered acceptable, since this analysis was performed to determine if there was another "hill" on the objective function within the physically possible range of the parameters.

Two parameters had deviations from the ideal curves of more than 1% and less than 5%: AK2fs and R2. The AK2fs parameter (upper to lower zone drainage coefficient under snow) had some deviations from the ideal curves. The deviations were up to 3% of the average peak flow for a station. These large deviations occurred at two streamflow stations for the low elevation, light forest class. The R2 parameter (river roughness coefficient) also had significant deviations from the ideal curves; the deviations were up to 5% of the average peak flow for a station. Again, however, the number of stations with large deviations was relatively small (1 station for river class 1, 5 stations for river class 2, and 4 stations for river

class 3). Therefore, for these two variables, the objective function was assumed to be mainly smooth, and it was assumed that the "hill" extended over the physically possible range for the parameters.

Two other parameters had large deviations from the ideal curves: MF and BASE. The MF parameter (melt factor) had one deviation of 16% from the ideal curve, and the rest of the deviations were less than 10% from the ideal curve. The BASE parameter (base temperature) was the most erratic parameter. The BASE parameter sets the temperature at which the snow in WATFLOOD/SPL begins to melt. For the high elevation forests and for the barren areas, there were large deviations from normal in the objective function (up to 30% of the average peak flow). In many cases, a second "ideal" BASE parameter value was visible, where the peak flows would not change if the base temperature were set to this value. Although a second optimal parameter value for peak flow generation may exist, using this value would significantly alter the hydrograph, as the timing of the snowmelt would change. The peak flows remained the same because the melt began sooner, and the same peak flow resulted. Because of the large deviations in the MF and BASE parameters, it was decided to perform a two-parameter analysis for the MF and BASE parameters. In addition, the effect on the hydrograph was also examined during the two-parameter analysis for MF and BASE.

Therefore, the investigation of the objective function showed that, for most parameters, the objective function was smooth, and the "hill" extended over the entire physically possible range. For most parameters, the response statistic was small near the calibrated value of the parameter, and larger as the distance from the calibrated value increased (or the response statistic remained constant as distance increased). However, the snowmelt parameters (MF and BASE) did not follow this pattern. The peak flows were

highly affected by the snowmelt parameters, particularly by the BASE parameter. For these parameters, the objective function investigation indicated that there may be additional "hills" within the physically possible range for the parameters.

A two-parameter analysis was performed to determine the types of interaction between the two snowmelt parameters. For each landcover combination of MF and BASE, a set of 36 five-year simulations were performed. Since there were five landcovers, there were 25 combinations of MF and BASE. A grid pattern with six values of the BASE parameter on the vertical axis and six values of the MF parameter on the horizontal axis was used to describe the effect of the two parameters. The six values for each parameter ranged from the lower end to the upper end of the physically possible range. An example of the variation in average difference between the five-year peak flows for Mica Dam is shown in Figure 5-17. This Figure used the Barren landclasses for both the MF parameter and the BASE parameter (Appendix D contains similar Figures for all combinations of landclasses for the MF and BASE parameters for Mica Dam). The Figure shows that there was a line where the average difference was equal to zero, and the difference was negative on one side of the line and positive on the other. The line indicated that several combinations of MF and BASE were optimum for peak flow generation. The Figure was similar for other streamflow stations, indicating a general pattern in the parameters. This indicated that the objective function contained a single optimal "ridge," and only a single "hill." The calibrated parameter values (indicated with a diamond symbol) are the optimum parameters (in hydrological terms) for snowmelt and accumulation. Therefore, the two parameter variation of the snowmelt parameters indicated that a single hill existed in the physically possible range for both

parameters, however, there was a ridge on the hill where the two parameters interacted with one another to produce equal peak flows.

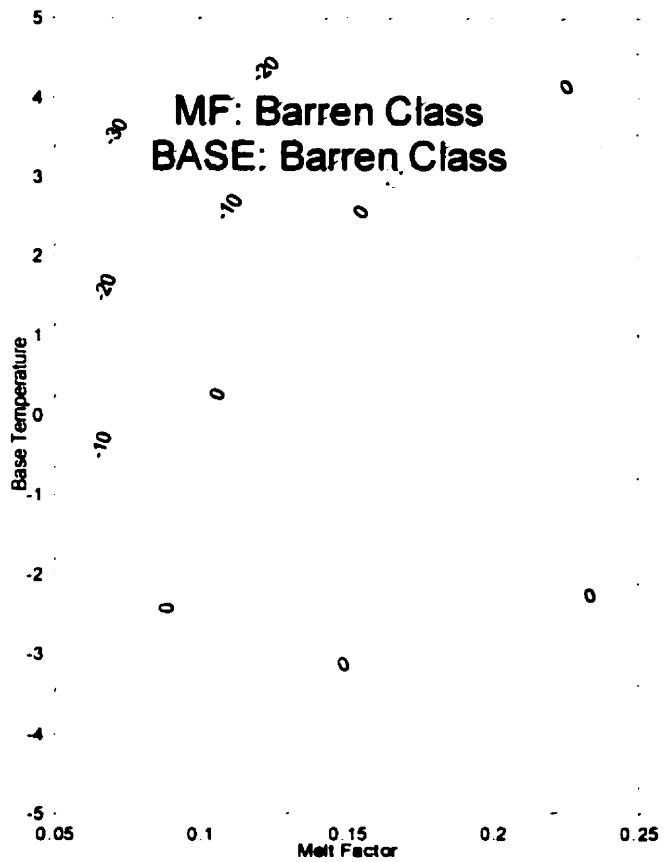


Figure 5-17 – Average difference between 5-year peaks (as percent of peak) for Mica Dam for the Barren classes of the Melt Factor and Base Temperature parameters

Figure 5-17 shows that an optimal "ridge" of combinations for the snowmelt parameters exists in the objective function, where the peak flows are unaffected by changes in the snowmelt parameters. Although the peak flows were unaffected, the shape of the hydrograph was significantly altered by changes in the snowmelt parameters. Several combinations of snowmelt parameters along the optimal ridge were chosen, and the root-mean-square (RMS) error for the Mica Dam basin was calculated. The average difference between the peak flows from the simulation and the peak flows from the calibration was also

calculated. The average difference was close to zero for all of the simulations (they were not exactly zero because the ridge in Figure 5-17 was calculated by interpolation and was not in the exact location). The RMS error shows that, although the peak flows were very similar with each simulation, the hydrograph was incorrect. The RMS was increased by changes in the snowmelt parameters, indicating that the simulated hydrograph was a poorer fit (than the hydrograph calculated with the calibrated parameter values). For the last simulation (BASE = 4°C, MF = 0.219 mm/°C), the RMS error decreased slightly (approx. 5 cms). The hydrograph for this simulation was compared to the hydrograph calculated with the calibrated parameters. The hydrographs were very similar to each other, and each simulation was a better match for the observed hydrograph for different parts of the hydrograph. The RMS errors for other streamflow stations showed that these two simulations were similar, with some stations having larger RMS error for the calibrated parameters and some stations having larger RMS error for the last simulation (BASE = 4°C, MF = 0.219 mm/°C). Therefore, in general, the RMS errors indicated that, although various combinations of snowmelt parameters can be used to generate the peak flows, the calibrated parameter values (or nearby) lead to the best hydrographs in terms of RMS error.

Table 5-9 – Calculation of RMS Error with several combinations of MF and BASE located on the optimal ridge (values for Mica Dam)

BASE	MF	Average difference in peaks (%)	RMS error 1981-1985 (compared to observed streamflow) (m^3/s)
-3	0.158	-0.01	255.7
-3	0.088	-0.50	234.1
-2	0.088	-0.03	224.5
-1	0.088	0.34	212.8
0	0.104	2.79	213.4
1	0.117	2.73	206.9
2	0.142	3.46	202.6
3	0.165	0 (calibrated value)	191.2
4	0.219	2.78	186.4

The objective function analysis indicated that the objective function contained a single "hill" for all of the parameters, with the exception of the snowmelt parameters. There was a single "ridge" of optimum parameter combinations for the snowmelt parameters. This analysis also indicated that the use of a single indicator of model calibration can be very misleading. It is stated in Section 3.3.2 that to consider a model calibrated, every possible indicator of model performance should be evaluated.

5.3.2 Choice of the Parameter Distributions

The parameter distributions were chosen based on the available information, the calibration philosophy of WATFLOOD/SPL and the results of the examination of the objective function. This section presents the parameter distributions.

Very little information was available regarding the distributions of the parameters. The results of the objective function analysis and the optimized value of each parameter were available, but inter-relationships between variables were not available. The parameters were therefore assumed independent for the purpose of this analysis. This assumption was known to be incorrect (e.g. the values of the snowmelt parameters depend on each other), however, the dependencies between variables were not clearly defined and therefore this assumption was used. It is likely that this assumption increased the variation in the model. Dependent variables would vary together in a pattern, whereas independent variables would vary across the entire range of physical possibility. This assumption, therefore, overestimated the parameter uncertainty. The information available for each parameter consisted of the results of the objective function investigation and the optimum parameter value.

The objective function analysis was used to determine how much of the physically possible range should be used in the distribution, to ensure that the Monte Carlo parameter sets remained on one "hill" of the objective function. The results showed that, with the exception of the snowmelt parameters, only one optimum parameter value existed within the physical limits of each parameter in terms of generating peak flows. In general, the objective function was smooth near the calibrated value, and either the absolute error increased as distance from the calibration point increased, or the absolute error remained constant as distance from the calibration point increased. For the snowmelt parameters, a "ridge" of optimum combinations of MF and BASE parameters existed, and as distance from the "ridge" increased, the absolute error increased. Therefore, the analysis indicated that a single multi-dimensional "hill" existed around the calibrated parameter values. Therefore, the entire physically possible range of each parameter was used in the distributions, since all combinations of the parameters were valid.

The calibrated parameter values are good estimates of the optimum parameter values, assuming that the full calibration and validation process of WATFLOOD/SPL has been performed. The calibrated parameters may vary somewhat from the optimum values due to uncontrollable uncertainties in the input data, such as variations in the quality of temperature and precipitation data, and therefore do contain some uncertainty. However, the optimum parameter values are the most likely values to be obtained from the calibration process. Therefore, the calibrated parameter values may be used as an estimate of the most likely parameter values, or the mode.

Lei and Schilling (1994) compared several parent distributions for parameters, and found that the parent distribution had little influence on the output of the Monte Carlo

analysis. Therefore, it was decided to use a simple distribution that only required the user to define the limits and the mode. The beta-1 distribution (e.g. Yevjevich, 1972) was used. The probability density function

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, \qquad B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \quad \text{.....(5-4)}$$

is a beta-1 distribution with boundaries zero and one. (The distribution was shifted and expanded to match the physical limits of each parameter.) The parameters α and β are used to define the shape of the distribution. To obtain a distribution with a mode and a probability of zero at both boundaries, α and β must both be greater than or equal to two. One of α and β must be defined, and the second can then be calculated from the mode. Since there was no other information to define the distribution, the values of α or β were chosen so that the largest possible standard deviation resulted. The probability density function requires that if the mode is between the lower limit and the midpoint, then α must be equal to 2.0 and β must be solved with the equation

$$mode = \frac{\alpha-1}{\alpha+\beta-2} \quad \text{.....(5-5)}$$

Similarly, if the mode is between the midpoint and the upper limit, then β must be equal to 2.0, and α is calculated from the above equation. With these values of α and β, the frequency distribution gave the widest possible distribution that had a defined mode at the calibrated value. This may result in an overestimation of the confidence limits, since some of the parameters are suspected to have a lower variance than the variance that is calculated with this equation. However, the probability distribution for each parameter has not been quantified, and therefore the widest possible distribution was used.

Some examples of the beta-1 distribution are shown for the BASE parameter in Figure 5-18. Three landclasses with different modes were selected. The calibrated value of the BASE parameter for the barren class was 3.0 °C; for the High Elevation Dense Forest class it was 1.0 °C; and for the Low Elevation Dense Forest class it was –2.0 °C. The distributions showed a mode at the calibrated value of the parameter, and the frequency dropped to zero at the physically possible limits. This distribution used all of the available information regarding the parameter distribution, but it did not add extra information. This parameter is an example of overestimation of the confidence limits due to a lower variance in the parameter than was used in the Monte Carlo. The BASE parameter values are not based on calibration but simply on the fact that (within one grid) the barren elevation is greater than the high elevation forest, which is greater than the low elevation forest. However, a single temperature value is used for the entire grid, and the base temperature represents the lapse rate to differentiate between the elevation bands within one grid. Therefore, the use of the entire range in the BASE temperature is an overestimation, but a more appropriate distribution has not yet been established.

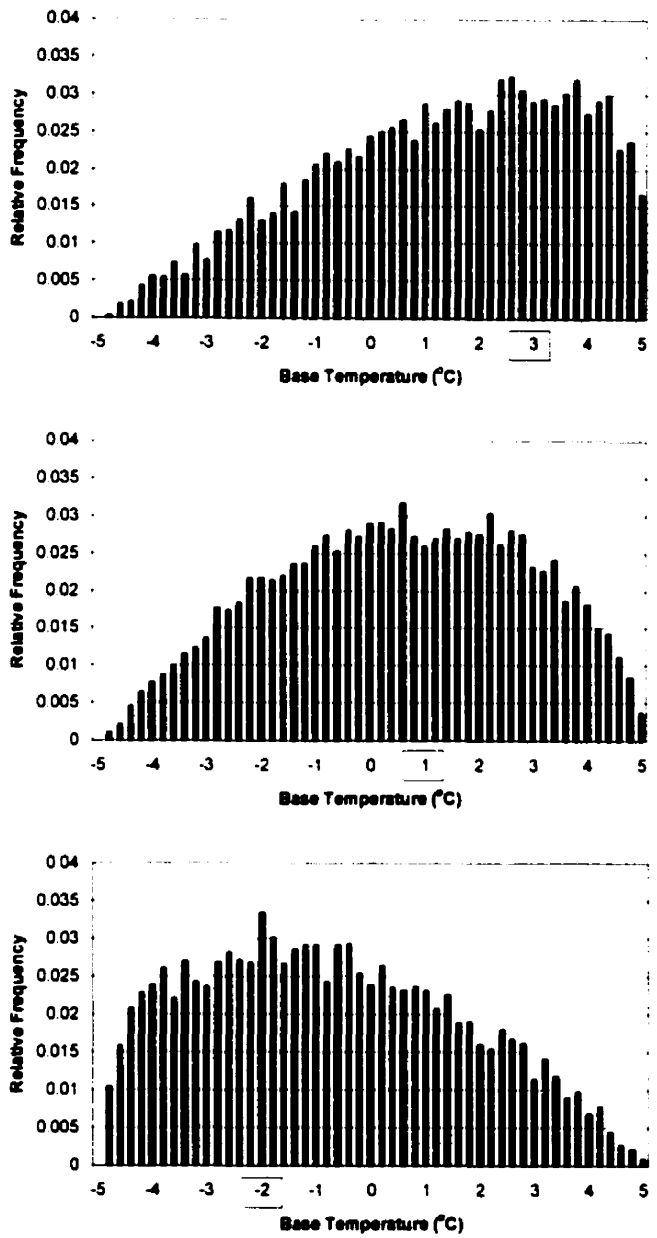


Figure 5-18 – Examples of histograms of the 10,000 values for BASE temperature, generated from the Beta-1 distribution (the mode is indicated with a red square). The three parameters are: a) BASE, Barren Class (mode = 3); b) BASE, High Elevation Dense Forest Class (mode = 1); BASE, Low Elevation Dense Forest Class (mode = -2)

5.3.3 Calculating the 95-year Confidence Limits

The Monte Carlo analysis was based on a five-year time series of streamflow. However, 96 years of meteorological input for WATFLOOD/SPL were available (one year

of data was discarded, yielding 95 years of simulated streamflow data. The large amount of computer simulation time (approximately three to four hours for one 95-year simulation on one CPU of an Origin 200 computer) forced the use of the smaller, five-year time series. The variation in the five-year time series was used as an analogue for the variation in the full time series. This section presents the method to convert the five-year confidence limits into 95-year confidence limits.

The Monte Carlo analysis was based on 10,000 runs of the hydrological model. According to Crosetto, *et al.* (2000), 100 runs per input factor are sufficient (as a rule of thumb) for a Monte Carlo. Since there were 70 parameters, a minimum of 7,000 simulations was required. Therefore, 10,000 sets of parameters were randomly chosen from the beta-1 distributions.

The hydrological model WATFLOOD/SPL was run for each of the 10,000 parameter sets. The meteorological data was the same for each run: the seven-year period from 1979-1985. These years were chosen because they represent a range of hydrological conditions from wet to dry. In addition, the largest observed storm occurred in 1983. The first two years (1979-1980) were discarded from the analysis, due to possible model spin-up errors. Therefore, the five years from 1981 to 1985 were used in the analysis. The maximum streamflow in each year of the five-year simulations was stored, so that the variation in peak flows was found. Therefore, for each streamflow station and B.C. Hydro dam inside the Columbia River Basin, a set of five peak flows for each of the 10,000 runs was available for analysis.

The 10,000 simulations were used to create 10,000 flood frequency curves (each made from five points). Figure 5-19 shows the results of the Monte Carlo for Mica Dam,

where the five peak flows of each simulation were ordered, and histograms were calculated for each ranking (plotted as scaled-down histograms at each probability level on Figure 5-19). The histograms for all of the probability levels were positively skewed. The skewness suggests that hydrological processes limit the lower boundary of the flood frequency curve, while the upper boundary is more difficult to define. The range for extremely low peak flows was smaller than the range for extremely high peak flows. This result was expected, since predictions of extremely high peak flows contain more uncertainty than predictions of low peak flows (Beven, 1993). In addition, the 2.5%, 50%, and 97.5% flow values of the histograms for each ranking were found (plotted as flood frequency curves on Figure 5-19). The behaviour of the histograms for each of the other 35 streamflow stations and B.C. Hydro dams was similar to the behaviour for Mica Dam.

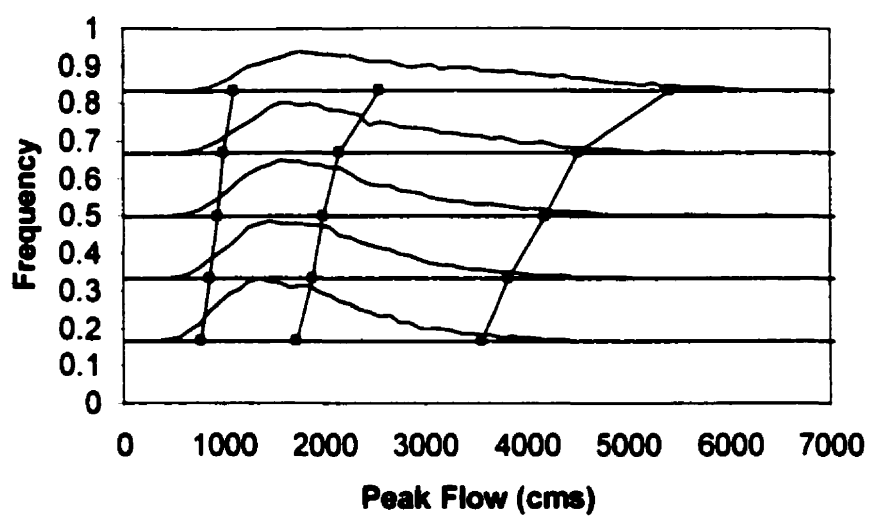


Figure 5-19 – Histograms of the 10,000 flood frequency curves for Mica Dam

Some parameter sets generated low peak flows, while others generated high peak flows. A particular simulation did not generate both extremely low peak flows and extremely high peak flows for a particular station. This indicated that a particular set of

parameters (one simulation) would only generate peak flows in a small range. The consistent behaviour of the hydrological model made it plausible to use the variation in the five-year frequency curves as an analogue for the 95-year frequency curves. The upper and lower frequency curves plotted on Figure 5-19 represent the values at 2.5% and 97.5% of the histogram for each peak flow (not a single simulation). Most simulations were entirely above or below these lines, but several simulations crossed these lines (i.e. some of the five peak flows were below the line, while others were above the line). The number of simulations that crossed a particular line formed from the histogram varied with the line chosen on the histograms (e.g. the 2.5% line versus the 97.5% line). The number of simulations that crossed the histogram lines varying from 1% to 99% was calculated and plotted in Figure 5-20 for Mica Dam. Relatively few simulations crossed a particular histogram line at the extreme flows, while a larger number crossed at medium flows. There were two possible reasons for this curve shape. First, there are many valid hydrological ways to calculate medium flows, but relatively few ways to calculate extreme flows (since several hydrological processes must all be at extreme values and co-operate together). Second, the parameter distributions specified that most of the simulations would occur near the calibrated parameter values. At Mica Dam, 203 of the 10,000 simulations crossed the 2.5% histogram line, while 362 crossed the 97.5% histogram line. Most of the other streamflow stations and B.C. Hydro dams also had more simulations crossing their 97.5% histogram line than their 2.5% histogram line. The shape of the curve for all 36 streamflow stations and B.C. Hydro dams was similar to the curve in Figure 5-20. Very few simulations crossed the histogram lines at the upper and lower probabilities, and a larger number of simulations crossed the histogram lines at the middle probabilities.

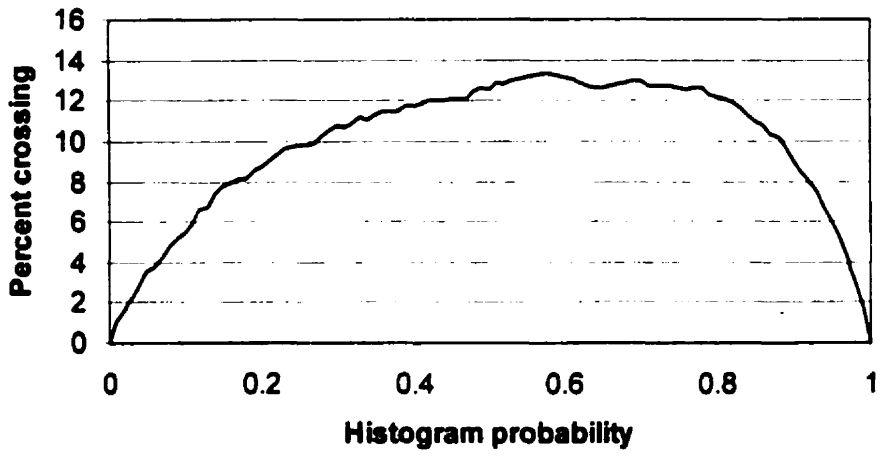


Figure 5-20 – Percent of 10,000 simulations whose flood frequency curves cross lines of the histogram for Mica Dam

The simulations that crossed the 2.5% or the 97.5% lines represented a set of parameters that generated extreme flows (either low or high) for the five-year time period from 1981 to 1985. For instance, each of the 203 simulations that crossed the 2.5% line represented a set of parameters that generated extremely low peak flow estimates. Likewise, each of the 362 simulations that crossed the 97.5% line represented a set of parameters that generated extremely high peak flow estimates. It was assumed that these same parameter sets could be used to generate extreme flows for the full 95-year time series. Therefore, the simulations that crossed the 2.5% or 97.5% histogram lines were found for all 36 streamflow stations and B.C. Hydro dams. Frequently, a simulation would cross the 2.5% line (or the 97.5% line) for several stations, further indicating that a particular set of parameters generated consistently low flows (or high flows). In total, 3464 simulations crossed either the 2.5% line or the 97.5% line at one or more of the 36 streamflow stations and B.C. Hydro dams. Comparatively few simulations (14 out of 3464) generated high flows at one station and low flows at another station. The 95-year simulated streamflow time series were

generated with the WATFLOOD/SPL model for each of the 3464 sets of parameters. There were 96 years of meteorological data available, however, the hydrological model has spin-up issues for the first two years. To allow 95 years of data to be used, the model was first run for a normal year (1971) so that the model would begin the 96-years with a normal watershed condition. Then the first year (1899) was discarded, leaving 95-years for analysis (1900 to 1994).

Since several simulations were performed for each station, there were several simulations available to represent the limits for each station. Therefore, it was necessary to develop "overall" confidence limits for each station. The 95-year simulated streamflow time series were used to generate frequency curves for each simulation for each station. The Wakeby distribution was used because it was also used to develop the regional flood frequency curve. Therefore, for each station, the 95-year frequency curves that corresponded to each of the simulations that crossed the 2.5% (or the 97.5%) histogram line were calculated. These frequency curves were aggregated to form an overall estimate of the upper and lower confidence limits.

The lower confidence limit was found as the average of all of the frequency curves for the station. This method was chosen because all of the simulations represent a set of parameters that generate peak flows near the 2.5% line of the histogram for that station, and therefore, they are all valid estimates of the lower confidence limits. For instance, 203 simulations crossed the 2.5% histogram line for Mica Dam. The 203 frequency curves generated from the 95-year simulation were averaged to obtain the lower confidence limit for Mica Dam. Figure 5-21 shows a comparison of the two estimates for the lower confidence limit for Mica Dam. The five points from the 2.5% histogram line of the Monte Carlo

analysis are compared with the average of the frequency curves calculated from the 203 95-year simulations. The two estimates nearly overlapped one another, and therefore the average of the frequency curves was able to extend the 2.5% histogram line.

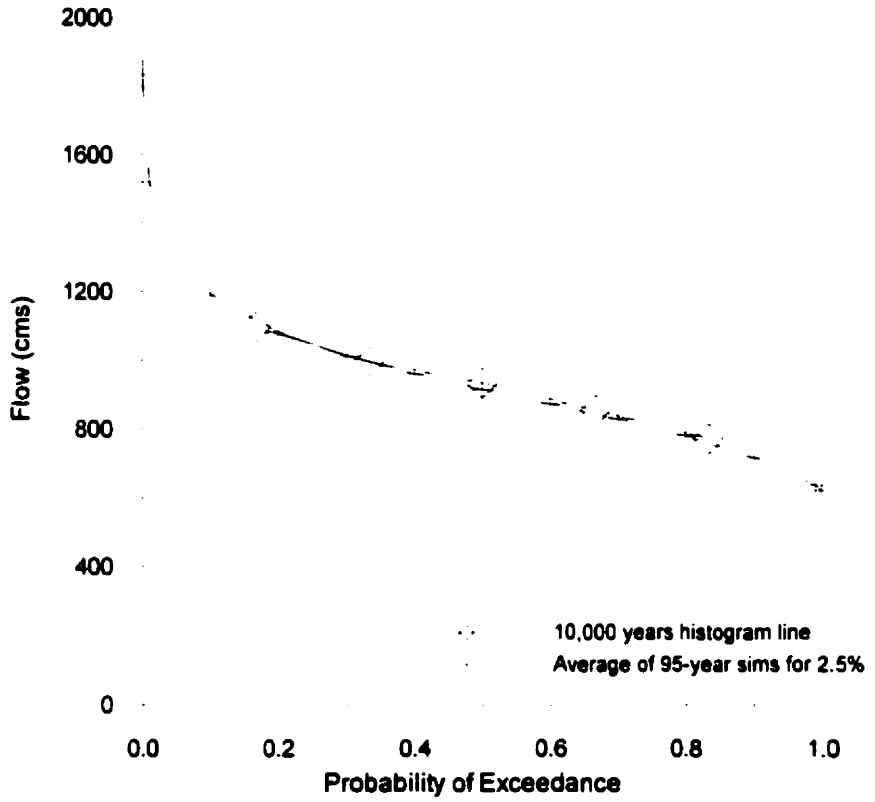


Figure 5-21 – Comparison of Lower Confidence Limits for Mica Dam calculated with the Monte Carlo and as the average of the 95-year simulations

For the upper confidence limit, the average of the frequency curves did not match the 97.5% histogram line from the Monte Carlo analysis (Figure 5-22). The average of the frequency curves was approximately 700 m^3/s too low (17% error at a probability of 0.5). However, the maximum of the frequency curves matched the 97.5% histogram line (Figure 5-22). There were two possible reasons why it was necessary to use the maximum of the frequency curves instead of the average of the frequency curves. The mean of the five-year peak flows was (on average) slightly higher than the mean of the 95-year peak flows. The

mean peak flow for 1981-1985 for Mica Dam was 10% larger than the mean peak flow for 1990-1994. A second reason is due to the fact that predictions of extremely high peak flows contain more uncertainty than predictions of low peak flows (Beven, 1993). The histograms created from the 10,000 simulations of the Monte Carlo analysis were positively skewed, and more simulations crossed the 97.5% line as opposed to the 2.5% line. Some of these simulations may have crossed the 97.5% line due to the larger variation, although they did not truly represent parameter sets that generated peak flows in the 97.5% range. These two factors together may have caused the average of the frequency curves to be lower than the 97.5% line of the histogram.

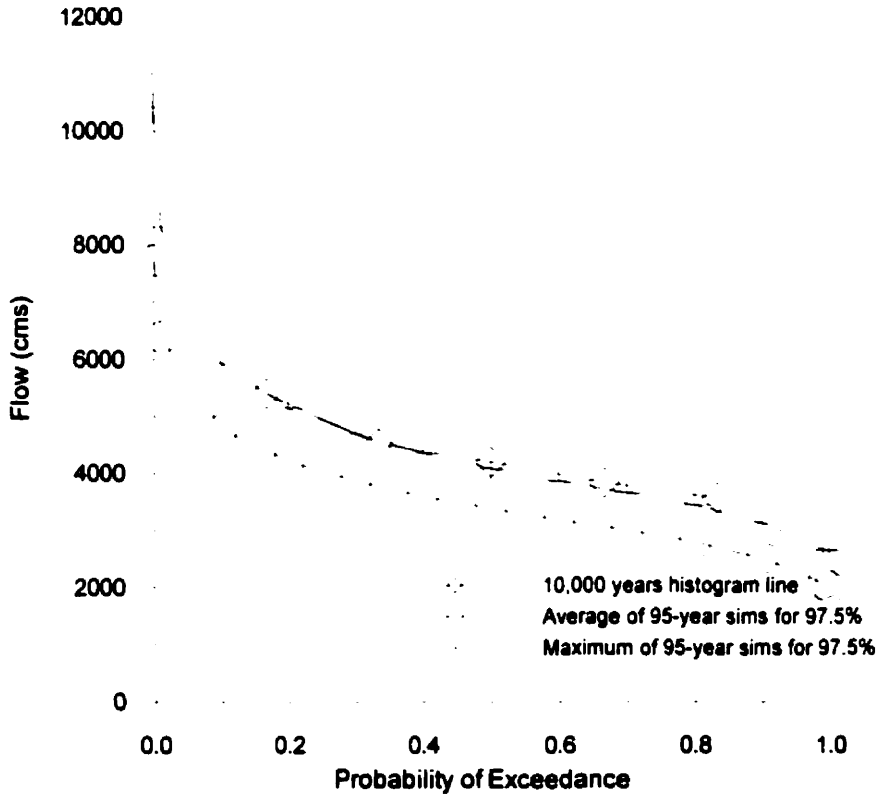


Figure 5-22 – Comparison of Upper Confidence Limits for Mica Dam calculated with the Monte Carlo and as the average and maximum of the 95-year simulations

The comparisons between the 2.5% and 97.5% histogram lines and the overall confidence limits for all of the stations are included in Appendix E. In most cases, the frequency curves derived from the 95-year simulations were coincident with the histogram lines. The 95-year simulations were able to "fill-in" the parts of the frequency curve not calculated by the Monte Carlo. This further indicated that the variation in the five-year simulations was a suitable analogue to determine the variation in the 95-year simulations.

5.3.4 Calculating the Regional Confidence Limits

The previous sections have described the development of 95% confidence limits for each individual streamflow station and B.C. Hydro reservoir. However, since the flood frequency curve (from simulated streamflow data, Section 5.2.2) was calculated with a regional flood frequency method (Hosking and Wallis, 1997), it was appropriate to also calculate a regional estimate of the confidence limits. The regionalization was based on regressions between the physiographic data (Solomon, 1968) and the individual frequency curves, using a method similar to that of Schaefer (1990).

The L-moments of the aggregated frequency curve, if they were calculated, would be fictitious L-moments, since they would not represent a single set of data. Therefore, the regionalization was based on the frequency curves for each station. The flow value for a probability of 0.5 was used to normalize the frequency curves for each station. The regressions were formed between the normalized confidence limit curves and the basin characteristics. The physiographic variables from Solomon (1968) were used for this purpose.

There were 36 upper confidence limit curves and 36 lower confidence limit curves (one for each streamflow station or B.C. Hydro dam). Each curve was calculated at discrete

probabilities of exceedance. Therefore, for each confidence limit, regression was required for each probability level. However, to reduce the possibility of regional confidence limits that appeared "jagged," a single regression was chosen for all probabilities of exceedance. Therefore, the best regression was the regression that had the best fit at all of the probability levels. Note that regression for the probability level of 0.5 was not required since this probability was used for normalizing, and the frequency was equal to one by definition.

