University of Alberta

The Analysis of Hydroclimatic Variability and Predictability in Western Canada

by

Adam Kenea Gobena
©

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Water Resources Engineering

Department of Civil and Environmental Engineering

Edmonton, Alberta

Spring 2008

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-45437-4
Our file Notre référence
ISBN: 978-0-494-45437-4

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

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

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

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

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

An analysis of the role of several large-scale climate anomalies in forcing low-frequency variability in the precipitation and streamflow data of Western Canada was carried out using an ensemble of statistical techniques. Wavelet and wavelet coherence analysis of precipitation data from stations across the region shows that at both regional and local scales, precipitation anomalies exhibit significant interannual oscillations that occurred haphazardly. At many stations, the temporal locations of these haphazard oscillations were different from those of large-scale climate anomalies, leading to weak and inconsistent relationships with climate indices. On a seasonal basis, winter precipitation shows modest correlations at 0- to 3-season lead times with the El Nino/Southern Oscillation (ENSO) and the Pacific Decadal Oscillation (PDO) indices. Although there were fewer stations with statistically significant interdecadal oscillations, the relationship with interdecadal climate anomalies was much more consistent than at the interannual scale.

Since the influence of ENSO on Western Canada's precipitation appears to be stronger than the other climate anomalies, it was used as a basis to investigate patterns in streamflow response across the region. From statistical significance testing and cluster analysis, Western Canada was zoned into five spatially coherent streamflow response regions. For each of these regions, a detailed correlation analysis was carried out to examine the value of various climate

indices for predicting seasonal streamflow anomalies. This analysis showed that some of the response regions show better correlations with ENSO indices while others exhibit higher correlations with PDO or other indices.

Based on the understanding derived from the diagnostic analysis, two seasonal ensemble streamflow forecasting models were developed and applied to the Bow and Castle rivers, both located in the headwaters of the South Saskatchewan River basin of Alberta. The first model is a statistical regression model based on the robust M-estimator and the nearest neighbor resampling algorithm, while the second one is based on the ensemble streamflow prediction (ESP) method. Results from the statistical model indicate that forecasts based on climate indices alone possess considerable skill (correlation of 0.65 and up for the Bow River) for forecasts issued early in the season and could thus extend the current forecast lead time by up to two months. In addition, ensemble forecasts were found to possess better skill than deterministic forecasts in terms of economic value because a wide range of forecast users with varying economic costs could potentially benefit from the probabilistic information contained in ensemble forecasts. While the ESP based forecasts also show promising results, they were generally found to be less skillful than the statistical forecasts partly because of hydrologic modeling uncertainties.

Acknowledgement

I am extremely grateful to Dr. Thian Yew Gan for his invaluable guidance, financial support and continuous encouragement during the course of my study at the University of Alberta. I am indebted to the members of my examination committee, Drs. Daryl McCartney, Ed P. Lozowski, Mark Loewen, Uldis Silins and Roger Barry for their valuable comments and suggestions.

I am grateful to the University of Alberta Graduate Assistantship program and the Queen Elizabeth II Doctoral Scholarship for providing financial assistance for this study. Financial support from the Natural Resources Engineering Council of Canada and the Canadian Water Network is also greatly appreciated. Thanks are due to Perry Fedun for his help in organizing the computing facilities of the Water Resources Engineering group.

I am forever indebted to my parents, who have been the source of constant encouragement and inspiration in my life. I am also grateful to my brothers for their guidance and support all along. I would like to thank my wife Chaltu for her patience and understanding.

Last but not least, I would like to express my gratitude to the many educators from the small town teachers at my elementary school in western Oromia (Ethiopia) to the professors at the University of Alberta, whose lessons have enriched my knowledge over the last two decades.

Table of contents

List of Tables

List of Figures

List of Notations and Acronyms

Notations

c	Tuning constant for the "Fair" function
C/L	Cost-loss ratio
C_δ	Reconstruction factor
d	Mahalanobis distance
E	Expected expense associated with using a forecast
f	Forecast probability of an event; also number of false alarms
F	False-alarm rate
$f(t)$	Real signal varying as a function of time
h	Number of hits
H	Hit rate
K, K'	Number of nearest neighbors
l	Dimension of a feature vector
M	Number of predictors in a model
m	Number of misses
N	Number of cases for a variable
o	Observation probability of an event (either zero or one)
$\bar{o}$	Climatological relative frequency
p	cumulative probability; also Pressure (Pa)
P	Precipitation (mm)
P'	Precipitation anomaly
$\overline{P}$	Precipitation climatology
r	Number of correct rejections
R^2	Coefficient of determination
$R^2(\xi)$	Wavelet coherence at scale ξ
$r_{xy\|z}$	Partial correlation between variables x and y given variable z
S	Sample variance
t	Time
T	Temperature (°C)

T'	Temperature anomaly
$\overline{T}$	Temperature climatology
u	Zonal wind speed (ms^{-1}); also uniform random number
v	Meridional wind speed (ms^{-1})
V	Relative economic value
Vx	Vertically integrated zonal moisture flux (Kg m^{-1}s^{-1})
Vy	Vertically integrated meridional moisture flux (Kg m^{-1}s^{-1})
W	Wavelet spectrum
W^2	Wavelet power spectrum
w_k	Weight assigned to the K^{th} nearest neighbor
w^r	Residual weight
w^x	Leverage weight
X	Independent variable (predictor)
x'	Wavelet band-pass filtered signal
y	Dependent variable (predictand)
$\hat{y}$	Conditional mean estimate of the predictand
$\tilde{y}$	Ensemble member
Z	Geopotential height
z	Random normal variate
α, η	Parameters for weight kernel
β	Regression coefficient
γ	Translation parameter
δj	Factor for scale averaging
δt	Sampling interval
λ	Gaussian kernel reference bandwidth
ξ	Scale parameter
ψ_0	Mother wavelet
ω	Frequency
ω_0	Nondimensional frequency (Morlet wavelet)

Acronyms

AGCM3	CCCma third generation atmospheric general circulation model
AHCCD	Adjusted historical Canadian climate data
BC	British Columbia
CCCma	Canadian Centre for Climate Modeling and Analysis
CNP	Central North Pacific sea level pressure anomaly index
CP	Canadian Prairies
CRPC1	Leading principal component of Central Region precipitation anomalies
CRR	Conceptual rainfall-runoff
DJF	December-January-February
ENSO	El Nino-Southern Oscillation
EP	East Pacific pattern
ERPC1	Leading principal component of Eastern Region precipitation anomalies
ESP	Ensemble streamflow prediction
GCM	General circulation model
GCV	Generalized cross-validation
GEM	Global Environmental Multiscale model
HFP	Historical Forecast Project
ITS	Index time series
K-NN	K nearest neighbors
LWP	Locally weighted polynomial
MLR	Multiple linear regression
MOS	Model output statistics
NCEP	National Centers for Environmental Prediction of the U.S.A.
Nino3	Sea surface temperature anomaly averaged over 150°W-90°W, 5°S-5°N
NOAA	National Oceanic and Atmospheric Administration of the U.S.A
NPO	North Pacific Oscillation
NWP	Numerical weather prediction

PC	Principal component
PCA	Principal components analysis
PDF	Probability density function
PDO	Pacific Decadal Oscillation
PNA	Pacific/North American pattern
RHBN	Reference Hydrometric Basin Network
ROC	Relative operating characteristic
RPS	Ranked probability score
RPSS	Ranked probability skill score
SAC-SMA	Sacramento soil moisture accounting model
SAWP	Scale-averaged wavelet power
SLP	Sea level pressure
SOI	Southern Oscillation index
SSA	Standardized streamflow anomalies
SSRB	South Saskatchewan River basin
SST	Sea surface temperature
SWE	Snow water equivalent
SWEI	Sub-basin snow water equivalent index
VIMF	Vertically integrated moisture flux
WP	West Pacific pattern
WRPC1	Leading principal component of Western Region precipitation anomalies

Chapter 1 Introduction

1.1. Statement of the Problem

Rivers in Western Canada are the primary sources of water for a variety of uses. For instance, Over 57% of the water allocated to date in Alberta originates in the South Saskatchewan River basin (SSRB) although the basin covers only about a quarter of the provincial land surface area (Alberta Environment, 2006). On the west coast, British Columbia generates about 90% of its electricity by hydropower (Hsieh and Tang, 2001). The availability of water in Western Canada's rivers depends mainly on the air-sea interaction over the surrounding ocean basins that form the source of moisture for precipitation occuring over the region. The primary source of moisture for precipitation during the winter season is the Pacific Ocean (Liu *et al.*, 2004). Even though snowfall constitutes a minor proportion of the annual precipitation in parts of Western Canada, spring snowmelt can contribute up to 80% of the annual runoff in some rivers in the Canadian Prairies (Granger and Gray, 1990). Thus, a diagnostic analysis of Western Canada's hydroclimatic variability and its dependence on the dominant modes of Pacific climate variability can provide useful information on the timing and quantity of water availability, which is vital to efficient water resources management.

The Pacific Ocean climate system exhibits strong low-frequency variability modes at both interannual and interdecadal scales. The most prominent source of interannual climate variability is the El Nino/Southern Oscillation (ENSO). ENSO is a coupled ocean-atmospheric circulation anomaly originating in the tropical Pacific but its impact extends into the extratropics. Other sources of interannual climate variability of interest to the Northern Hemisphere include teleconnection patterns that originate in the North Pacific basin such as the Pacific/North America (PNA) pattern, West Pacific (WP) pattern, and East Pacific (EP) pattern.

The primary mode of Pacific interdecadal climate variability is the Pacific Decadal Oscillation (PDO), which is related to sea surface temperature (SST) anomalies of the Pacific basin north of 20°N.

During the last several decades, various studies have documented the relationships between ENSO and interannual hydroclimatic variability in the western regions of USA and Canada (e.g. Yarnal and Diaz, 1986; Ropelewski and Halpert, 1987; Kiladis and Diaz, 1989; Kahya and Dracup, 1993; Shabbar *et al.*, 1997). According to Shabbar *et al.* (1997), ENSO affects precipitation variability over a large portion of southern Canada extending from British Columbia, through the Prairies, and into the Great Lakes region. Annual streamflows across Western Canada have been correlated with two indices that are commonly used as a measure of the strength of ENSO, namely the Nino3 SST index (Coulibaly and Burn, 2004) and the Southern Oscillation Index (SOI) (Woo and Thorne, 2003). However, possible spatial patterns in the ENSO-streamflow relationships, and periods during which the ENSO information may provide predictive information for Western Canada's streamflow has yet to be objectively identified.

The PNA pattern is characterized by opposite geopotential height anomaly centers over the Aleutian Islands and Western Canada. The anomalies associated with different phases of PNA alter the normal upper air flow, thus affecting hydroclimatic patterns over the North American region (Shabbar *et al.*, 1997). The PNA pattern has been related to variations in Western Canadian streamflow (e.g. Moore, 1996; Moore and McKendry, 1996; Woo and Thorne, 2003; Coulibaly and Burn, 2004) and snow water equivalent (Hsieh and Tang, 2001).

Interdecadal climatic variations in western North America have been related to interdecadal variations in the North Pacific climate, with the oceanic component described by the PDO (Mantua *et al.*, 1997). Cayan *et al.* (1998) found that decadal precipitation variability in western North America could account for up to 50% of the variance of annual precipitation. McCabe and Dettinger (1999) showed that the strength of the relationships between western US precipitation and SOI closely resembles decadal variations in the PDO index. Similarly, Brown and Comrie (2004) found that spatial variations in the precipitation-ENSO relationships in western US are synchronized with the PDO phase shift. The PDO regime has also been shown to modulate winter temperature responses during El Nino and La Nina events across most of Western Canada (Bonsal *et al.*, 2001). However, to date little attention has been given to the study of decadal to interdecadal variations in other hydroclimatic variables in Western Canada.

While previous studies focusing on Western Canada have provided significant insights on various aspects of the region's hydroclimatic variability, there are still several areas which have not been addressed adequately. Most of the studies reported in the literature have heavily depended on stationary techniques such as harmonic analysis, event-based compositing and/or direct correlation analysis to establish relationships between anomalous ocean/atmosphere circulation patterns and hydroclimatic variables (e.g. Yarnal and Diaz, 1986; Ropelewski and Halpert, 1987; Kahya and Dracup, 1993; Moore, 1996; Shabbar *et al.*, 1997; Woo and Thorne, 2003) even though many large-scale climate variability modes such as ENSO are non-stationary (e.g. Torrence and Compo, 1998). Currently, there is a growing trend of using non-stationary approaches such as wavelet transforms for the analysis of hydroclimatic variability in different parts of the world. Wavelet transforms have recently been used to study the interannual variability of Canadian annual and seasonal streamflows and their relationships with some large-scale climate patterns (Coulibaly and Burn, 2004, 2005).

Past research has primarily focused on establishing links between hydroclimatic variations and large-scale climate anomalies where climate anomalies are treated as independent of one another. This is generally not true because two climate anomalies with different characteristic time scales may have a synergetic or destructive effect on hydroclimatic variables depending on whether they are in phase or out of phase with each other. Thus, it is necessary to understand the degree to which Western Canada's hydroclimatic responses to ENSO are affected by extratropical Pacific climate modes (e.g. PNA, PDO, etc). Understanding the interaction between variability at different scales is important because synergetic interactions between inter-decadal and interannual modes of variability may produce extreme hydroclimatic events such as droughts and floods.

This research deals with a systematic analysis of interannual to interdecadal scale precipitation and streamflow variability in Western Canada, with the ultimate objective of developing long-range streamflow forecasting models that incorporate large-scale climate information. Essentially, the following questions will be addressed: Are there dominant spatial and/or temporal modes of low frequency variability in the precipitation and streamflow of Western Canada? If so, how are these modes related to large-scale ocean-atmosphere climate modes? What is the relative importance of each mode? How do hydroclimatic responses at one frequency interact with those at another frequency? To address these questions, the statistical characteristics of precipitation, streamflow and indices of various climate modes will be analysed using wavelets, statistical significance testing, compositing and correlation analysis. From these findings, the feasibility of using large-scale climate information for accurate prediction of basin-scale seasonal streamflow volumes will be examined. Two watersheds in the South Saskatchewan River Basin (SSRB) will be used to test and verify the skills of the various streamflow forecast models to be developed in this study.

1.2. Research Objectives

The objectives of this study were:

1. Identify the dominant spatial and/or temporal modes of low-frequency precipitation and streamflow variability across Western Canada;

2. Relate any detected precipitation and streamflow signals to large-scale climate modes using frequency and time domain analyses;
3. Investigate the relative influence of the various large-scale climate modes on Western Canada's hydroclimatic variability;
4. Develop and verify long-range streamflow forecast models that incorporate predictors identified from large-scale climate modes relevant to the Western Canadian region.

1.3. Climatology of Western Canada

For the purpose of this study, Western Canada refers to the southern portions of the three Canadian Prairie provinces of Alberta, Saskatchewan and Manitoba, and British Columbia (BC) (Figure 1.1). The climate along the BC west coast is controlled by Pacific westerlies, which must rise and condense upon encountering the Coast Mountains of BC. The Pacific coast of BC experiences moderate temperatures and high precipitation year round as a result of this airstream. Mean annual precipitation exceeds 3000 mm in some areas of the Pacific coast, with most of the precipitation occuring during the autumn and winter seasons. Although the westerly flow is still common in the interior of BC, precipitation decreases quickly as one moves eastwards, averaging about 400 mm annually, due to the rain-shadow effect of the coast mountains (Phillips, 1990). While the mountains and higher plateau surfaces in the interior of BC are moist, the valley systems are dry, because the Pacific westerlies must descend to reach them (Hare and Thomas, 1974).