The regressions were calculated, and 11 variables were required in order to obtain suitable regressions. They explained 50% or more of the variation in the normalized frequency curves across all 36 streamflow stations and B.C. Hydro dams (Table 5-10 and Table 5-11). Two regressions were not able to explain 50% or more of the variation in the normalized frequency curves; they were both regressions for low flows of the upper confidence limit and were of less interest than the regressions for high flows. The regionalization altered the flow values for extreme drought and flood probabilities. In both cases, the confidence limits were not significantly changed in the mid-probability range, but there were small changes at the extreme tails of the confidence limits. The low flows were decreased and the high flows were increased for the lower confidence limit. For the upper confidence limit, the low flows were decreased, while the high flows remained unchanged. For instance, at a return interval of 1 in 1000, the lower confidence limit increased from 1832 m^3/s to 1947 m^3/s (Table 5-10), while the upper confidence limit decreased from 10413 m^3/s to 10355 m^3/s (Table 5-11). These changes in the extreme tails of both confidence limits caused the 95% confidence limits to become 2% narrower at a return interval of 1 in 1000.

Table 5-10 – Regression Results for Lower Confidence Limit at Mica Dam

Probability of exceedance	R	Individual		Regional		Variables
		m^3/s	Normalized	Normalized	m^3/s	
0.001	0.74	1832	2	2.12	1947	Average Julian Date of Peak
0.01	0.79	1520	1.66	1.73	1585	Ratio of Peak to Mean
0.1	0.85	1191	1.3	1.34	1232	Azimuth
0.2	0.85	1082	1.18	1.21	1111	DTO-NW
0.3	0.86	1014	1.11	1.13	1032	DTO-SE
0.4	0.86	962	1.05	1.06	971	DTO-SW
0.5	N/A	917	1	1	917	SHE-NE
0.6	0.88	875	0.95	0.94	866	SHE-SW
0.7	0.89	831	0.91	0.89	813	BH-NE
0.8	0.9	782	0.85	0.82	752	BH-NW
0.9	0.88	719	0.78	0.74	676	BH-SW
0.99	0.81	637	0.69	0.63	576	
0.999	0.77	623	0.68	0.59	541	

Table 5-11 – Regression Results for Upper Confidence Limit at Mica Dam

Probability of exceedance	R	Individual		Regional		Variables
		M^3/s	Normalized	Normalized	m^3/s	
0.001	0.81	10413	2.55	2.54	10355	Area
0.01	0.69	8311	2.04	1.86	7572	Average Julian Date of Peak
0.1	0.67	5931	1.45	1.43	5811	Ratio of Peak to Mean
0.2	0.69	5152	1.26	1.26	5153	Azimuth
0.3	0.7	4687	1.15	1.16	4724	DTO-NE
0.4	0.67	4347	1.07	1.07	4377	DTO-SW
0.5	N/A	4077	1	1	4077	SHE-NE
0.6	0.7	3855	0.95	0.93	3772	SHE-SW
0.7	0.74	3651	0.9	0.84	3444	BH-NW
0.8	0.77	3433	0.84	0.76	3093	BH-SE
0.9	0.76	3120	0.77	0.64	2622	BH-SW
0.99	0.59	2657	0.65	0.56	2283	
0.999	0.53	2630	0.64	0.56	2286	

The regionalized confidence limits (Table 5-10 and Table 5-11) were combined with the frequency curves shown in Figure 5-15 to create Figure 5-23. In general, the confidence interval widened as the return interval increased. This was expected, because the Monte Carlo histograms were positively skewed and the range of high peak flows was much larger than the range of low peak flows. This indicated that the model parameterization contained

greater uncertainty for higher peak flows. This agreed with the findings of Beven (1993), who found that uncertainty increased for peak flows.

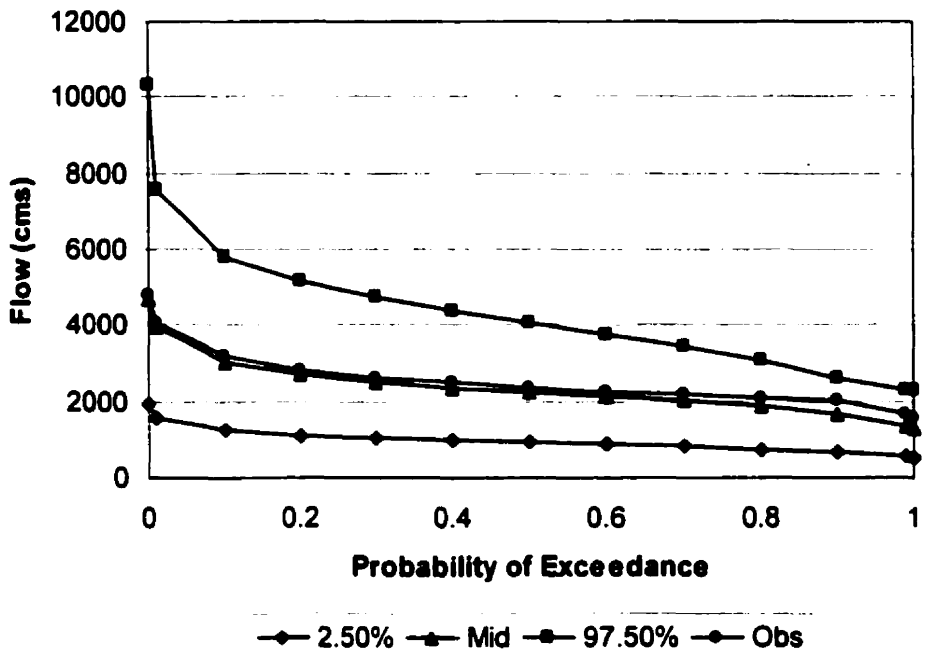


Figure 5-23 – Regional Frequency Curve (Schaefer, 1990, method) for Mica Dam Combined with the Regional Confidence Limits Determined from the Monte Carlo Analysis

The confidence limits compared well with other parameter uncertainty studies. Relative to the simulated frequency curve, the flows of the lower confidence limit were one-third to one-half of the magnitudes of the frequency curve. The upper confidence limit varied from approximately 1.5 to three times the magnitude of the frequency curve. One parameter uncertainty study by Beven (1993) calculated 5% and 95% confidence limits for TOPMODEL on a small watershed (3.5 km^2). Beven (1993) found that the upper confidence limit was approximately twice the observed peak streamflow, while the lower confidence limit was approximately 80% of the observed peak streamflow. Similarly, Binley, *et al.* (1991) calculated 5% and 95% confidence limits for IHDM on a 3.9 km^2 watershed and

found that the upper confidence limit was approximately 1.5 times the observed peak streamflow, while the lower confidence limit overestimated the observed peak streamflow. In both of these studies, the probability of exceedance of the peak streamflow in the simulated time series was not reported, and therefore cannot be compared to the present study. Therefore, the upper confidence limit derived for the frequency curve calculated with simulated streamflow data from WATFLOOD/SPL was similar in magnitude to other studies by other authors. The lower confidence limit, however, was lower than those presented in literature.

The simulated data were derived from a hydrological model that was subject to parameter uncertainty. The confidence limits are wider than those that would be calculated for frequency curves derived from observed data, and indicate large uncertainty (particularly for extremely large flows). Although the model was properly calibrated to ensure good streamflow prediction, the parameter uncertainty remained significant. The confidence limits may in fact overestimate the parameter uncertainty of WATFLOOD/SPL, due to the use of the widest-possible distribution for each parameter. However, A. Mailhot, et al. (1997), also found that parameter uncertainty remained large even after calibration. Some of this uncertainty may be due to the model parameterization, and not the parameters. For instance, the equation for modeling snowmelt has uncertainty, and at some point, greater accuracy in the parameters will no longer decrease the model uncertainty. The model parameterization is an important factor to remember during analysis to decrease parameter uncertainty. The parameter uncertainty for this model may be lowered mainly through reducing the width of the physical limits of each parameters. Therefore, the physically-based limits for the parameters should be examined to determine if they might be altered to lower the parameter

uncertainty. One landcover may have different parameter limits from another, for instance. One method for this would be to use the "physically probable" parameter limits (the limits for the parameters that are likely to occur), instead of the "physically possible" limits. Alternately, when the certainty of the calibrated value of the parameter warrants it, the beta-1 distribution could be altered so that the dispersion decreases. Some effort should be performed to lower the parameter uncertainty of WATFLOOD/SPL.

It was known that certain parameters have narrower distributions than those used in the Monte Carlo. For instance, the BASE parameter should have a narrow distribution, since the base temperature accounts for the within-grid elevation range between the different landcover elevations. This, in turn, sets the value of the MF parameter. The establishment of more narrow distributions for these parameters would lower the parameter uncertainty by up to 25%, as determined from a separate Monte Carlo analysis. A 1000-run Monte Carlo analysis that varied only these 10 parameters (5 landcovers for each of MF and BASE) was performed. The same 5-year analysis was performed, and the 2.5% and 97.5% histogram lines were determined. The 2.5% and 97.5% histogram lines were added to Figure 5-23 to create Figure 5-24. The snowmelt parameters accounted for approximately 25% of the variation around the calibrated value. This represented a significant amount of the overall width of the confidence interval, and suggested that the distributions of these two parameters should be the first to quantify to reduce the parameter uncertainty.

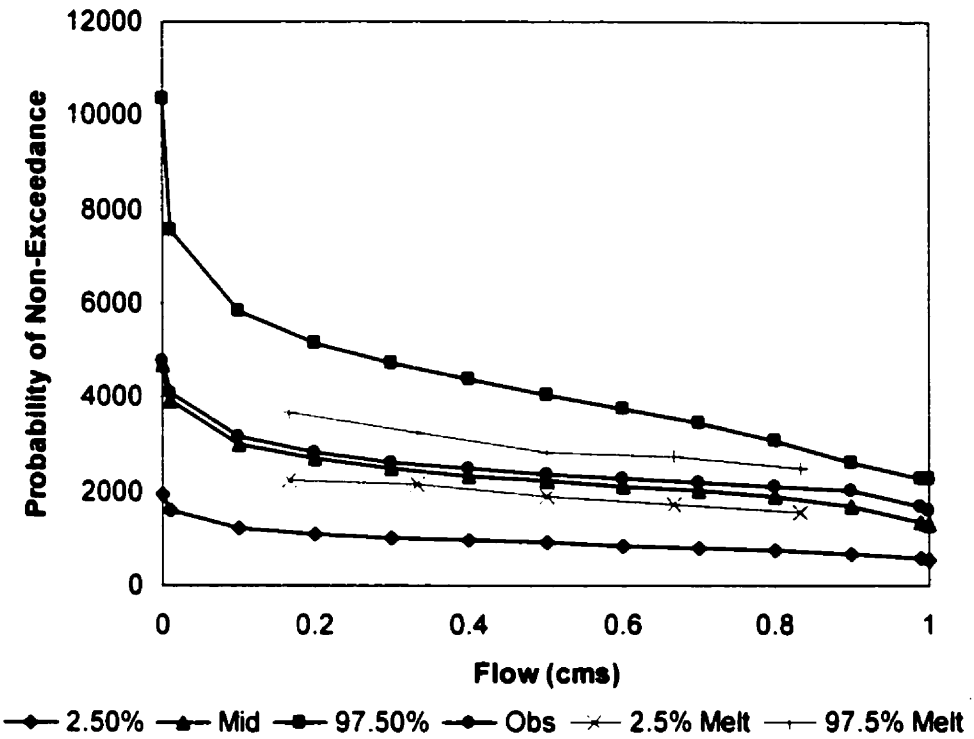


Figure 5-24 – Regional Frequency Curve for Mica Dam, showing Confidence Limits and amount of variation due to snowmelt parameters (MF and BASE)

5.4 Chapter Summary

This chapter has explored the use of a long continuous historical series of simulated streamflow data to develop frequency curves. This application was the first to use a simulated streamflow time series that was longer than the observed record of streamflow. Meteorological data over the Columbia River basin was available for a 96-year period from 1899 to 1994. The distributed physically-based hydrological model WATFLOOD/SPL was used to develop simulated streamflow data.

The output from the WATFLOOD/SPL model was validated for the production of frequency curves. It was found that the model produced reasonable estimates for mid-range probabilities, but relatively poor estimates for the high-flow, low-probability range.

However, there was better agreement for streamflow stations with longer time series, which indicated that the observed frequency curves may be in error for short time series stations. Therefore, the output from the WATFLOOD/SPL was suitable for use in predicting flood frequency curves. In addition, the longer time series data could potentially improve the accuracy of frequency curves, especially for short time series stations.

The simulated streamflow data were used with regional frequency analysis to generate a frequency curve for the Mica Dam basin. The Schaefer (1990) method for modifying the Hosking and Wallis (1997) method generated the best frequency curves.

Finally, the 95% confidence limits for the frequency curve due to the parameter uncertainty of WATFLOOD/SPL were calculated with a Monte Carlo method. The method was developed specifically for the WATFLOOD/SPL model. The parameters within WATFLOOD/SPL are robust and are known to be hydrologically correct. Most Monte Carlo methods use a normal distribution for the parameters, which would be inappropriate for the WATFLOOD/SPL model. Therefore, the parameter distributions were set up so that only combinations of parameters that were hydrologically possible were used. The method also allowed a shorter time series to be used instead of the full time series. The parameter uncertainty was quite large, but within the range of other hydrological models. Establishing tighter boundaries on the parameter distributions would decrease the uncertainty. Several of the parameters are known to have a narrower dispersion than what was used in the Monte Carlo.

Chapter 6 will use the frequency curve developed in this chapter to compare the floods from Chapter 4.

6 Comparing the Probabilities of the Floods

This research has been used to develop a physically-based flood (Chapter 4) and an improved frequency curve using simulated streamflow data (Chapter 5). These two pieces of information are important tools for safety analysis and risk-based analysis of dams. However, in this section, the frequency curve is used to compare the floods generated by the two estimates of maximum precipitation. The return intervals of the floods were used as a validity check to determine if the floods were in the presumed size range of a PMF.

The Probable Maximum Precipitation (PMP) was calculated with the World Meteorological Organization method (WMO, 1986). The Probable Maximum Storm (PMS) was calculated with the MC2-PMS model. The two storms were used with various antecedent conditions to produce floods (Figure 4-15). In all cases, the PMS flood was significantly less than the flood from the PMP. Instead of comparing the magnitudes of the floods, however, the return intervals of the floods were used to compare the floods.

According to Smith (1998), the return period of the Probable Maximum Flood (PMF) may range from 10,000 to 1,000,000 years (Annual Exceedance Probability, AEP, of 10^{-6} to 10^{-4}). The large range of presumed AEP for the PMF is due to widespread disagreement over correct methods of calculating the PMF, and possible errors in frequency curve extrapolation.

The frequency curve in Figure 5-23 was used to calculate the Annual Exceedance Probability (AEP) of the peak flow from the different floods in Figure 4-15. The AEP are compared in Table 6-1. The frequency curve and its confidence limits were extrapolated out to a probability of 10^{-7}. This value was chosen because the desired PMF range was 10^{-4} to 10^{-6}, and extrapolation to 10^{-7} would allow evaluation of flood peaks in the PMF range. Further extrapolation was considered unreasonable.

Table 6-1 – Comparing the Probabilities of the Different Floods (Streamflow on June 1 for each simulation)

Simulation #	Flood	AEP	95% Range
1	1972	0.72	0.001 to >0.99
2	1972 with PMS	0.02	$10^{-5.4}$ to 0.56
3	PMF scenario(with PMP): 100-yr snowpack, 100-yr melt temperatures, and PMP	$<<10^{-7}$	$<<10^{-7}$ to $10^{-3.9}$
4	PMF scenario(with PMS): 100-yr snowpack, 100-yr melt temperatures, and PMS	$<10^{-7}$	$<<10^{-7}$ to 0.005

The AEP of the floods for the two simulations that used the regular 1972 meteorology (simulations 1 and 2) could be calculated with no extrapolation of the frequency curves. The AEP of the flow on June 1, 1972 was slightly below the mean (72% chance of exceedance). The range of AEP for the June 1, 1972 streamflow indicated that 1972 was well within the "normal" range (simulation 1). After the PMS storm was added to June 1, 1972 (simulation 2), the probability of the streamflow dropped to a 2% chance of exceedance (but it may be as high as a 56% chance of exceedance). The 1972 data with the PMS was not within the desired range of the PMF, however, the data indicated that the flood was close to the desired range. Although the hydrograph was not shown in Figure 4-15, the PMP storm was also added to June 1, 1972. The PMP caused a streamflow on June 1, 1972 that had a probability of exceedance of less than 10^{-7} (with a range up to $10^{-3.0}$). These data indicated that the PMP storm may be overestimated, since "wet" antecedent conditions were not required to generate streamflow that was higher than floods in the PMF range. The PMS, however, was in a size range that required slightly "wetter" antecedent conditions (than the conditions that occurred in 1972) and/or some snowmelt to generated a flood in the probability range of the PMF. In addition, the PMS presented in Chapter 4 is preliminary, and further study into the assumptions of the PMS module and the MC2 model is recommended. Therefore, the PMS may increase, and result in a flood in the size range of a PMF.

The AEP of the floods for the two simulations that used the 100-year snowpack and the 100-year melt temperatures (simulations 3 and 4) could not be calculated, even after extrapolation of the frequency curves. The use of the 100-year snowpack and 100-year temperature sequence alone resulted in a streamflow on May 31 that had a probability of exceedance of $10^{-5.5}$ (may be as high as a 9% chance of exceedance). When the PMP was added to the antecedent conditions (simulation 3), the probability was much less than 10^{-7} (up to $10^{-3.9}$ chance of exceedance), whereas when the PMS was added to the antecedent conditions (simulation 4), the probability was somewhat less than 10^{-7} (up to 0.5% chance of exceedance). The addition of the PMP to the antecedent conditions resulted in streamflow that had a very low probability (the 95% confidence limits were just above the desired range for a PMF). Since the confidence limits presented in Chapter 5 may represent an overestimation of the parameter variation within WATFLOOD/SPL (due to the assumptions regarding the variance in the parameters), these results indicate that the combination of PMP and antecedent conditions may fall entirely outside the desired range of probability for a PMF. However, the addition of the PMS resulted in a more suitable flood hydrograph: the probability was still very low ($<10^{-7}$), but not as low as the flood calculated with the PMP. The combination of the PMS and the 100-year snowpack and 100-year melting temperatures may also be an overestimate of the PMF, but its upper confidence limit may fall within the PMF range after the overestimation in the confidence limits was corrected.

The relative probabilities of the floods indicated that the PMS was able to generate floods that were somewhat larger than a reasonable size range for a PMF, but which may be appropriate given the modeling uncertainty. However, the PMP generated floods that were

much larger than a reasonable size range for a PMF (even when the modeling uncertainty was taken into account).

7 Conclusions

The method presented in this research for developing the Probable Maximum Precipitation (PMP) was to use the MC2-PMS model to develop a physically-based PMP (called the PMS). This research determined that an atmospheric maximum precipitation does in fact exist, and that it can be calculated with the model. The method is less subjective than the World Meteorological Organization (WMO) method, and accounts implicitly for topography. The precipitation occurred mainly in valleys and mountain passes, and very little precipitation occurred at higher elevations. The maximum precipitation event was lower than the PMP, and therefore the WMO method may overestimate the PMP in mountainous terrain. In addition, the flood produced by the PMS was significantly lower than the flood produced by the PMP. This difference in magnitude may affect the design of safety structures for dams.

The WATFLOOD/SPL model calculated a long simulated streamflow time series that was able to generate reasonable frequency curves. The increased time series length decreased the sampling uncertainty associated with the frequency curve. However, the frequency curve became subject to modeling uncertainty.

An investigation into the behavior of the hydrologic model showed that there was only one multi-dimensional hill within the limits of the parameter space, with the calibrated parameters near the peak of the hill. Therefore, the parameters within WATFLOOD/SPL are robust and the calibration process leads to the true optimum parameter values. The parameter distributions used the physically possible boundaries for limits, and the calibrated parameter values as the mode. A Monte Carlo analysis was performed to determine the parameter uncertainty within the model.

The confidence limits for the flood frequency curve grew wider as the return period increased. Although the width of the confidence limits was large and consequently an analysis of the return intervals of the floods was difficult, the width was within the range of published values for confidence limits. Approximately 25% of the variation was due to the snowmelt parameters, indicating that these parameters may be most significant for reduction of the variation within the model.

The Annual Exceedance Probability of the different floods indicated that the antecedent conditions (100-year snowpack and 100-year melt temperature sequence) may overestimate the PMF. In addition, the addition of the PMP to the antecedent conditions resulted in very large estimates of the PMF, with a very low probability of exceedance. The addition of the PMS resulted in a more reasonable estimate of the PMF, and the flood produced by the PMS was closer to the desired range for a PMF.

7.1 Contributions

This research has contributed in two major areas. The first area was in the development of a physically-based maximum precipitation. The second area was in the development of frequency curves and their associated confidence limits with simulated streamflow data. To summarize, the contributions are:

- Procedures for use of the MC2-PMS model for the development of extreme precipitation events and determining that a maximum precipitation event does exist. This approach is an improvement over the traditional method of estimation of the PMP.
- Demonstration that the maximum storm and flood (from physically-based techniques) were smaller than the PMP and PMF, supporting the view that the traditional techniques for PMP and PMF estimation are flawed in mountainous regions.

- Validating the use of a long simulated streamflow time series from WATFLOOD/SPL model for the development of frequency curves and using these data to develop a regional frequency curve.
- Development of an efficient method to derive confidence limits for flood frequency curves that are based on streamflow data generated by a hydrologic model with a large number of parameters.
- Application of the methods to assess the effect of parameter uncertainty on the range of output for extreme events for WATFLOOD/SPL.

8 Recommendations for Further Research

The MC2 model has been shown capable of developing a maximum atmospherically possible precipitation. However, further research to check the atmospheric assumptions within MC2 is needed to ensure that they are suitably conservative for calculation of maximum precipitation. In addition, the characteristics of the storm should be examined further; the storm type, the date that the storm is embedded into, and other variables such as sea temperature should all be examined to ensure the true maximum precipitation is being developed.

The use of a long continuous simulated streamflow time series improved the estimate of the frequency curve by decreasing the sampling uncertainty. Meteorological data can be derived with atmospheric models over much of the world, which can be converted into streamflow. This would lengthen the time series for both gauged and ungauged areas, and increase the accuracy of the frequency curves. Further research is needed to develop the meteorological data with atmospheric models, and convert it to streamflow using physically-based hydrological models that require only minimal calibration.

However, simulated streamflow data contain uncertainty due to the hydrological model (and/or the meteorological model). Therefore, the parameters within the hydrological models should be examined to determine their behavior, and to minimize the width of the confidence interval wherever possible. For the WATFLOOD/SPL model, more accurate descriptions of the parameters' behavior should be developed to better represent the information that is known about them.

This research should be expanded by comparing the frequency characteristics of the volume of inflow at Mica Dam. The traditional PMF should be compared to the new method of developing the PMF, in terms of the probability of exceedance for the inflow volume.

Finally, other methods to test and/or validate the findings in this research should be performed. For instance, paleohydrology would help determine the largest historic flood.

9 References

Abbs, Deborah J., 1999, "A Numerical Modeling Study to Investigate the Assumptions used in the Calculation of Probable Maximum Precipitation," *Water Resources Research*, Vol. 35(3), pp. 785-796.

Adamowski, Kaz, Younes, Alila, and Pilon, Paul J., 1996, "Regional rainfall distribution for Canada," *Atmospheric Research*, Vol. 42, pp. 75-88.

American Society of Civil Engineers, 1969, "Design and Construction of Sanitary and Storm Sewers," *Manuals and Reports of Engineering Practice*, No. 37, New York.

Anderson, E.A., 1976, *A Point Energy and Mass Balance Model of a Snow Cover*, NOAA Technical Report, NWS 19, 150 pages.

Benoit, Robert, Desgagné, Michel, Pellerin, Pierre, Pellerin, Simon, Chartier, Yves, and Desjardins, Serge, 1997a, "The Canadian MC2: A Semi-Lagrangian, Semi-Implicit Wideband Atmospheric Model Suited for Finescale Process Studies and Simulation," *Monthly Weather Review*, Vol. 125, pp. 2382-2415.

Benoit, Robert, Pellerin, Pierre, Larocque, Yvan, and Desgagné, Michel, April 1997b, *High-Resolution Modelling of Theoretical Meteorological Storm: Probable Maximum Storm (PMS) Model*, A Preliminary Report prepared for BC Hydro, Maintenance, Engineering and Projects, Recherche en Prévision Numérique.

Beven, Keith, 1989, "Changing Ideas in Hydrology – The Case of Physically-Based Models," *Journal of Hydrology*, Vol. 105, pp. 157-172.

Beven, Keith, 1993, "Prophecy, reality and uncertainty in distributed hydrological modelling," *Advances in Water Resources*, Vol. 16, pp. 41-51.

Bingeman, Allyson K., Kouwen, Nicholas, and Soulis, Eric D., 2001, "Validation Experiments with a Distributed, Physically-Based Hydrological Model," Submitted to Water Resources Research, May 11, 2001.

Binley, A. M., Beven, K. J., Calver, A., and Watts, L.G., 1991, "Changing Responses in Hydrology: Assessing the Uncertainty in Physically Based Model Predictions," *Water Resources Research*, Vol. 27, No. 6, pp. 1253-1261.

Boyle, Douglas P., Gupta, Hoshin V., and Sorooshian, Soroosh, 2000, "Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods," *Water Resources Research*, Vol. 36(12), pp. 3663-3674.

Burn, Donald H., Zrinji, Zolt, and Kowalchuk, Michael, 1997, "Regionalization of Catchments for Regional Flood Frequency Analysis," *Journal of Hydrologic Engineering*, Vol. 2(2), pp. 76-82.

Calder, Nicola, October 1999, *EnSim Hydrologic Help Manual – Draft*, Canadian Hydralics Centre, National Research Council: Ottawa, Ontario, 50 pages.

Carlaw, Shari M., 2000, *Soil Moisture Accounting in Distributed Hydrologic Modelling*, M.A.Sc. Thesis, University of Waterloo, Waterloo, 168 pages.

Cameron, D.S., Beven, K.J., Tawn, J., Blazkova, S., and Naden, P., 1999, "Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty)," *Journal of Hydrology*, Vol. 219, 169-187.

Côté, Jean, Gravel, Sylvie, Méthot, André, Patoine, Alain, Roch, Michel, and Staniforth, Andrew, 1998, "The Operational CMC-MRB Global Environmental Multiscale (GEM) Model. Part I: Design Considerations and Formulation and Part II: Results," *Monthly Weather Review*, Vol. 126, pp. 1373-1418.

Cranmer, A, Kouwen, N., and Mousavi, S.F., 2001, "Proving WATFLOOD: Modelling the Non-Linearities of Hydrologic Response to Storm Intensities," *Canadian Journal of Civil Engineering*, Accepted for publication January 18, 2001.

Crosetto, Michele, Tarantola, Stefano, and Saltelli, Andrea, 2000, "Sensitivity and uncertainty analysis in spatial modelling based on GIS," *Agriculture, Ecosystems and Environment*, Vol. 81, pp. 71-79.

Dalrymple, Tate, 1960, "Flood Frequency Analyses," *Water Supply Paper 1543-A*, U. S. Geological Survey, Reston, Virginia.

Danard, Maurice, 1971, "A Simple Method of Computing the Variation of Annual Precipitation over Mountainous Terrain," *Boundary Layer Meteorology*, Vol. 2, pp. 188-206.

Danard, Maurice, Galbraith, Jack, and Jorgenson, Georg, March 1993, *A High-Resolution Variational or Feedback Model for Prediction of Precipitation in Complex Terrain*, Prepared by Atmospheric Dynamics Corp. for B.C. Hydro, Hydrotechnical Department, March 1993.

Danard, Maurice, and Galbraith, Jack, December 1994, *Boundary-Layer Models for Analysing Precipitation, Maximum and Minimum Temperature, and Snow Water Equivalent*, Prepared by Atmospheric Dynamics Corp. for B.C. Hydro, Hydrotechnical Department.

Danard, Maurice and Galbraith, Jack, February 1995, *Distributed PMP's, Prestorms, Extreme Snowpacks, Critical Melt Temperatures, and Background Conditions for the Upper Columbia River Watershed*, Prepared by Atmospheric Dynamics Corporation for the BCHydro Hydrotechnical Department, Hydroelectric Engineering Division.

Danard, Maurice, Galbraith, Jack, and Davies, Keith, January 1996a, *High Resolution Analyses of Precipitation, Maximum and Minimum Temperatures, and Snow Water Equivalent Using Boundary-Layer Models*, Prepared by Atmospheric Dynamics Corp. for B.C. Hydro, Hydrotechnical Department.

Danard, Maurice, and Galbraith, Jack, February 1996b, *High Resolution Analyses of Daily Precipitation Amounts and Temperatures in the Upper Columbia Basin from 1899 to 1994*, Prepared by Atmospheric Dynamics Corp. for B.C. Hydro, Hydrotechnical Department.

Danard, Maurice, and Galbraith, Jack, September 1997, *High Resolution Analyses of Daily Precipitation Amounts and Temperatures in Southwestern British Columbia 31 Oct 1971 to 29 June 1995*, Prepared by Atmospheric Dynamics Corp. for B.C. Hydro, Hydrotechnical Department.

Daviau, J.-L., Adamowski, K., and Patry, G.G., 2000, "Regional Flood Frequency analysis using GIS, L-moment and geostatistical methods," *Hydrological Processes*, Vol. 14, pp. 2731-2753.

Desjardins, Serge, Benoit, Robert, and Swail, Val, 1998, "The Influence of Mesoscale Features of the Sea Surface Temperature Distribution on Marine Boundary Layer Winds off the Scotia Shelf during the Superstorm of March 1993," *Monthly Weather Review*, Vol. 126, pp. 2793-2808.

Donald, J.R, E.D. Soulis, N. Kouwen and A. Pietroniro, 1995, "A land cover-based snow representation for distributed hydrologic models," *Water Resources Research*, 31(4), pp. 995-1009.

Federal Energy Regulatory Commission, 1993, *Engineering Guidelines for the Evaluation of Hydropower Projects*, Washington, D.C.

Filho, Otto Correa Rotunno, 1995, *Soil Moisture Mapping using Remote Sensing and Geostatistics Applied to Rainfall-Runoff Models*, Ph.D. Thesis, University of Waterloo, Waterloo, 396 pages.

Gal-Chen, T., and Somerville, R. C., 1975, "On the use of a coordinate transformation for the solution Navier-Stokes equations," *Journal of Computation Physics*, Vol. 17(2), pp. 209-228.

Grayson, Rodger B., Moore, Ian D., and McMahon, Thomas A., 1992, "Physically Based Hydrologic Modeling: 2. Is the Concept Realistic?" *Water Resources Research*, Vol. 26(10), pp. 2659-2666.

Green, W.H., and Ampt, G.A., 1911, "Studies in Soil Physics. 1. Flow of Air and Water Through Soils," *Journal of Agricultural Science*, Vol. 4, pp. 1-24.

Greenwood, J. A., Landwehr, J. M. Matalas, N. C. and Wallis, J. R., 1979, "Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form," *Water Resources Research*, Vol. 15(5), pp. 1049-1054.

Gupta, Hoshin Vijay, Sorooshian, Soroosh, Yapo, and Patrice Ogou, 1999, "Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration," *Journal of Hydrologic Engineering*, Vol. 4(2), pp. 135-143.

Gupta, Hoshin Vijay, Sorooshian, Soroosh, Yapo, and Patrice Ogou, 1998, "Toward Improved calibration of hydrologic models: Multiple and noncommensurable measures of information," *Water Resources Research*, Vol. 34(4), pp. 751-763.

Hargreaves, G.H. and Samani, Z.A., 1982, "Estimating Potential Evapotranspiration," *Journal of the Irrigation and Drainage Division*, ASCE, 108 (IR3), pp. 225-230.

Hosking, J.R.M., and Wallis, J.R., 1985, "Estimation of the Generalized Extreme Value Distribution by the Method of Probability-Weighted Moments," *Technometrics*, Vol. 27(3), pp. 251-261.

Hosking, J.R.M., and Wallis, J.R., 1987, "Parameter and Quantile Estimation for the Generalized Pareto Distribution," *Technometrics*, Vol. 29(3), pp. 339-349.

Hosking, J.R.M., and Wallis, J.R., 1997, *Regional Frequency Analysis: An Approach Based on L-Moments*, Cambridge University Press: New York, New York, 224 pages.

Jarrett, R.D., 1990a, "Paleohydrologic techniques used to define the spatial occurrence of floods", *Geomorphology*, Vol. 3, pp. 181-195.

Jarrett, R.D., 1990b, "Hydrologic and Hydraulic Research in Mountain Rivers", *Water Resources Bulletin*, Vol. 26(3), pp. 419-429.

Jarrett, R.D., and Costa, J.E., 1988, *Evaluation of the flood hydrology in Colorado Front Range using Precipitation, Streamflow, and Paleoflood data*, U.S. Geological Survey, WRIR 87-4117, 37 pages.

Jarrett, R.D., and Crow, L.W., 1988, "Experimental Marvin Windshield Effects on Precipitation Records in Leadville, Colorado," *Water Resources Bulletin*, Vol. 24(3), pp. 615-626.

Jarrett, Robert D., and Tomlinson, Edward M., 2000, "Regional interdisciplinary paleoflood approach to assess extreme flood potential," *Water Resources Research*, Vol. 36(10), pp. 2957-2984.

Klemeš, Vit, 2000, "Tall Tales about Tails of Hydrological Distributions: I and II," *Journal of Hydrologic Engineering*, Vol. 5(3), pp. 227-239.