The Canadian Prairies (CP) experience a continental climate with bitterly cold winters, short and warm summers, and light precipitation regimes (Hare and Thomas, 1974). The Prairies are mainly flat and as such topography is not a key factor in climatic control. As a result, they are exposed to a variety of air masses originating in the Arctic (cold, dry air), Pacific (cool, moist air), American (dry, continental air) and Gulf of Mexico (humid air). The interactions between these air masses are responsible for the wide variation of weather observed in this semi-arid region. The Prairies generally receive less than 500 mm of annual precipitation. In winter, most of the moisture for precipitation in the Prairies comes from the Pacific Ocean (Liu *et al.*, 2004). Although the Western Cordillera significantly reduces the frequency of cyclonic precipitation, some Pacific air streams succeed in penetrating past this obstruction, bringing a considerable fraction of the water vapor and heat from the Pacific to the Prairies. Winter precipitation is usually light, amounting to less than 25 mm during each of the winter months (Hare and Thomas, 1974). In summer, the primary moisture source for precipitation is the Gulf of Mexico (Liu *et al.*, 2004). Precipitation is maximum in late spring or in summer when a thermal low is often present over the warm, sunny Prairies (Phillips, 1990). The eastern Prairies receive higher annual precipitation than the western parts due to a near humid type of climate resulting from frequent influxes of moist air from the Gulf of Mexico (Hare and Thomas, 1974).

1.4. Thesis Outline

The thesis is organized in seven chapters. Chapter 2 deals with diagnostic analysis of low frequency variability and teleconnection of precipitation data obtained from Western Canadian station observations. Chapter 3 focuses on the analysis of teleconnections between Western Canadian streamflow variability and large-scale climate anomalies. In Chapter 4, a basin-scale hydroclimatic variability and predictability analysis is carried out for the Alberta portion of the SSRB. On the basis of the understanding of the climate diagnostics gained from the previous three chapters, two different methods to forecast the seasonal streamflow of the SSRB are attempted in the next two chapters. In Chapter 5, a statistical modeling approach is explored, where a robust regression and a nearest neighbor resampling algorithm are used to generate ensemble seasonal streamflow forecasts. In Chapter 6, a hydrologic model based approach is explored, where a technique for incorporating seasonal climate forecasts from numerical weather prediction models into the ensemble streamflow prediction (ESP) system is developed. Finally, the overall conclusions and suggestions for future work are presented in Chapter 7.

References

Alberta Environment. (2006). *Water in Alberta.* Alberta Environment website http://www.environment.gov.ab.ca (accessed August 2006).

Bonsal, B. R., Shabbar, A., and Higuchi, K. (2001). Impacts of low frequency variability modes on Canadian winter temperature. *Int. J. Climatol.,* 21, 95-108.

Brown, D. P., and Comrie A. C. (2004). A winter precipitation 'dipole' in the western United States associated with multi-decadal ENSO variability. *Geophys. Res. Lett.,* 31, L09203, doi:10.1029/2003GL018726.

Cayan, D. R., Dettinger, M. D., Diaz, H. F., and Graham, N. E. (1998). Decadal variability of precipitation over western North America. *J. Climate*, 11, 3148-3166.

Coulibaly, P., and Burn, D. H. (2004). Wavelet analysis of variability in annual Canadian streamflows. *Water Resour. Res.,* 40, W03105, doi:10.1029/2003WR002667.

Coulibaly, P., and Burn, D. H. (2005). Spatial and temporal variability of Canadian seasonal streamflows. *J. Climate,* 18(1), 191–210.

Granger, R. J., and Gray, D. M. (1990). A net radiation model for calculating snowmelt in open environments. *Nordic Hydrology,* 21, 217-234.

Hare, F. K., and Thomas, M. K. (1974). *Climate Canada*, Wiley, Toronto, 256pp.

Hsieh, W., and Tang, B. (2001). Interannual variability of accumulated snow in the Columbia basin, British Columbia. *Water Resour. Res.,* 37, 1753-1759.

Kahya, E., and Dracup, J. A. (1993). U.S. Streamflow Patterns in Relation to the El Nino/Southern Oscillation. *Water Resour. Res.,* 29, 2491-2503.

Kiladis, G. N., and Diaz, H. F. (1989). Global climatic anomalies associated with extremes of the southern oscillation. *J. Climate,* 2, 1069-1090.

Liu, J., Stewart, R. E., and Szeto, K. K. (2004). Moisture transport and other hydrometeorological features associated with the severe 2000/01 drought over the western and central Canadian Prairies. *J. Climate,* 17, 305-319.

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., and Francis, R. C. (1997). A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.,* 78, 1069–1079.

McCabe, G. J., and Dettinger, M. D. (1999). Decadal variations in the strength of ENSO teleconnections with precipitation in the western United States. *Int. J. Climatol.,* 19, 1399-1410.

Moore, R. D. (1996). Snowpack and runoff responses to climatic variability, southern Coast Mountains, British Columbia. *Northwest Science,* 70, 321-333.

Moore, R. D., and McKendry, I. G. (1996). Spring snowpack anomaly patterns and winter climatic variability, British Columbia, Canada. *Water Resour. Res.,* 32, 623-632.

Phillips, D. (1990). *The climates of Canada*, edited by D. Francis, D. Gullett, and E. Truhlar, Canadian Government Publishing Centre, 176pp.

Ropelewski, C. F., and Halpert, M. S. (1987). North American temperature and precipitation patterns associated with the El Nino/Southern Oscillation (ENSO). *Mon. Weather Rev.,* 114, 2352-2362.

Shabbar, A., Bonsal, B., and Khandekar, M. (1997). Canadian precipitation patterns associated with the Southern Oscillation. *J. Climate,* 10, 3016-3027.

Torrence, C., and Compo, G. P. (1998). A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.,* 79(1), 61-78.

Woo, M.-K., and Thorne, R. (2003). Comment on 'Detection of hydrologic trends and variability'. *J. Hydrol.,* 277, 150-160.

Yarnal, B., and Diaz, H. F. (1986). Relationships between extremes of the Southern Oscillation and the winter climate of Anglo-American Pacific coast. *J. Climatol.,* 6, 197-219.

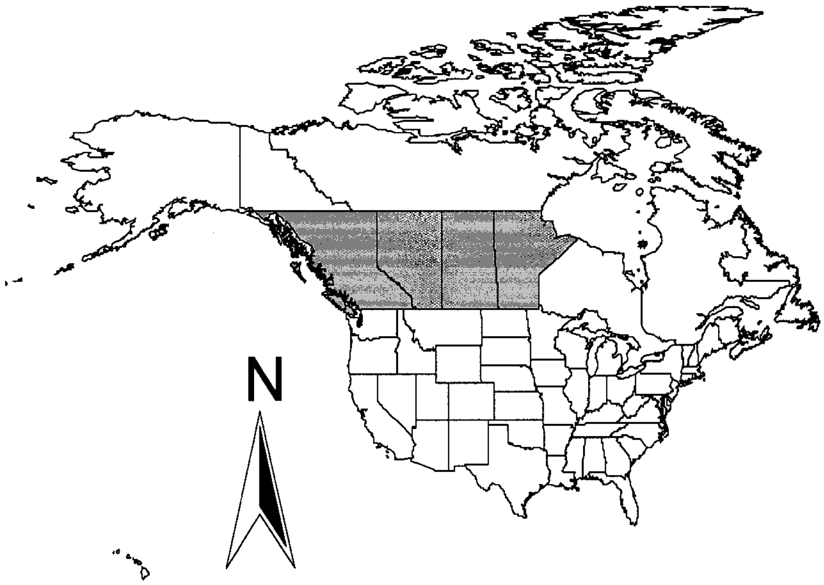


Figure 1.1. Location of the four Western Canadian provinces relative to the North American continent.

Chapter 2 Wavelet and Teleconnection Analysis of Western Canadian Precipitation

2.1. Introduction

Studies based on instrumented records show the existence of inter-decadal precipitation variability in western North America, e.g., regional dipole or seesawing of precipitation pivoting near 40°N (Dettinger *et al.*, 1998; Cayan *et al.*, 1998), Canadian Prairies (CP) (Bonsal *et al.*, 1999), and others. Low frequency fluctuations at decadal and higher time scales could give rise to large floods and extreme droughts that bring severe damages and grave economic losses. Examples of such extremes in western North America include droughts in the CP (Gan, 1998) and California (Rogers, 1994), and flooding in the west coast (Lott *et al.*, 1997). North America experiences low-frequency climate variability partly because of the Pacific Ocean that has sufficient storage for long-term memory.

Precipitation processes are highly variable spatially and temporally and their properties, such as the intensity-duration-frequency relationship, are often based on the probability theory of stochastic processes that assume an infinite fractal dimension. However, recent discoveries of correlation dimensions (D_2) of precipitation (e.g. Islam *et al.*, 1993), and multifractal features of precipitation (Lovejoy and Schertzer, 1985; 1990; Svensson *et al.*, 1996; Gan *et al.*, 2002), suggest that precipitation more likely follows a chaotic rather than a stochastic system. Precipitation exhibits multifractal properties whenever its D_2 is associated with certain density functions $h(D_2)$, which alone does not describe its spatial or temporal singularities. Similarly, the classical Fourier power spectrum can identify the dominant frequencies in a stationary time series but cannot provide any information about the temporal locations of the dominant events. It is now well established that certain large-scale climate variability modes such as the

El Nino/Southern Oscillation (ENSO) are non-stationary (e.g. Torrence and Compo, 1998), and these non-stationary features are also manifested in hydroclimatic time series (e.g. Coulibaly and Burn, 2004; Mwale *et al.*, 2004). Wavelets use fast decaying oscillating waveforms for signal decomposition, which makes them suitable for studying the power fluctuation in non-stationary time series in the frequency and time domains simultaneously. Wavelets have been used to study daily streamflow characteristics (Smith *et al.*, 1998), daily rainfall-runoff relationships (Labat *et al.*, 2000; Lafreniére and Sharp, 2003), annual and seasonal streamflow variability (Coulibaly and Burn, 2004; 2005), seasonal rainfall and sea surface temperature (SST) variability (Mwale *et al.*, 2004, Mwale and Gan, 2005), and the temporal characteristics of the southern oscillation index (SOI) and the Nino3 SST index (Torrence and Compo, 1998).

This study has three objectives: (1) To identify the dominant oscillations of precipitation data from Western Canada and their temporal variations using wavelet transforms; (2) To relate the detected precipitation signals to some large-scale climate anomalies or prominent teleconnection patterns found over the Pacific using both frequency and time domain analyses; and (3) from the above findings, attempt to address issues such as whether different precipitation time scales correspond to different teleconnection patterns, whether decadal or higher-level precipitation variations arise from climate dynamics that are separate from interannual or lower-level variations, and the feasibility of seasonal precipitation predictions by teleconnection with climate indices.

2.2. Data Description

2.2.1. Precipitation

The precipitation data used in the wavelet analysis were obtained from the Adjusted Historical Canadian Climate Data (AHCCD) database, which contains station observations statistically adjusted for known problems such as missing values, instrument changes, trace events, wind under-catch, evaporation and wetting loss (Mekis and Hogg, 1999). Monthly precipitation data for 21 stations with long historical records were extracted from the AHCCD database (see Table 2.1 and Figure 2.1). To avoid getting results that suffer from low elevation bias, we selected stations with elevations ranging from 8 m at Quatsino, British Columbia (BC) to 1073 m at Calgary, Alberta (see Table 2.1). The wavelet analysis was carried out on standardized seasonal precipitation anomalies of the 21 stations for 1914–2001. Since our data cover both the CP and BC, we can compare results from precipitation under the direct influence of the Pacific westerly wind and that under the rain shadow effect of the Cordillera (Canadian Rockies).

The CP consists of undulated grassy or cultivated plains with countless small ponds and few low hills. BC generally experiences much wetter climate because of the wet westerly flow from the Pacific. Some parts of BC receive over 1500 mm of annual precipitation while the Prairies generally get less than 500 mm of precipitation because cyclonic precipitation rarely reaches these places from either west or east coasts, and partly because of the frequent visit of the dry Arctic air. Some Pacific air streams succeed in penetrating past the Cordillera, bringing a considerable fraction of the water vapor and heat from the Pacific to the Prairies. Precipitation is usually light in winter and maximum in late spring or in summer when a thermal low is often present over the warm, sunny Prairies, e.g., the Alberta lows ("Alberta clippers"), where there could be up to 30 days of thunderstorms per year (Phillips, 1990).

Among the 21 climate stations selected, Princeton, BC, located in the Okanagan River basin receives the lowest mean annual precipitation of about 370 mm. Annual precipitation increases north and westward to well over 2000 mm at the Pacific coast stations of Quatsino, BC and Prince Rupert, BC and eastward to 598 mm at Winnipeg, Manitoba. Agassiz and Bella Coola, BC, which are under the influence of moist Pacific westerly winds, also receive well over 1000 mm of annual precipitation (Table 2.1). In winter, northern Arctic air sweeps down over the Prairies. At a relatively high altitude, southern Alberta has relatively less of the Arctic air than the Chinook (a foehn-type wind that occurs several times a year). While winter and autumn are wet in parts of BC, February is usually the driest month in the Prairies, with precipitation amounting to less than 25 mm per month. The eastern Prairies receive higher annual precipitation than the western parts due to a near humid type of climate resulting from frequent influxes of moist air from the Gulf of Mexico (Hare and Thomas, 1974).

2.2.2. Large-scale Climate Indices

Globally, ENSO is by far the most widely documented source of interannual climate variability. ENSO has its origin in the equatorial Pacific basin. The warm phase of ENSO (known as 'El Nino') is associated with an eastward spreading of warm ocean waters due to weakening and/or reversal of trade winds and ocean currents. El Nino occurrences affect fishing along the coast of Peru because the warm waters displace nutrient-rich cold water off the west coast of South America. There are also significant changes in precipitation patterns because of an eastward shift in the location of the equatorial convective loop. During the cold phase of ENSO (called 'La Nina'), anomalously cool ocean waters move from east to west and the convective loop shifts further westward.

ENSO occurrences are near global climate phenomena because their effects have been linked to various climatic anomalies worldwide, e.g., precipitation patterns in western and continental USA, and east of the Rockies (Ropelewski and Halpert, 1986), Caribbean and tropical America (Rogers, 1988), and north-east Brazil (Kane, 1997). ENSO teleconnections to the extratropical Northern Hemisphere occur during winter, when its "mature phase" is often associated with anomalously deep central North Pacific lows (Rasmusson and Carpenter, 1982). Albeit ENSO occurs every 3 to 7 years, it also undergoes mysterious decadal variations – with a relatively large intensity in the late 1800s to early 1900s, followed by a relatively low intensity between 1920 and 1950, and then an increased activity after 1960 (Torrence and Compo, 1998).

In this study, we investigate the effects of ENSO through teleconnection of the Nino3 index, a time series of equatorial Pacific SST anomalies averaged over the equatorial Pacific (5°S–5°N, 150°W–90°W), and the Southern Oscillation Index (SOI), a time series of normalized monthly differences in sea level pressure (SLP) at Tahiti (≈ 150°W, 18°S) and Darwin (≈ 130°E, 13°S) (Allan *et al.*, 1991), which are commonly used as a measure of the strength of ENSO. We also study the effects of extratropical Northern Hemisphere ocean/atmosphere circulation through teleconnections of the Pacific/North America pattern (PNA), West Pacific pattern (WP), East Pacific pattern (EP), Central North Pacific sea level pressure index (CNP) and Pacific Decadal Oscillation (PDO). The PNA pattern represents a quadripole of 700 mbar geopotential height anomalies, with opposite anomalies centered over the Aleutian Low and Western Canada, and between the Hawaiian Islands and southeastern US (Wallace and Gutzler, 1981),

$$PNA = Z(170^\circ \mathrm{W}, 20^\circ \mathrm{N}) - Z(165^\circ \mathrm{W}, 45^\circ \mathrm{N}) + Z(115^\circ \mathrm{W}, 58^\circ \mathrm{N}) - Z(98^\circ \mathrm{W}, 30^\circ \mathrm{N}) \qquad (2.1)$$

where Z is the geopotential height. PNA exhibits a wide range of variability in the Northern Hemisphere extratropics, and has been teleconnected to the Great Salt Lake levels (Moon and Lall, 1996) and streamflow along the US west coast (Cayan and Peterson, 1989).