Klemeš, Vit, 1986, "Operational testing of hydrological simulation models," *Journal of Hydrological Sciences*, Vol. 31(1), pp. 13-24.

Kouwen, N., E.D. Soulis, A. Pietroniro, and R.A. Harrington, 1990, "Remote Sensing to Access Vegetative Cover Effects for Flood Forecasting," *International Conference on River Flood Hydraulics 17-20 September, 1990*. Paper M2, pp. 437-446.

Kouwen, N., E.D. Soulis, A. Pietroniro, J. Donald and R.A. Harrington, 1993, "Grouped Response Units for Distributed Hydrologic Modelling," *Journal of Water Resources Planning and Management*, Vol. 119(3), pp. 289-305.

Kouwen, Nicholas, 2000, *WATFLOOD/SPL9: Hydrological Model & Flood Forecasting System*, University of Waterloo: Waterloo, Ontario. Available from http://www.watflood.ca/ .

Kouwen, N., Danard, M., Bingeman, A., Luo, W., Seglenieks, F.R., and Soulis, E.D., 2000, "Watershed Modeling with Numerical Weather Model Data," Submitted to the ASCE Journal of Hydrological Engineering on June 9, 2000.

Kouwen, N., and J. Innes, 2000, *Application of Hydrologic Models for RAPHAEL (Runoff and Atmospheric Processes for Flood Hazard Forecasting and Control): Final Report*, Coordinated by Prof. B. Bacchi, University of Brescia, Department of Civil Engineering, Brescia, Italy, pp. 2.4.84 – 2.4.94.

Kouwen, N. and J. Innes, 2001, "Coupled MC2-WATFLOOD/SPLFlood Forecasting," *Hydrological Aspects in the Mesoscale Alpine Project (MAP) – SOP Experiment*, Edited by R. Ranzi and B. Bacchi, University of Brescia, Department of Civil Engineering, Brescia, Italy, Technical Report Number 10.

Kuczera, George, and Parent, Eric, 1998, "Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm," *Journal of Hydrology*, Vol. 211, pp. 69-85.

Lackmann, Gary M., Gyakum, John R., and Benoit, Robert, 1998, "Moisture Transport Diagnosis of a Wintertime Precipitation Event in the Mackenzie River Basin," *Monthly Weather Review*, Vol. 126, pp. 668-691.

Lamb, Robert, 1999, "Calibration of a conceptual rainfall-runoff model for flood frequency estimation by continuous simulation," *Water Resources Research*, Vol. 35(10), pp. 3103-3114.

Lange, J., Leibundgut, C., Greenbaum, N., and Schick, A.P., 1999, "A noncalibrated rainfall-runoff model for large, arid catchments," *Water Resources Research*, Vol. 35(7), pp. 2161-2172.

Lei, Jian Hua, and Schilling, Wolfgang, 1996, "Preliminary Uncertainty Analysis – A Prerequisite for Assessing the Predictive Uncertainty of Hydrologic Models," *Water, Science and Technology*, Vol. 33(2), pp. 79-90.

Lei, J., and Schilling, W., 1994, "Parameter Uncertainty Propagation Analysis for Urban Rainfall Runoff Modeling," *Water, Science and Technology*, Vol. 29(1-2), pp. 145-154.

Linsley, R.K. Jr., Kohler, M.A., and Paulhus, J.L.H., 1949, *Applied Hydrology*. McGraw-Hill: New York, New York, 689 pages.

Mailhot, A., Gaume, É., and Villeneuve, J. P., 1997, "Uncertainty Analysis of Calibrated Parameter Values of an Urban Storm Water Quality Model Using Metropolis Monte Carlo Algorithm," *Water, Science, and Technology*, Vol. 36(5), pp. 141-148.

Mailhot, Jocelyn, Bélair, Stéphane, Benoit, Robert, Bilodeau, Bernard, Delage, Yves, Fillion, Luc, Garand, Louis, Girard, Claude, and Tremblay, André, May 1998, *Scientific Description of RPN Physics Library, Version 3.6*, Recherche en Prévision Numérique, Atmospheric Environment Service, Dorval, Québec. Accessible from http://www.cmc.ec.gc.ca/rpn.

Mailhot, J., Sarrazin, T., Bilodeau, B., Brunet, N., and Pellerin, G., 1997, "Development of the 35-km Version of the Canadian Regional Forecast System," *Atmosphere-Ocean*, Vol. 35(1), pp. 1-28.

Miller, John F., March 1993, *Probable Maximum Precipitation Estimates for Columbia River Basin above the Lower Border Dam Site below confluence with the Pend D'Oreille River*, Prepared for the BCHydro Hydrotechnical Department, Hydroelectric Engineering Division, Report No. H2719.

Mimikou, M.A., Baltas, E., Varanou, E., and Pantazis, K., 2000, "Regional impacts of climate change on water resources quantity and quality indicators," *Journal of Hydrology*, Vol. 234, pp. 95-109.

Mousavi, Sayed-Farhad, and Nicholas Kouwen, 2000, Substituting MODFLOW by a Simple Lower-Zone Function in Hydrological Modelling of a Small Watershed, *Journal of Hydrology*, Submitted in August 2000.

Mroczkowski, Marek, G. Paul Raper and George Kuczera, 1997, "The quest for more powerful validation of conceptual catchment models," *Water Resources Research*, Vol. 33(10), pp. 2325-2335.

National Research Council (NRC), 1988, *Estimating Probabilities of Extreme Floods, Methods and Recommended Research*, National Academy: Washington D.C., 141 pages.

Neff, T.A.M., 1996, *Mesoscale Water Balance of the Boreal Forest Using Operational Evapotranspiration Approaches in a Distributed Hydrologic Model*, M.A.Sc. Thesis, University of Waterloo, Waterloo, 235 pages.

Nuss, Wendell A., and Anthes, Richard, A., 1987, "A Numerical Investigation of Low-Level Processes in Rapid Cyclogenesis," *Monthly Weather Review*, Vol. 115, pp. 2728-2743.

Pao, Yoh-Han, 1989, *Adaptive pattern recognition and neural networks*, Addison-Wesley: Don Mills, Ontario, 309 pages.

Pellerin, Pierre, 2000, Meteorologist, Recherche en Prévision Numérique, personal communication.

Philip, J.R., 1954, "An Infiltration Equation with Physical Significance," *Soil Science*, Vol. 77, pp. 153-157.

Pilon, Paul J., and Adamowski, K., 1992, "The value of regional information to flood frequency analysis using the method of L-moments," *Canadian Journal of Civil Engineering*, Vol. 19, pp. 137-147.

Pinty, Jean-Pierre, Benoit, Robert, Richard Evelyne, and Laprise, René, 1995, "Simple Tests of a Semi-Implicit Semi-Lagrangian Model on 2D Mountain Wave Problems," *Monthly Weather Review*, Vol. 123, pp. 3042-3058.

Priestley, C.H.B. and R.J. Taylor, 1972, "On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters," *Monthly Weather Review*, Vol. 100(2), pp. 81-92.

Pruess, J., Wohl, E.E., and Jarrett, R.D., 1998, "Methodology and Implications of Maximum Paleodischarge Estimates for Mountain Channels, Upper Animax Basin, Colorado, USA," *Arctic and Alpine Research*, Vol. 30(1), pp. 40-50.

Recherche en Prévision Numérique, 1997, *PMS_model user's manual*, Prepared for BCHydro Company, Hydroelectric Engineering Division by Recherche en Prévision Numérique, Atmospheric Environment Service, Dorval, Québec.

Refsgaard, Jens Christian, and Knudsen, Jesper, 1996, "Operational validation and intercomparison of different types of hydrological models," *Water Resources Research*, Vol. 32(7), pp. 2189-2202.

Refsgaard, Jens Christian, 1997, "Parameterisation, calibration and validation of distributed hydrological models," *Journal of Hydrology*, Vol. 198, pp. 69-97.

Schaefer, M.G., 1990, "Regional analysis of precipitation annual maxima in Washington State," *Water Resources Research*, Vol. 26(1), pp. 119-131.

Schaefer, M.G., (1997) *Magnitude-Frequency Characteristics of Precipitation Annual Maxima in Southern British Columbia*, Prepared for B.C. Hydro, Hydrotechnical Department. 48 pages.

Seglenieks, F.R., Soulis, E.D., Snelgrove, K.R., Kouwen, N., Lee, M., and Solomon, S.I., 1998, "Integrated Hydrologic Modelling of the Mackenzie River Basin for MAGS," paper presented at the Canadian Geophysical Union Conference, Quebec City, Quebec, May 18-20, 1998, p. A-169.

Smith, C.D., 1998, "The PMF Does Have a Frequency," *Canadian Water Resources Journal*, Vol. 23(1), pp. 1-8.

Solomon, S.I., Denouvilliez, J.P., Chart, E.J., Woolley, J.A., and Cadou, C., 1968, "The use of a square grid system for computer estimation of precipitation, temperature and runoff," *Water Resources Research*, Vol. 4(5), pp. 919-929.

Tao, T. and N. Kouwen, 1989, "Remote Sensing and Fully Distributed Modelling for Flood Forecasting," *Journal of Water Resources Planning and Management*, Vol. 115(6), pp. 809-823.

Thyer, Mark, Kuczera, George and Bates, Bryson C., 1999, "Probabilistic optimization for conceptual rainfall-runoff models: A comparison of the shuffled complex evolution and simulated annealing algorithms," *Water Resources Research*, Vol. 35(3), pp. 767-773.

Vertessy, Robert A., and Elsenbeer, Helmut, 1999, "Distributed modeling of storm flow generation in an Amazonian rain forest catchment: Effects of model parameterization," *Water Resources Research*, Vol. 35(7), pp. 2173-2187.

Viessman, Warren Jr., Lewis, Gary L., and Knapp, John W., 1989, *Introduction to Hydrology: Third Edition*, Harper Collins Publishers: New York, New York, 778 pages.

Wong, Janet, 2000, personal communication, B. C. Hydro, 6911 Southpoint Dr., Burnaby, B.C., V3N 4X8, janet.wong@bchydro.bc.ca.

World Meteorological Organization, 1986, *Manual for Estimation of Probable Maximum Precipitation: 2nd Edition*, Operational Hydrology Report No. 1, WMO No. 332, Geneva, Switzerland, 269 pages.

Yevjevich, Vujica, 1972, *Probability and Statistics in Hydrology*, Water Resources Publications: Fort Collins, Colorado, 302 pages.

Yu, Wei, Lin, Charles A., Benoit, Robert, and Zawadzki, Isztar, 1998, "High Resolution Model Simulation of Precipitation and Evaluation with Doppler Radar Observation," *Water, Science, and Technology*, Vol. 37(11), pp. 179-186.

Yu, Z., Lakhtakia, M.N., Yarnal, B., White, R.A., Miller, D.A., Frakes, B., Barron, E.J., Duffy, C., and Schwartz, F.W., 1999, "Simulating the river-basin response to atmospheric forcing by linking a mesoscale meteorological model and hydrologic model system," *Journal of Hydrology*, Vol. 218, pp. 72-91.

Appendix A

This appendix contains a time series of geopotential height plots, obtained from MC2-PMS. (This storm was calculated at the original location, with parameters ax=15, bx=15, s=10.) The plots were obtained every three hours for the first 48 hours of the storm simulation. The first plot, at 0 hours, shows the pressure wave of the perturbation. The fifth plot, at 12 hours, shows a developed cyclone with a depression of 32 mb, just to the west of Vancouver Island. The storm dissipates over the next 18-24 hours. The low pressure "arm" of the storm passes over the Columbia River Basin between hours 6 and 30. This corresponds to the time period where the most rainfall occurred. (Note: since three hours at the beginning of the simulation are lost with each "cascade", the 24 hours of precipitation referred to in the text corresponds to hours 6 to 30 in this time series of plots.)