The WP pattern consists of a north-south dipole of anomalies centered over Kamchatka Peninsula, portions of southeastern Asia and the lower part of western North Pacific. The EP pattern reflects a north-south dipole of height anomalies over the eastern North Pacific, with its northern center located around Alaska and the west coast of Canada (Wallace and Gutzler, 1981). The CNP index is SLP anomalies averaged over central north Pacific window (35°N–55°N, 170°E–150°W) (Cayan and Peterson, 1989). The PDO, which represents inter-decadal oscillations in the extratropical north Pacific climate system, is represented by a time series of the leading principal component of the North Pacific SST anomalies poleward of 20°N (Mantua and Hare, 2001).

Some of the selected climate indices are significantly correlated to one another, e.g., winter season PDO, CNP, PNA and WP indices are all significantly correlated to SOI and Nino3 (Table 2.2). Warm ENSO episodes are associated with negative mid-tropospheric geopotential height anomalies over the North Pacific, which should lead to strong SLP anomalies in the central and North Pacific regions (Horel and Wallace, 1981). In winter strong SLP anomalies at the central North Pacific consistently correspond to ENSO events with Nino3 ≥ 2 (Cayan and Peterson, 1989). The PNA index is significantly correlated to CNP for all seasons and to PDO except during summer. Although PDO is an oceanic (SST) variability mode, its signature extends through the depth of the troposphere and is manifested in the mid-troposphere as persistence in the PNA (Mantua and Hare, 2001). WP is significantly correlated to Nino3 and SOI in winter and spring, to CNP in spring and autumn and to PNA in spring. EP is linked to Nino3, SOI and CNP in spring.

2.3. Research Methodology

To extract the dominant oscillations for Western Canadian precipitation, we used the continuous Morlet wavelet. The relations between the detected oscillations and some large-scale climate anomalies that are known to exert influences on the climate of western North America were investigated using principal components, wavelet coherence and multi-scale correlation analyses. The relations between precipitation and climate indices were further explored on seasonal basis using composite and correlation analysis. The wavelet analysis is briefly described below.

In the classical Fourier transform, one has only the frequency ω as a parameter in the basis functions of the form $e^{i\omega t}$. These basis functions are globally uniform in time and as such the Fourier transform does not contain any time dependence of the signal. A wavelet transform overcomes the limitations of the Fourier transform by using fast decaying waveforms that can be dilated and translated to reveal oscillations at both high and low frequencies. The wavelet transform is mathematically defined as

$$W(\xi,\gamma) = \xi^{-1/2} \int f(t)\psi^{*}\left(\frac{t-\gamma}{\xi}\right)dt \qquad \xi > 0,\ \gamma \in R \qquad (2.2)$$

where *f(t)* is a real signal, $\psi(x)$ is the basic wavelet function satisfying certain admissibility conditions, ξ is a scale or dilation parameter, γ is a translation parameter, and ψ^{*} is the complex conjugate of ψ (Lau and Weng, 1995). Visually, the dilation parameter ξ controls the width and rate of the local oscillation of the basic wave and can intuitively be thought of as controlling the frequency scale, whereas the translation parameter γ moves the wavelet through the time domain. By varying ξ and γ, one can construct a picture of how the wavelet power spectrum of a signal varies in the time-frequency domain, thus making it suitable for studying non-stationarity in time and in frequency. Since the frequency resolution depends on the scale ξ, both high and low frequency fluctuations of a process can be captured by varying either the frequency or time scale. The wavelet power spectrum is defined as $|W_t(\xi,\gamma)|^2$. When comparisons of the power spectra across several time series are desirable, the wavelet power spectrum may be normalized by the variance of the original time series.

An ideal wavelet for climate signal detection should have the ability to detect both time-dependent amplitude and phase changes for different frequencies in the time series (Lau and Weng, 1995). Due to its complex nature, the Morlet wavelet is adapted to such oscillatory behaviors and it has been employed in this study. The Morlet wavelet consists of a plane wave modulated by a Gaussian,

$$\psi_0(t) = e^{i\omega_0 t} e^{-t^2/2} \qquad (2.3)$$

Figure 2.2 shows the Morlet wavelet transform of the seasonal precipitation anomaly (Figure 2.2a) at Calgary, Alberta. The thick contour lines in Figure 2.2b enclose regions of statistically significant wavelet power in the time-frequency space at the 5% significance level of a white noise process. Significant interannual (≈ 2-8 year) oscillations at Calgary occurred in the early 1900s, 1930s and 1950s, while inter-decadal (≈ 10-32 year) oscillations were active from 1900s to 1960s. Since the wavelet transform is performed in the Fourier space, the ends of the time series are padded with zeros to bring it to the next higher power of two to reduce wraparound effects (Torrence and Compo, 1998). The wavelet power outside the cone of influence (dashed line in Figure 2.2b) is suppressed due to the zero-padding and should be interpreted with caution.

Various quantities can be derived from the wavelet transform so as to condense the vast quantity of information contained in the wavelet spectrum and enhance further interpretation and statistical analysis. One such quantity is the global wavelet spectrum, which is the time average of all the local wavelet power spectra for each scale,

$$\overline{W}^2(\xi) = \frac{1}{N}\sum_{t=0}^{N-1} |W_t(\xi,\gamma)|^2 \tag{2.4}$$

The global wavelet spectrum only shows dominant scales with no temporal information. The global wavelet spectrum for Calgary (Figure 2.2c) shows statistically significant oscillations at 1, 12 and about 22 year scales.

The scale-averaged wavelet power (SAWP) is used to examine the spatial and temporal fluctuations of the wavelet power in a given scale band. SAWP is defined as,

$$\overline{W}^2(t) = \frac{\delta j \delta t}{C_\delta}\sum_{j=j_1}^{j_2} \frac{|W_t(\xi,\gamma)|^2}{\xi_j} \tag{2.5}$$

where δj is a factor for scale averaging, C_δ is a reconstruction factor, δt is the sampling interval and j_1 and j_2 are the upper and lower cutoff scales.

Scale bands and time periods within which precipitation exhibits co-variance with climate indices can be identified from wavelet coherence, which is defined as (Torrence and Webster, 1999),

$$R_t^2(\xi) = \frac{\left|\left\langle \xi^{-1} W_t^{XY}(\xi,\gamma)\right\rangle\right|^2}{\left\langle \xi^{-1}\left|W_t^X(\xi,\gamma)\right|^2\right\rangle \cdot \left\langle \xi^{-1}\left|W_t^Y(\xi,\gamma)\right|^2\right\rangle} \tag{2.6}$$

where $W_t^{XY}(\xi,\gamma)$ is the cross-wavelet spectrum of X and Y, $\langle . \rangle$ is a smoothing operator, and $0 \leq R_t^2(\xi) \leq 1$. Equation 2.6 is an accurate representation of the normalized covariance between two time series since the wavelet transform conserves variance. It has been used to study the relationships between Indian rainfall and ENSO (Torrence and Webster, 1999) and Baltic Sea ice conditions and the Arctic and North Atlantic oscillation indices (Jevrejeva *et al.*, 2003).

Pearson's correlations between the principal components (PC) of band-pass filtered signals of precipitation and climate indices at multiple scale bands were also assessed. The band-passed signal *x'(t)* over a subset of scales can be recovered from the wavelet coefficients by the equation,

$$x'(t) = \frac{\delta j\, \delta t^{1/2}}{C_\delta \psi(0)} \sum_{j=j_1}^{j=j_2} \frac{\Re\{W_t(\xi_j)\}}{\xi_j^{1/2}} \tag{2.7}$$

where $\psi(0)$ is a factor to remove the energy scaling (Torrence and Compo, 1998).

In addition to wavelets, composite and correlation analysis were also used to examine the relationships between winter precipitation in Western Canada and large-scale climate anomalies. A composite is constructed by averaging individual cases of a quantity (e.g. precipitation) for years during which a selected climate anomaly has occurred (Barry and Perry, 1973). Composite analysis assumes that the selected climate anomaly is primarily responsible for any extreme conditions observed in the quantity of interest for those years included in the composite.

2.4. Discussions of Results

2.4.1. Wavelet Analysis of Precipitation

The wavelet power spectra for all 21 stations are presented using SAWP and time-longitude plots (also called Power Hovmöller) in Figure 2.3. The SAWP was computed for three scale bands (2–3 year, 3–8 year and 8–30 year). The precipitation variability in the first two bands could be related to the high and low frequency components of ENSO and extratropical modes of low frequency climate variability whereas the last band corresponds to decadal to inter-decadal oscillations that could be associated with decadal variations in the North Pacific climate. The solid contours in Figure 2.3 enclose statistically significant SAWP at the 5% level of a white noise process.

At individual stations, the 2–3 year scale (Figure 2.3a(i)) accounts for up to 38% of the variance (e.g. in the 1910s at Regina and Indian Head, Saskatchewan; in the 1980s at Vavenby, BC). At the regional scale, there appears to be some spatial coherency in the early part of the record, with the activities around the 1920s and 1930s occurring at several stations across the Prairies and BC (Figure 2.3a(i)). This is followed by a period of little activity during 1940s to late 1950s with the exception of Victoria, BC in 1950. Significant activities since the late 1950s are observed for stations west of 117°W and east of 105°W, while those in eastern BC, Alberta, and most of Saskatchewan show little activity. Considering the space-averaged power, the proportion of variance explained by the 2–3 year band ranges from as low as 5% in the 1950s to about 15% in the 1910s (Figure 2.3a(ii)). This is far below the maximum variance explained at individual stations, showing that significant precipitation activities tend to be not well organized in both space and time. Hence, significant precipitation activities in the 2–3 year scale band generally have short lifetime and seem to be haphazard in nature.

The 3–8 year band (Figure 2.3b(i)) accounts for up to 30% of the precipitation variance at some stations (e.g. Prince Albert, Saskatchewan around 1920;

Quatsino, BC around 1950). As opposed to the 2–3 year band, the 1950s is the time during which the precipitation across the study area shows the largest significant power for the 3–8 year band. There are also scattered activities before and after 1950s, such as those in Saskatchewan in the 1920s, and in BC in the 1960s and 1990s. Similar to the 2–3 year scale, these fluctuations have short lifetimes and show less spatial coherence (Figure 2.3b(ii)). The variability at the 8–30 year scale (Figure 2.3c(i)) accounts for up to 20% of the variance at a few stations (e.g. from 1910 to 1930 at Calgary, Alberta; from 1920 to 1950 at Victoria, BC; from 1910 to 1940 at Agassiz, BC; and from 1985 to end of record at Creston, BC). Although there appears to be no spatial coherence except for that at Agassiz and Victoria, BC during 1920–1940, this scale shows persistent variability at those few stations exhibiting significant precipitation activity.

2.4.2. Regional Precipitation Activities

Visual inspection of the spatial climate pattern of the Hovmoller diagrams for the 2–3 year and 3–8 year scales in Figure 2.3 suggests that the 21 precipitation stations can be grouped into 3 regions: (1) west of 117°W (9 stations), (2) 117°W to 107°W (6 stations), and (3) east of 105°W (6 stations). For simplicity, these three regions will be referred to as western, central and eastern regions, respectively in subsequent discussions. Stations in the western region tend to show more variability in the second half of the record period (i.e. after 1950) than in the first half of the record period whereas those in the central region essentially exhibit an opposite pattern, with very little activity during the second half of the record, especially for the 3–8 year scale. Stations in the eastern region appear to exhibit moderate activity in both halves of the record period.

2.4.3. Wavelet Analysis of Climate Indices

The wavelet power spectra of Nino3, SOI, CNP, PDO, PNA and WP indices are shown in Figure 2.4a-f. In agreement with previous studies (Torrence and Compo, 1998), the Nino3 index exhibits interannual (2-8 year scale) oscillations of large amplitude during pre-1920 and post-1960 periods, and a reduced level of activity in between (Figure 2.4a). Nino3 also shows inter-decadal oscillations in the early 1900s that persisted with decreasing power until 1940s. Similarly, SOI shows interannual variability of 2-8 year scale during pre-1920s, and post-1970s, and few significant oscillations in between (Figure 2.4b). There is also some inter-decadal variability in SOI that persisted up to the 1920s, and resumed in the 1980s.

CNP shows scattered interannual oscillations in the 1910s, 1940s, 1950s and 1980s, and a strong 15 year oscillation mode during 1916-1947 (Figure 2.4c). Even though PDO shows relatively strong power near the 20 year scale, most of its power is concentrated at scales of 42 and 62 years, which are outside the cone of influence (Figure 2.4d). PNA shows statistically significant oscillations of 8 to 12 year modes in the 1960s, and a 4 year oscillation from the second half of 1970s to mid 1980s (Figure 2.4e). According to Graham (1994), the PNA pattern changed from a negative phase in 1964-1967 to a positive phase in 1976-1988, which probably agrees with the southward shift and intensification of the Aleutian low noted in the mid-1970s. Its negative phase again dominated during 1989-1990, followed by a prolonged positive phase from fall 1991 to spring 1993. WP shows significant interannual oscillation during 1950-1965, and 1985-1990s (Figure 2.4f).

2.4.4. Influence of Large-scale Climate Anomalies

2.4.4.1. Wavelet Coherence and Phase Difference

A visual comparison of Figure 2.3 and Figures 2.4a-f indicates that the relationships between the wavelet powers of precipitation anomalies of individual stations and climate indices appear to be highly unstable. Using the wavelet coherence analysis, we could measure the links between regional precipitation signals and climate indices. Since the leading principal components of precipitation anomalies for the western, central and eastern regions account for a considerable percentage of the variance in regional precipitation (41.8%, 42.7% and 53.7%, respectively), we used these leading principal components as surrogates for regional precipitation signals. In subsequent discussions, the western, central and eastern leading principal components will be referred to as WRPC1, CRPC1 and ERPC1, respectively. Their wavelet power spectra are shown in Figure 2.4g-i.

Figures 2.5a-f show the wavelet coherence between the WRPC1 and six climate indices. The thick contours in Figure 2.5 enclose periods of statistically significant coherence of a red noise process as determined by a Monte Carlo experiment (Jevrejeva *et al.*, 2003). The phase differences between the two signals for coherences greater than 0.5 are plotted as vectors in Figure 2.5, where a right pointing arrow indicates that the two signals are in-phase while a left pointing arrow indicates an anti-phase relationship. Arrows deviating from the horizontal indicate that there is a lagged relationship between the two signals.

At the interannual scale, WRPC1 and Nino3 show scattered coherences of over 0.8 in the 1950s, 1970s, 1980s and 1990s (Figure 2.5a). Note that the existence of significant coherence between the two signals does not necessarily depend on the existence of significant wavelet power in both signals. For instance, both Nino3 (Figure 2.4a) and the WRPC1 (Figure 2.4g) do not have significant power near the 5-year scale in the 1950s but still show significant coherence. Similarly, the WRPC1 power near the 2 year scale in the 1980s, albeit relatively strong, is not statistically significant but the coherence between WRPC1 and Nino3 for that time period and scale is statistically significant. The phase difference shows that there is generally an anti-phase relationship between Nino3 and WRPC1 for time periods of significant coherence. The phase distribution outside periods of significant coherence is less consistent. For instance, while the phase difference near the 15 year scale is anti-phase before 1950s, it is close to in-phase after 1980's.

For the post-1950 period, the coherence between SOI and WRPC1 (Figure 2.5b) mostly mirrors that of Nino3. Periods of scattered but high coherence between

WRPC1 and SOI are also observed near the 2-year scale during 1910s to 1920s, 1930's and late 1940s. For time periods with significant coherence, the WRPC1 and SOI are generally in phase. For CNP, the strongest coherence with the WRPC1 occurs at the interdecadal scale centered near 15 years up until 1960's (Figure 2.5c). This is not surprising since both CNP and WRPC1 have significant wavelet power at that scale (Figures 2.4c and 2.4g). For time periods of significant coherence, the CNP and WRPC1 are generally in phase, the only exception being the anti-phase relationship in the 2-3 year scale in the 1980s. The strength of the 15 year scale relationship between WRPC1 and the North Pacific ocean-atmosphere system is also supported by the relatively strong coherence between WRPC1 and PDO up until the 1940s (Figure 2.5d).

Most of the high coherence between PNA and WRPC1 is found in the 1-2 year scale (Figure 2.5e). However, interannual scale coherence values of over 0.8 are also observed during the early 1970s, 1980s and late 1990s. A notable feature of the relation between WRPC1 and PNA is the phase change from anti-phase in the 1970's to in-phase in the 1980s and then back to nearly anti-phase in the 1990's. Statistically significant coherence between WP and WRPC1 is observed during the early 1970s, 1980s and late 1990s (Figure 2.5f), with the relation being out of phase for all three periods.

The CRPC1 shows relatively weak interannual scale coherence with Nino3 and SOI, with the only noticeable periods of significant coherence being the 1920s and 1980s at the 4-year scale (Figures 2.6a-b). Both Nino3 and SOI show relatively strong coherence with CRPC1 in the 1-2 year scale. The CNP index also has weak interannual scale coherence with the CRPC1 except in the 1920s and 1960s (Figure 2.6c). An interesting feature of Figures 2.6c-d is the anti-phase (in-phase) relationship between CNP (PDO) and CRPC1 for the interdecadal scale centered near 15 years for the earlier part of the record. This observation is in direct contrast to the phase distribution for the WRPC1 (Figure 2.5c-d). The PNA and WP indices appear to show better overall interannual scale coherence with the

CRPC1 than the remaining climate indices (Figures 2.6e-f). A phase shift in the coherence between WP and CRPC1 around 1980 is evident from Figure 2.6f.

Significant interannual scale coherence between the ENSO indices and ERPC1 exists only in the 1920s, 1950s and 1970s (Figure 2.7a-b). CNP and PDO lead the ERPC1 at the interdecadal scale centered near 15 years (Figure 2.7c-d). Again PNA appears to exhibit better interannual scale coherence with the ERPC1 compared to the other indices (Figure 2.7e). For WP, the only periods of significant coherence occurred in the 1970s and 1980s, with the phase difference showing anti-phase relations between the two signals (Figure 2.7f).

The above results show that the strength and consistency of interannual scale relations between Western Canadian seasonal precipitation anomalies and large-scale climate indices changes in both the time and frequency domains. This is partly due to the haphazard nature of interannual oscillations in the seasonal precipitation anomalies (Figure 2.3 and Figure 2.4g-i). In other words, as compared to climate indices like Nino3 and SOI, seasonal precipitation anomalies have low overall signal-to-noise ratio, leading to scattered periods of high coherence with climate indices.

Notwithstanding variations in phase differences seen in the coherence maps, decadal to inter-decadal precipitation variability in Western Canada appears to be linked mainly to CNP and PDO. Low-frequency precipitation variability should be linked to SST variations since the latter has distinct decadal fluctuations (Trenberth and Hurrell, 1994; White *et al.*, 1997). Cayan *et al.* (1998) found that the Prairies' precipitation fluctuations are associated with extensive shifts of SLP and SST anomalies, which they suggested are components of low-frequency precipitation variability from global scale climate processes. Weaver *et al.* (1991) showed that freshwater flux could excite decadal and inter-decadal oceanic variability that may be important in the observed decadal/ inter-decadal variability in our climate system. Ghil and Vautard (1991) suggested that inter-decadal oscillation of global surface air temperature could be related to changes in the extratropical ocean circulations. From a coupled ocean-atmosphere model and observations, Latif and Barnett (1994) found that about one-third of the low frequency variability (mode of 20 year) over the North Pacific Ocean and North America can be attributed to unstable air-sea interactions between the sub-tropical re-circulation in the North Pacific and the Aleutian low system. They showed that a correlation of -0.5 to -0.6 exists between anomalies of atmospheric pressure south of the Aleutians and Prairies' air temperature.

2.4.4.2. Correlations at Multiple Scales

We also explored the relations between the PC of band-passed seasonal precipitation and band-passed climate indices at multiple scales. If a climate index consistently exerts significant influence on the regional precipitation at a given scale, we expect the band-passed climate index and the band-passed precipitation PC scores to show strong correlations. Table 2.3 shows Pearson's correlations between the first three band-passed precipitation PC scores of each region and climate indices for the 2-3 year, 3-8 year and 8-30 year scale bands, with statistically significant correlations at the 5% level indicated in bold text. Statistical significance of the correlations was estimated by the bootstrap re-sampling approach as follows (Efron and Tibshirani, 1993). Given N data points, the bootstrap procedure for a single realization involves randomly re-sampling N rows of precipitation PC and climate index pairs with replacement, and computing the correlation coefficient. We estimated the 5th and 95th percentiles of 10,000 bootstrap correlations. Positive (negative) correlations for which the 5th (95th) percentile is greater (less) than the 5% significance level from the standard significance test for correlation were deemed significant. Correlations significant at the 5% level are indicated in bold text in Table 2.3.

From the correlations in Table 2.3, it is clear that no single climate index can explain more than 30% of interannual precipitation variability in Western Canada. The rather weak correlations for interannual scale bands corroborate the inconsistencies observed in the wavelet coherence and phase difference presented in Section 2.4.4.1. In general, it appears that regional precipitation shows better correlation with climate indices in the 2-3 year and 8-30 year scale bands than the 3-8 year band.

2.4.4.3. Composite Analysis

Here, compositing is used to explore the impacts of the extreme phases of ENSO, PDO, CNP, PNA, WP and EP on the winter (December-January-February (DJF)) precipitation of Western Canada. The El Nino and La Nina composites were based on years during which the 5-month moving average of SOI remained in the lower (higher) 25% of the distribution for a period of 5 months or longer (Shabbar *et al.*, 1997). High (low) PNA, WP and EP phases were defined using a threshold of ±0.5 standard deviations based on the DJF standardized index. High (low) CNP phases were defined as those years where the standardized DJF CNP index is above (below) ±0.75 standard deviations. The years used for ENSO, CNP, PNA, WP and EP composites are listed in Table 2.4. Observational studies indicate that 20th century PDO regime shifts have occurred around 1924/25, 1946/47, 1976/77 and 1998/99 (Mantua and Hare, 2001). Therefore, our warm PDO composites are for 1925-1946 and 1977-1998 while the cool PDO composites are for 1914-1924 and 1947-1976. It should be noted that the change in Pacific climate in 1976/77 period has been identified as a regime shift (Nitta and Yamada, 1989; Trenberth, 1990; Miller *et al.*, 1994).

The composite precipitation (e.g. for El Nino) for a given station was computed as the ratio of the mean of winter precipitation for anomalous years (e.g. the 21 El Nino years in Table 2.4) relative to the long-term mean winter precipitation of that station. The long-term means were computed from the period 1914-2001 for ENSO and PDO composites, 1914-1990 for CNP composites, 1948-2001 for PNA composites, and 1950-2001 for WP and EP composites. Figure 2.8 shows the composites for all 21 stations, where a composite value of greater than unity means that the climate anomaly is associated with positive winter precipitation anomaly, and vice versa. Albeit the composite magnitudes show considerable variation from station to station, La Nina (El Nino), cool PDO (warm PDO), low PNA (high PNA) and low WP (high WP) are typically associated with positive (negative) winter precipitation anomalies across Western Canada (Figures 2.8a-c; Figure 2.8e). In addition, high (low) EP years are associated with positive (mainly negative) precipitation anomalies in the western region, though lately changes to low EP years is unclear (Figure 2.8d).

In general, ENSO, PDO, PNA and WP appear to distinguish negative and positive precipitation anomalies better than CNP and EP. Low CNP winters (anomalously low central North Pacific SLP) are mostly associated with negative precipitation anomalies across the study area (Figure 2.8f). On the other hand, high CNP winters are associated with weak positive precipitation anomalies in the eastern half of the study area but mostly (seven of the nine stations) with negative precipitation anomalies in the western half.

The composites for stations in the western, central and eastern regions were also averaged separately and for the entire study area in order to assess the characteristic response of regional precipitation to the extreme phases of the climate anomalies (see Table 2.5). Prince Rupert and Quatsino, BC were excluded from the regional aggregate composites since their ENSO and PNA composites had opposite signs to those of the remaining seven stations in the western region (see stations 1 and 2 of Figures 2.8a and 2.8b). Considering the average of the remaining 19 stations, El Nino (La Nina) is associated with a 14% decrease (20% increase) of winter precipitation, high (low) PNA is associated with a 12% decrease (9% increase), warm (cool) PDO is associated with a 8% decrease (9% increase), high (low) CNP is associated with a 11% decrease (4% increase), high (low) WP is associated with a 8% decrease (9% increase), and high (low) EP is associated with a 12% increase (1% decrease) relative to the long-term mean winter precipitation. In general, regional precipitation responses to the opposite phases of a climate anomaly are not mirror images of each other.

The precipitation ENSO/PNA relations detected in this study generally concur with previous studies, e.g. Shabbar *et al.* (1997) showed that El Nino (La Nina) events are associated with below (above) normal precipitation anomalies across Western Canada. Cayan *et al.* (1998) indicated that ENSO could explain about 10–20% of the annual precipitation variance for the CP, with the variance explained increasing to 20–30% towards the BC coast. Moore and McKendry (1996) indicated that winters dominated by an enhanced PNA pattern are associated with below normal spring snowpack in BC. Hsieh and Tang (2001) showed that El Nino (La Nina) and high (low) PNA years are associated with lighter (heavier) than normal April 1 snow water equivalent in the Columbia River basin, BC.

Provided that the climatology of winter precipitation in Western Canada is influenced by large-scale climate anomalies, atmospheric circulation patterns that prevailed during anomalous winter seasons should provide supporting physical evidence for the relations revealed by the composite analysis. For example, during warm PDO phases, the SST field tends to be anomalously cool over the central North Pacific basin, but anomalously warm SST prevails along the west coast of North America (Mantua and Hare, 2001). These SST anomalies are accompanied by variation in the placement and strength of the Aleutian low during winter (Bond and Harrison, 2000). A deeper than normal Aleutian low favors enhanced cyclonic winds, which could reinforce the cool SST anomalies over the central North Pacific and inhibit latent and sensible heat release. Bond and Harrison (2000) showed that surface latent and sensible heat fluxes in the central North Pacific basin are suppressed (enhanced) by about 15 W/m^2 during warm (cool) PDO phases for both ridge and trough events. They also showed that the meridional temperature gradient in the planetary boundary layer across the Pacific roughly north of 40°N is suppressed (enhanced) during warm (cool) PDO phases. These conditions would likely be associated with disturbances in the location and intensity of moisture bearing westerly flows over western North America.

Mature El Nino winters are typically associated with deepening of the Aleutian low and an amplification of the western Canadian ridge. The ridge causes enhanced anticyclones and a northward shift in the mid-latitude jet stream, leading to relatively dry conditions in Western Canada (Shabbar *et al.*, 1997). On the other hand, mature La Nina winters are typically associated with an erosion of the western Canadian ridge and enhanced westerly flow, resulting in relatively wet conditions in Western Canada. Like El Nino winters, high PNA winters are associated with deeper than normal Aleutian low and an enhanced ridge over Western Canada. However, the height anomalies over North America during El Nino winters exhibit pronounced meridional gradients whereas the gradients during high PNA winters are mainly zonally oriented (Straus and Shukla, 2002).

Figure 2.9 shows the DJF 300-mbar geopotential and wind anomaly patterns associated with anomalous CNP winters. During low CNP winters (Figure 2.9a), the orientation of height anomalies around the Aleutian low has some superficial resemblance to that during mature El Nino winters (not shown). A closer scrutiny of the composite fields reveals that the Aleutian anomaly center during El Nino winters is shifted eastwards by 5 to 10 degrees relative to that during low CNP winters. In addition, the highs over western North America extend further south during low CNP winters than during El Nino winters. These differences are evident from Figure 2.9b, which is a difference field between El Nino and low CNP composites. The positive (negative) height fields over the central North Pacific (western North America and eastern Pacific) in Figure 2.9b shows that the anomalies near the centers of the Aleutian low and the western Canadian high are much stronger during low CNP winters than during El Nino winters. The difference in the wind patterns in Figure 2.9b suggests a more enhanced meridional flow during El Nino winters than during low CNP winters though both types of winters appear to be devoid of westerlies; hence both are associated with negative precipitation anomalies over Western Canada.

Figure 2.9c shows the 300-mbar level circulation patterns associated with high CNP winters. Again, the orientation of the height anomalies around the Aleutian low during high CNP winters has some similarity to that during La Nina winters (not shown). On the contrary, the orientation of the height anomalies over Western Canada during high CNP winters is considerably different from that during La Nina winters. First, the upper level flow during high CNP winters is predominantly meridionally oriented, whereas that during La Nina winters is mostly zonally oriented. Second, there is a southwestward displacement of the western Canadian anomaly center during high CNP winters compared to La Nina winters. The difference field between the wind patterns during La Nina and high CNP winters (Figure 2.9d) suggests an enhanced westerly flow over Western Canada during La Nina winters relative to high CNP winters, possibly due to a southward shift in the position of geotropic westerlies during the latter. To substantiate this argument further, we examined the 300-mbar zonal (U-) wind anomaly patterns during ENSO and CNP winters in Figure 2.10.

The zonal wind patterns during El Nino (Figure 2.10a) and La Nina (Figure 2.10b) winters show the familiar ENSO related upper level circulation (i.e. a southward displacement of the subtropical jet stream during El Nino winters; enhanced westerlies over western North America during La Nina winters). The zonal wind patterns associated with low CNP winters (Figure 2.10c) are broadly similar to those during El Nino winters, with the only major difference being an apparent weakening of the jet stream over the subtropical eastern Pacific during low CNP winters compared to El Nino winters. On the other hand, there is a clear southward displacement of the jet stream over western North America during high CNP winters (Figure 2.10d) compared to La Nina winters. In addition, the jet stream axis during high CNP winters is aligned in the southwest-northeast direction, as opposed to a northwest-southeast orientation during La Nina winters. Hence, with much of the storm track expected to be to the south, the western part of our study area would experience negative precipitation anomalies during high CNP winters. A significant implication of the non-linearity of Western Canadian precipitation responses to anomalous CNP winters is that anomalous atmospheric circulation over the central North Pacific alone is not a sufficient condition for the prevalence of moisture deficiency (or surplus) over Western Canada.

As noted earlier, the impacts of the EP pattern are more likely confined to the western parts of the study area. This result is also supported by composites of geopotential height and wind anomalies for high EP (Figure 2.11a) and low EP (Figure 2.11b) winters. Figure 2.11a exhibits the signature of a strong positive phase of the EP pattern, with a deeper than normal trough in the vicinity of the Bering Sea/ Alaska, and positive height anomalies to the northeast of Hawaii (Horel and Wallace, 1981). This phase of EP is associated with enhanced westerly flow over southwestern BC and the US Pacific Northwest.

During low EP winters, negative height anomalies cover a large portion of the subtropical Pacific extending to the west coast of the US (Figure 2.11b). A similar anomaly center to that of the Pacific is also observed over a large portion of the polar region and the high latitudes (not shown). Positive height anomalies are centered over the Gulf of Alaska. The flow over the eastern central North Pacific is dominated by easterlies, which could lead to a split and/or displacement of the jet stream over North America (Figure 2.11b). This would lead to drier than normal conditions over southern BC. A similar circulation pattern dominated the region during the winter of 1992-1993, leading to drier than normal conditions across Canada, and wetter than normal conditions across southwestern and central US (Bell and Basist, 1994).

2.4.4.4. Correlations with Raw Precipitation

To assess the usefulness of large-scale climate anomalies for predicting seasonal precipitation, we computed Pearson's correlations between climate indices at zero to 3 season lead times and the dominant modes of regional precipitation. The regional precipitation signals for the western, central and eastern regions were separately represented by the leading PC's (PC1) of seasonal precipitation anomalies from stations in each region. These leading PC's explain a large portion of the variance of regional precipitation. For the western region, the PC1 of each season (excluding data at Prince Rupert and Quatsino, BC) explains 50% of the variance for winter, 37% for spring, 53% for summer and 54% for autumn. The amount of variance explained by PC1 of the central region stands at 49% for winter, 34% for spring, 42% for summer and 49% for autumn. For the eastern region, PC1 explains 55% of the variance for winter, 52% for spring, 47% for summer and 58% for autumn.