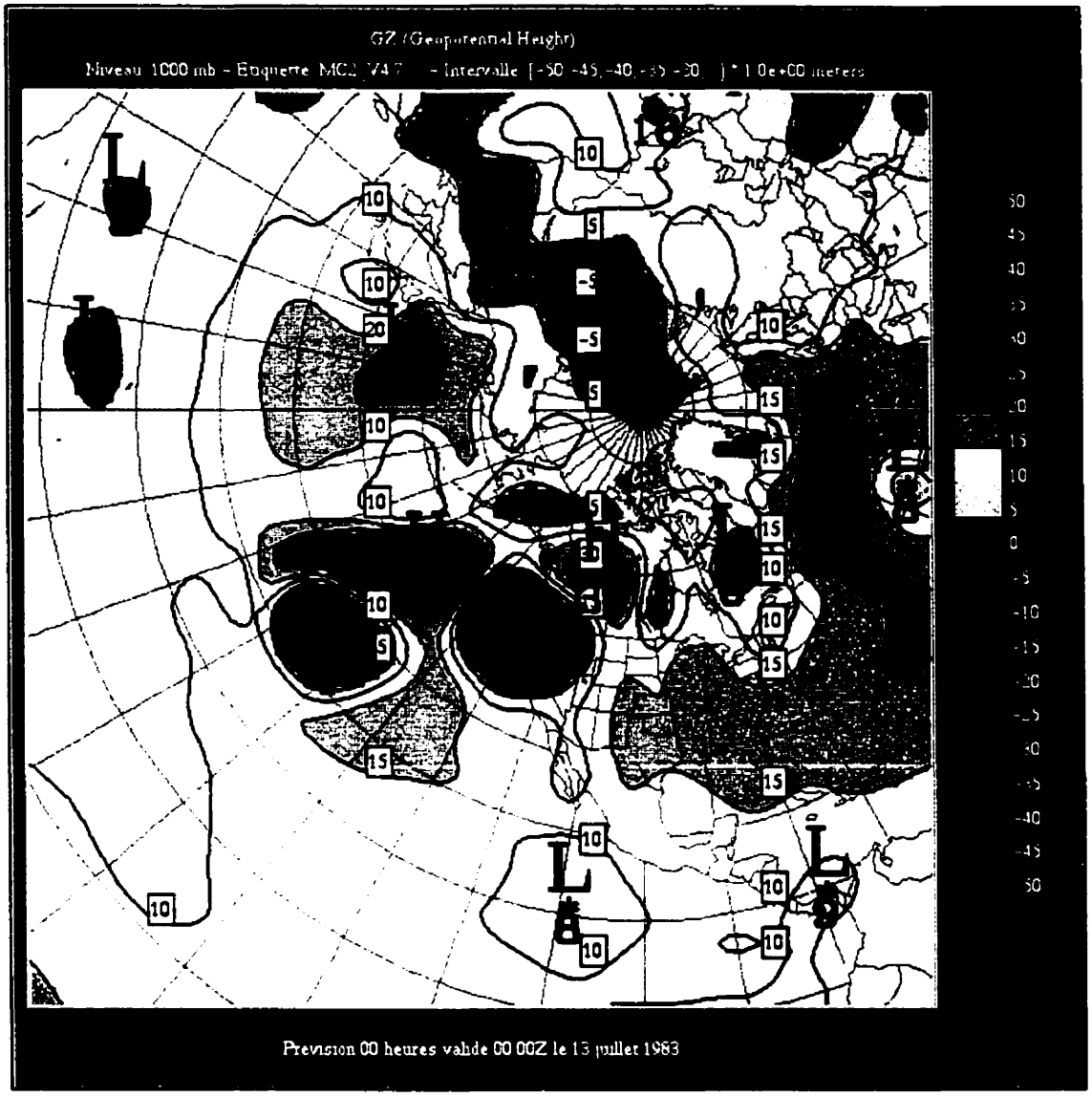


Figure A-1 – Geopotential Height (m) at Hour 0

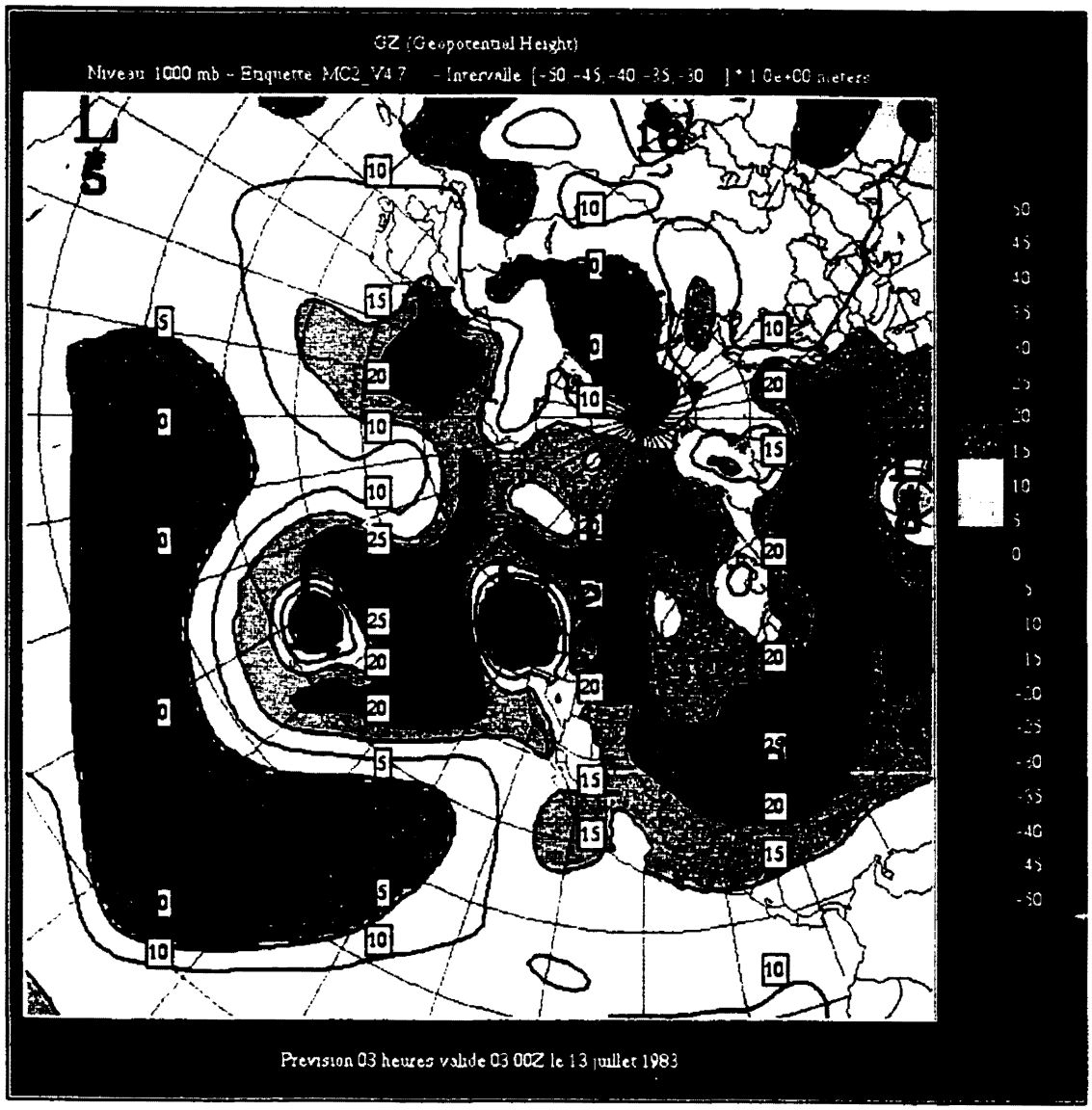


Figure A-2 – Geopotential Height (m) at Hour 3

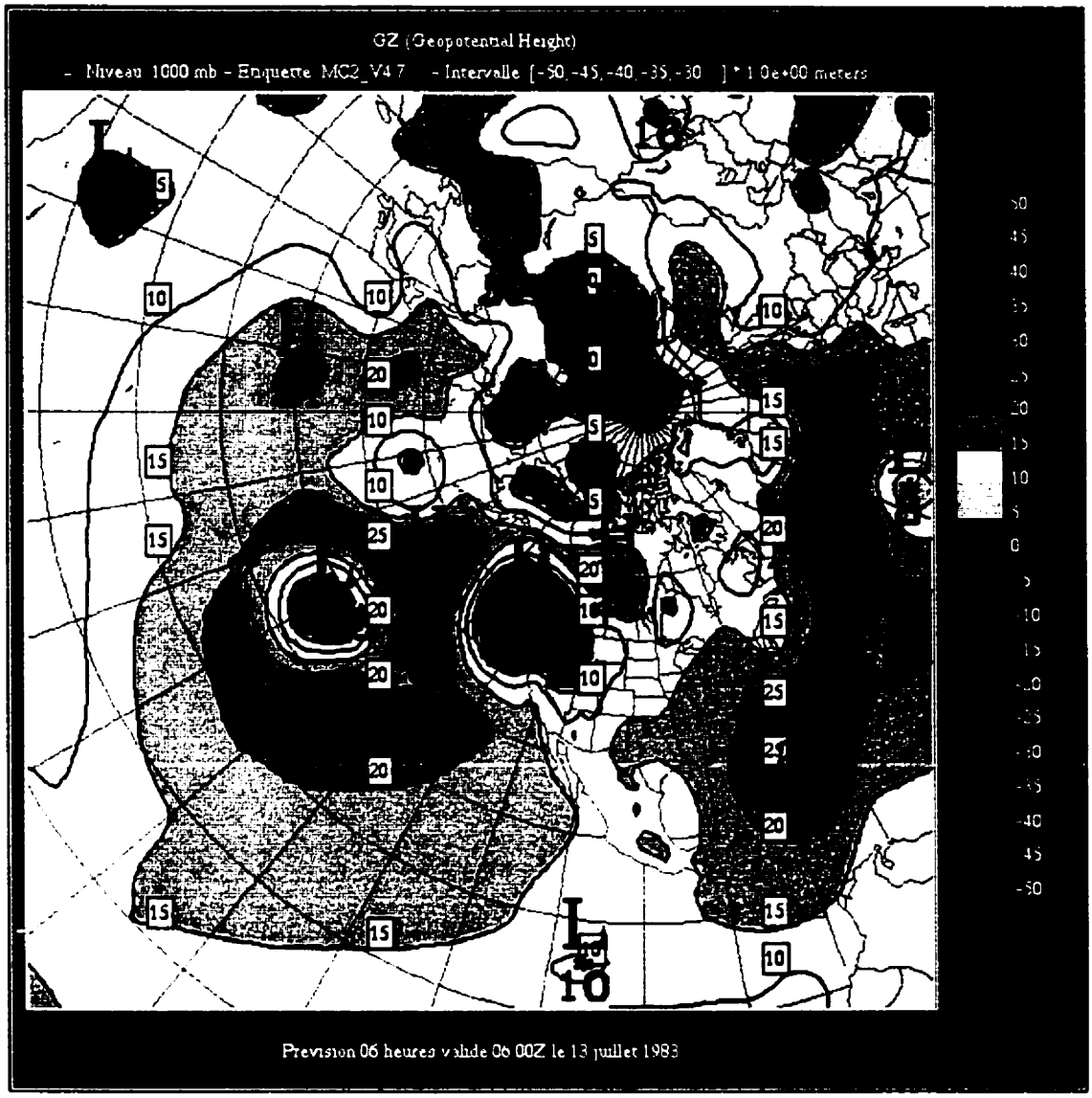


Figure A-3 – Geopotential Height (m) at Hour 6

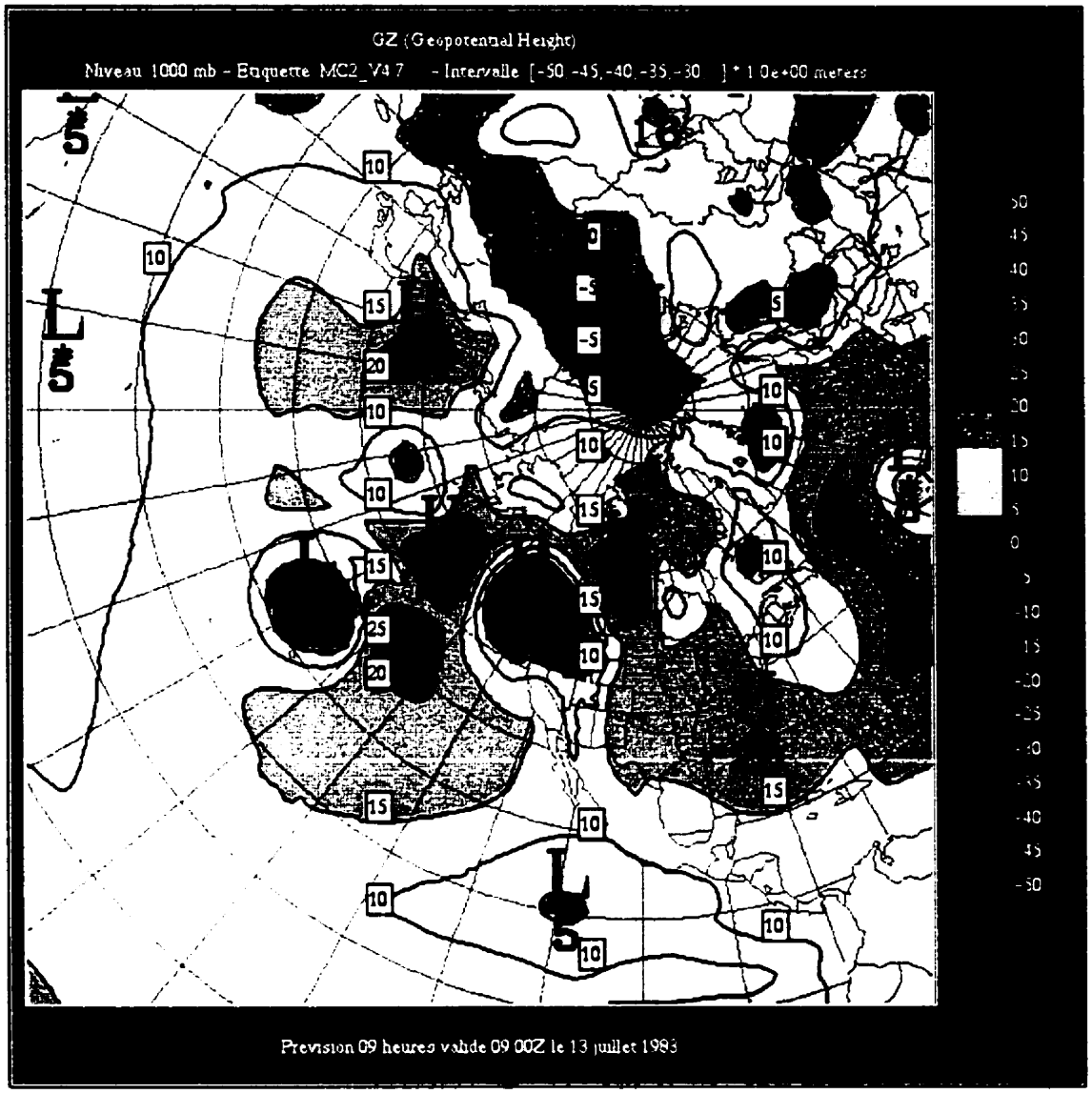


Figure A-4 – Geopotential Height (m) at Hour 9

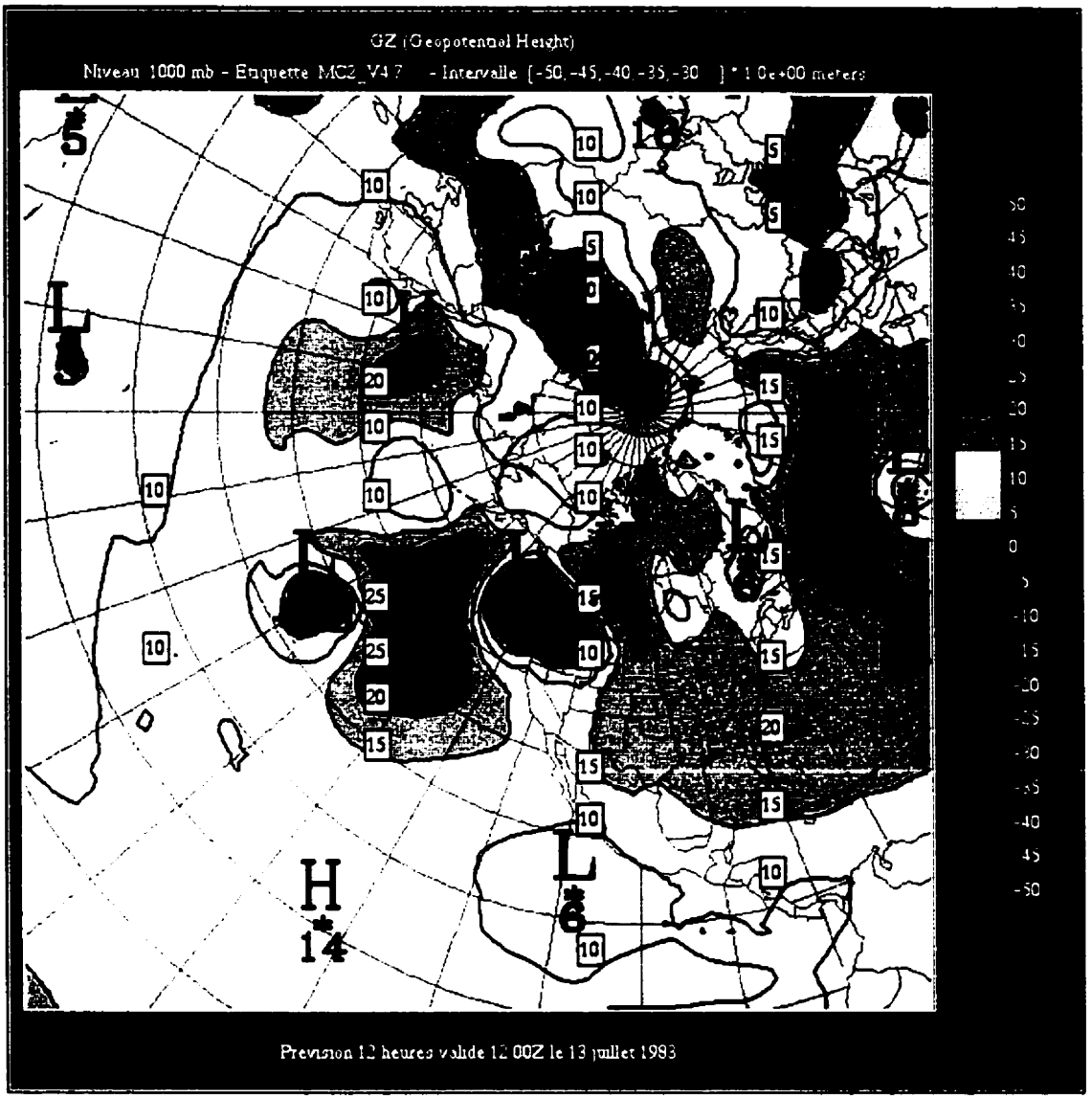


Figure A-5 – Geopotential Height (m) at Hour 12

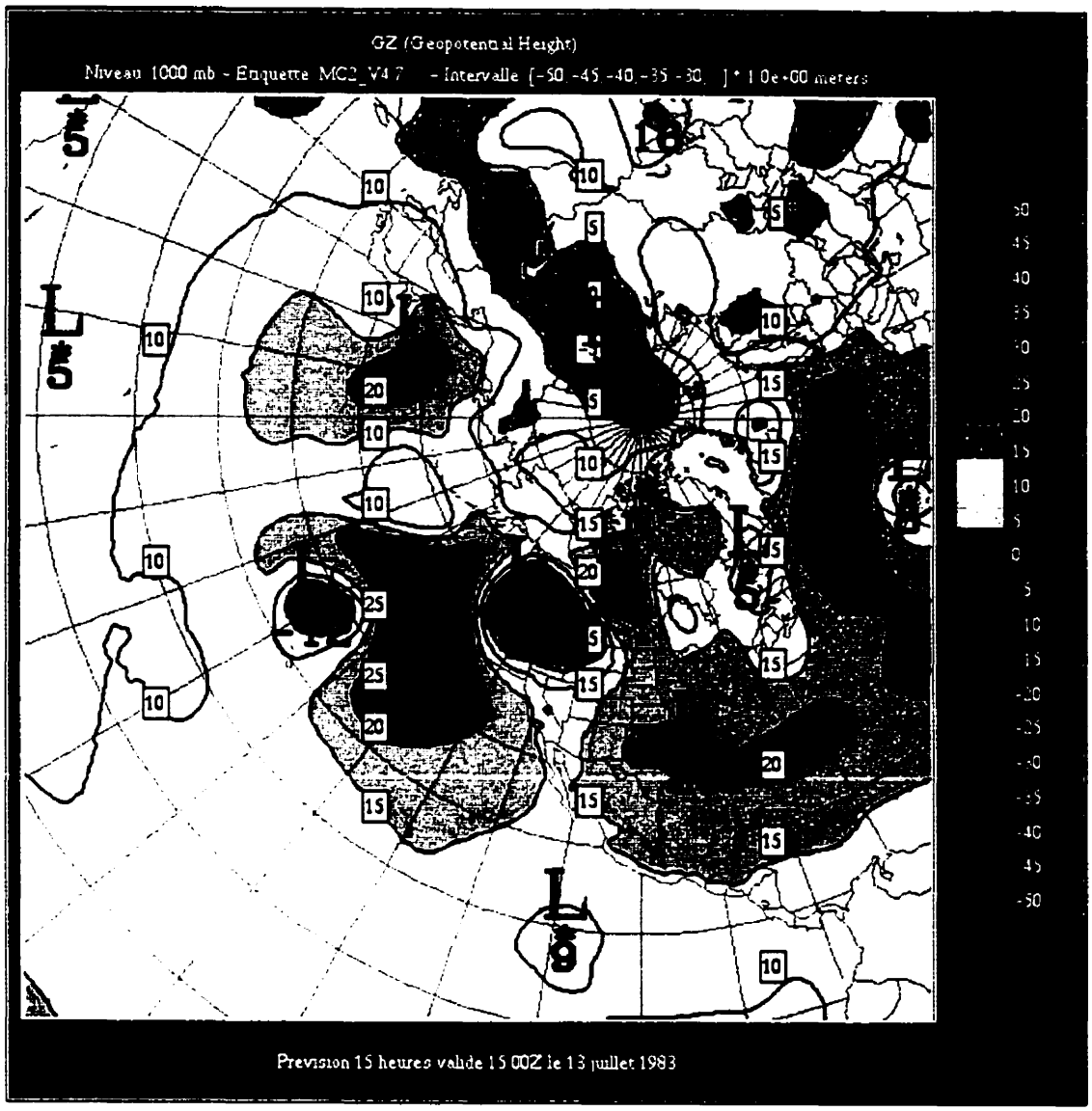


Figure A-6 – Geopotential Height (m) at Hour 15

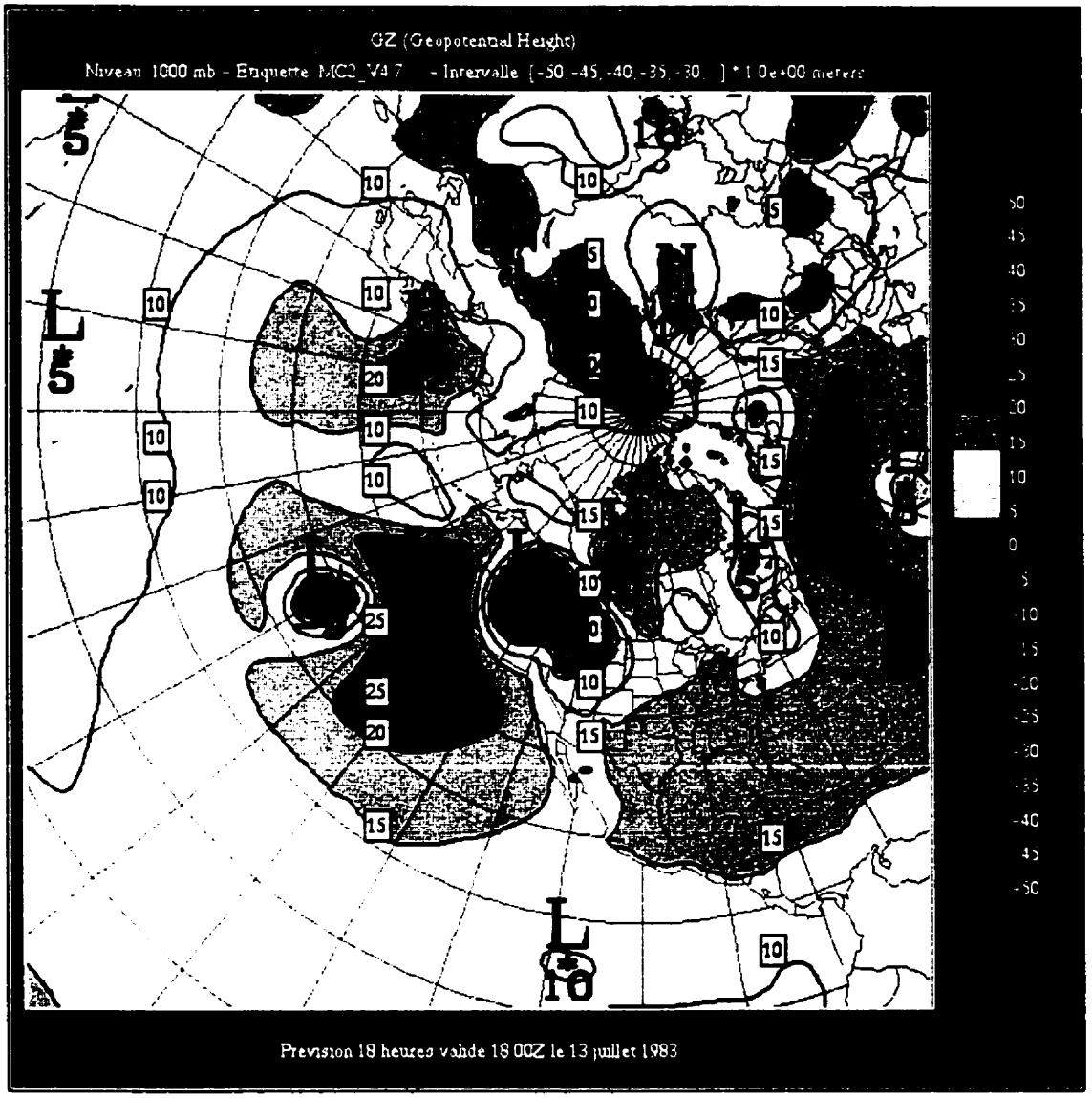


Figure A-7 – Geopotential Height (m) at Hour 18

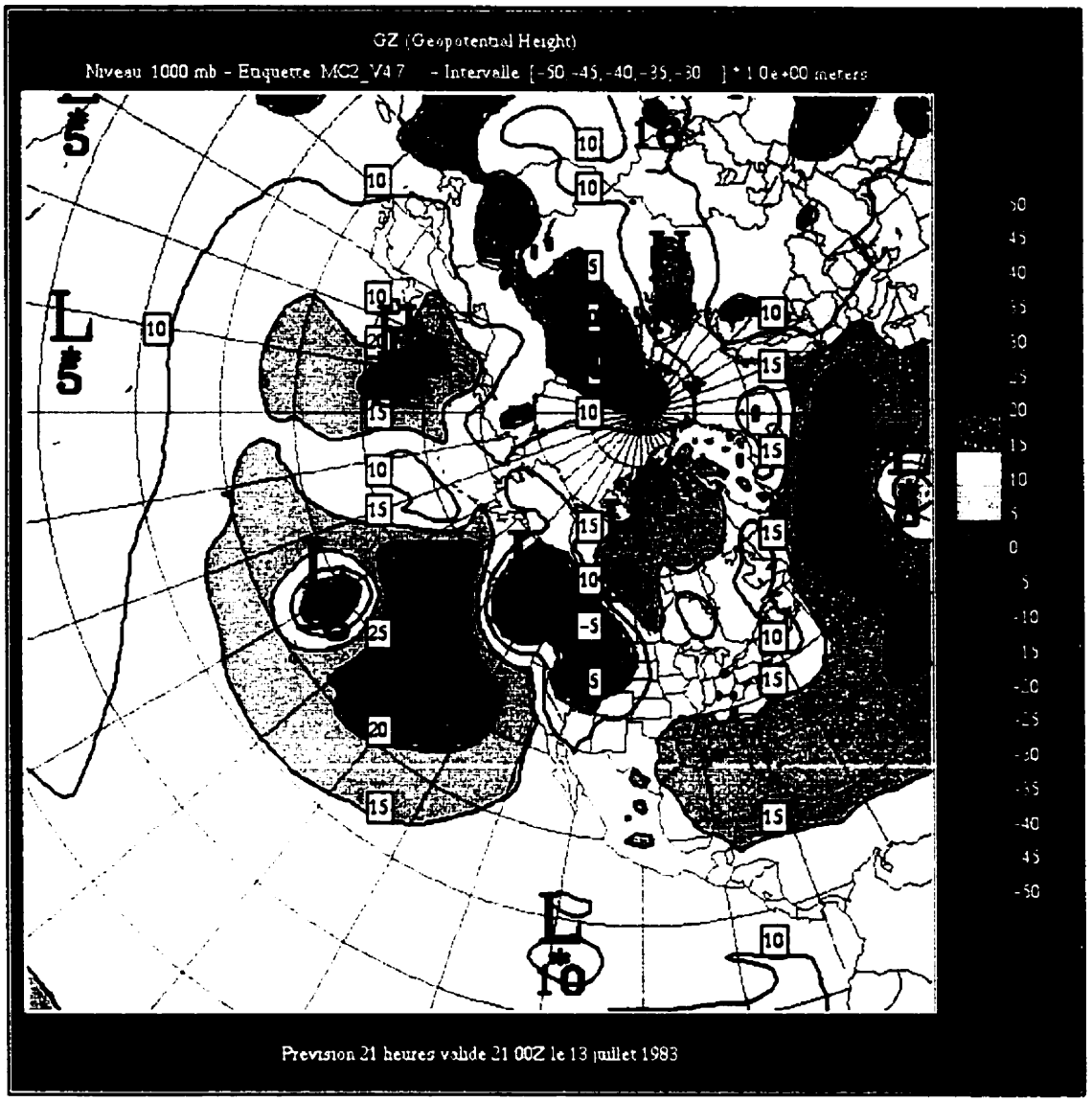


Figure A-8 – Geopotential Height (m) at Hour 21

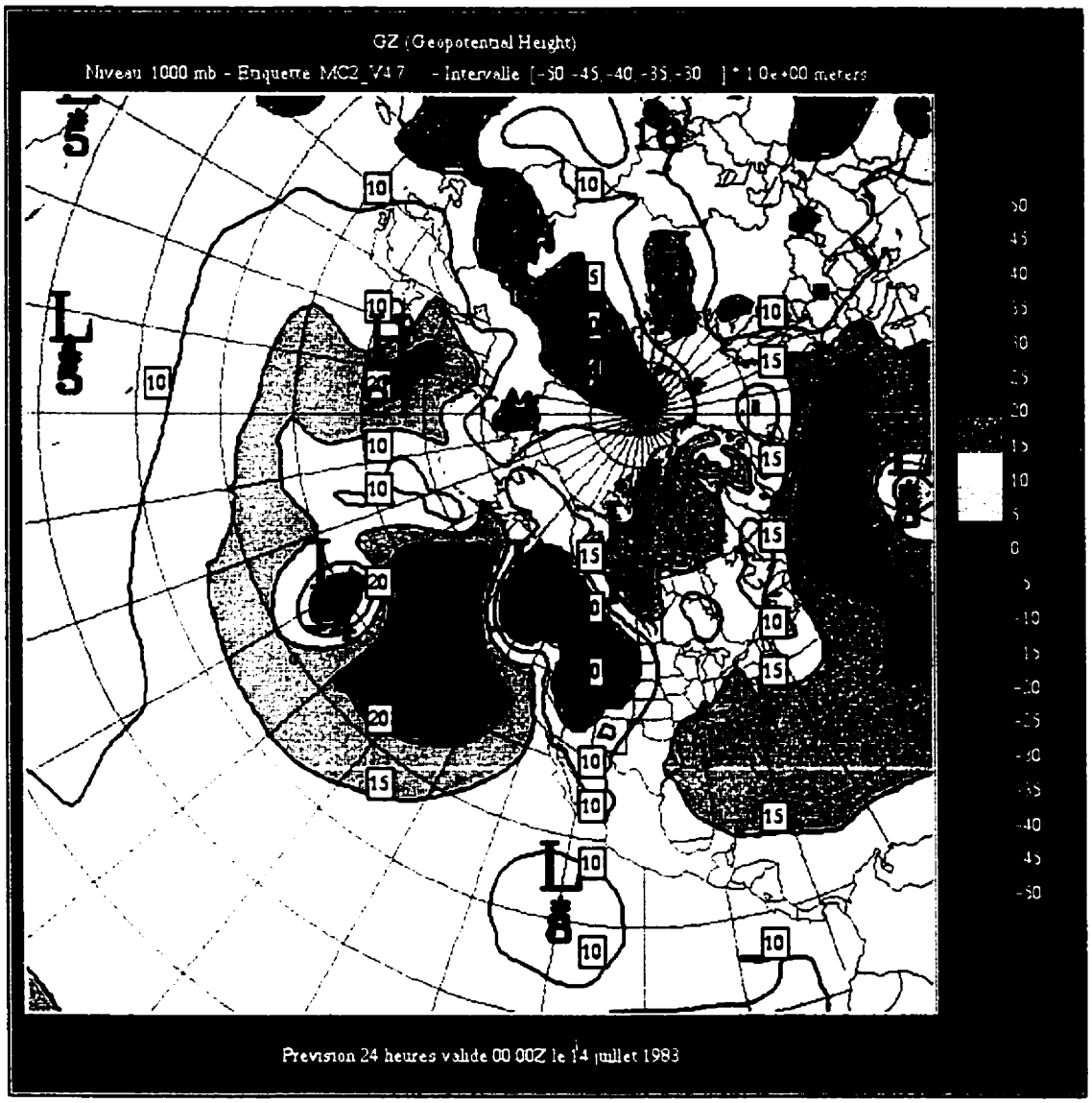


Figure A-9 – Geopotential Height (m) at Hour 24

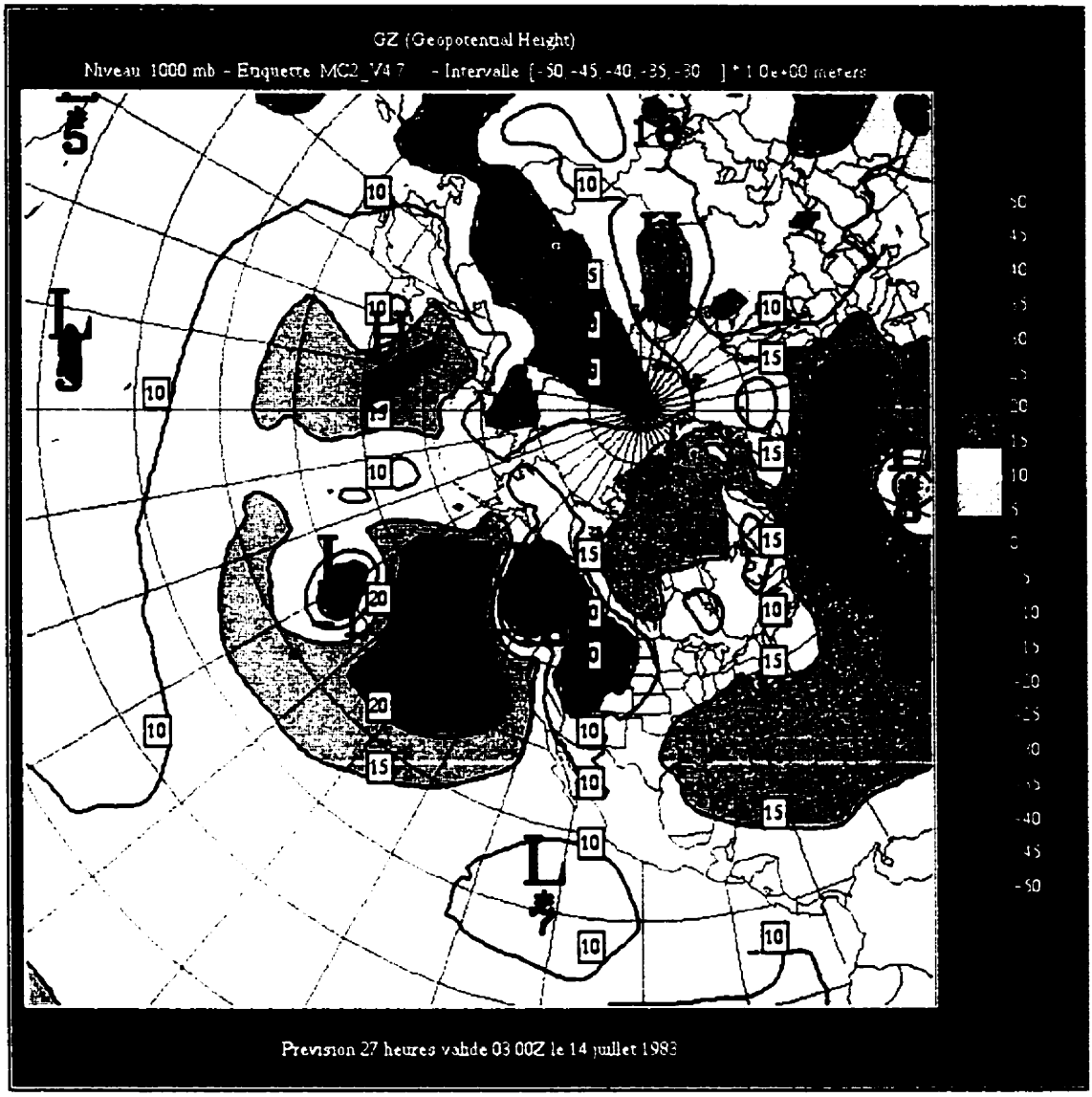


Figure A-10 – Geopotential Height (m) at Hour 27

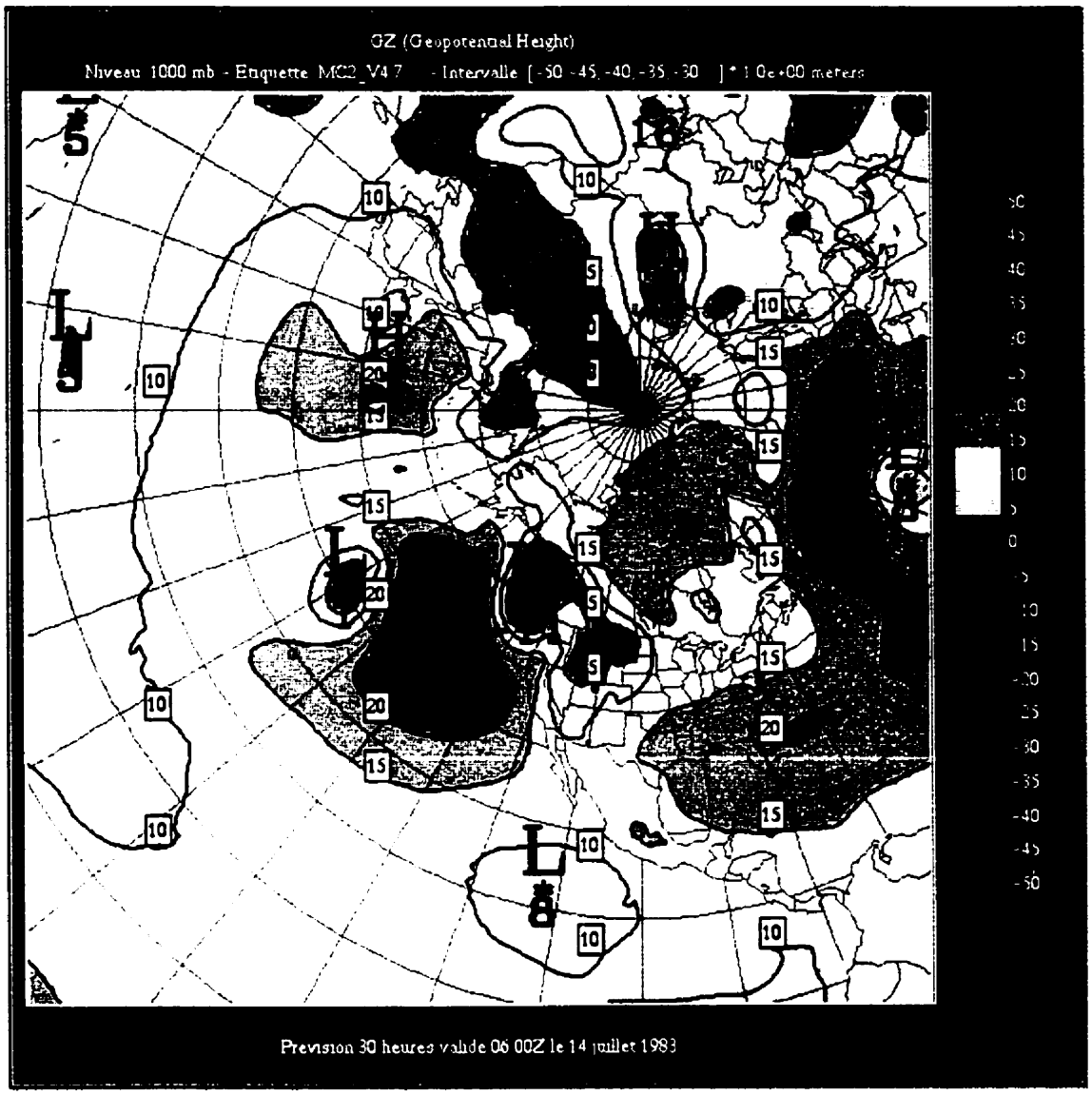


Figure A-11 – Geopotential Height (m) at Hour 30

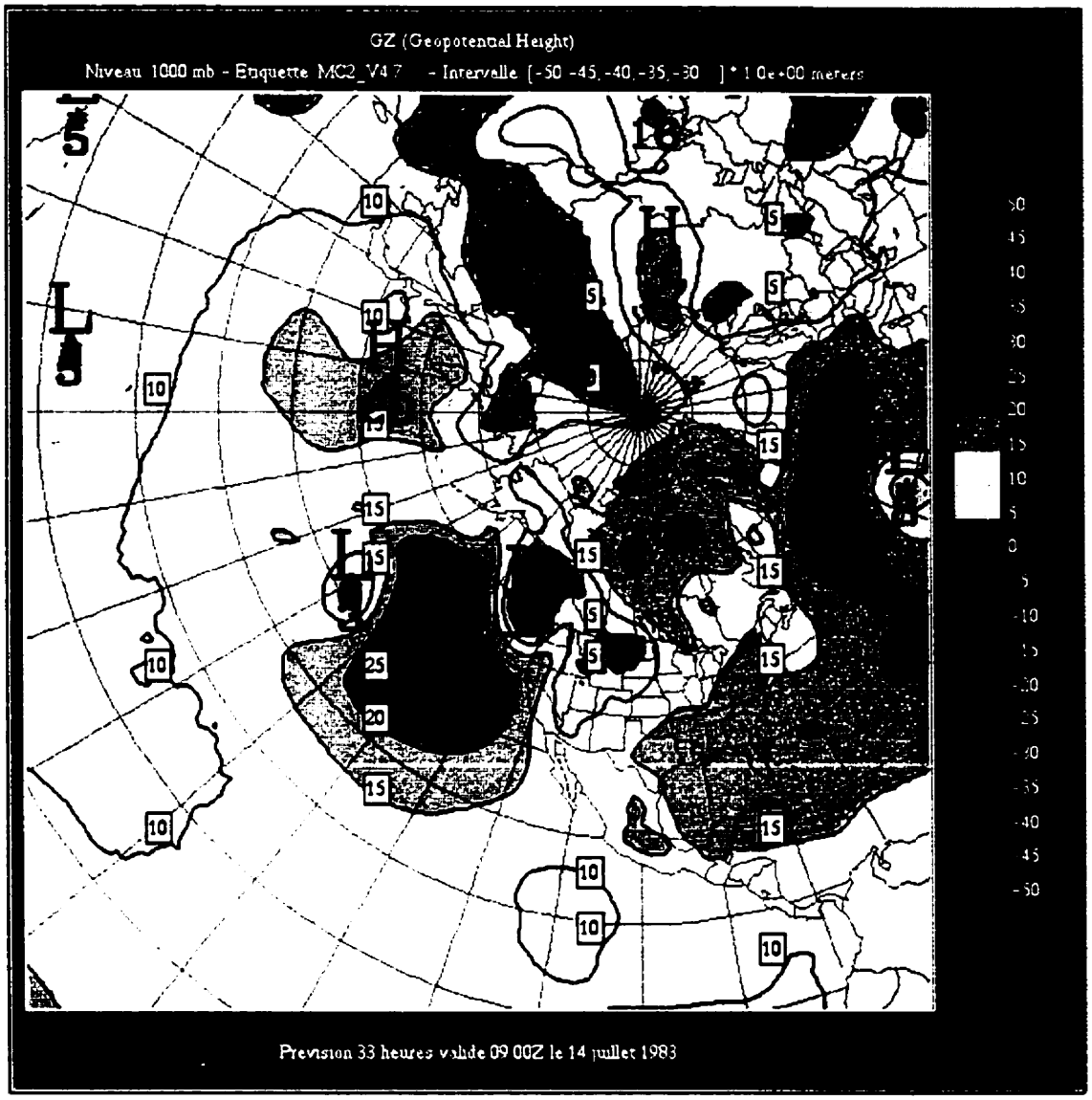


Figure A-12 – Geopotential Height (m) at Hour 33

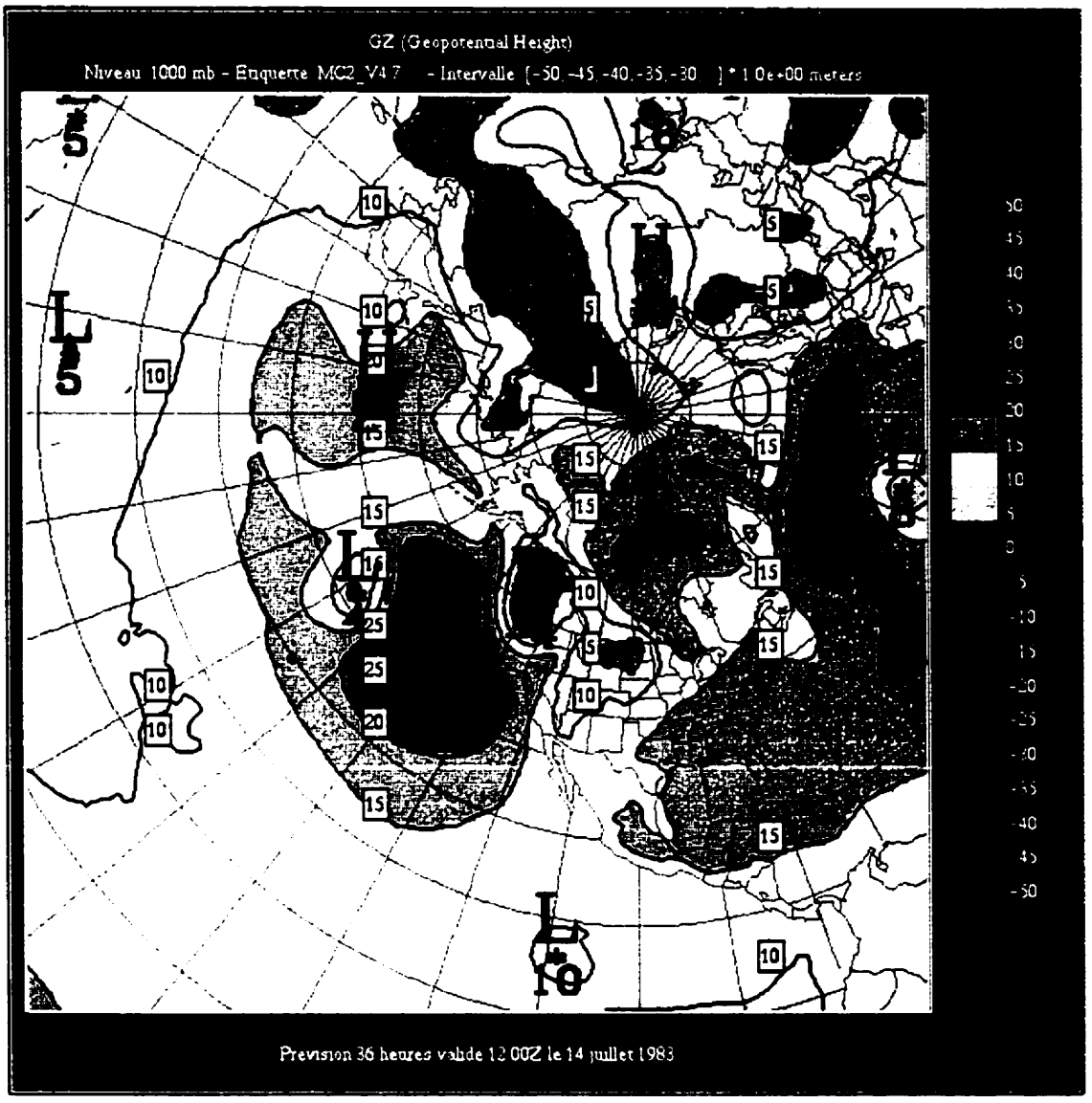


Figure A-13 – Geopotential Height (m) at Hour 36

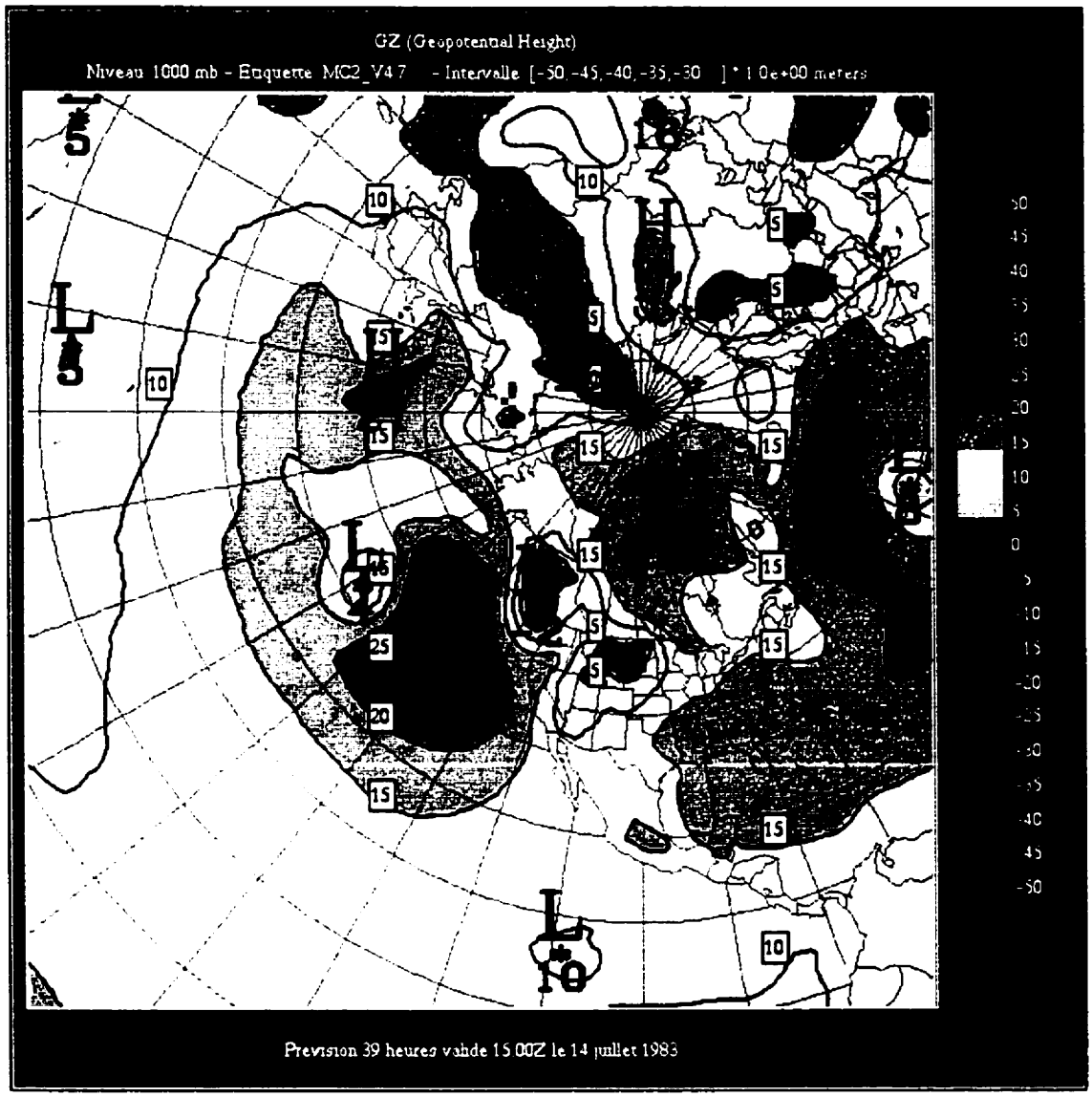


Figure A-14 – Geopotential Height (m) at Hour 39

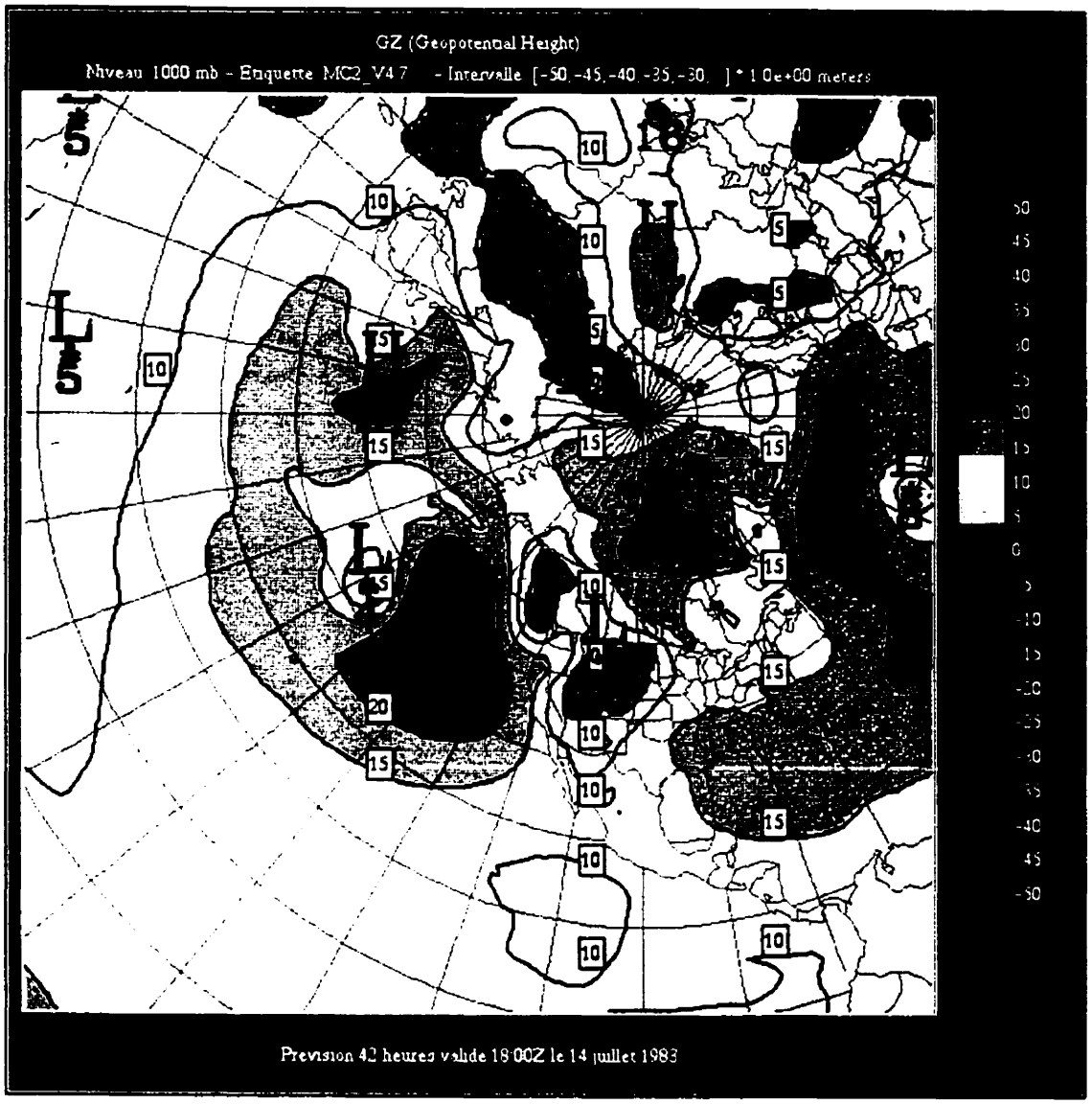


Figure A-15 – Geopotential Height (m) at Hour 42

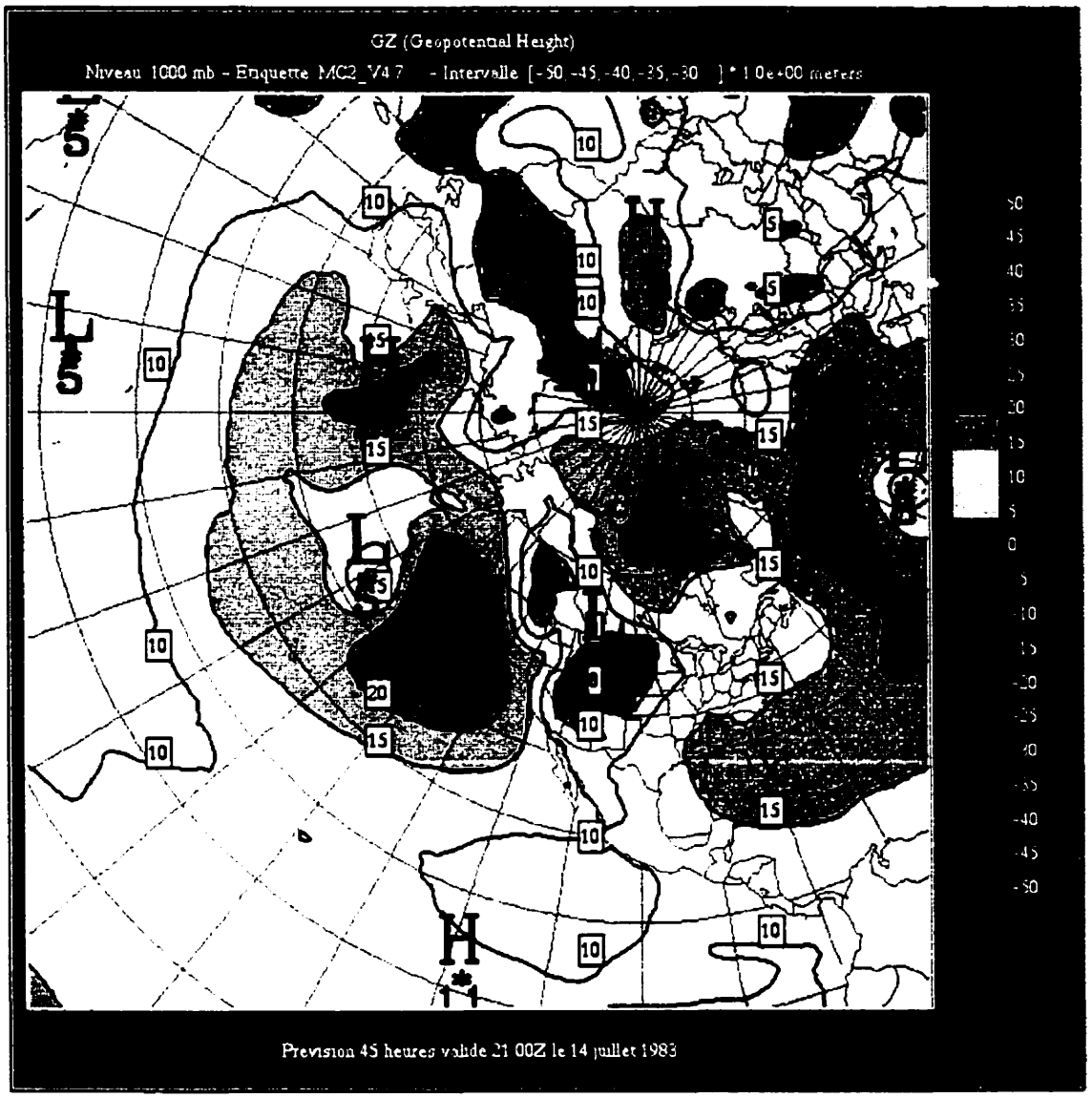


Figure A-16 – Geopotential Height (m) at Hour 45

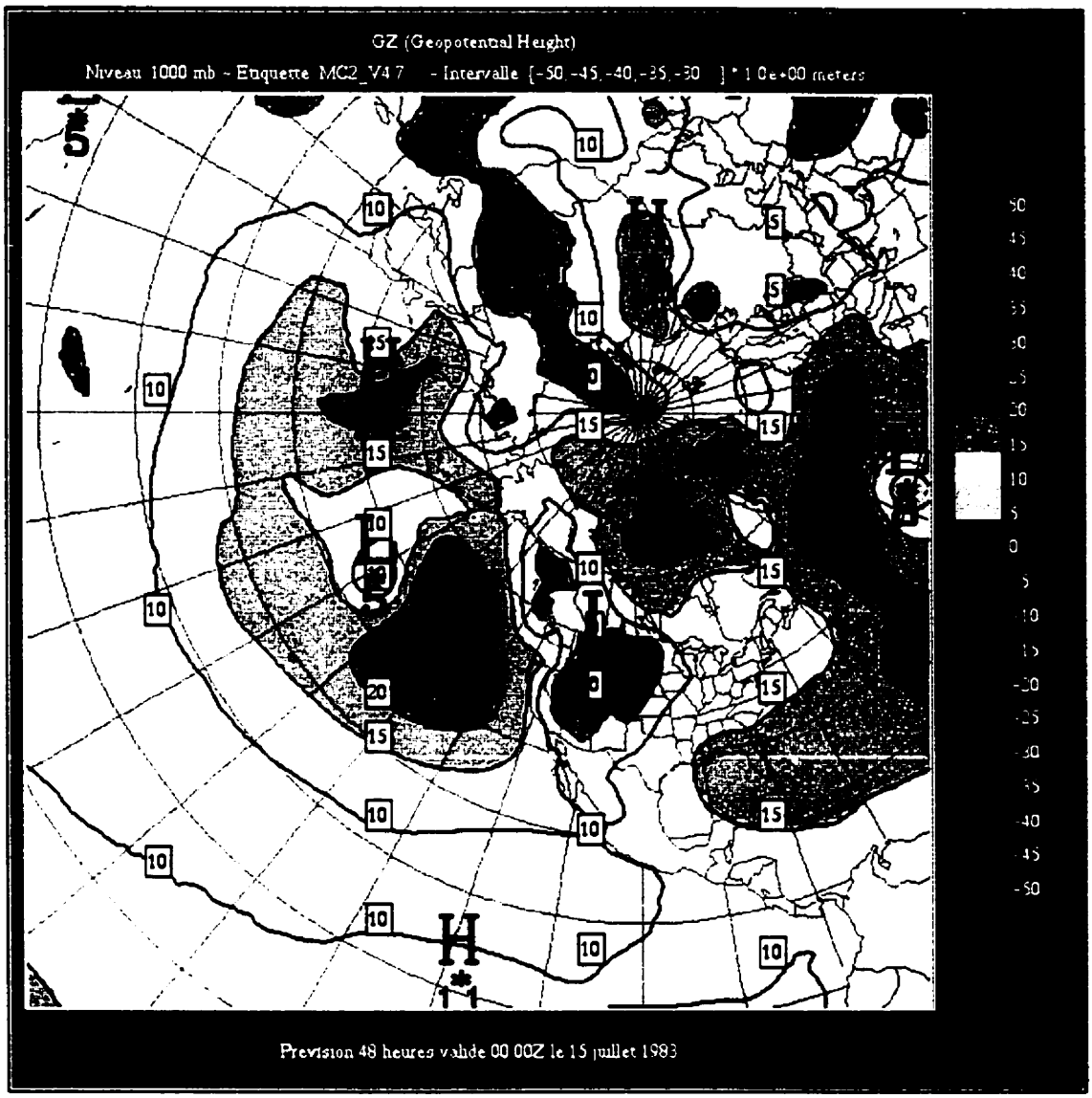


Figure A-17 – Geopotential Height (m) at Hour 48

Appendix B