The winter season correlations, when teleconnections in the extratropical Northern Hemisphere are known to be relatively strong (Wallace and Gutzler, 1981; Shabbar *et al.*, 1997), are given in Table 2.6 (where correlations significant at the 5% level are indicated in bold text). The correlations between Nino3 and PC1 of the eastern and central regions increase for one season lead time, and those between SOI and the PC1 of all three regions increase up to one or two-season lead times. Comparatively, the PDO and PNA indices show the strongest contemporaneous (lag-0) correlations with PC1 of all three regions. The strongest lag-0 correlation of PDO and PNA is with the central region PC1 (-0.54 for PDO and -0.52 for PNA). These results are in agreement with the composite analysis, i.e. El Nino (La Nina), warm PDO (cool PDO), high PNA (low PNA), low CNP (high CNP), high WP (low WP) and low EP (high EP) are generally associated with negative (positive) winter precipitation anomalies. However, for prediction purposes, only ENSO and PDO seem to provide useful information at one- to three-seasons lead-time.

During spring, only CNP and PNA show significant lag-0 correlations with the central region PC1 (ρ=0.23 with CNP and ρ=-0.38 with PNA) and the eastern region PC1 (ρ=0.25 with CNP and ρ=-0.41 with PNA). During summer, CNP shows significant correlation at lag-0 with the central region PC1 (ρ=0.25), and at lag-1 with the western region PC1 also (ρ=0.36). In autumn, significant lag-0 correlations exist between the western region PC1 and PNA (ρ=-0.37), the western region PC1 and CNP (ρ=0.22), the central region PC1 and PNA (ρ=0.44) and the eastern region PC1 and WP (ρ=0.30). It appears that ENSO events affect Western Canadian precipitation mainly during winter but the influence of CNP, PNA and WP could extend to spring and autumn.

It seems that the proportion of precipitation variance explained by individual climate indices is generally not strong enough to achieve consistently accurate prediction of precipitation anomalies in Western Canada. The unstable relations between climate indices and precipitation may partly be because teleconnection patterns mainly capture the large-scale features of variability while local changes in their anomaly centers can result in large differences in western North American surface climate (Yarnal and Diaz, 1986). In addition, some climate indices (e.g. CNP), being of limited area averages, may be too simplistic to capture the (non-linear) dynamical links between large-scale circulation and Western Canadian precipitation.

Inconsistencies in the strength of the links between a climate index and precipitation may also reflect the complicated relations between the tropical Pacific SST forcing and the internal dynamics of the North Pacific during the mature phase of ENSO (Bell and Basist, 1994). In particular, precipitation activities may be affected by constructive (destructive) interactions between two or more climate patterns operating at different characteristic time scales. For instance, using coupled general circulation model (GCM) experiments, Yeh and Kirtman (2004) argued that decadal scale central North Pacific SST variability can influence the midlatitude ENSO response on interannual time scales. Using historical data, several studies have shown that western North American surface climate and ENSO relationships exhibit interdecadal variations that are synchronized with PDO (McCabe and Dettinger, 1999; Gobena and Gan, 2006). Similar PDO-ENSO interactions are evident in Western Canadian precipitation response, with El Nino (La Nina) winters during warm (cool) PDO regimes being associated with enhanced negative (positive) precipitation anomalies relative to the mean El Nino (La Nina) response (Figure 2.12). On the other hand, precipitation anomalies during El Nino (La Nina) winters conditioned on cool (warm) PDO regimes remain close to normal, suggesting that the opposite phases of these two climate anomalies could negate each other. The differences between the conditioned and unconditioned ENSO responses in Figure 2.12 are all statistically significant at the 5% level of the two sample *t* test.

Precipitation climate relations could also be influenced by interactions between interannual modes such as ENSO and PNA. While there is a higher tendency for anomalous PNA years to be associated with mature ENSO phases (Trenberth and Hurrell, 1994), Table 2.4 shows that extreme phases of PNA can also occur without the ENSO forcing. Using ensemble GCM experiments, Straus and Shukla (2002) showed that tropical SST anomalies during El Nino events primarily force midlatitude circulation patterns that are distinctly different from the PNA pattern. From the analysis of Western Canadian streamflow anomalies, Gobena and Gan (2006) found that streamflow responses to high PNA years conditioned on Non-El Nino years were not symmetrical with streamflow responses to low PNA years conditioned on non-La Nina years. While the former produced responses that are similar to those during El Nino years, the latter were associated with streamflow responses that were significantly weaker than those during La Nina years.

Precipitation processes are also affected by local factors such as topography, surface heating and friction, which will contribute significantly to precipitation spatial variability. In particular, orographic controls could exert significant influence on stations located in the valleys of the rugged terrains of inland BC, where the Pacific air stream has to descend to reach the valley stations. Canada's precipitation has been shown to be a high dimensional chaotic process with D_2 of 8 to 9 (Gan *et al.*, 2002), implying that many independent variables are needed to describe the precipitation process.

2.5. Summary and Conclusions

From the wavelet and composite analysis of 21 stations of long-term precipitation data of Western Canada, and their teleconnection to large-scale climate anomalies for atmospheric circulation and SST of the Pacific, the conclusions are:

1) Using the Morlet wavelet, statistically significant (α=0.05) interannual oscillations that occurred haphazardly (even for stations located in the same region) have been detected in the precipitation data of Western Canada. Inter-decadal fluctuations tend to be more persistent but significant oscillations were detected only at a few stations.

2) Based on similarities in precipitation wavelet power at interannual scales, three regions were identified and the precipitation activities in each region were related to large-scale atmospheric circulation patterns (via climate indices). The wavelet coherence and phase difference between climate indices and the leading PC of precipitation for each region was found to be highly inconsistent in both time and frequency. In addition, Pearson's correlations between the PC's of band-pass filtered precipitation and climate indices were generally found to be weak, indicating that the strength of teleconnections changes in both time and frequency.

3) In general, ENSO (Nino3 and SOI) exerted relatively strong influence on the winter precipitation data of Western Canada as compared to the other climate anomalies. At regional level, El Nino (La Nina) leads to a 14% decrease (20% increase) in mean winter precipitation. PNA, PDO, CNP, WP and EP also have their shares in forcing the precipitation over the region. Strong positive (negative) PNA leads to a 12% decrease (9% increase) in mean winter precipitation. The influence of PNA also extends to the spring and autumn seasons. PDO is associated with an 8% decrease (9% increase) whereas high (low) CNP is associated with a 11% decrease (4% increase) in mean winter precipitation. Strong positive (negative) WP leads to an 8% decrease (9% increase) in mean winter precipitation. The impact of EP tends to be confined to the western region

of the study area only, with strong positive (negative) EP leading to a 14% increase (4% decrease) in mean winter precipitation in that region. The detected teleconnections could occur at interannual or inter-decadal levels (depending on which anomaly index is used), and their strength changes in time and space, making their applications for seasonal precipitation prediction unreliable. Further, climate indices alone are too simplistic to capture the precipitation variability consistently.

Gan *et al.* (2007) indicated that Western Canada's precipitation data are characterized as a high dimensional multifractal process with correlation dimensions (D_2) of 8 to 9. This result corroborates the unstable interannual oscillations revealed by the wavelet analysis. Given the high D_2 and multifractal properties identified, we expect the climate systems for precipitation prediction to be very sensitive to the initial conditions, a basic characteristic of a chaotic system. In other words, any slight perturbation or errors will amplify exponentially in successive iterations. Furthermore, compounded by their haphazard low frequency oscillations, it will be difficult to get consistent seasonal predictions of the highly nonlinear precipitation processes in Western Canada by only teleconnecting with climate indices. Some selected SST and/or SLP predictor fields from the Pacific Ocean could be useful. The prediction result will likely be season and region dependent, with better results expected during winter (or autumn) seasons. This is also supported by the fact that lagged correlations between climate indices and precipitation that are significant are mainly found during the winter.

Acknowledgements

This project was partly funded by the NSERC of Canada. The AHCCD monthly precipitation data were obtained from the Climate Research Branch of the Meteorological Service of Canada (MSC). The software codes for wavelet analysis were made available by Torrence and Compo (http://paos.colorado.edu/research/wavelets/), and A. Grinsted (http://www.ac.uk/home/research/wavelet coherence).

References

Allan, R. J., Nicholls, N., Jones, P. D., and Butterworth, I. J. (1991). A further extension of the Tahiti-Darwin SOI, early ENSO events and Darwin pressure. *J. Climate,* 4, 743-748.

Barry, R. G., and Perry, A. H. (1973). *Synoptic Climatology: Methods and Applications*. London: Methuen & Co Ltd.

Bell, G. D., and Basist, A. N. (1994). Seasonal climate summary: The global climate of December 1992- January 1993: Mature ENSO conditions continue in the tropical Pacific, California drought abates. *J. Climate,* 7, 1581-1605.

Bond, N. A., and Harrison, D. E. (2000). The Pacific Decadal Oscillation, air-see interaction and central north Pacific winter atmospheric regimes. *Geophys. Res. Lett.,* 27(5), 731-734.

Bonsal, B. R., Zhang, X., and Hogg, W. D. (1999). Canadian Prairie growing season precipitation variability and associated atmospheric circulation. *Clim. Res.,* 11, 191-208.

Cayan, D. R., and Peterson, D. H. (1989). The influence of North Pacific atmospheric circulation on streamflow in the west. In D. H. Peterson (Ed.), *Aspects of Climate Variability in the Pacific and Western America* (pp. 375-397). Geophysical Monograph Series, 55. AGU, Washington, DC.

Cayan, D. R., Dettinger, M. D., Diaz, H. F., and Graham, N. E. (1998). Decadal variability of precipitation over western North America. *J. Climate,* 11, 3148-3166.

Coulibaly, P., and Burn, D. H. (2005). Spatial and temporal variability of Canadian seasonal streamflows. *J. Climate,* 18(1), 191–210.

Coulibaly, P., and Burn, D.H. (2004). Wavelet analysis of variability in annual Canadian streamflows. *Water Resour. Res.,* 40, W03105, doi:10.1029/2003WR002667.

Dettinger, M. D., Cayan, D. R., Diaz, H. F., and Meko, D. M. (1998). N-S precipitation patterns in North America in interannual-to-decadal time series. *J. Climate,* 11, 3095-3111.

Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap.* San Francisco: Chapman & Hall.

Gan T. Y., Wang, Q., and Seneka, M. (2002). Correlation dimensions of climate subsystems and their geographic variability. *J. Geophys. Res.,* 107 (D23), p. 4728.

Gan, T. Y. (1998). Hydroclimatic trends and possible climatic warming in the Canadian Prairies. *Water Resour. Res.,* 34, 3009-3015.

Gan, T. Y., Gobena, A. K., and Wang, Q. (2007). Precipitation of southwestern Canada – Wavelet, scaling, multifractal analysis, and teleconnection to climate anomalies. *J. Geophys. Res.*, 112, D10110, doi:1029/2006JD007157.

Ghil, M., and Vautard, R. (1991). Inter-decadal oscillations and the warming trend in global temperature time series. *Nature,* 350, 324-327.

Gobena, A. K., and Gan, T. Y. (2006). Low-frequency variability in Southwestern Canadian streamflow: Links with large-scale climate anomalies. *Int. J. Climatol.,* 26, 1843-1869.

Graham, N. E. (1994). Decadal-scale climate variability in the tropical and North Pacific during the 1970s and 1980s: observations and model results. *Clim. Dynam.,* 10, 135-162.

Hare, F. K., and Thomas, M. K. (1974). *Climate Canada*, Wiley, Toronto, 256pp.

Horel, J. D., and Wallace, J. M. (1981). Planetary-scale atmospheric phenomena associated with the southern Oscillation. *Mon. Weather Rev.,* 109, 813-829.

Hsieh, W., and Tang, B. (2001). Interannual variability of accumulated snow in the Columbia basin, British Columbia. *Water Resour. Res.,* 37, 1753-1759.

Islam, S., Bras, L. B., and Rodriguez-Iturbe, I. (1993). A possible explanation for low correlation dimension estimates for the atmosphere. *J. Appl. Meteorol.,* 32, 203-208.

Jevrejeva, S., Moore, J. C., and Grinsted, A. (2003). Influence of the Arctic oscillation and El Nino-Southern oscillation (ENSO) on ice conditions in the Baltic Sea: The wavelet approach. *J. Geophys. Res.,* 108(D21), 4677, doi:10.1029/2003JD003417.

Kane, R. P. (1997). Prediction of droughts in North-East Brazil: Role of ENSO and use of periodicities. *Int. J. Climatol.,* 17, 655-665.

Labat, D., Ababou, R., and Mangin, A. (2000). Rainfall-runoff relations for karstic springs. Part II: continuous wavelet and discrete orthogonal multiresolution analyses. *J. Hydrol.,* 238, 149–178.

Lafrenière, M., and Sharp, M. (2003). Wavelet analysis of interannual variability in the runoff regimes of glacial and nival stream catchments, Bow Lake, Alberta. *Hydrological Processes,* 17, 1093–1118.

Latif, M., and Barnett, T. P. (1994). Causes of decadal climate variability over the North Pacific and North America. *Science,* 266, 634-637.

Lau, K.-M., and Weng, H. (1995). Climate signal detection using wavelet transform: How to make a time series sing. *Bulletin Amer. Meteorol. Soc.*, 76(12), 2391-2402.

Lott, N., Ross, D., and Sittel, M. (1997). *The winter of 96-97 West Coast flooding,* Research Customer Service Group Tech. Rep. 97-01, National Climate Data Center, 23 pp.

Lovejoy, S., and Schertzer, D. (1985). Generalized scale invariance in the atmosphere and fractal models of rain. *Water Resour. Res.*, 21(8), 1233-1250.

Mantua, N., and Hare, S. (2001). The Pacific Decadal Oscillation. *J. Oceanography,* 58, 35-44.

McCabe, G. J., and Dettinger, M. D. (1999). Decadal variations in the strength of ENSO teleconnections with precipitation in the western United States. *Int. J. Climatol.,* 19, 1399-1410.

Mekis, É., and Hogg, W. D. (1999). Rehabilitation and analysis of Canadian daily precipitation time series. *Atmos.-Ocean,* 37(1), 53-85.

Miller, A. J., Cayan, D. R., Barnett, T. P., Graham, N. E., and Oberhuber, J. M. (1994). The 1976-77 climate shift in the Pacific Ocean. *Oceanography,* 7, 21-26.

Moon, Y., and Lall, U. (1996). Atmospheric flow indices and interannual Great Salt Lake variability. *J. Hydrol. Engrg.*, 1(2), 1-8.

Moore, R. D., and McKendry, I. G. (1996). Spring snowpack anomaly patterns and winter climatic variability, British Columbia. Canada. *Water Resour. Res.*, 32, 623-632.

Mwale, D., and Gan, T.Y. (2005). Wavelet analysis on variability, teleconnectivity and predictability of the September-November East African Rainfall. *J. Applied Meteorology,* 44, 256-269.

Mwale, D., Gan, T.Y., and Shen S. S. P. (2004). A new analysis of variability and predictability of seasonal rainfall of central southern Africa for 1950–94. *Int. J. Climatol.,* 24, 1509–1530.

Nitta, T., and Yamada, S. (1989). Recent warming of tropical sea surface temperature and its relationship to the northern hemisphere circulation. *J. Meteor. Soc. Japan*, 67, 375-383.

Phillips, D. (1990). *The climates of Canada*, edited by D. Francis, D. Gullett, and E. Truhlar, Canadian Government Publishing Centre, 176pp.

Rasmusson, E. M., and Carpenter, T. H. (1982). Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Nino. *Mon. Weather Rev.,* 110, 354-384.