This appendix contains a series of Figures that compare the frequency curves calculated from observed data and frequency curves calculated from simulated data (shortened to match the observed time series). Each figure is labeled with the station name and number, and the number of years of data for the curves. A frequency curve was calculated for each individual station, using the L-moments for that station, and the Wakeby distribution. The line marked "Observed" refers to the frequency curve calculated from the observed streamflow time series (from WSC or B.C. Hydro). The lines marked "Simulated, short series" and "Simulated with paf, short series" refer to the simulated streamflow time series calculated by WATFLOOD/SPL (with the HRBL data as forcing data), without and with the Precipitation Adjustment Factor (PAF) field correction, respectively. The simulated streamflow time series was shortened to match with the observed time series, so that the same years were used in both analyses.

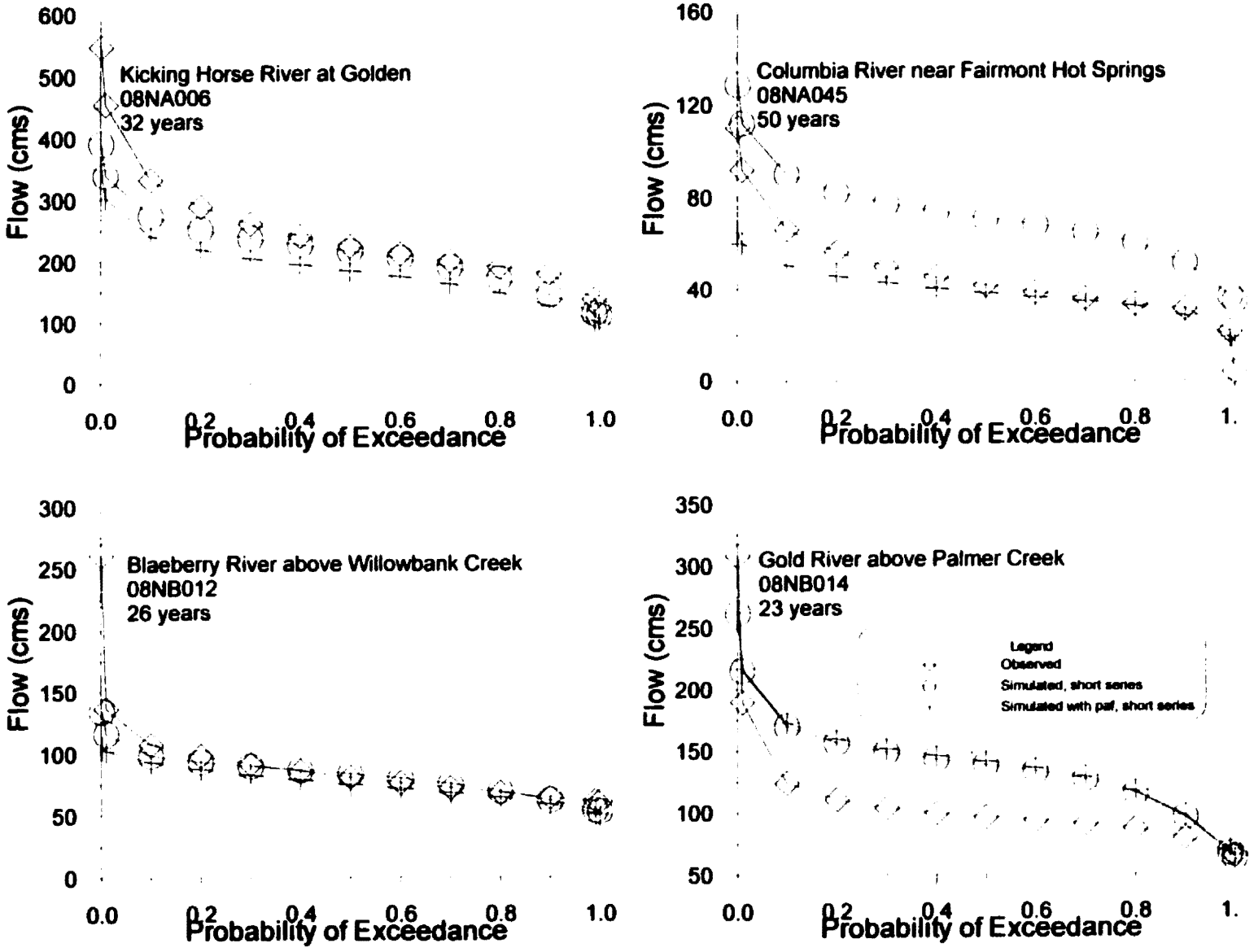


Figure B-1 – Comparison of Observed and Simulated (with and without PAF) Frequency Curves, Part 1

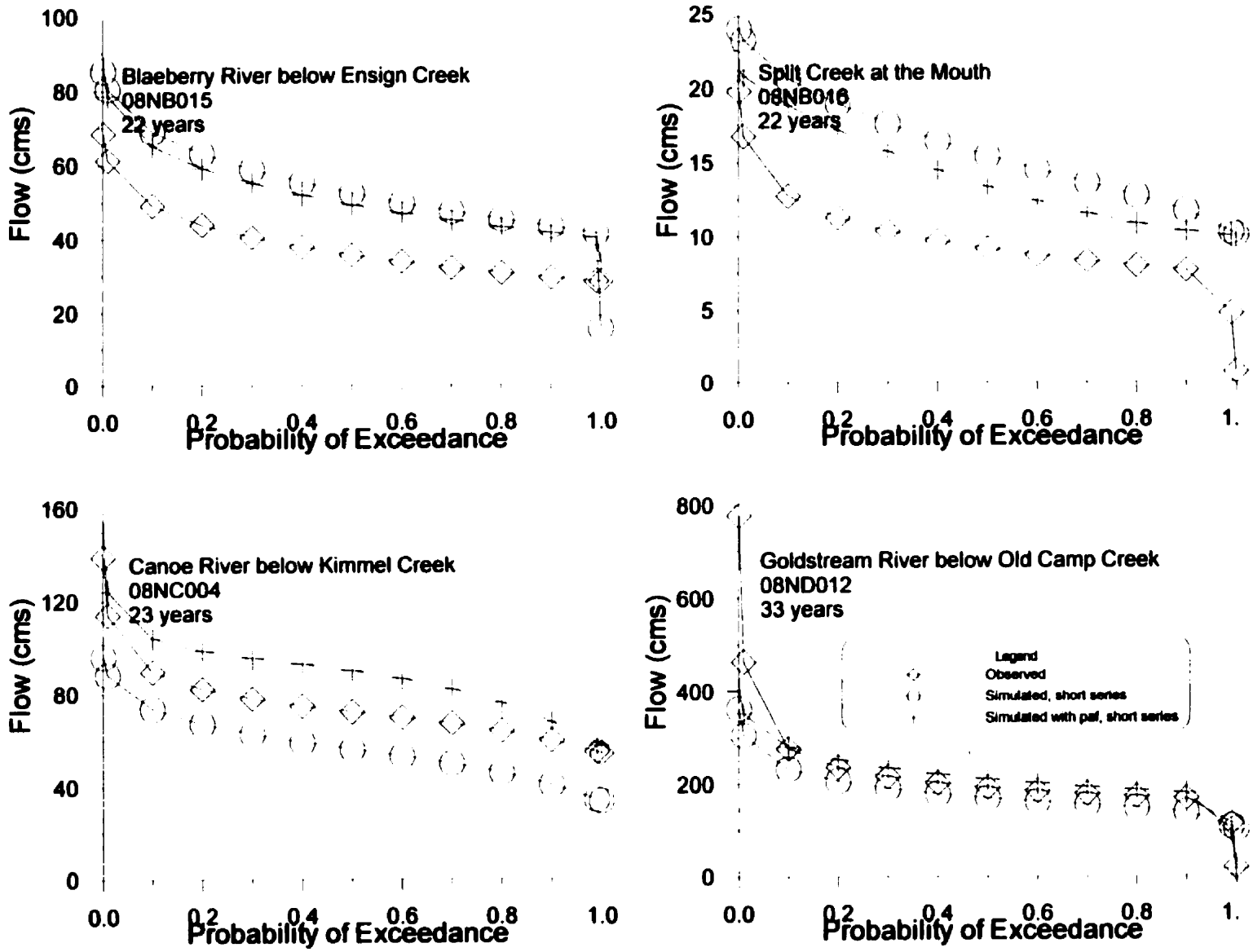


Figure B-2 – Comparison of Observed and Simulated (with and without PAF) Frequency Curves, Part 2

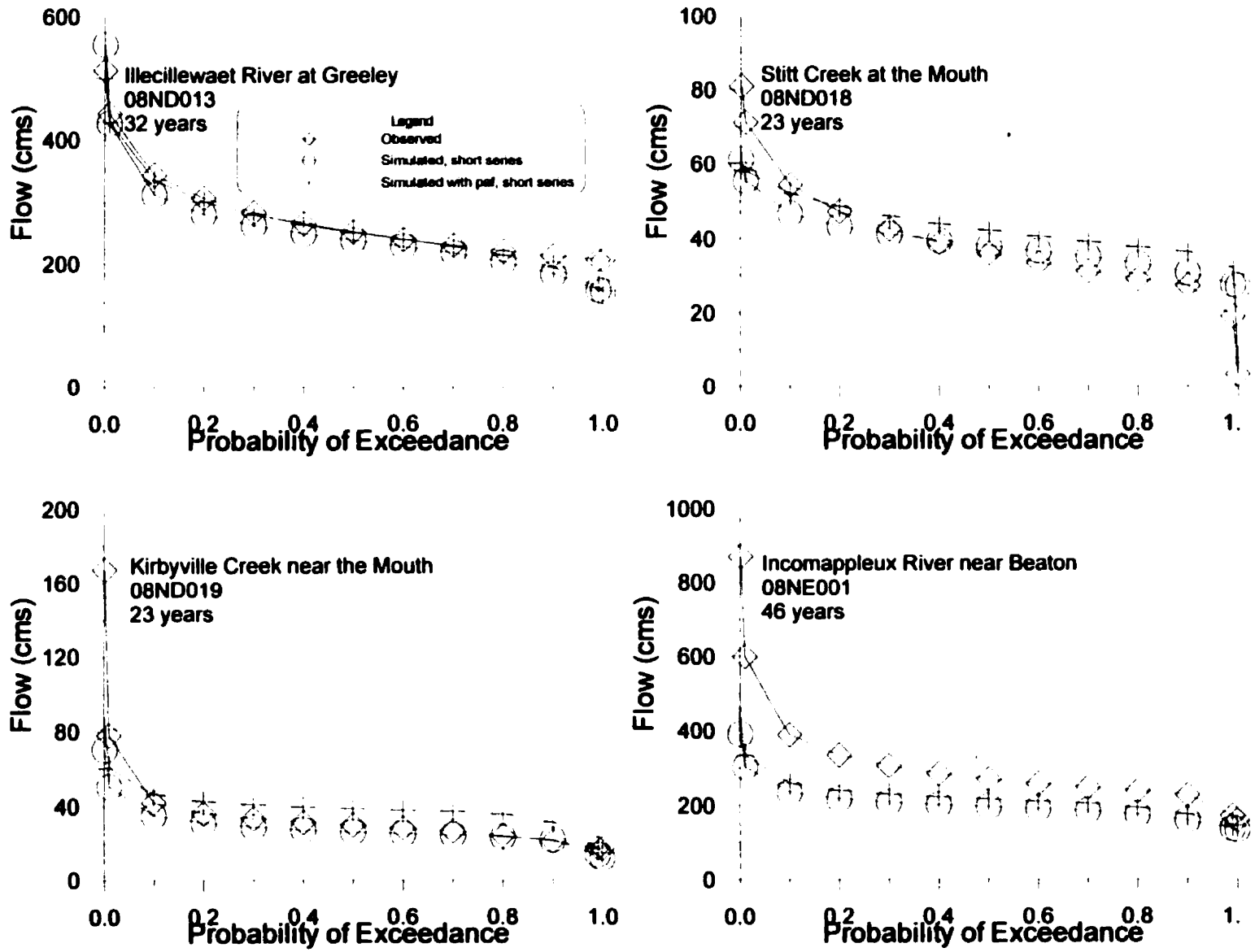


Figure B-3 – Comparison of Observed and Simulated (with and without PAF) Frequency Curves, Part 3

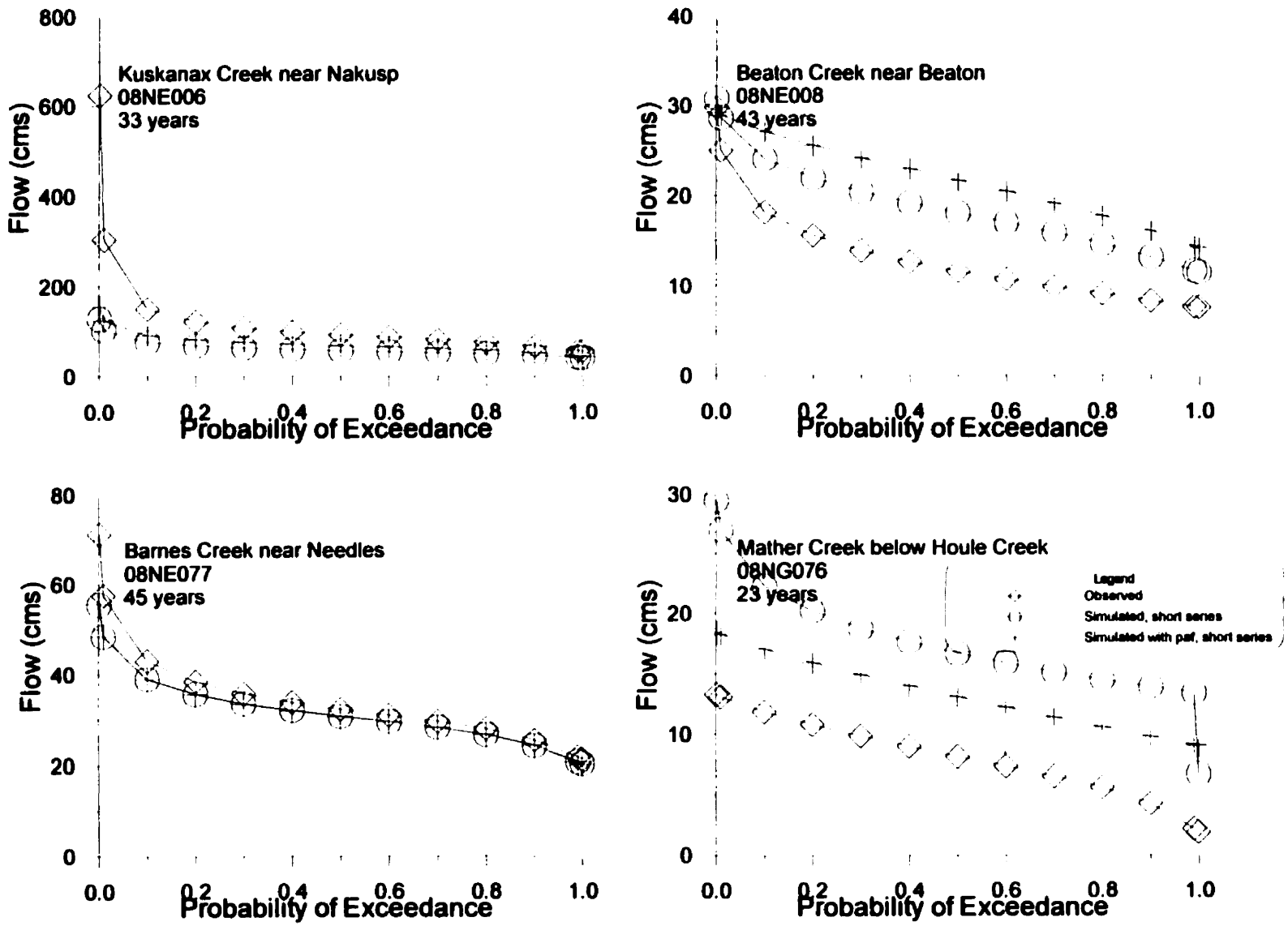


Figure B-4 – Comparison of Observed and Simulated (with and without PAF) Frequency Curves, Part 4

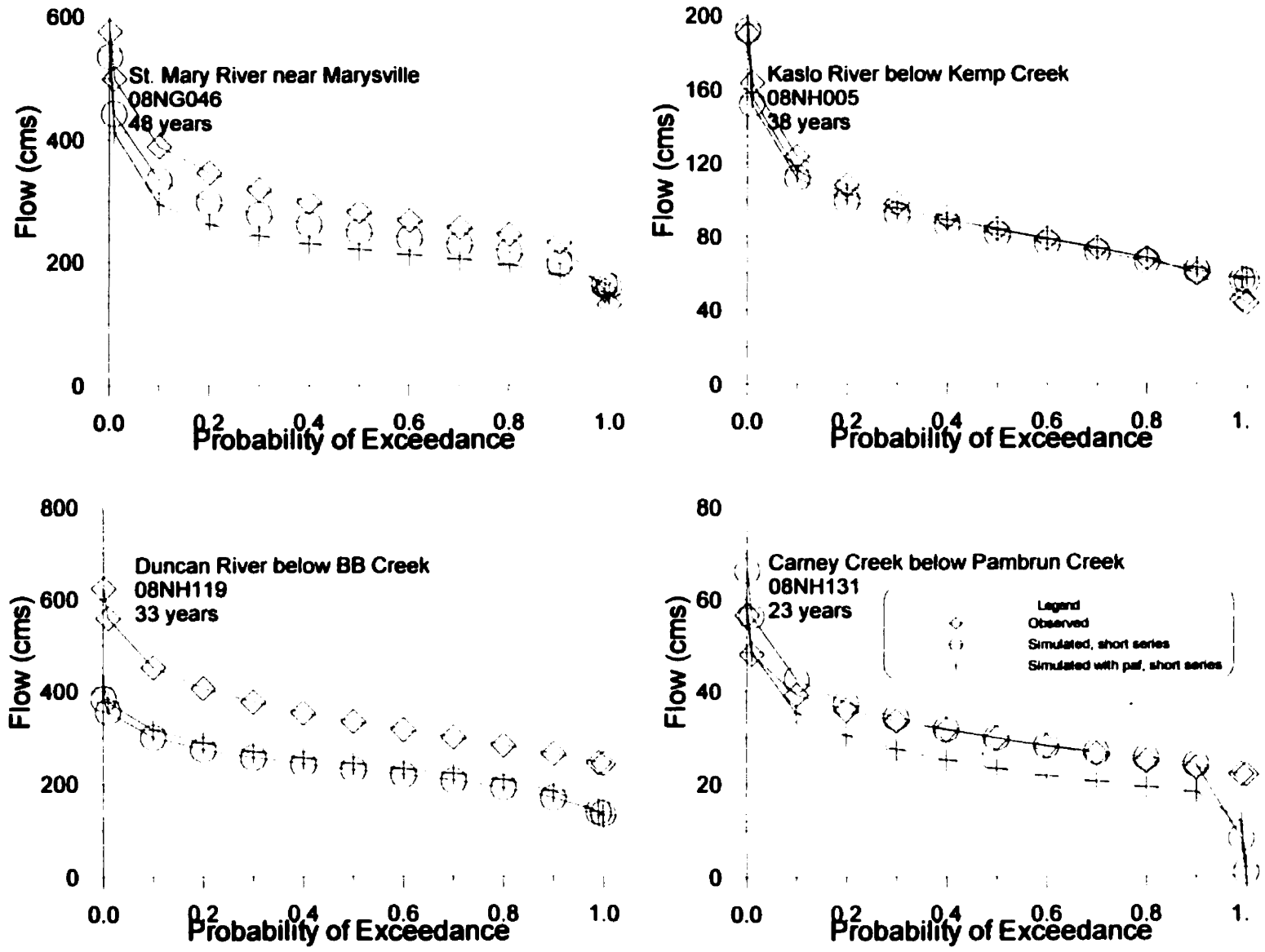


Figure B-5 – Comparison of Observed and Simulated (with and without PAF) Frequency Curves, Part 5

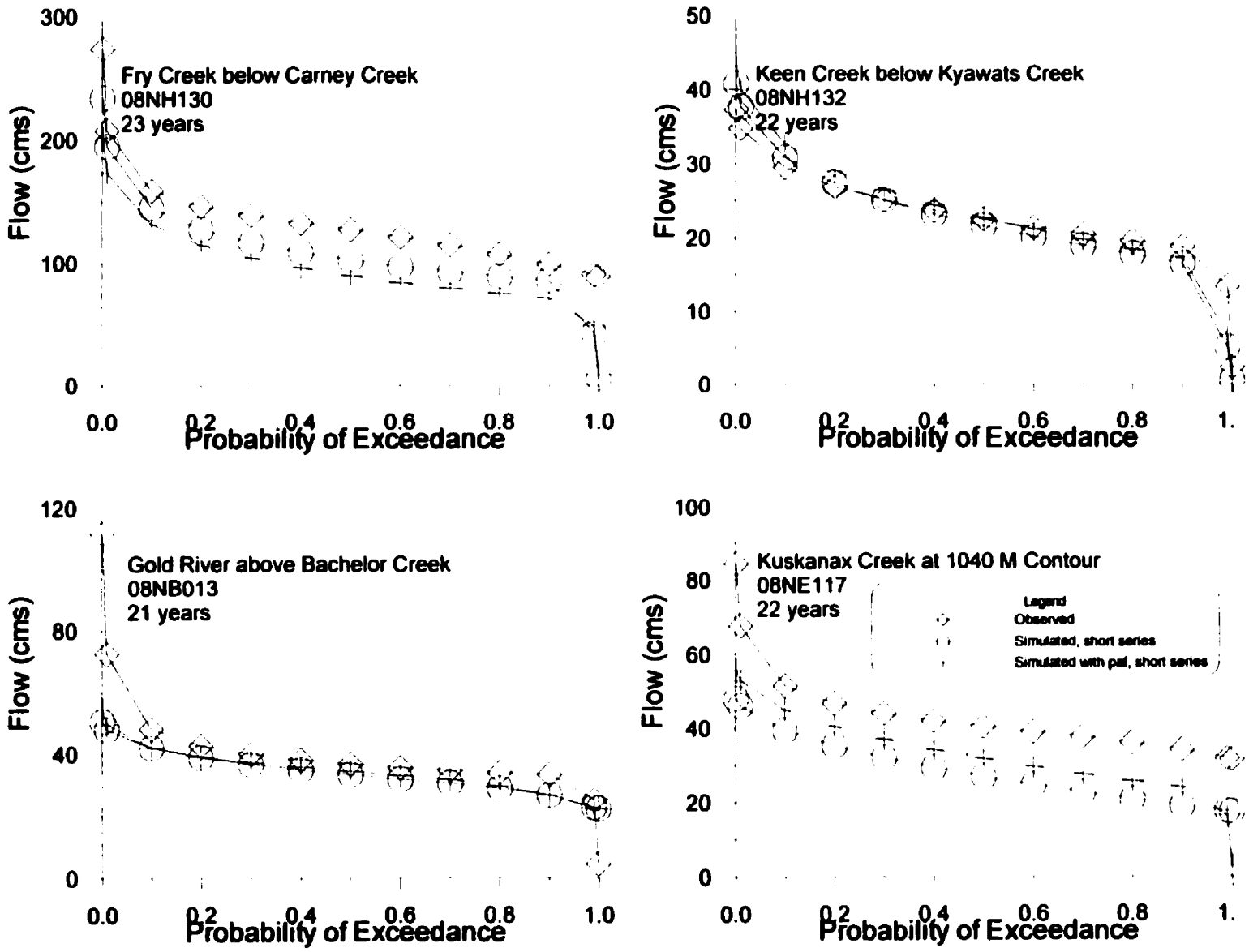


Figure B-6 – Comparison of Observed and Simulated (with and without PAF) Frequency Curves, Part 6