Rogers, J. C. (1988). Precipitation variability over the Caribbean and Tropical Americas Associated with the Southern Oscillation. *J. Climate,* 1, 172-180.

Rogers, P. (1994). Assessing the socioeconomic consequences of climate change on water resources. *J. Climatic Change,* 28, 179-208.

Ropelewski, C. F., and. Halpert, M. S. (1986). North American precipitation and temperature patterns associated with the El Nino/Southern Oscillation (ENSO). *Mon. Weather Rev.,* 114, 2352-2362.

Shabbar, A., Bonsal, B., and Khandekar, M. (1997). Canadian precipitation patterns associated with the Southern Oscillation. *J. Climate,* 10, 3016-3027.

Smith, L.C., Turcotte, D.L., and Isacks, B.L. (1998). Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrological Processes,* 12, 233–249.

Straus, D. M., and Shukla, J. (2002). Does ENSO force the PNA? *J. Climate,* 15, 2340-2358.

Svensson, C., Olsson, J., and Berndtsson, R. (1996). Multifractal properties of daily rainfall in two different climates. *Water Resour. Res.,* 32, 2463-2472.

Torrence, C., and Compo, G. P. (1998). A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.,* 79(1), 61-78.

Torrence, C., and Webster, P. J. (1999). Interdecadal changes in the ENSO-monsoon system. *J. Climate,* 12, 2679-2690.

Trenberth, K. E. (1990). Recent observed interdecadal climate changes in the northern hemisphere. *Bull. Amer. Meteor. Soc.*, 71, 988-993.

Trenberth, K. E., and Hurrell, J. W. (1994). Decadal atmospheric-ocean variations in the Pacific. *Clim. Dynam.,* 9, 303-319.

Wallace, J. M., and Gutzler, D. S. (1981). Teleconnections in the geopotential height field during northern hemisphere winter. *Mon. Weather Rev.,* 109, 784-812.

Weaver, A. J., Sarachik, E. S., and Marotze, J. (1991). Freshwater flux forcing of decadal and inter-decadal oceanic variability. *Nature,* 353, 836-838.

White, W. B., Lean, J., Cayan, D. R., and Dettinger, M. D. (1997). Response of global upper ocean temperature to solar irradiance anomalies. *J. Geophys. Res.,* 102, 3255-3266.

Yarnal B., and Diaz, H. F. (1986). Relationships between extremes of the Southern Oscillation and the winter climate of Anglo-American Pacific coast. *J. Climatol.,* 6, 197-219.

Yeh, S.-W., and Kirtman, B. P. (2004). Decadal North Pacific sea surface temperature variability and the associated global climate anomalies in a coupled general circulation model. *J. Geophys. Res.*, 109, D20113, doi:10.1029/2004JD004785.

Table 2.1. Summary of the precipitation stations used in the study.

Station No.	Station Index	Station name and province	Lat, deg	Long, deg	Elevation, m	Mean annual Precipitation, mm
		Western Region				
1	1066481	Prince Rupert A., BC	54.30	130.43	34	2588
2	1036570	Quatsino, BC	50.53	127.65	8	2464
3	1060840	Bella Coola, BC	52.37	126.68	18	1555
4	1092970	Fort St James, BC	54.45	124.25	695	462
5	101HFEE	Victoria Phyllis, BC	49.45	123.27	8	780
6	1096630	Quesnel A., BC	53.03	122.52	545	543
7	1100120	Agassiz CDA, BC	49.25	121.77	15	1734
8	1126510	Princeton A., BC	49.47	120.52	700	371
9	1168520	Vavenby, BC	51.58	119.78	447	446
		Central Region				
10	1173210	Golden, BC	51.30	116.97	787	473
11	1142160	Creston, BC	49.10	116.52	597	560
12	3031093	Calgary Int'l A., AB	51.12	114.02	1071	472
13	3012208	Edmonton M. A., AB	53.57	113.52	668	497
14	3034480	Medicine Hat A., AB	50.02	110.72	713	383
15	4048520	Waseca, SK	53.10	109.50	644	439
		Eastern Region				
16	4057120	Saskatoon, SK	52.23	106.62	491	419
17	4056240	Prince Albert A., SK	53.22	105.68	426	480
18	4016560	Regina A., SK	50.43	104.67	574	455
19	4013480	Indian Head, SK	50.53	103.67	583	496
20	5010485	Brandon CDA, MB	49.87	99.98	361	532
21	5023222	Winnipeg Int'l A., MB	49.90	97.23	237	598

Table 2.2. Cross-correlations between winter season climate indices

	Nino3[a]	SOI[b]	PDO[c]	CNP[d]	PNA[e]	WP[f]	EP[f]
Nino3	1.00	**-0.81**	**0.42**	**-0.31**	**0.55**	**0.63**	-0.18
SOI		1.00	**-0.38**	**0.26**	**-0.50**	**-0.55**	0.23
PDO			1.00	**-0.54**	**0.70**	**0.38**	-0.03
CNP				1.00	**-0.76**	-0.13	-0.23
PNA					1.00	0.24	0.13
WP						1.00	0.13
EP							1.00

[a] Statistically significant correlations at the 5% level are indicated in bold fonts.

[a] Data period used for correlation analysis is 1914-1997.

[b] Data period used for correlation analysis is 1914-1999.

[c] Data period used for correlation analysis is 1914-2001.

[d] Data period used for correlation analysis is 1914-1990.

[e] Data period used for correlation analysis is 1948-2001.

[f] Data period used for correlation analysis is 1950-2001.

Table 2.3. Pearson's correlations between the PC scores of band-passed precipitation and band-passed climate indices for selected scale bands [a]

Scale	PC	Variance, %	Nino3	SOI	CNP	PNA	WP	PDO
Western Region								
2-3 year	1	46.9	**-0.39**[b]	**0.47**[b]	**0.26**[b]	**-0.30**	**-0.42**[b]	**-0.26**
2-3 year	2	18.0	0.00	**-0.19**	0.00	**0.25**	**-0.33**	0.04
2-3 year	3	11.8	-0.12	0.07	**0.23**	**-0.42**	-0.18	**-0.23**
3-8 year	1	46.3	-0.14	0.07	0.14	-0.15	**-0.36**	-0.12
3-8 year	2	15.7	**-0.37**	**0.33**	**0.51**	**-0.61**	**-0.56**	**-0.34**
3-8 year	3	9.9	-0.09	0.05	0.15	**-0.28**	**-0.26**	-0.17
8-30 year	1	46.7	**-0.32**	**0.21**	**0.85**[c]	0.14	-0.17	-0.16
8-30 year	2	11.7	-0.03	**0.18**	0.07	0.18	**0.29**	0.04
8-30 year	3	10.8	0.07	**0.18**	**-0.42**	**0.34**	**-0.37**	**0.37**
Central Region								
2-3 year	1	50.8	-0.17	**0.21**	**0.34**[†]	-0.15	**-0.53**[b]	**-0.29**
2-3 year	2	22.0	**0.41**[†]	**-0.42**[†]	-0.09	**0.46**	**0.33**	**0.30**
2-3 year	3	9.8	0.04	-0.13	0.09	-0.01	**0.24**	**0.23**
3-8 year	1	49.0	0.01	-0.16	-0.01	-0.14	0.21	0.08
3-8 year	2	19.6	-0.04	0.02	-0.04	0.02	-0.19	-0.05
3-8 year	3	10.7	**-0.32**	**0.37**	0.13	-0.03	-0.11	**-0.27**
8-30 year	1	33.4	-0.16	-0.14	**-0.41**	**-0.24**	**0.50**	-0.14
8-30 year	2	25.6	**-0.20**	0.16	-0.01	**-0.44**	-0.13	**-0.32**
8-30 year	3	20.2	0.15	-0.02	0.19	-0.11	**0.32**	0.12
Eastern Region								
2-3 year	1	59.9	**-0.46**[b]	**0.23**[b]	0.12	**-0.43**	-0.21	**-0.32**
2-3 year	2	13.3	**-0.26**	-0.05	0.13	0.00	-0.20	**-0.20**
2-3 year	3	11.0	-0.03	0.04	**-0.43**	-0.02	0.09	**0.21**
3-8 year	1	59.4	**-0.41**	**0.23**	0.16	**-0.36**	**-0.24**	**-0.24**
3-8 year	2	13.9	0.03	0.02	-0.03	-0.11	-0.12	-0.09
3-8 year	3	11.6	-0.02	0.16	-0.06	-0.06	0.07	-0.12
8-30 year	1	61.6	**-0.49**	0.05	0.05	**-0.43**[c]	0.21	**-0.31**
8-30 year	2	17.7	-0.14	0.05	**0.46**	0.22	-0.15	-0.10
8-30 year	3	8.4	-0.12	0.04	0.12	**-0.36**	-0.04	**-0.39**

[a] Correlations significant at the 5% level based on 10,000 bootstrap samples are shown in bold fonts.

[b] Higher correlation at one year lead time than at no lead time.

[c] Higher correlation at one year lead time than at no lead time.

Table 2.4. Years included in composite analysis of winter precipitation for the extreme phases of ENSO, CNP, PNA, WP and EP patterns.

Composite	Years included in composites
El Nino	1912, 1913, 1915, 1919, 1920, 1926, 1927, 1930, 1931, 1940,1942, 1952, 1954, 1958, 1959, 1966, 1970, 1973, 1977, 1983, 1987, 1992,1998
La Nina	1917, 1918, 1925, 1929, 1939, 1951, 1956, 1957, 1965, 1971,1972, 1974, 1976, 1989, 1996, 1999
High CNP	1915, 1916, 1922, 1932, 1937, 1949, 1950, 1952, 1955, 1956, 1968, 1969, 1972, 1979, 1985, 1989
Low CNP	1914, 1926, 1927, 1929, 1936, 1939, 1940, 1941, 1942, 1944, 1946, 1953, 1958, 1963, 1964, 1970, 1978, 1980, 1986, 1987
High PNA	1953, 1958, 1961, 1963, 1970, 1977, 1978, 1981, 1983, 1986, 1987, 1992, 1998
Low PNA	1948, 1949, 1950, 1952, 1956, 1957, 1965, 1966, 1969, 1971, 1972, 1979, 1982, 1988, 1989, 1997, 1999
High WP	1964, 1966, 1975, 1979, 1983, 1987, 1988, 1989, 1992, 1998, 2001
Low WP	1950, 1956, 1957, 1961, 1962, 1963, 1965, 1968, 1971, 1974, 1981, 1986, 1991, 1996, 1997
High EP	1953, 1954, 1964, 1967, 1971, 1974, 1975, 1999, 2000
Low EP	1957, 1969, 1978, 1979, 1991, 1992, 1993, 1994, 1995, 1997

Table 2.5. Aggregate composites of winter precipitation for the western, central and eastern regions.

Region	ENSO		PNA		WP		EP		CNP		PDO	
	La Nina	El Nino	Low	High	Low	High	Low	High	High	Low	Warm	Cool
Western	1.20	0.89	1.09	0.89	1.06	0.92	0.96	1.14	0.98	0.86	0.94	1.07
Central	1.25	0.86	1.13	0.87	1.12	0.84	0.99	1.12	1.05	0.92	0.90	1.12
Eastern	1.16	0.84	1.05	0.89	1.09	0.99	1.03	1.09	1.09	0.88	0.92	1.09
All	1.20	0.86	1.09	0.88	1.09	0.92	0.99	1.12	1.04	0.89	0.92	1.09

Table 2.6. Pearson's correlations at zero to 3-season lags between selected climate indices and winter precipitation PC1 time series of western, central and eastern regions [a]

	Nino3	SOI	CNP	PDO	PNA	WP	EP
Western Region PC1							
Lag-0	**-0.41**	**0.37**	**0.27**	**-0.44**	**-0.50**	-0.24	**0.31**
Lag-1	**-0.38**	**0.43**	**0.26**	**-0.25**	-0.18	-0.10	0.10
Lag-2	**-0.29**	**0.41**	0.11	**-0.27**	0.22	-0.07	0.12
Lag-3	**-0.33**	**0.22**	0.09	**-0.27**	-0.01	0.14	0.19
Central Region PC1							
Lag-0	**-0.39**	**0.34**	**0.25**	**-0.54**	**-0.52**	**-0.46**	0.09
Lag-1	**-0.42**	**0.39**	**0.24**	**-0.37**	-0.18	-0.21	-0.01
Lag-2	**-0.32**	**0.42**	0.08	**-0.33**	0.12	0.18	0.01
Lag-3	**-0.39**	0.20	0.00	**-0.37**	-0.23	-0.13	0.10
Eastern Region PC1							
Lag-0	**-0.40**	**0.26**	**0.28**	**-0.51**	**-0.44**	**-0.35**	0.09
Lag-1	**-0.42**	**0.26**	**0.25**	**-0.33**	**-0.28**	**-0.34**	-0.07
Lag-2	**-0.34**	**0.35**	0.10	**-0.25**	0.13	0.03	0.00
Lag-3	-0.21	0.03	-0.02	**-0.26**	-0.31	-0.16	0.19

[a] Statistically significant correlations at the 5% level are indicated in bold fonts.

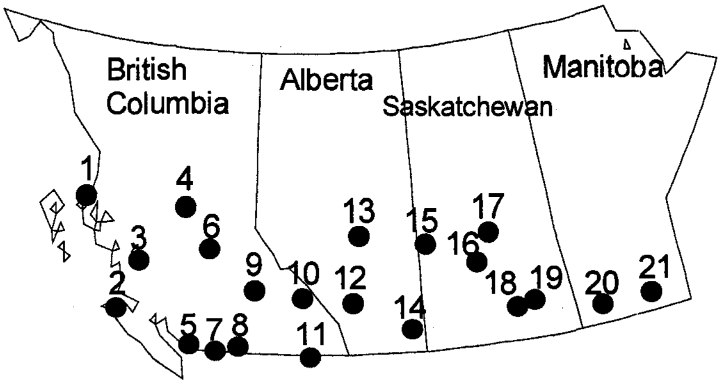


Figure 2.1. Location of the 21 precipitation stations selected for the study. The station numbers correspond to those given in column 1 of Table 2.1.

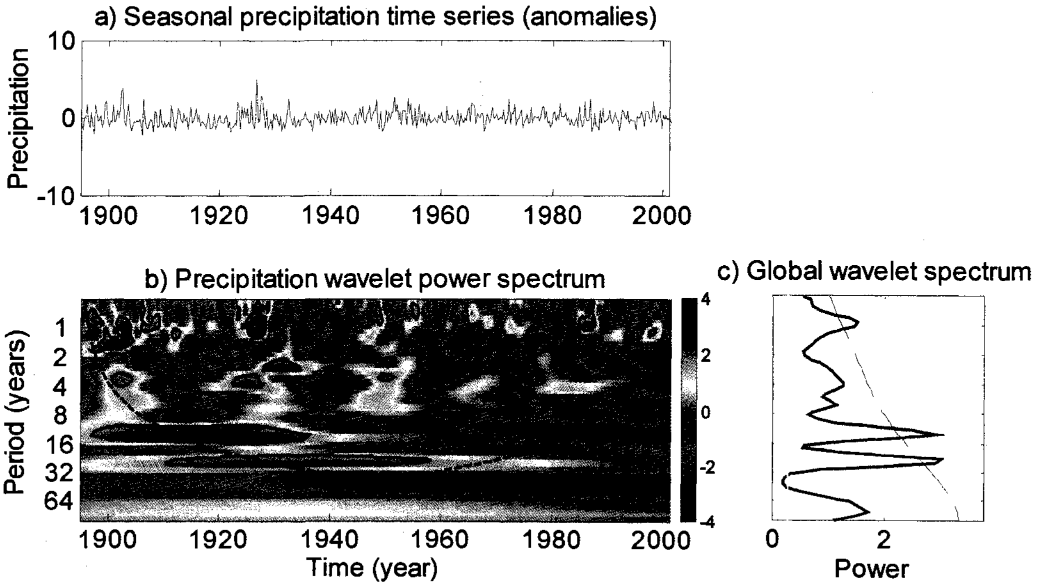


Figure 2.2. The continuous wavelet spectrum of seasonal precipitation at Calgary, Alberta. (a) Time series of standardized seasonal precipitation anomalies, (b) Morlet wavelet power spectrum of (a). The thick black contours depict the 95% confidence level of local power relative to a white noise background. The dashed line is the cone of influence beyond which the energy is contaminated by the effect of zero-padding. (c) Global wavelet power spectrum (solid line) with the 95% confidence level (dashed line).