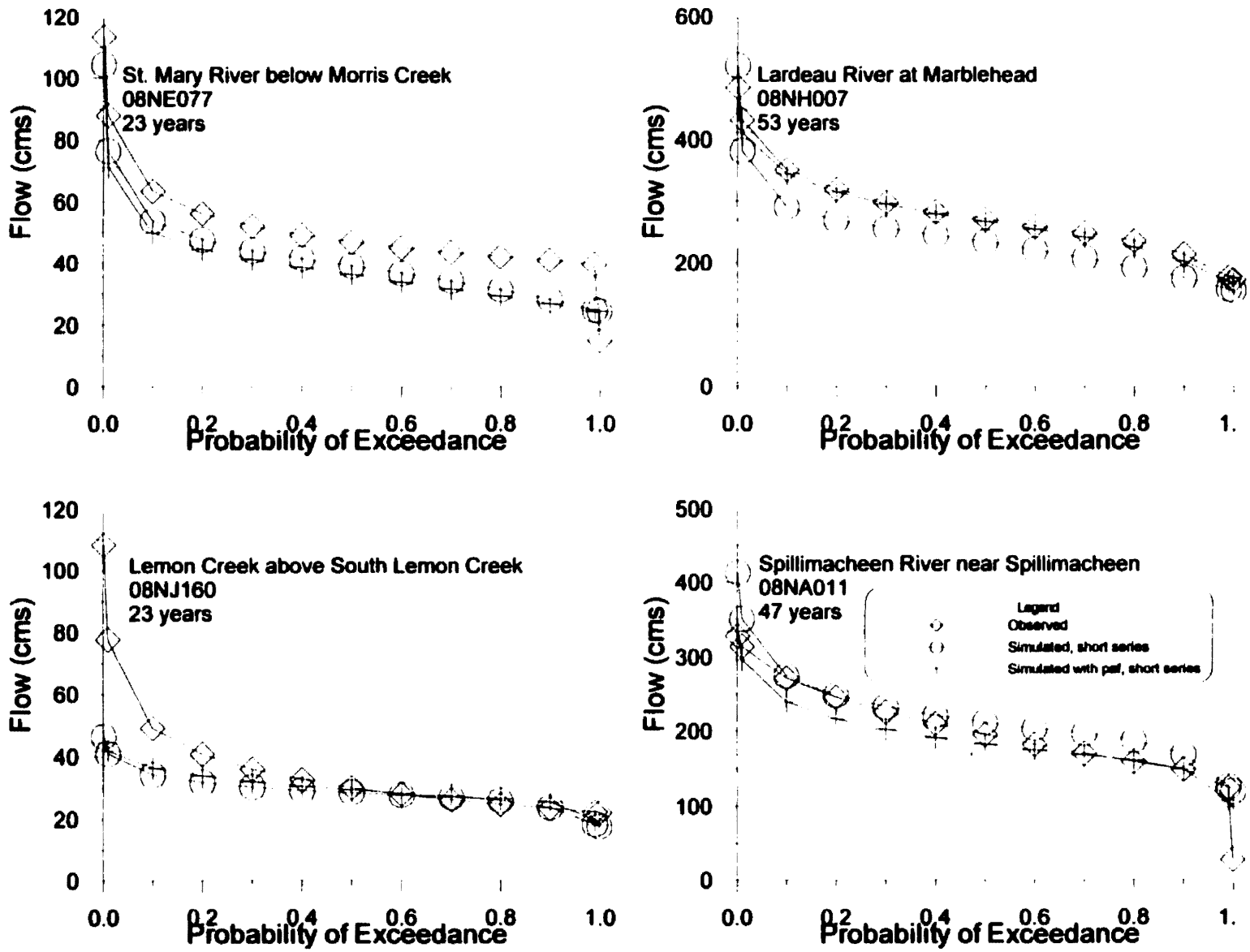


Figure B-7 – Comparison of Observed and Simulated (with and without PAF) Frequency Curves, Part 7

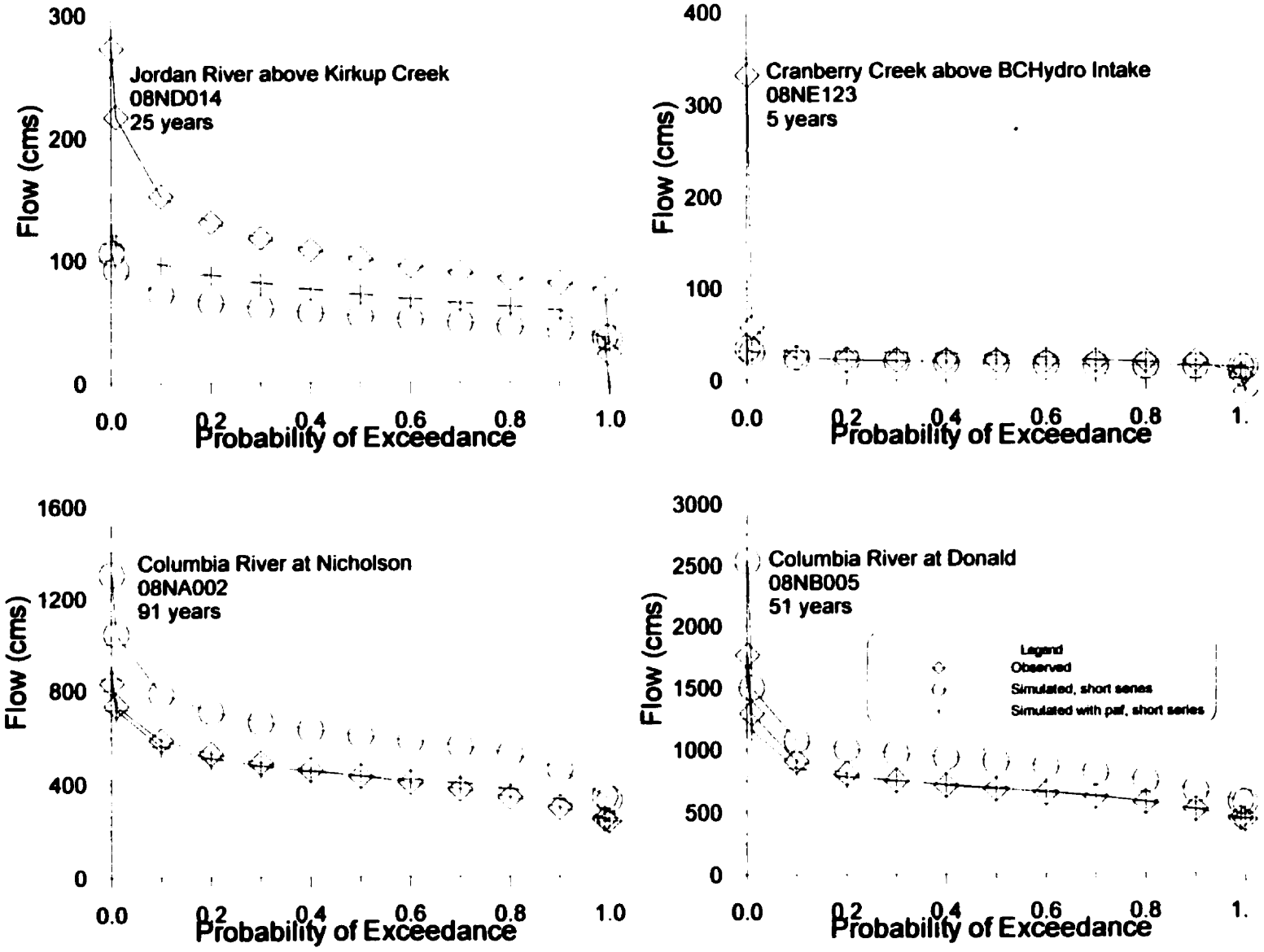


Figure B-8 – Comparison of Observed and Simulated (with and without PAF) Frequency Curves, Part 8

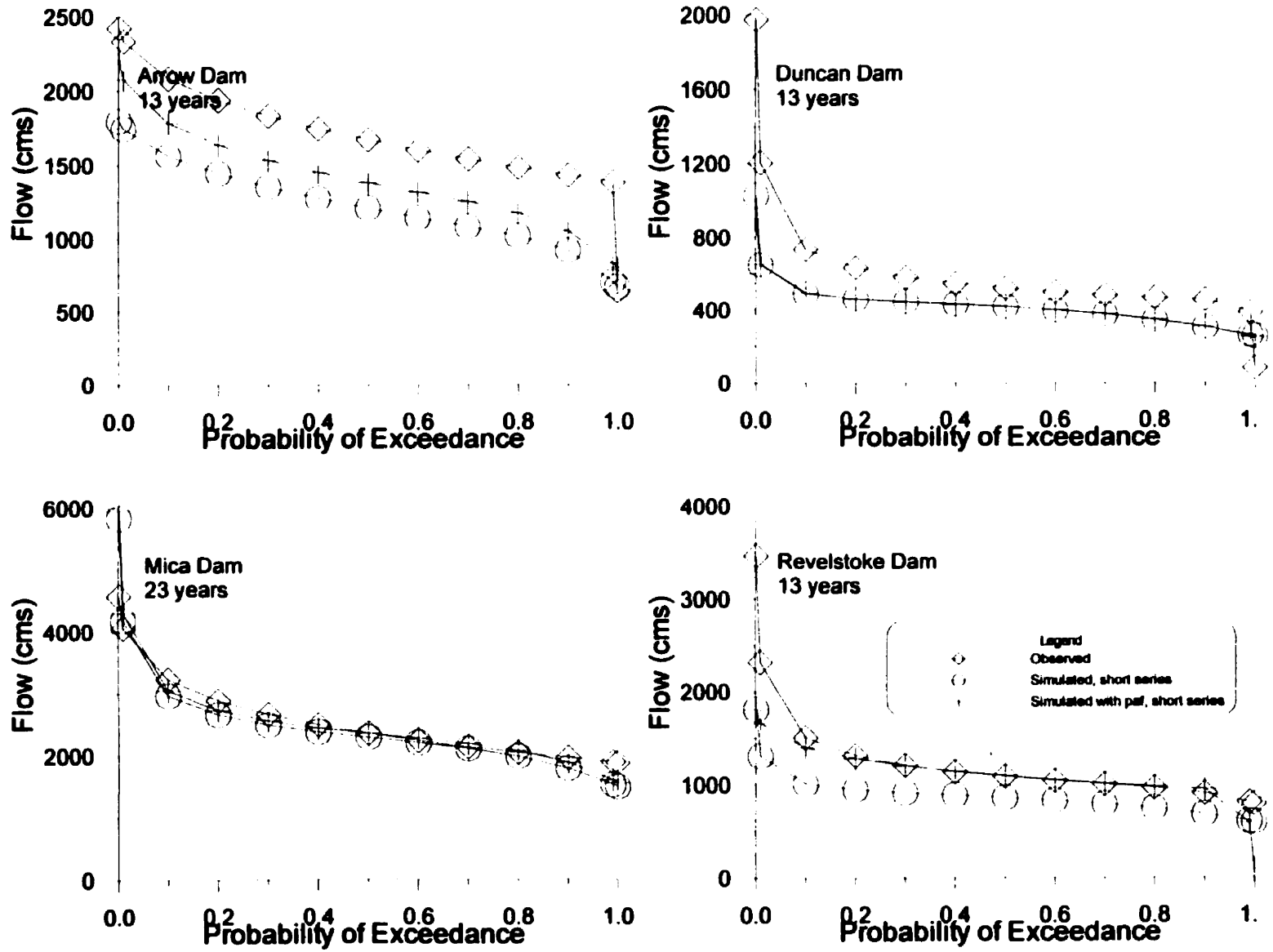


Figure B-9 – Comparison of Observed and Simulated (with and without PAF) Frequency Curves, Part 9

Appendix C

This appendix contains a series of 16 Figures, which show the response surface curves for each parameter for the Mica Dam Basin. Brief descriptions of each Figure are included above the Figure. If the percent response remained below 1% for the entire range of the parameter, the value of the parameter was assumed to be insignificant in the generation of the peak flows. The meaning of each parameter can be found in the following table.

Code	Full Name	Algorithm it is used in
A5	Unsaturated Zone Moisture Coefficient	Infiltration
AK	Surface Permeability	Infiltration
AKfs	Surface Permeability under snow	Infiltration
REC	Interflow storage-discharge coefficient	Interflow
R3	Overland flow conveyance parameter	Surface Runoff
R3fs	Overland flow conveyance parameter under snow	Surface Runoff
RETN	Soil retention coefficient	Evaporation
AK2	Upper to lower zone drainage coefficient	Groundwater recharge
AK2fs	Upper to lower zone drainage coefficient under snow	Groundwater recharge
LZF	Lower zone drainage function	Groundwater flow
PWR	Lower zone drainage function exponent	Groundwater flow
R2	River roughness coefficient	River Routing
MF	Melt factor	Snowmelt
BASE	Base temperature	Snowmelt
FPET	Potential Evapotranspiration Factor	Evaporation
FTALL	Evapotranspiration Factor for Tall Vegetation	Evaporation

The order of the landcovers and river classes can be seen in the following table.

Landcover	Name	River Class	Name
1	Barren	1	Valley
2	High Elevation, Dense Forest	2	High Elevation Mountain
3	Low Elevation, Dense Forest	3	Low Elevation Mountain
4	High Elevation, Light Forest		
5	Low Elevation, Light Forest		

The A5 and AK parameters were both insignificant for peak flow generation at Mica Dam. The A5 parameter was insignificant for all of the streamflow stations, while the AK parameter was significant for some stations. Where the AK parameter was significant, the response surface curves had similar shapes to the ones below (with different magnitudes). A value of AK greater than 10 mm/hour does not change the flow because the precipitation would only rarely be larger than 10 mm/hour.

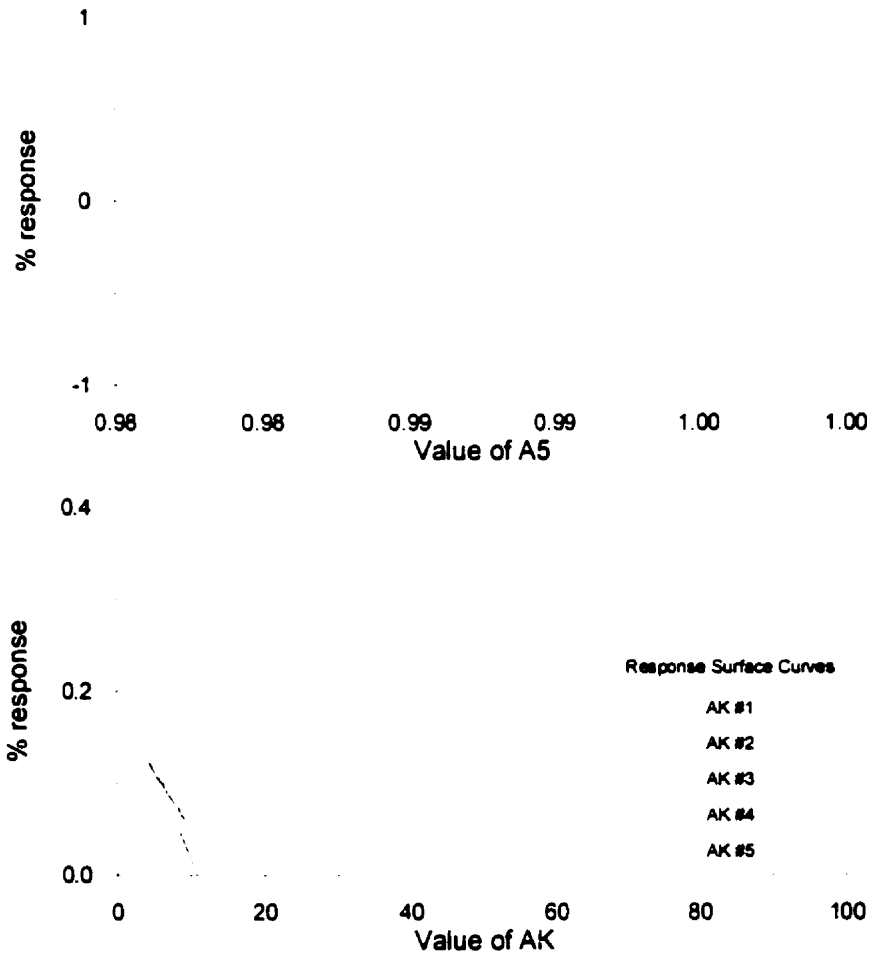
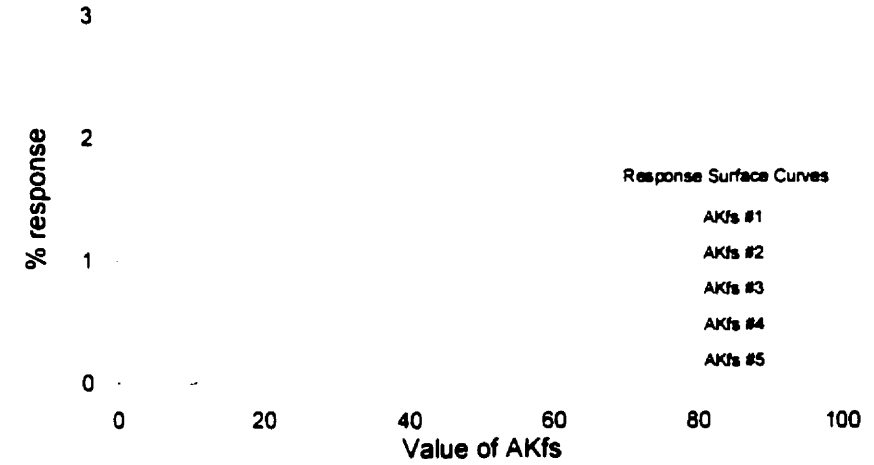


Figure C-1 – Response Curve Shape for A5 and AK

The AKfs parameter was significant for the barren landcover, but not significant for the others. The barren areas are at higher elevations where snow cover lasts longer, and therefore, infiltration can control peak flows. Beyond a certain value, however, the peak flows are unaffected, because the precipitation would only rarely be larger than 10 mm/hour. Similar results were found for other streamflow stations. The REC parameter was significant for all landcovers. A monotonically increasing curve was found, indicating that a single correct value of the parameters may be found. Other streamflow stations were similar.



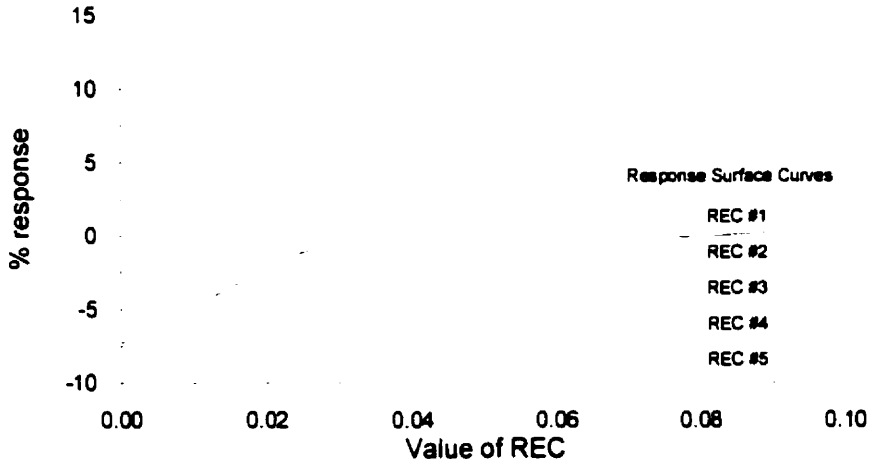


Figure C-2 – Response Curve Shape for AKfs and REC

The R3 and R3fs parameters were both insignificant for generation of peak flows at Mica Dam, and the other streamflow stations. The roughness of the ground did not affect peak flows, since the roughness affects mainly very large amounts of surface runoff.

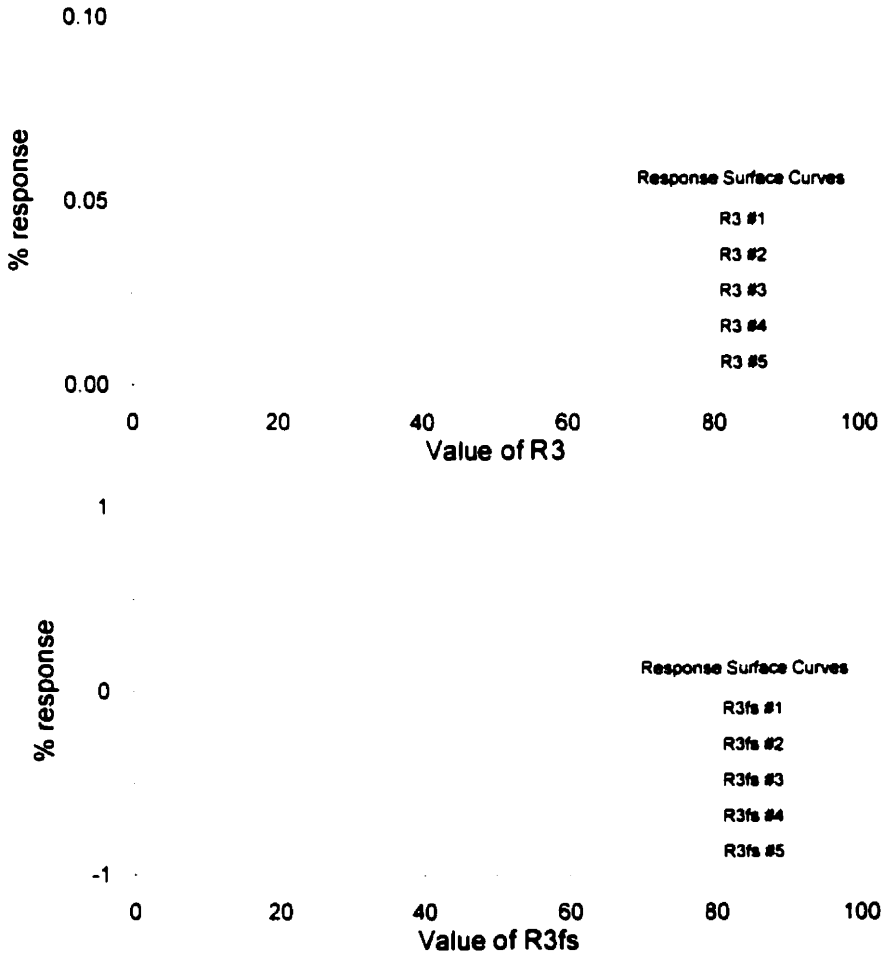
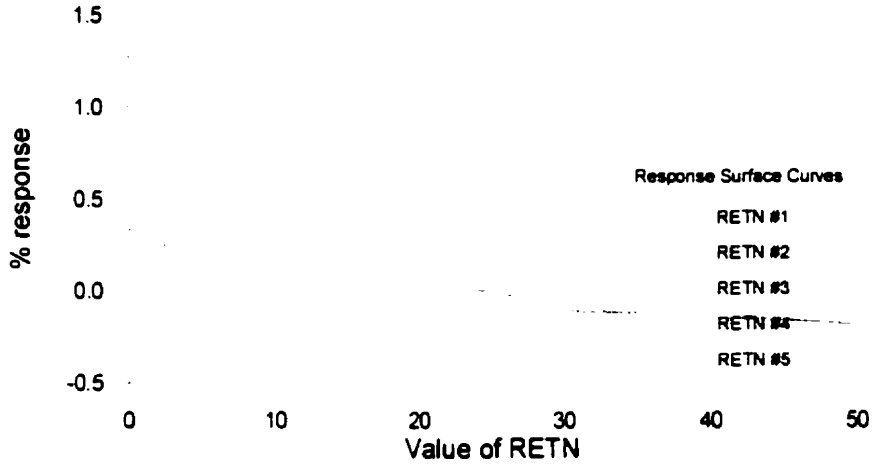


Figure C-3 – Response Curve Shape for R3 and R3fs

The RETN parameter was significant for two of the landcovers (low elevation dense and light forests), and insignificant for the others. When the parameter was significant (for a combination of landcover and streamflow station), the response surface was generally monotonically decreasing. The AK2 parameter was significant in generating the peak flows for all landcovers, and most streamflow stations. The parameter was generally monotonically decreasing, showing that as greater drainage to the lower zone occurred, there was less peak flow.



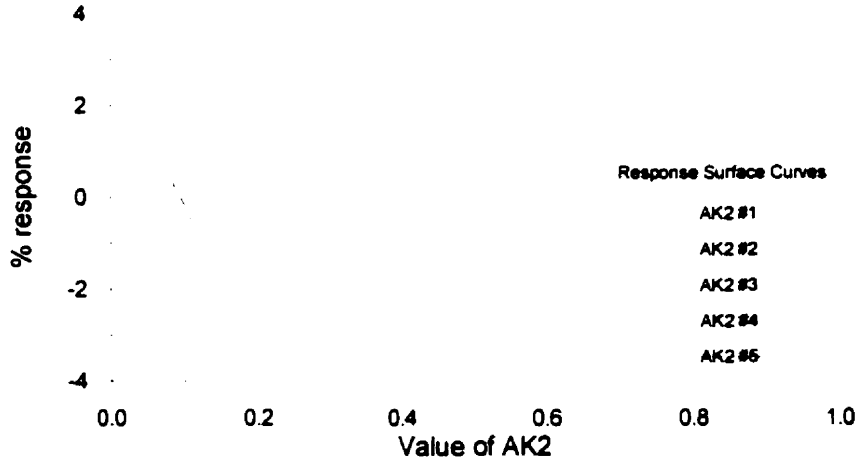


Figure C-4 – Response Curve Shape for RETN andAK2

The AK2fs parameter was significant for most combinations of landcover and streamflow. The curve monotonically decreased for three of the landcovers, but increased for the other two (low elevation forests). The hydrological behaviour changes between high and low elevation drainage to the lower zone. The LZF parameter was most significant for the high elevation mountain river class, and mildly significant for the other two classes: the LZF valley parameter was insignificant for most of the stations in the valley class. The surface was monotonically increasing for all three river classes.

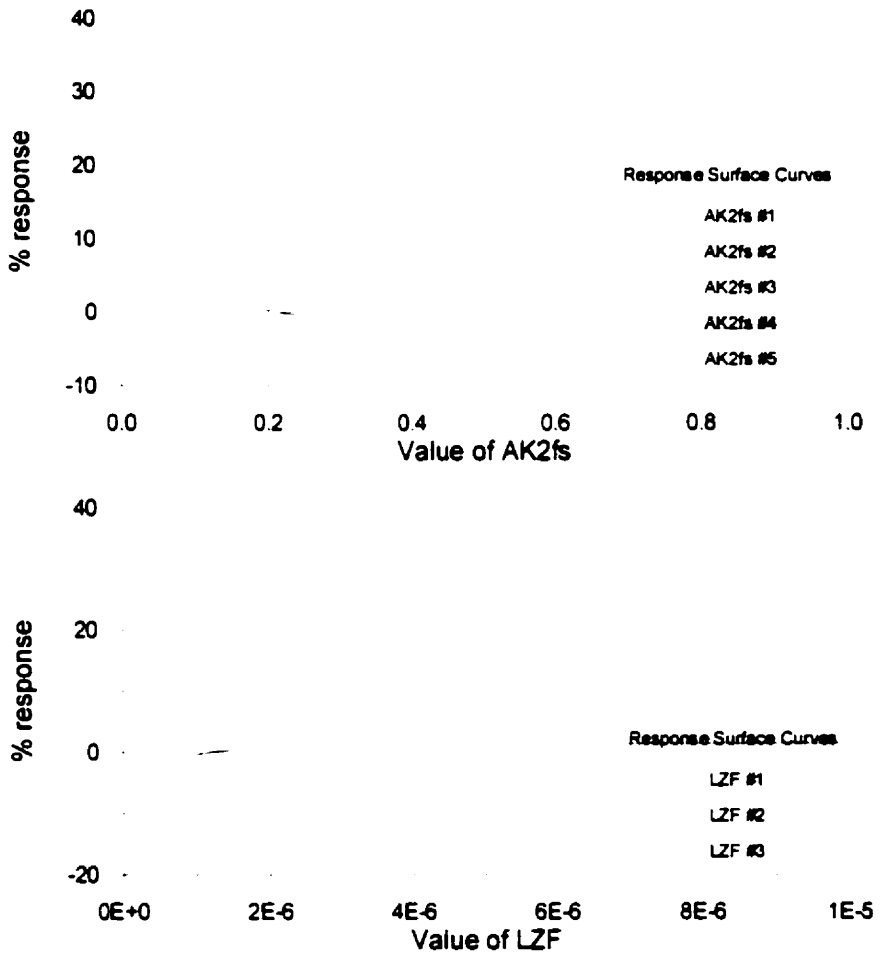


Figure C-5 – Response Curve Shape for AK2fs and LZF

The PWR parameter was most significant for the high elevation mountain river class, and somewhat significant for the other two. The PWR parameter for the valley river class was insignificant for most stations. The PWR parameter was monotonically increasing. The R2 parameter for the valley river class was also insignificant for most stations (but not Mica Dam Basin), indicating that most stations have very little valley river, which is accurate. The R2 parameter was monotonically decreasing for some stations, and convex for others, when significant.

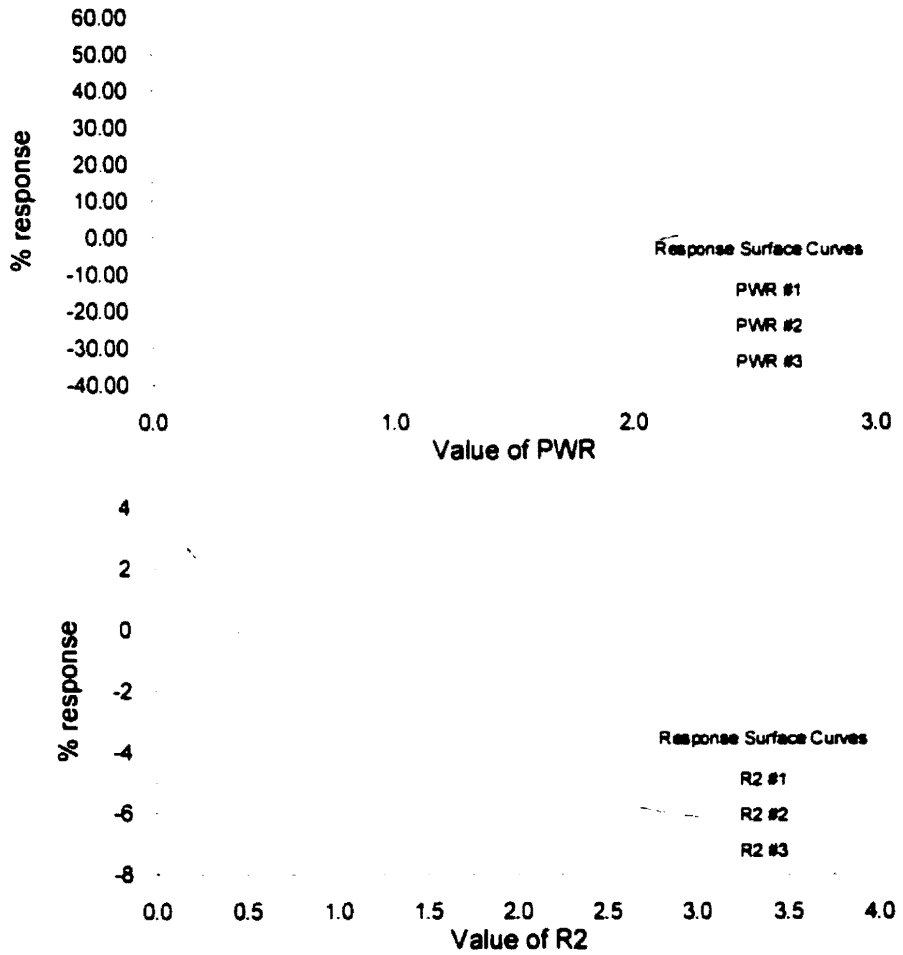
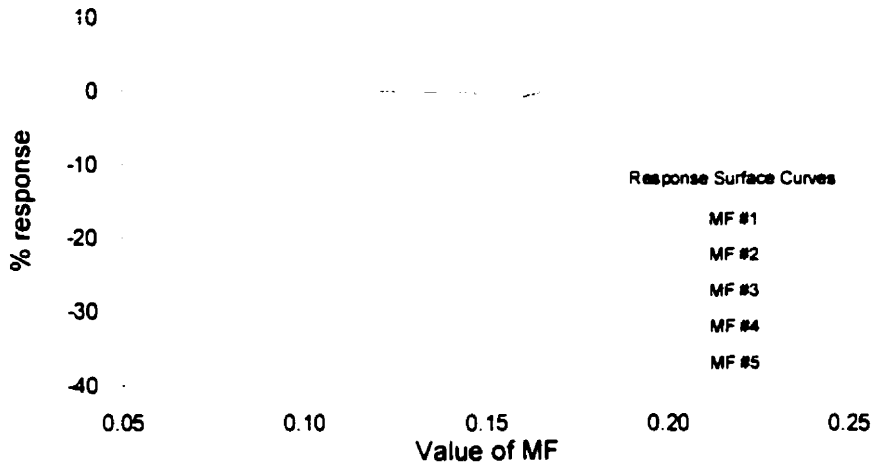


Figure C-6 – Response Curve Shape for PWR and R2

The MF parameter, like the AK2fs parameter, shows a change in hydrological behaviour between high and low elevation. For high elevations, the peak flow increases as the parameter increases, but the opposite is true for low elevations. Curves vary in shape (concave, convex, and monotonic), and there are some large deviations from "ideal" curves. The BASE parameter had various shapes of response curves. Generally, the high-elevation landcovers had non-ideal response curves, while most of the low-elevation response curves were acceptable.



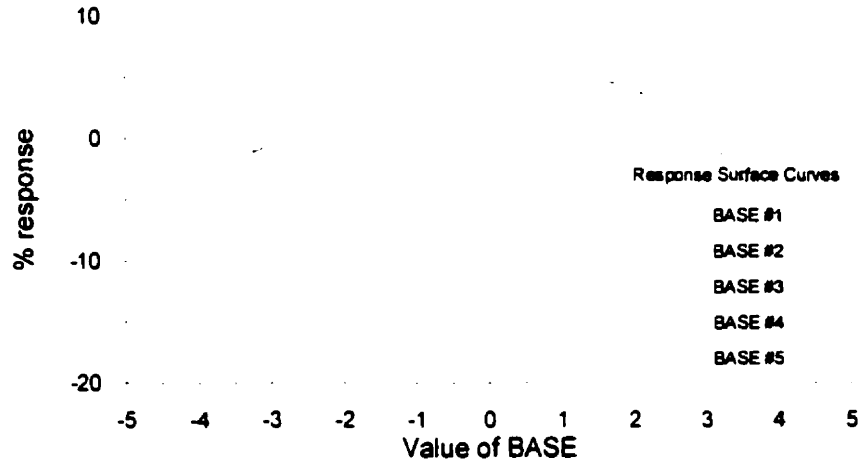
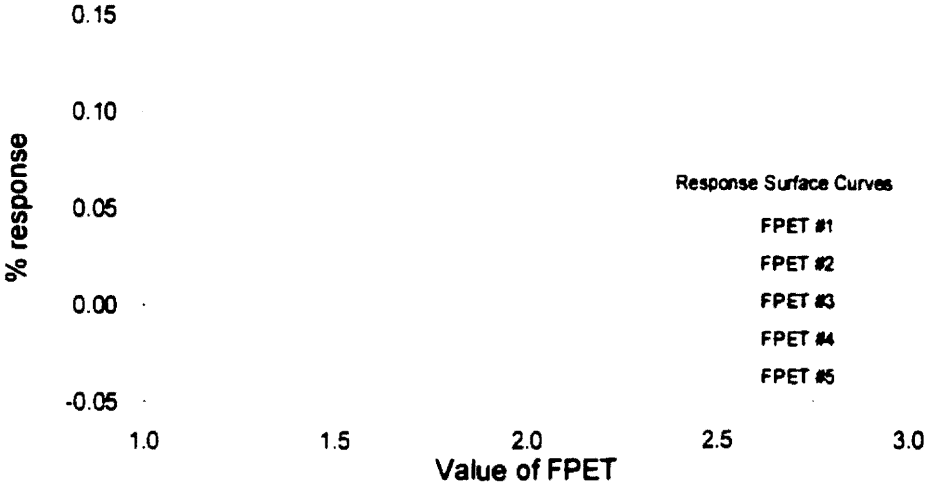


Figure C-7 – Response Curve Shape for MF and BASE

The FPET parameter was insignificant for generation of peak flows for all combinations of streamflow stations and landcovers. The FTALL parameter was insignificant for many combinations of streamflow stations and landcovers, and was monotonically decreasing for combinations that were significant.



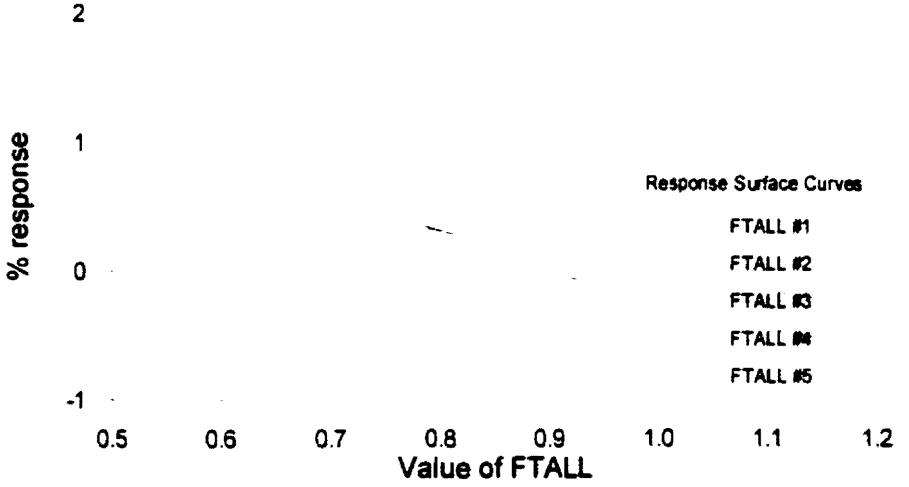


Figure C-8 – Response Curve Shape for FPET and FTALL

Appendix D

This appendix contains a series of 25 figures, which show the two variable interaction between the MF and BASE parameters for the Mica Dam. Each figure was generated from 36 simulations that were performed with different combinations of the MF and BASE parameters. The parameters were varied in a grid with six values of the BASE parameter on the vertical axis and six values of the MF parameter on the horizontal axis. The average difference between the five peak flows for each simulation and the reference peak flows was calculated and used to create the figure. One figure was generated for each combination of landcovers for the MF and BASE parameter (five landcovers for each parameter). Each figure shows a smooth variation in the objective function with a single line where the average difference equaled zero. The line indicated that several combinations of MF and BASE were optimum for peak flow generation. Only one such line existed, forming a "ridge" on the multi-dimensional hill of the objective function. The figures indicated that only one "hill" existed within the physically possible range. The calibrated parameter values are indicated on each figure with a diamond symbol. As expected, the calibrated parameter values were on the zero line for each figure (interpolation effects caused the diamond to be slightly off the line in some figures).

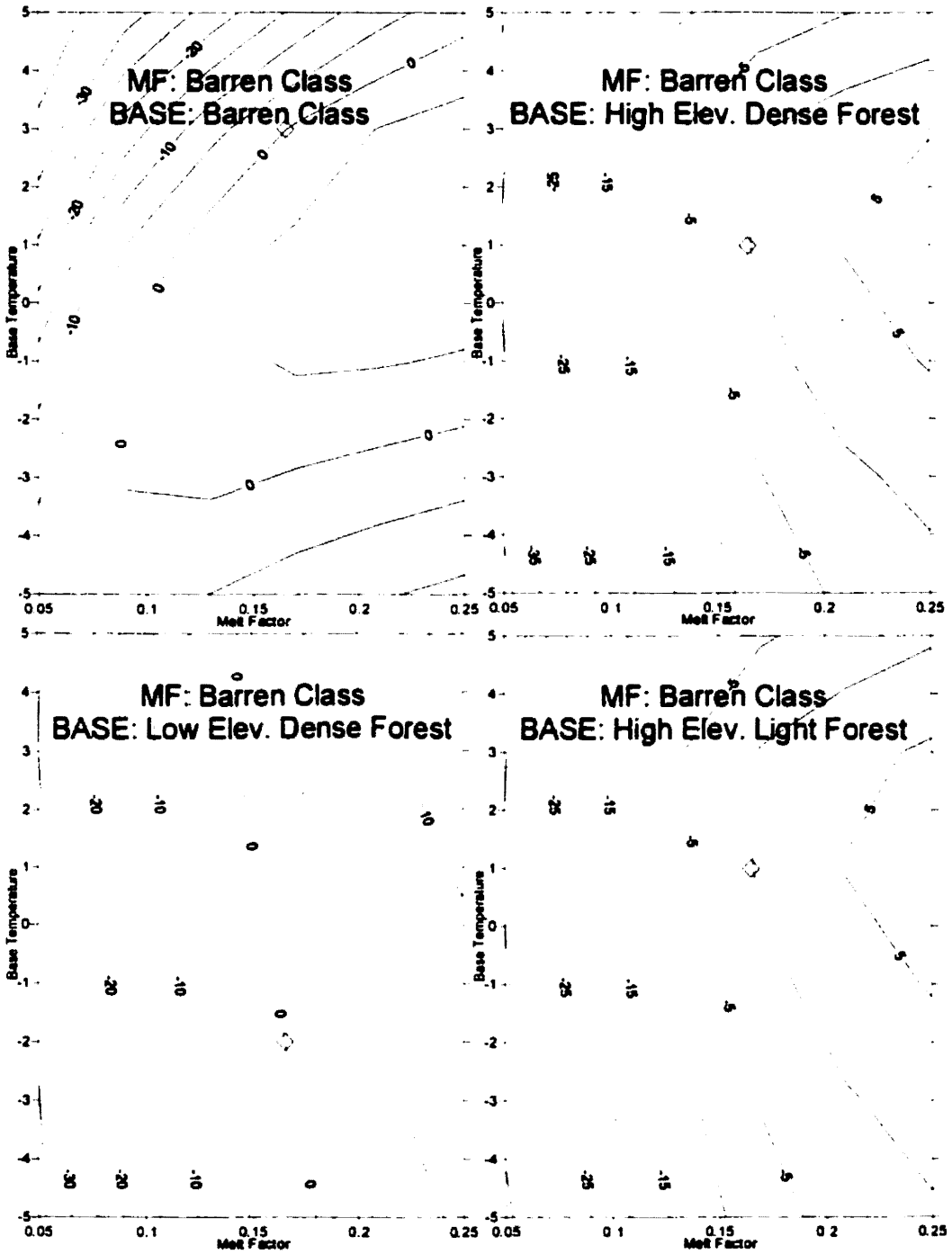


Figure D-1 – Two-parameter variation plots for MF and BASE, Part 1

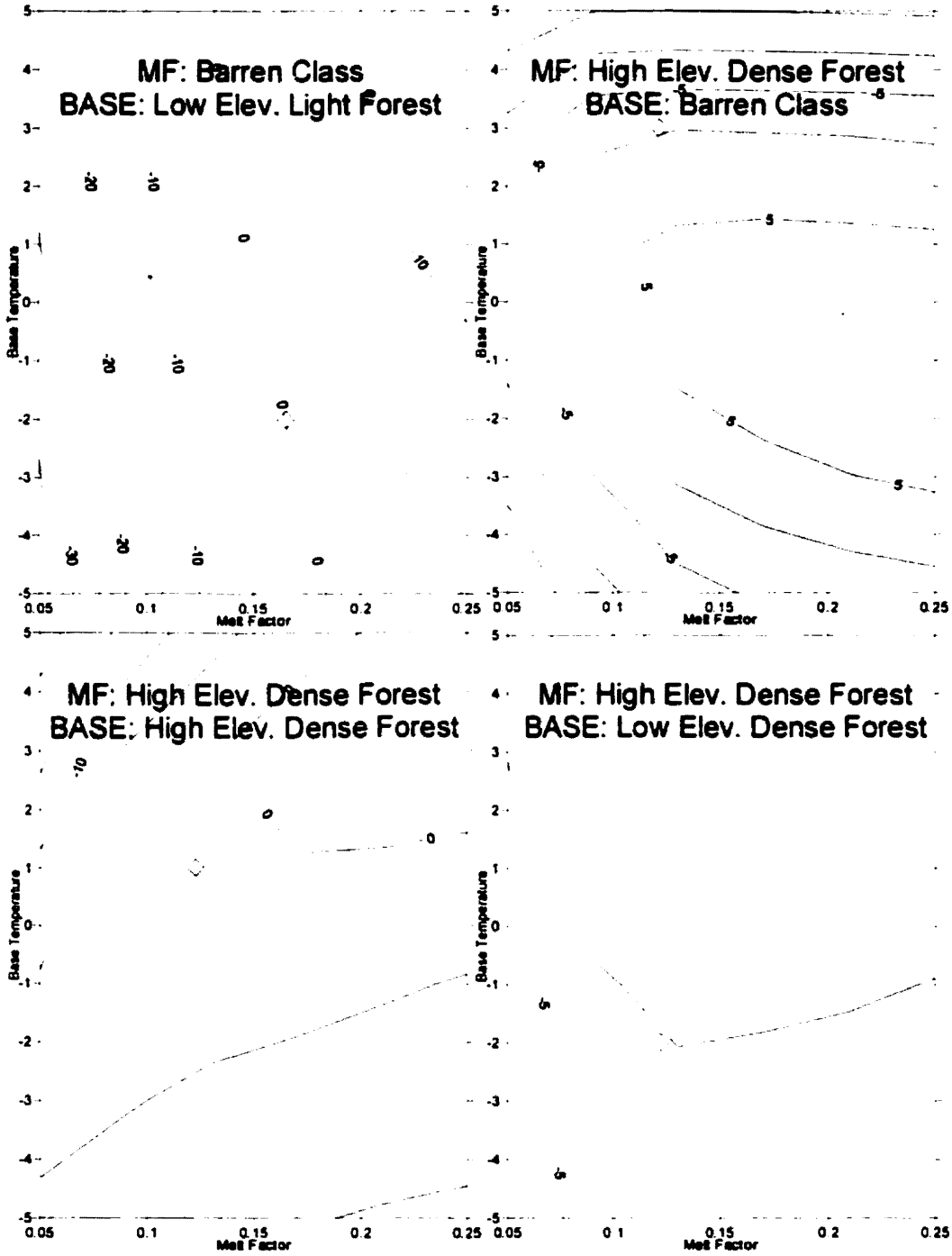


Figure D-2 – Two-parameter variation plots for MF and BASE, Part 2

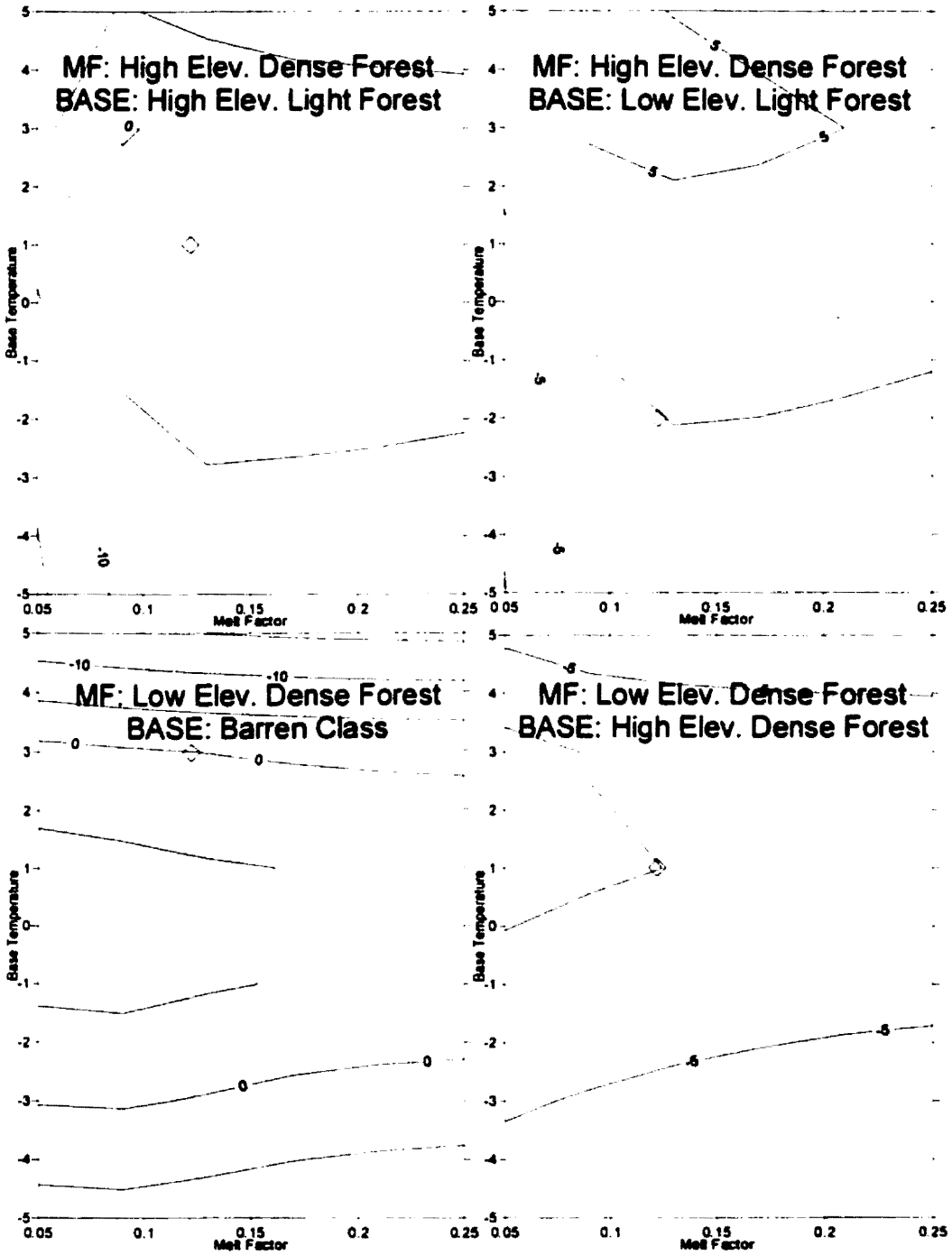


Figure D-3 – Two-parameter variation plots for MF and BASE, Part 3

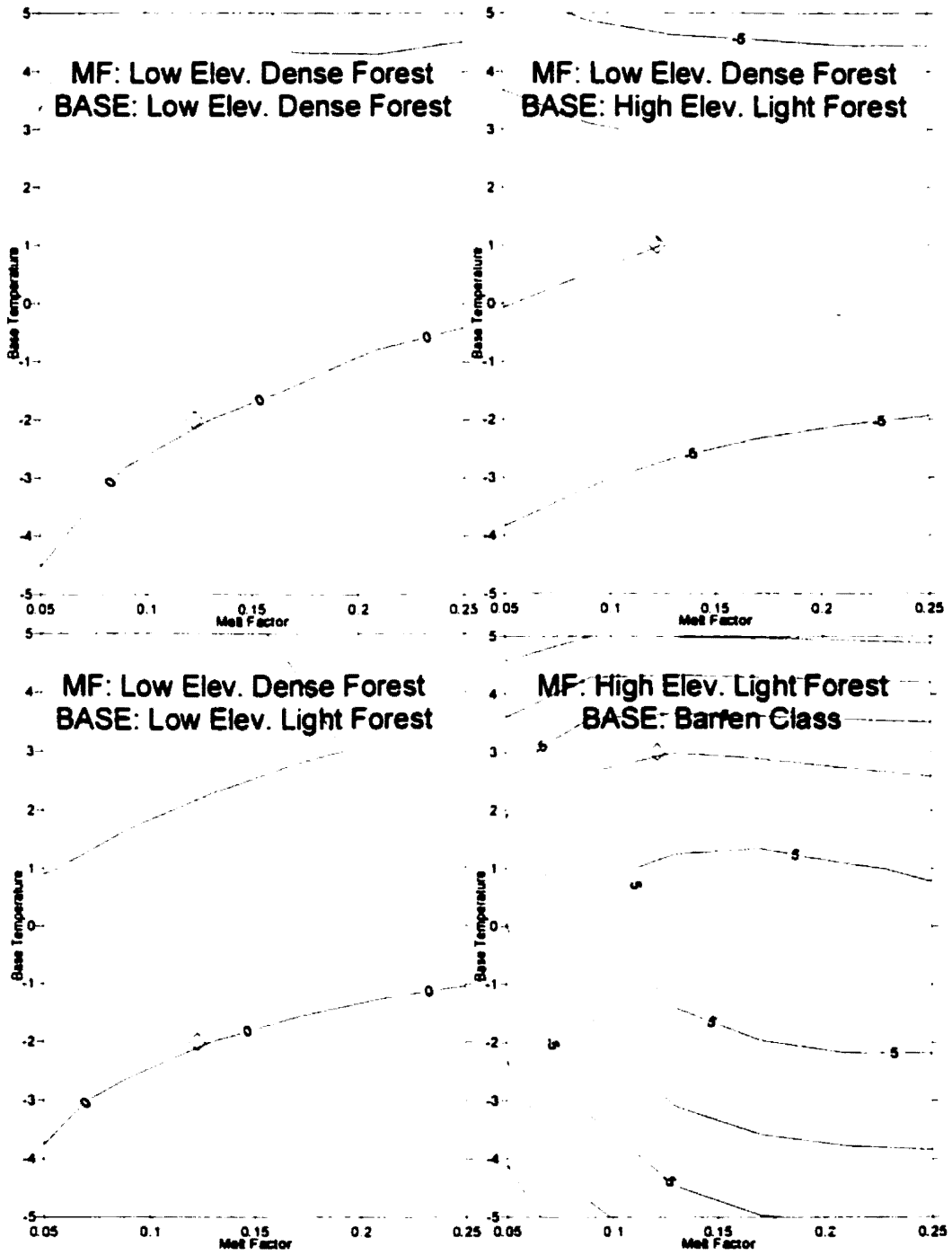


Figure D-4 – Two-parameter variation plots for MF and BASE, Part 4

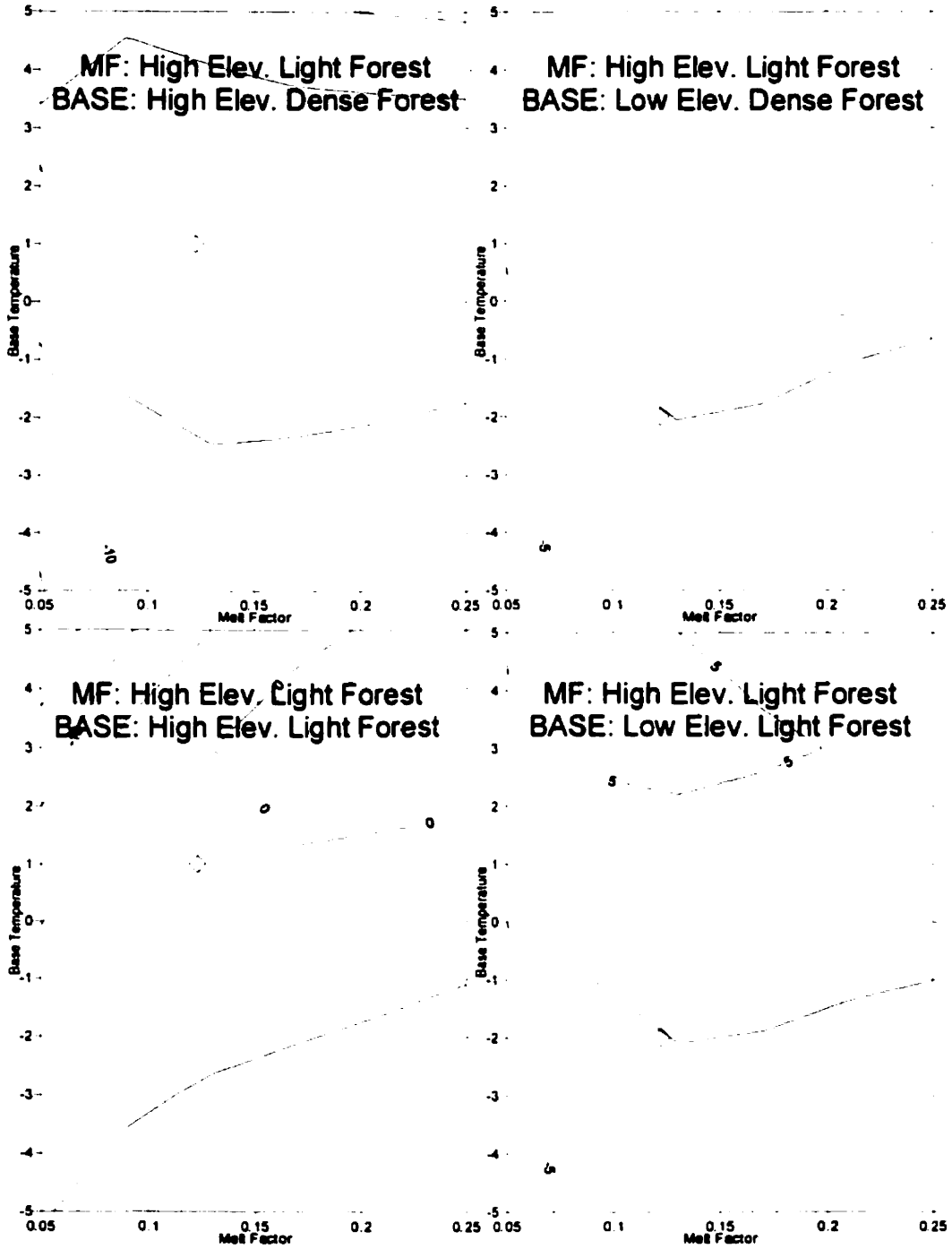


Figure D-5 – Two-parameter variation plots for MF and BASE, Part 5

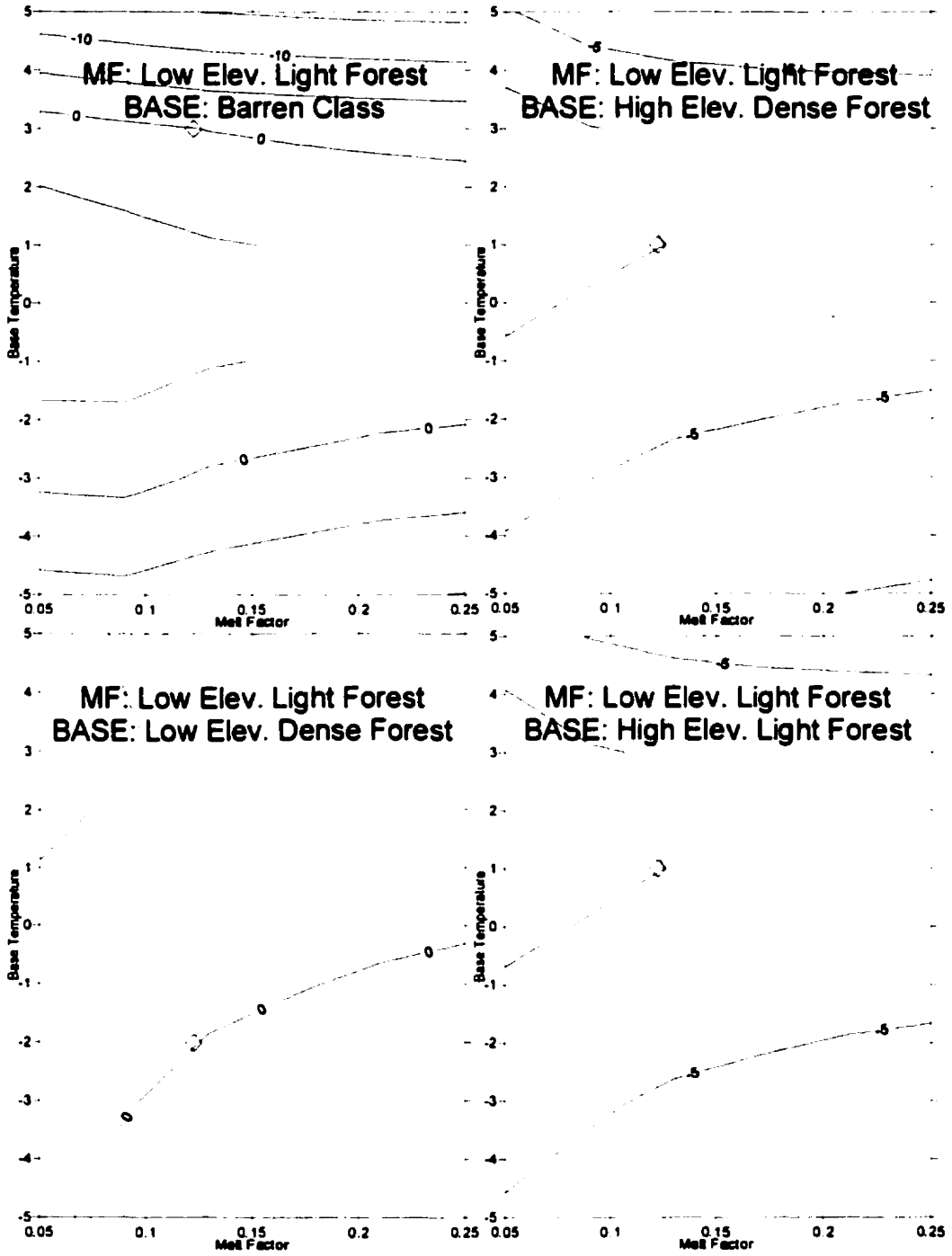


Figure D-6 – Two-parameter variation plots for MF and BASE, Part 6

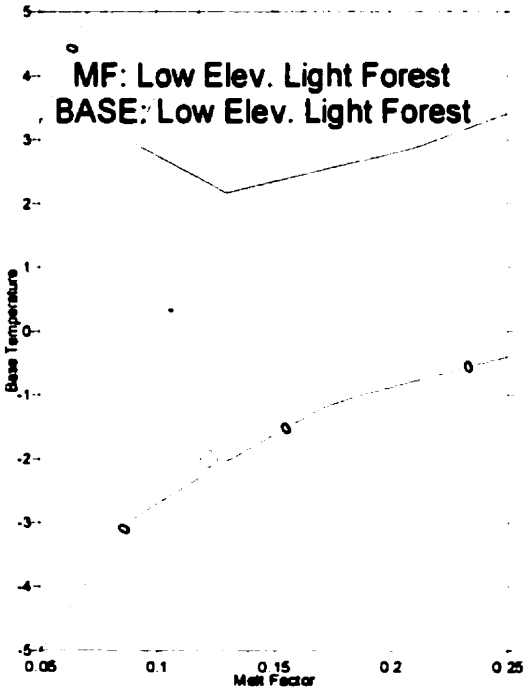


Figure D-7 – Two-parameter variation plots for MF and BASE, Part 7

Appendix E

This appendix contains a series of figures that compare the confidence limits calculated by the two methods in this research. The first method was to use the Monte Carlo analysis. The Monte Carlo analysis allowed all of the parameters to vary simultaneously in 10,000 simulations. The five-year peaks for each simulation were ordered and five histograms were calculated (one for each peak). The 2.5% and 97.5% points of each histogram were found and connected to form an estimate of the 95% confidence limits. The second method was to refine the estimate of the confidence limits by calculating the full 95-year time series in WATFLOOD/SPL for the simulations that crossed the 2.5% and 97.5% histogram lines. The average of the frequency curves calculated by the 95-year simulations is plotted in the figures (scaled by the average peak flow for the 2.5% curve, and scaled by the maximum peak flow for the 97.5% curve). The figures show that the two estimates for the confidence limits were very similar.