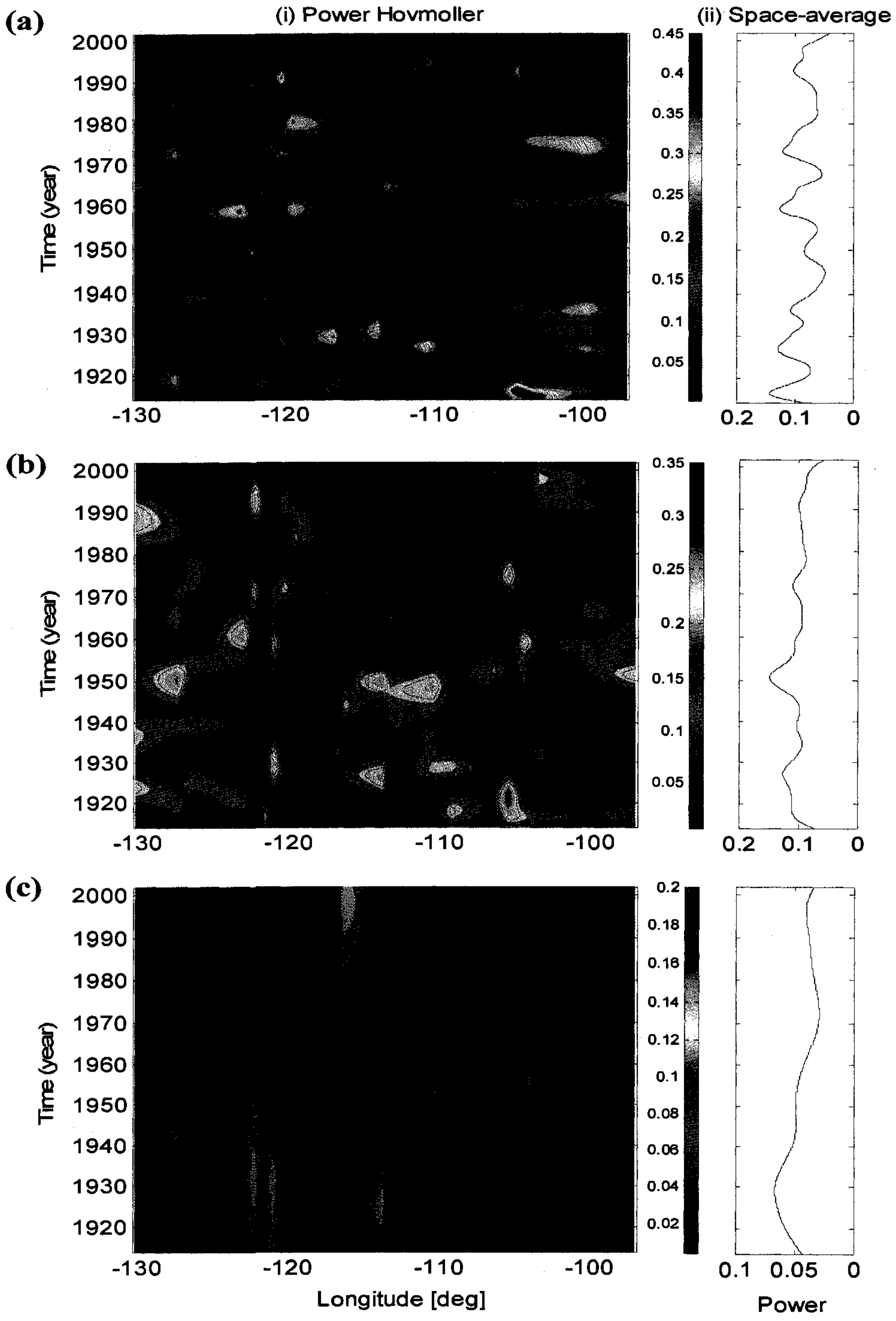


Figure 2.3. SAWP (i) and space-averaged SAWP (ii) of seasonal precipitation anomalies at the 21 stations across SW Canada: (a) 2–3-year scale band. (b) 3–8-year scale band, (c) 8–30-year scale band. The solid contours enclose periods of statistically significant SAWP relative to white noise at the 5% significance level. The vertical line in the Power Hovmöller corresponds to the boundary between precipitation stations from the Prairies (right) and BC (left).

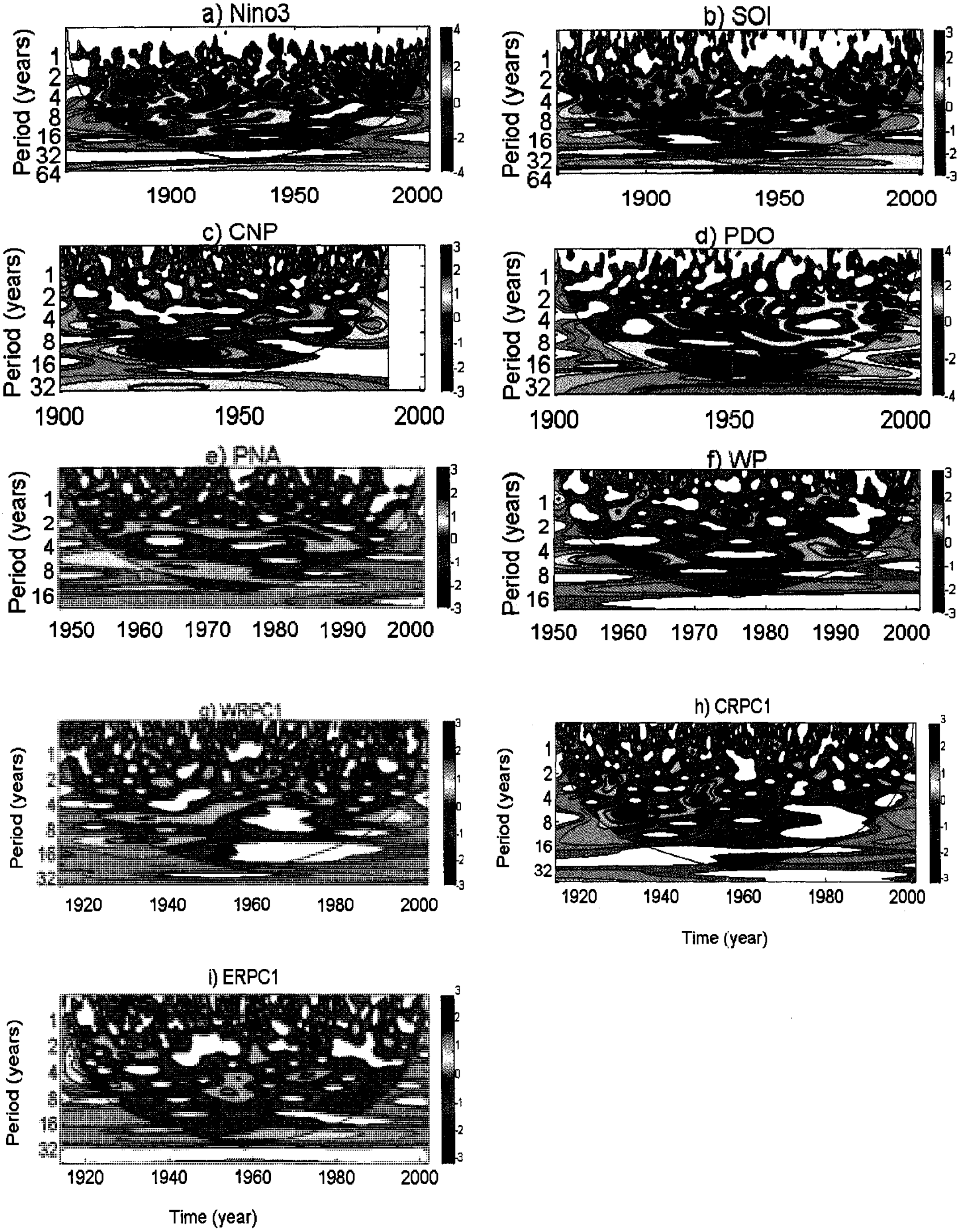


Figure 2.4. Wavelet power spectra of seasonal climate indices: (a) Nino3, (b) SOI, (c) CNP, (d) PDO, (e) PNA and (f) WP. Wavelet power spectra of the leading principal components (PC1) of seasonal precipitation anomalies: (g) western, (h) central and (i) eastern region. The thick black contours enclose statistically significant wavelet power at the 5% level of a red noise process, and the dashed line is the cone of influence.

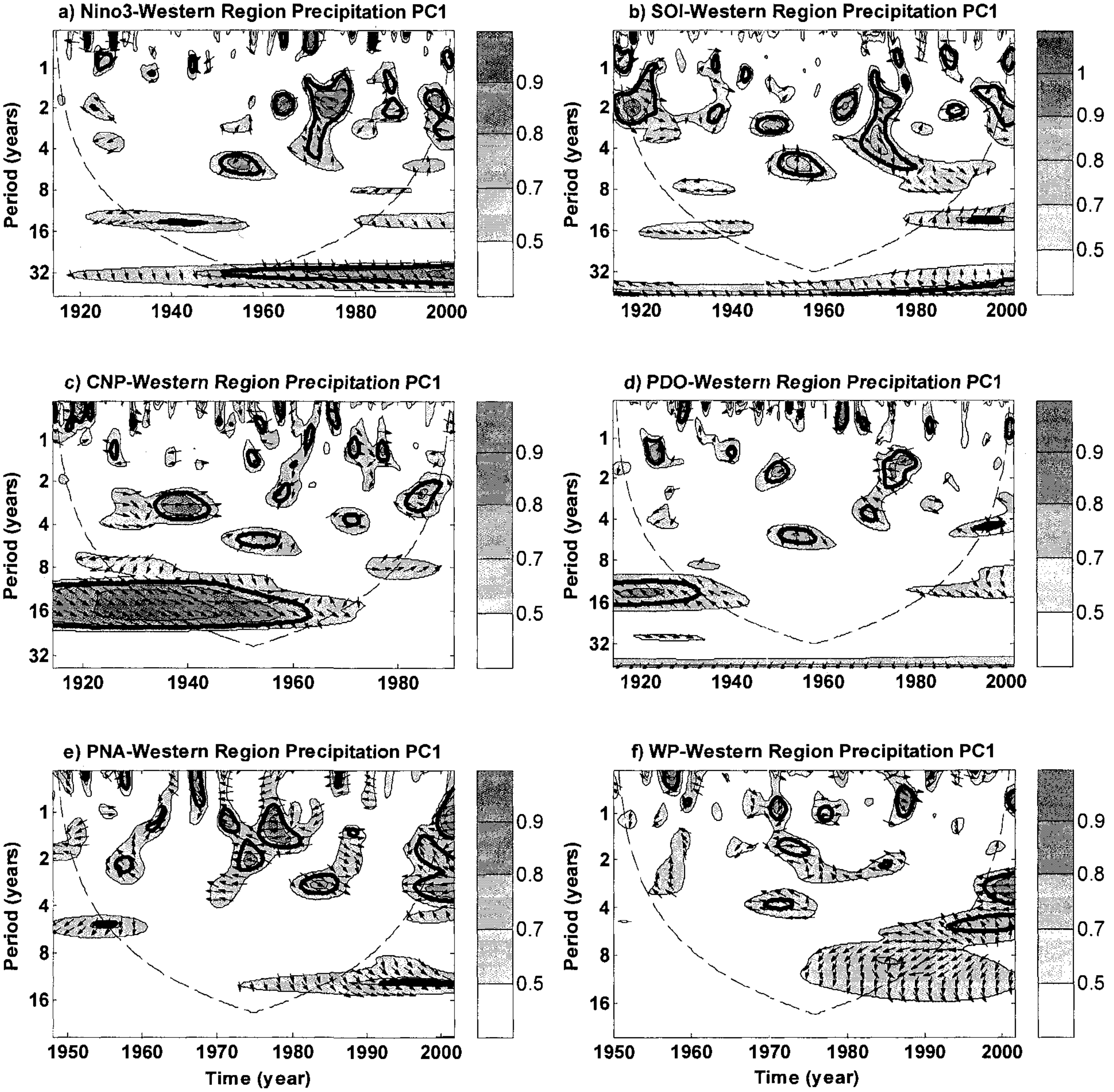


Figure 2.5. Wavelet coherence and phase difference between the western region precipitation PC1 and (a) Nino3, (b) SOI, (c) CNP, (d) PDO, (e) PNA, (f) WP. The thick black contours enclose periods of statistically significant coherence at the 5% level of a red noise process, and the dashed line is the cone of influence. The phase difference is plotted only for time periods and scales with coherence over 0.5. Right-pointing arrows indicate that the two signals are in-phase while left-pointing arrows are for anti-phase signals.

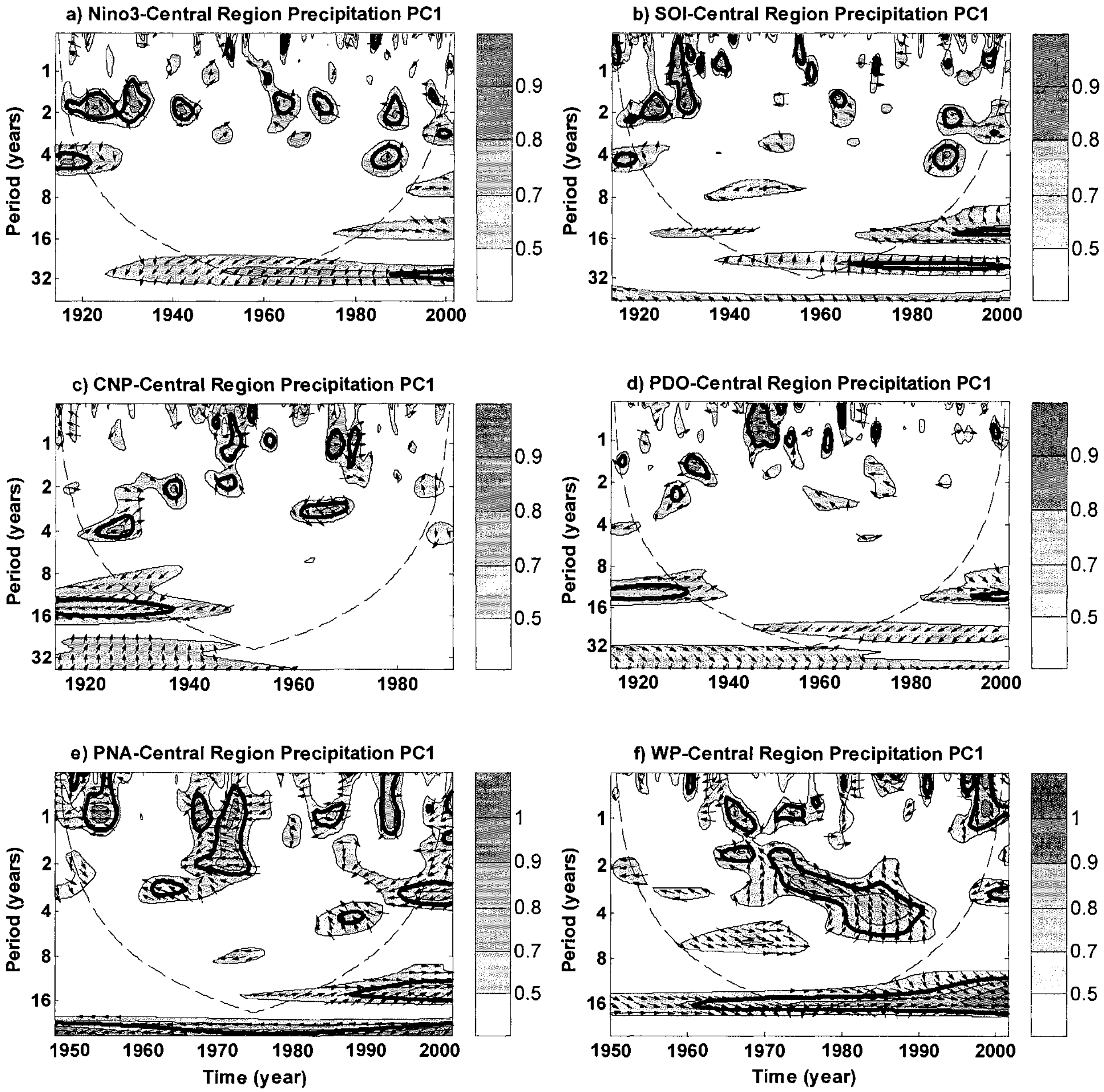


Figure 2.6. Wavelet coherence and phase difference between the central region precipitation PC1 and (a) Nino3, (b) SOI, (c) CNP, (d) PDO, (e) PNA, (f) WP. All features are the same as in Figure 2.5.