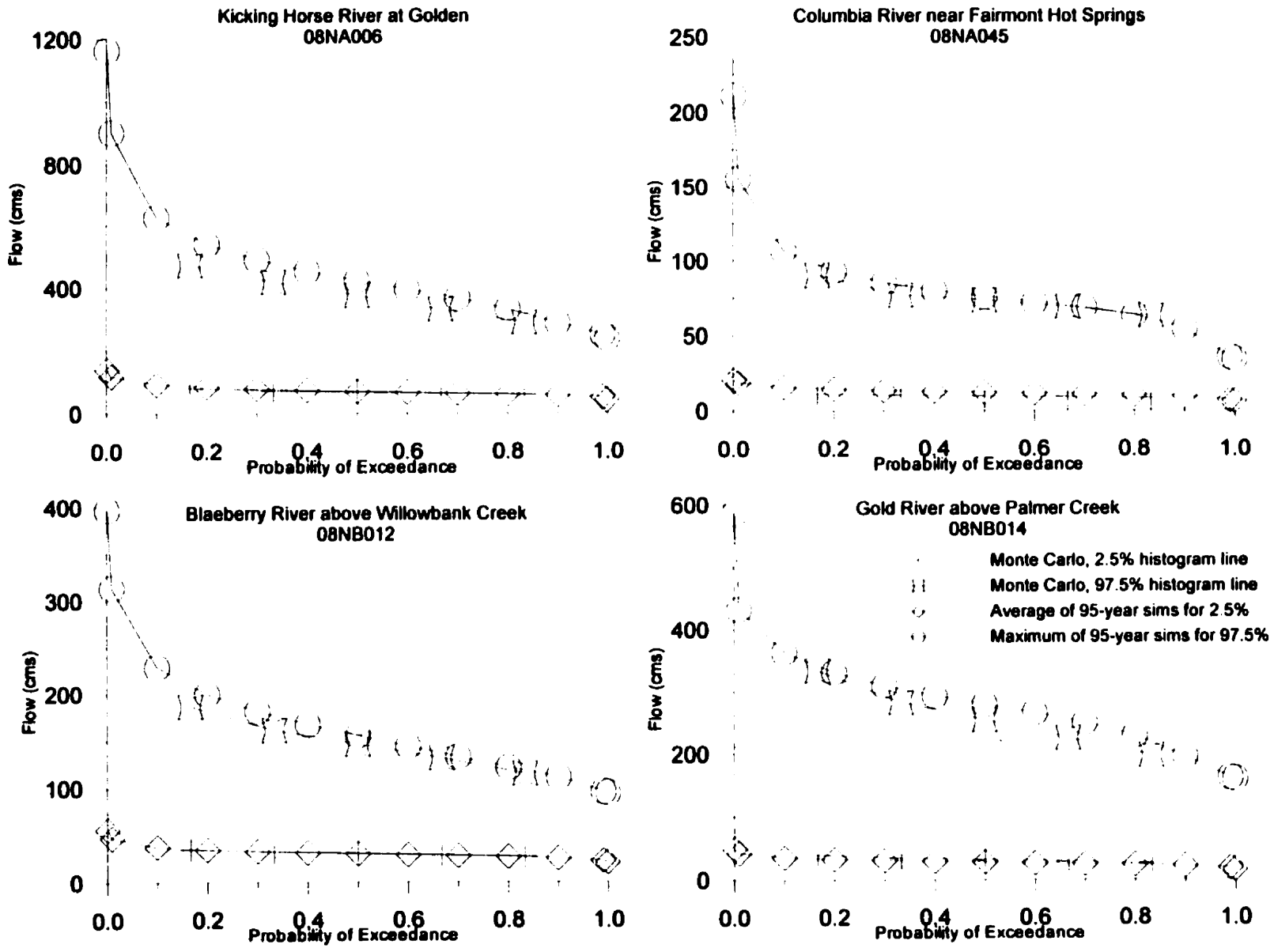


Figure E-1 – Comparison of Confidence Limits – Part 1

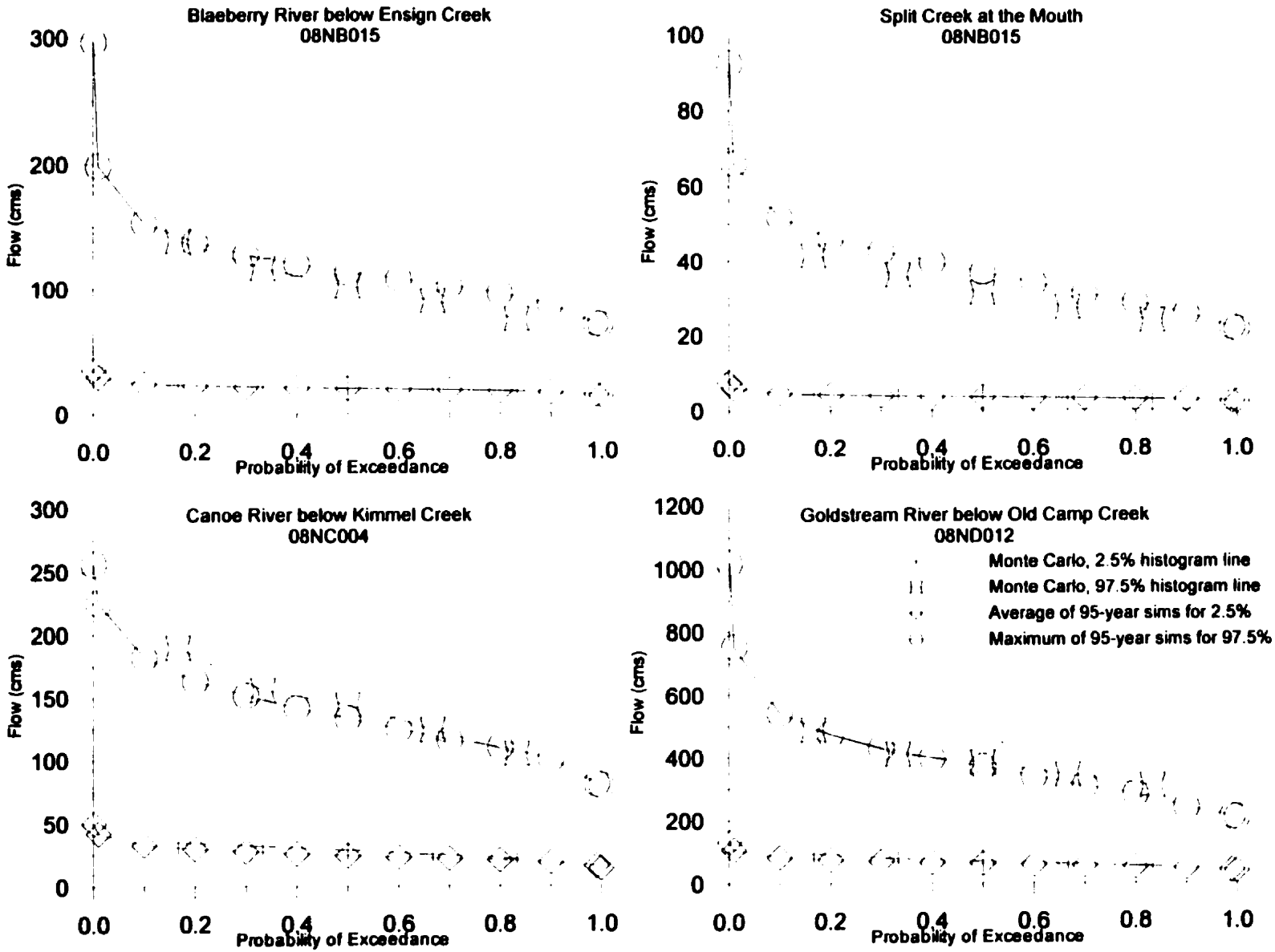


Figure E-2 – Comparison of Confidence Limits – Part 2

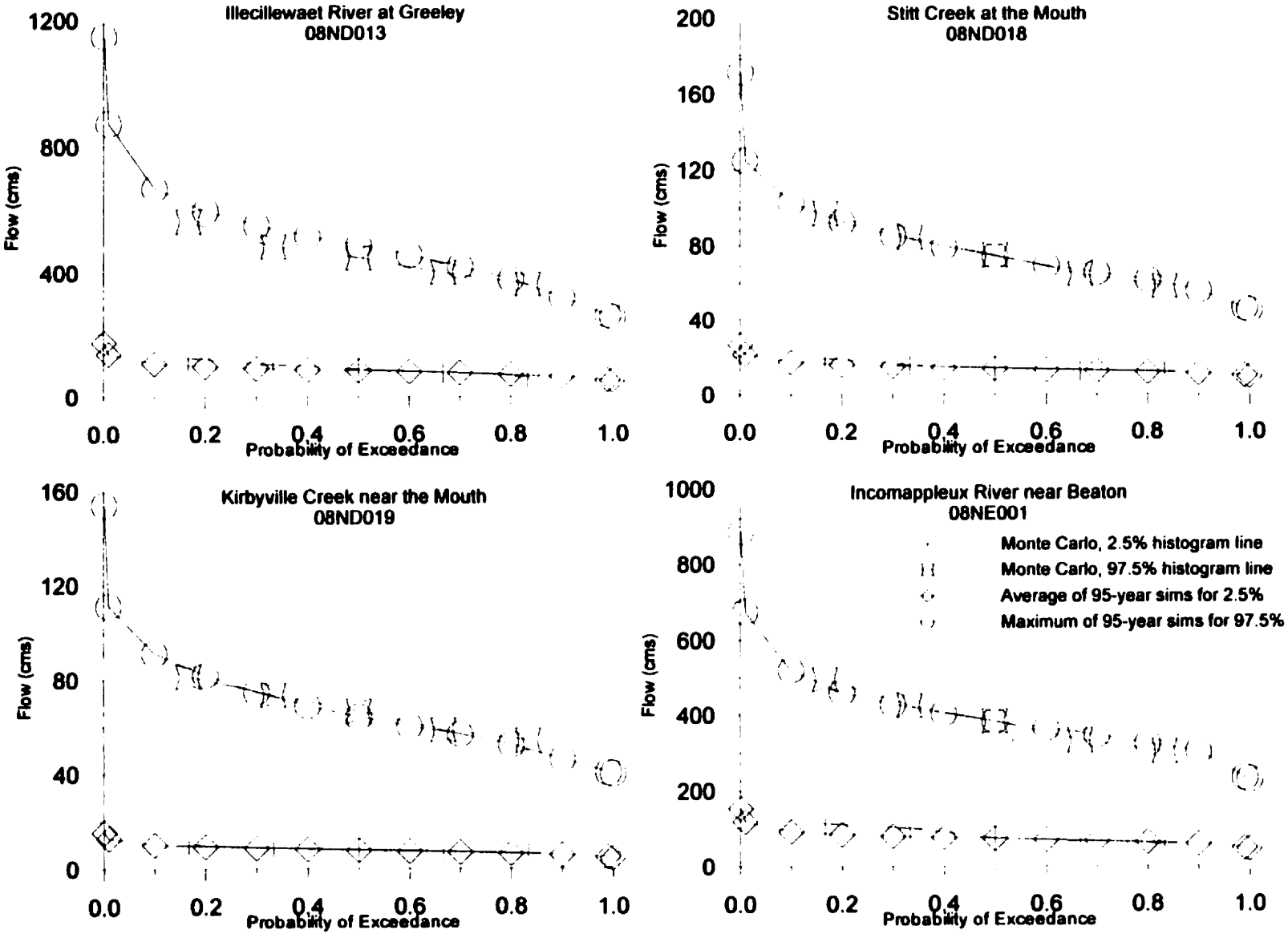


Figure E-3 – Comparison of Confidence Limits – Part 3

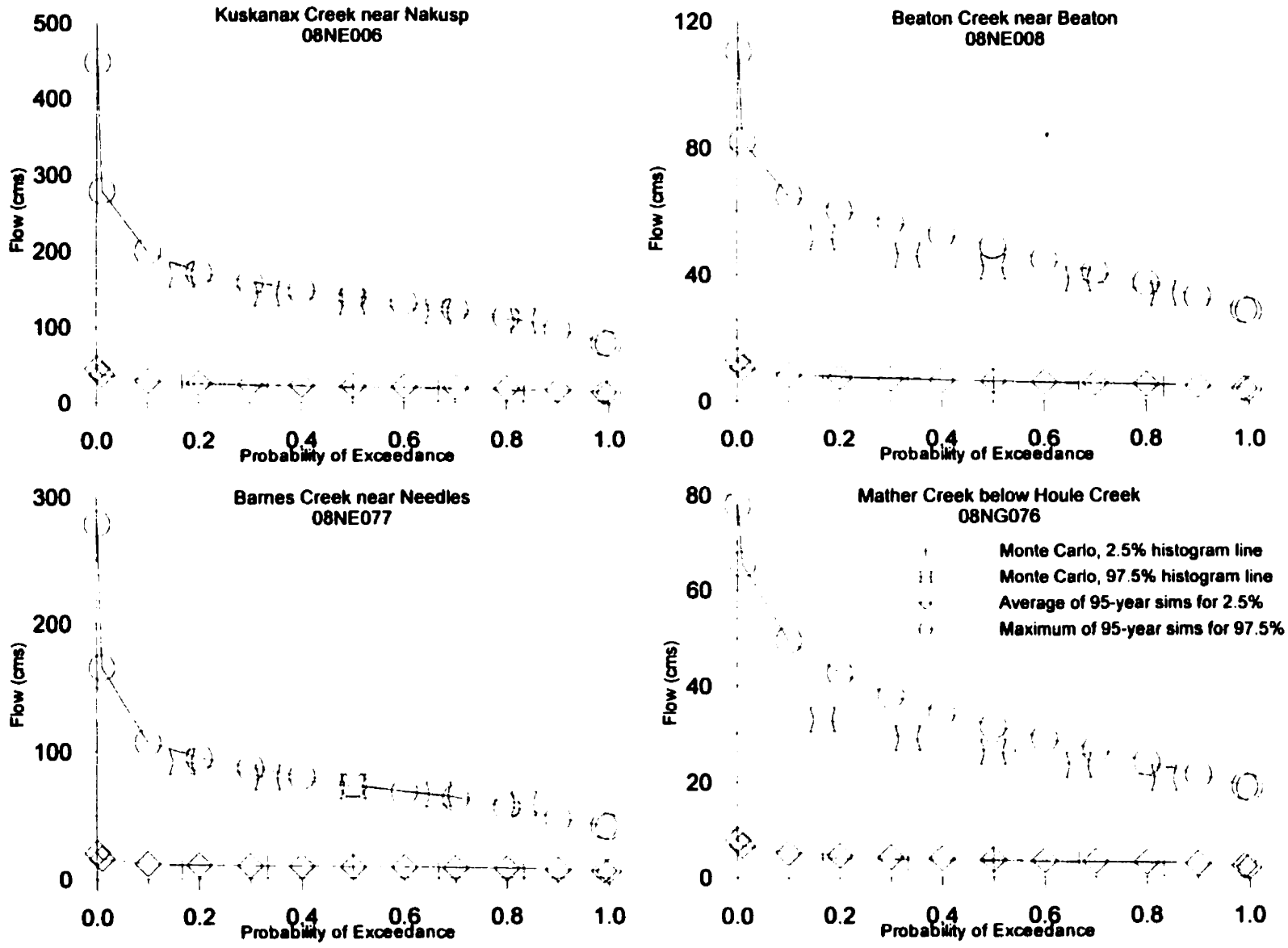


Figure E-4 – Comparison of Confidence Limits – Part 4

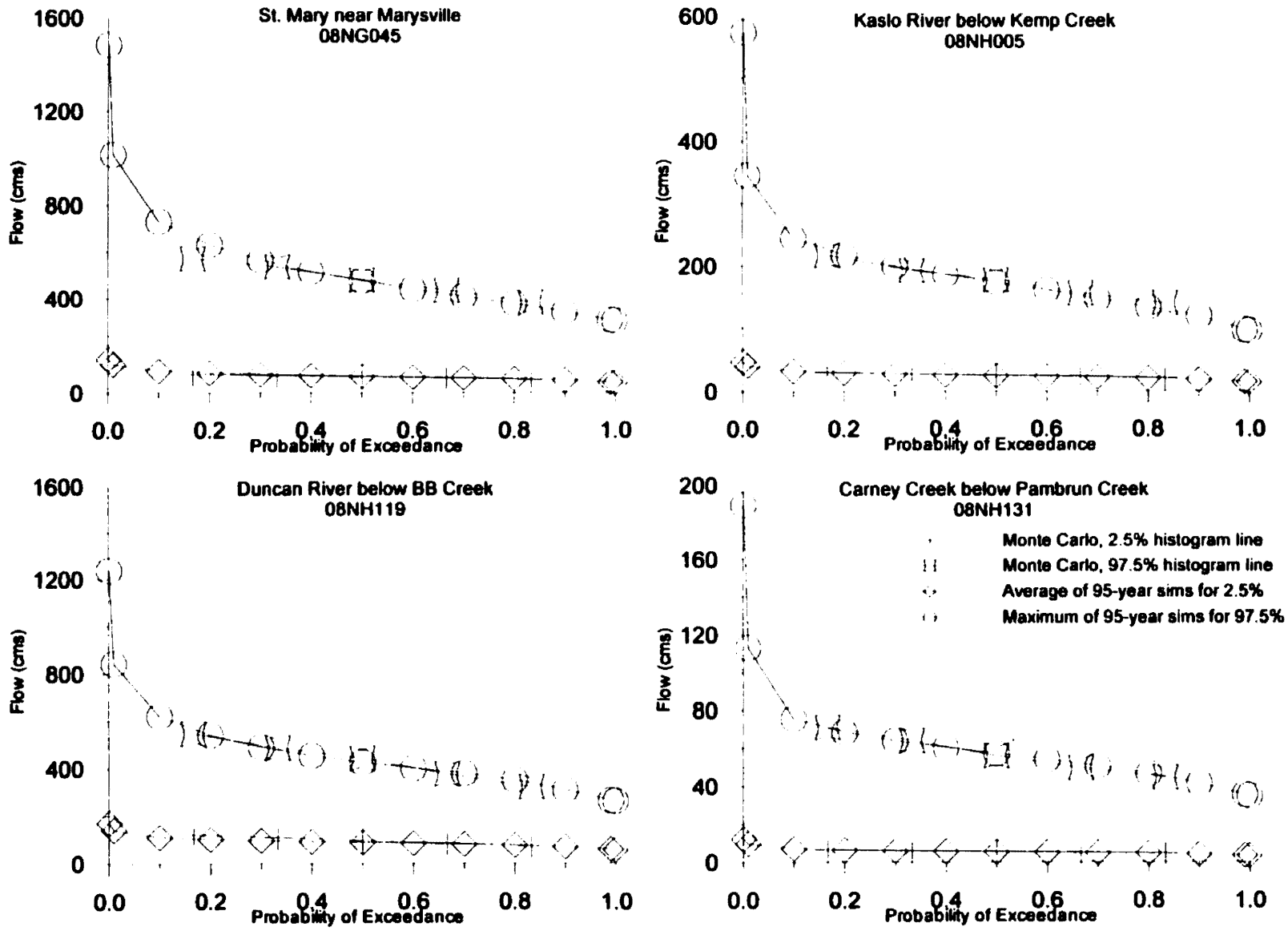


Figure E-5 – Comparison of Confidence Limits – Part 5

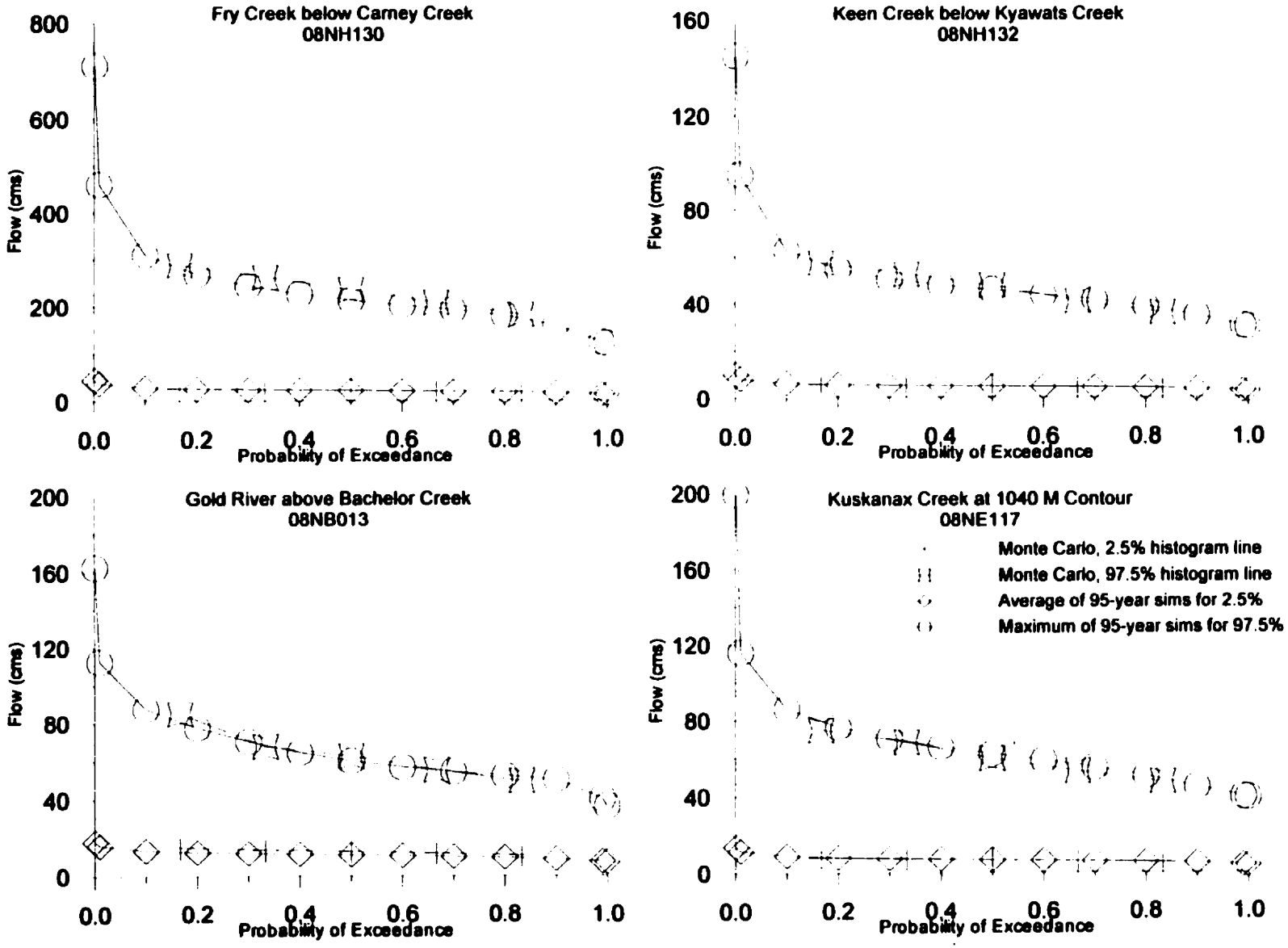


Figure E-6 – Comparison of Confidence Limits – Part 6

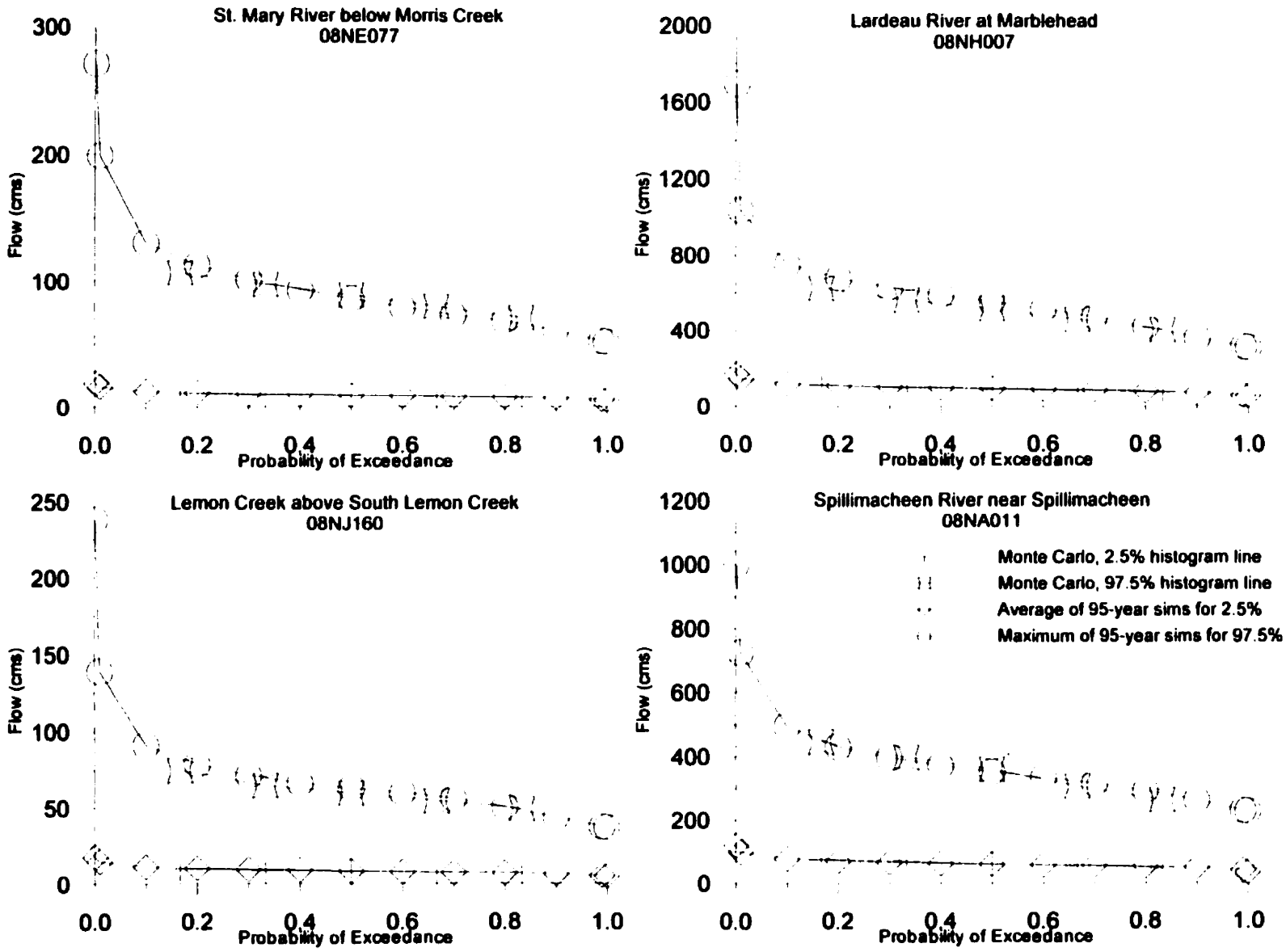


Figure E-7 – Comparison of Confidence Limits – Part 7

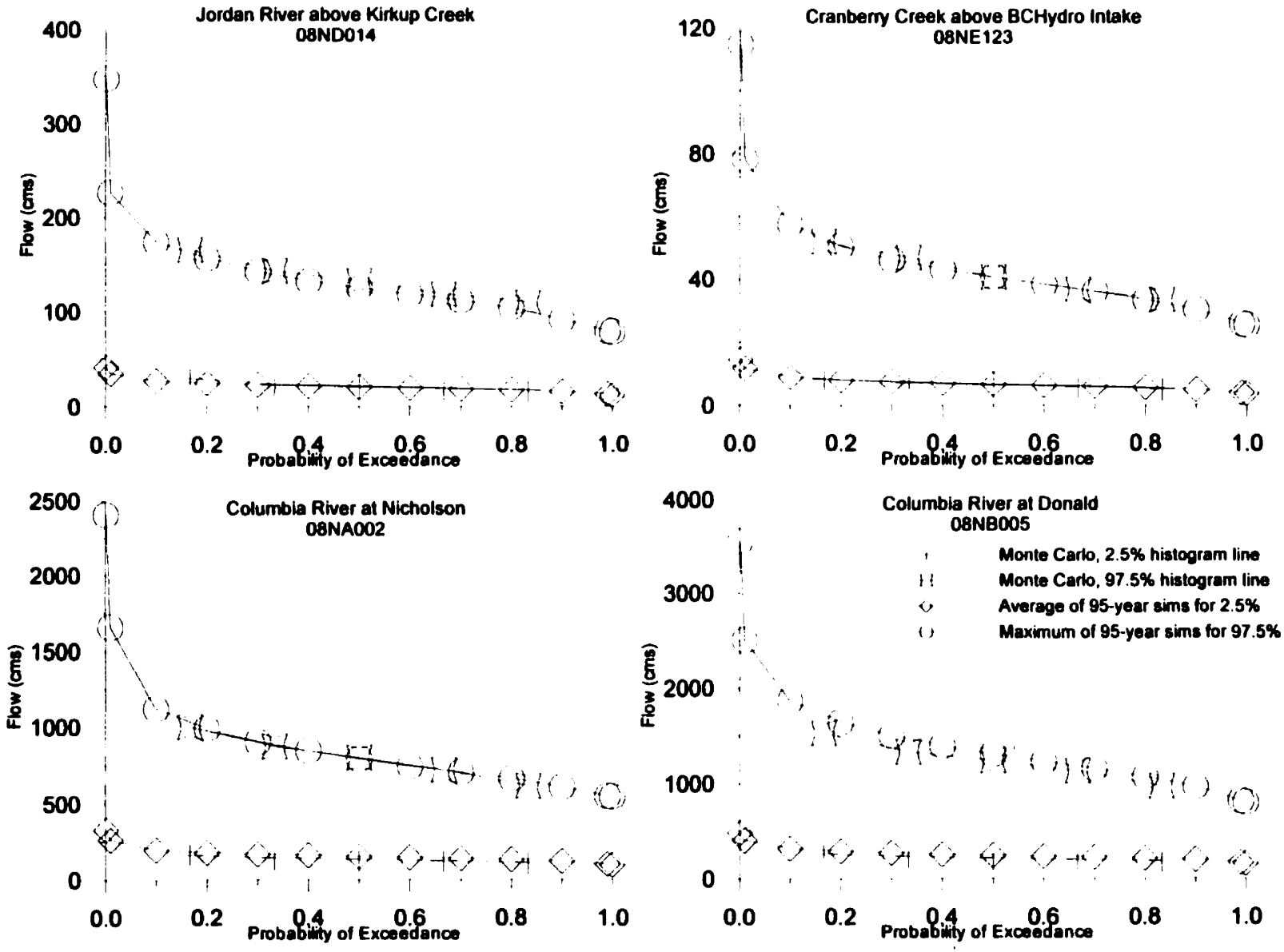


Figure E-8 – Comparison of Confidence Limits – Part 8

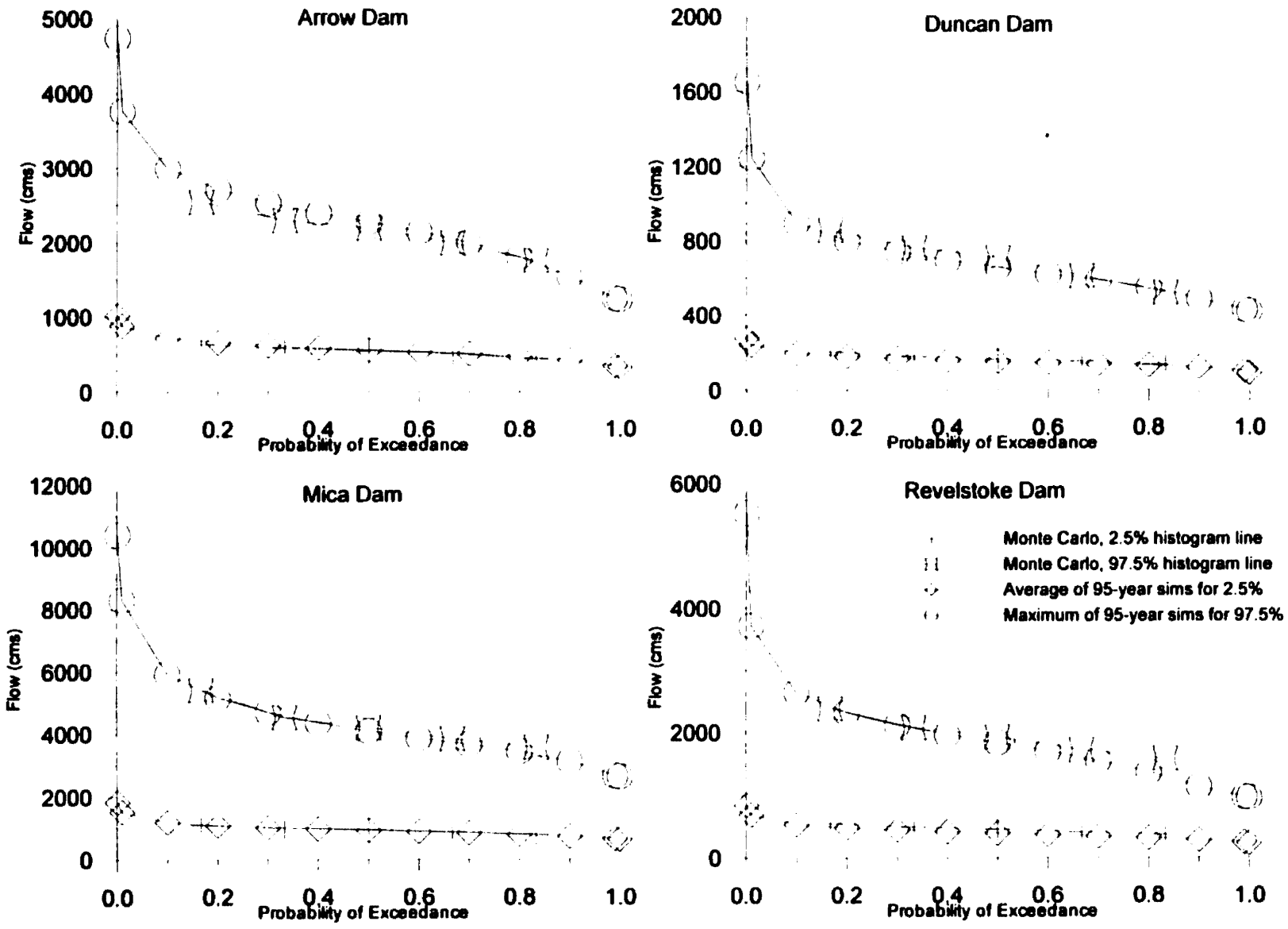


Figure E-9 – Comparison of Confidence Limits – Part 9