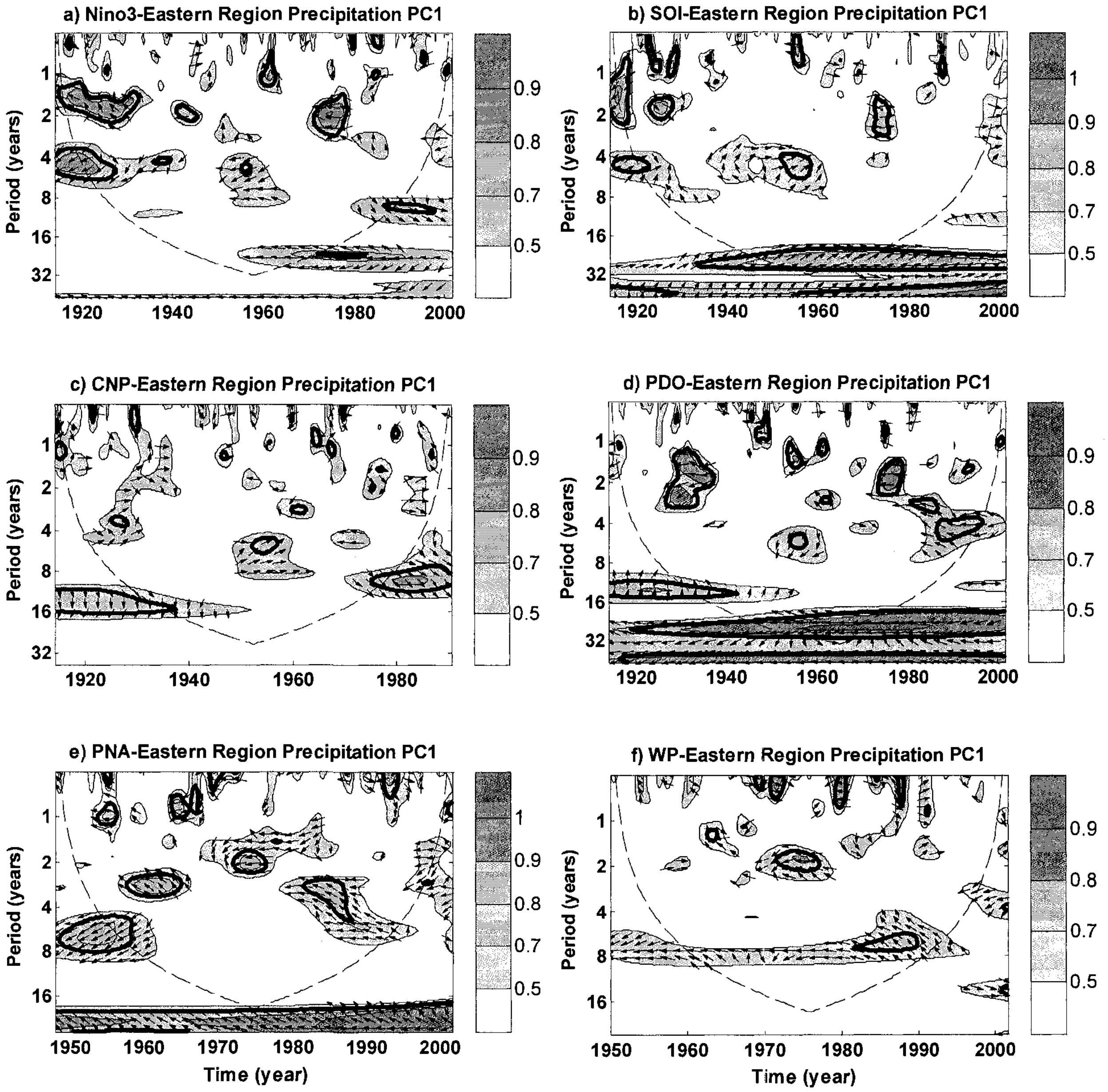


Figure 2.7. Wavelet coherence and phase difference between the eastern region precipitation PC1 and (a) Nino3, (b) SOI, (c) CNP, (d) PDO, (e) PNA, (f) WP. All features are the same as in Figure 2.5.

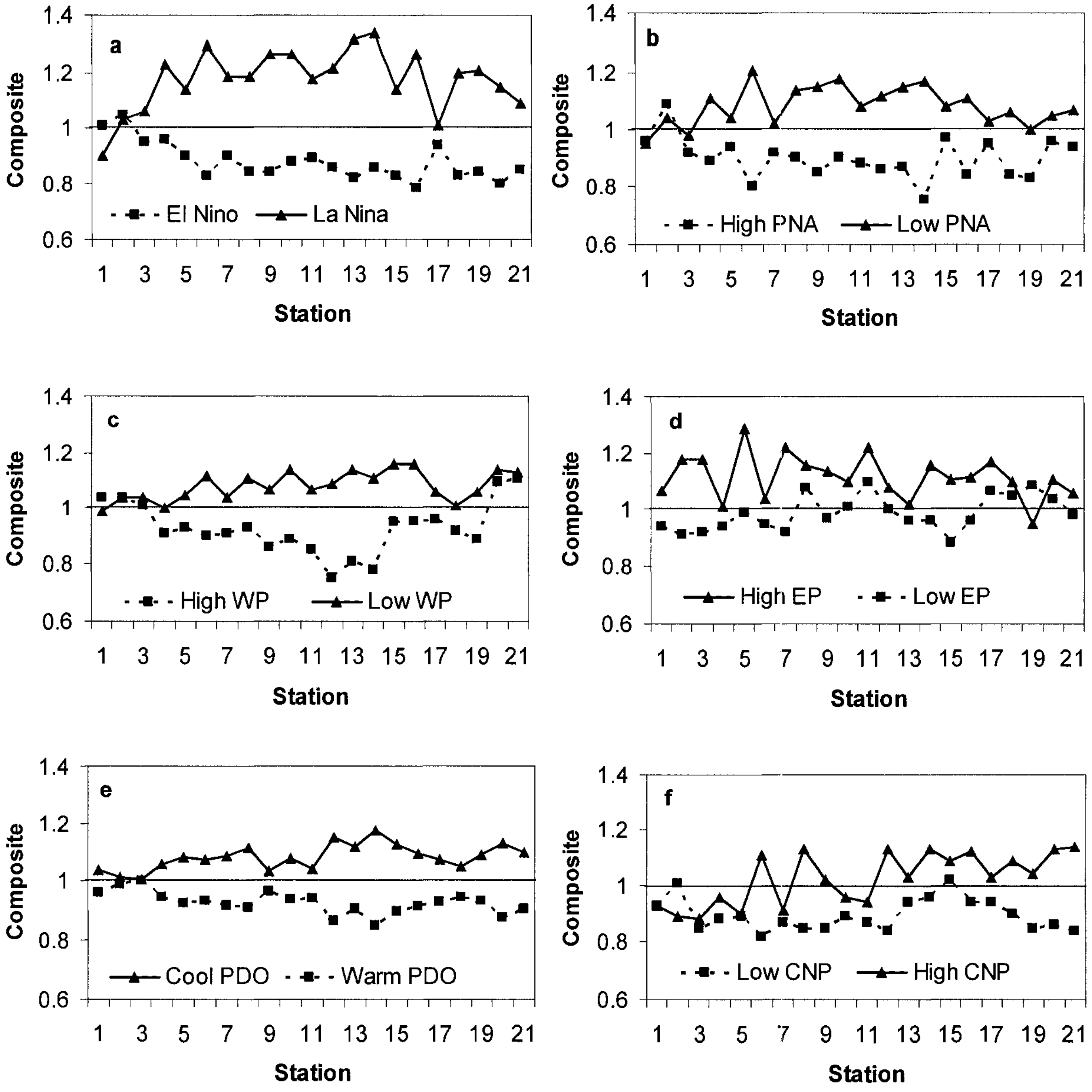


Figure 2.8. Composite winter precipitation associated with (a) El Nino and La Nina years. (b) High and low PNA years. (c) High and low WP years. (d) High and low EP years. (e) Warm and cool PDO years. (f) High and low CNP years. The composite for each station is computed as the ratio of the mean winter precipitation during anomalous years to the long-term mean winter precipitation. Station numbers correspond to those given in column 1 of Table 2.1.

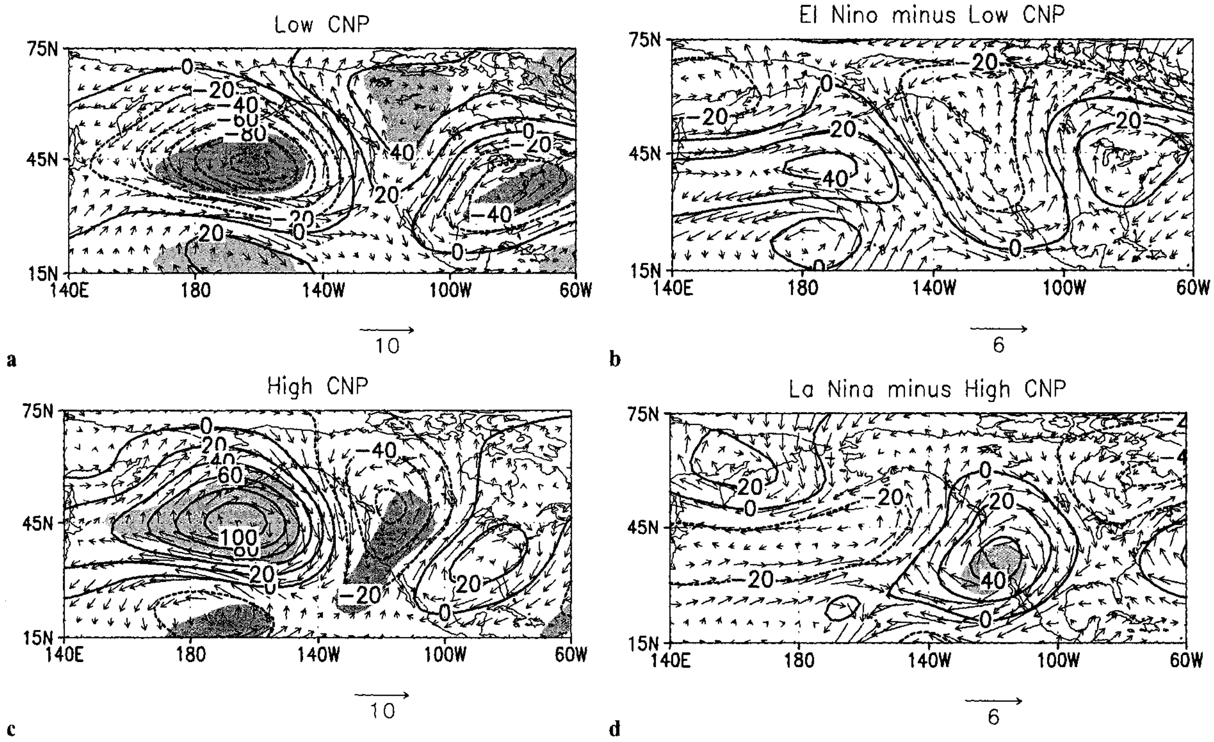
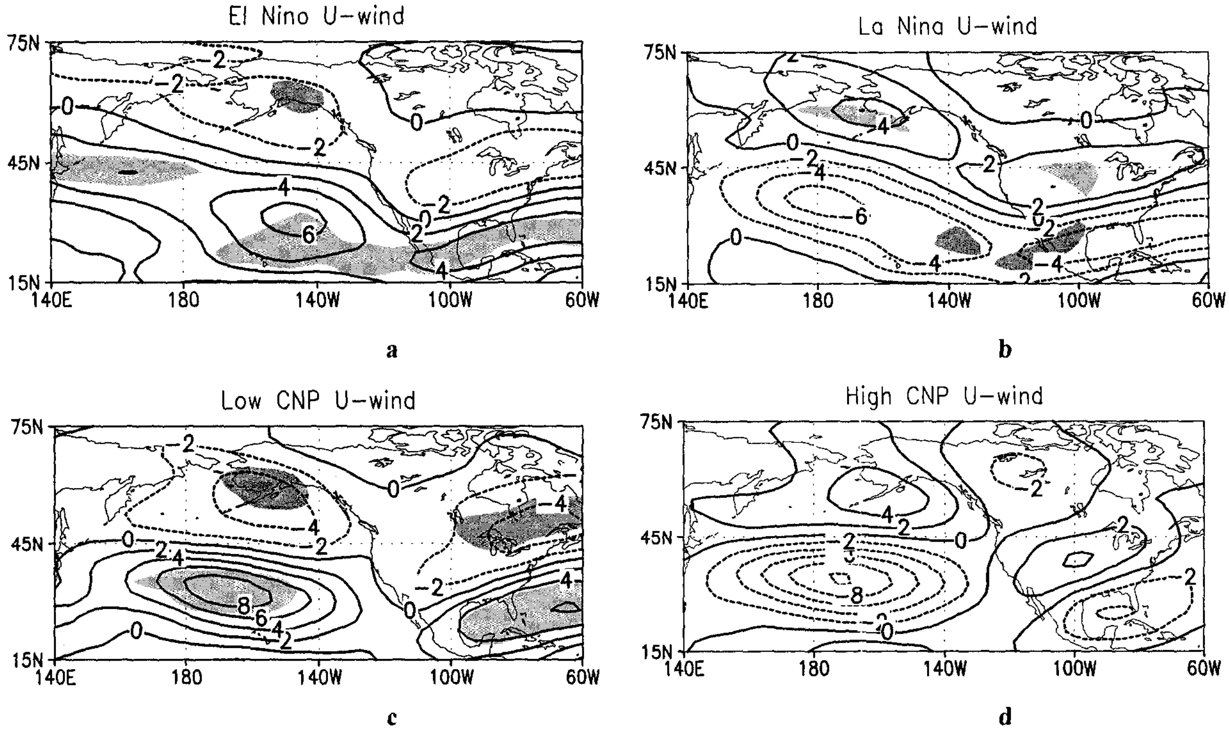


Figure 2.9. The composite DJF 300-mbar geopotential and vector wind anomalies associated with (a) Low CNP winters. (b) El Nino minus Low CNP winters. (c) High CNP winters. (d) La Nina minus High CNP winters. In (a) and (b) height anomalies significant at the 1% level are shaded. In (c) and (d) difference fields significant at the 1% level are shaded. The units are m for geopotential height and ms^{-1} for wind speed.


El Nino U-wind

La Nina U-wind

Low CNP U-wind

High CNP U-wind

a

b

c

d

Figure 2.10. The composite DJF 300-mbar zonal wind anomaly patterns associated with (a) El Nino winters, (b) La Nina winters, (c) Low CNP winters, (d) High CNP winters. Anomalies significant at the 1% level are shaded. Wind speed is in ms^{-1}.

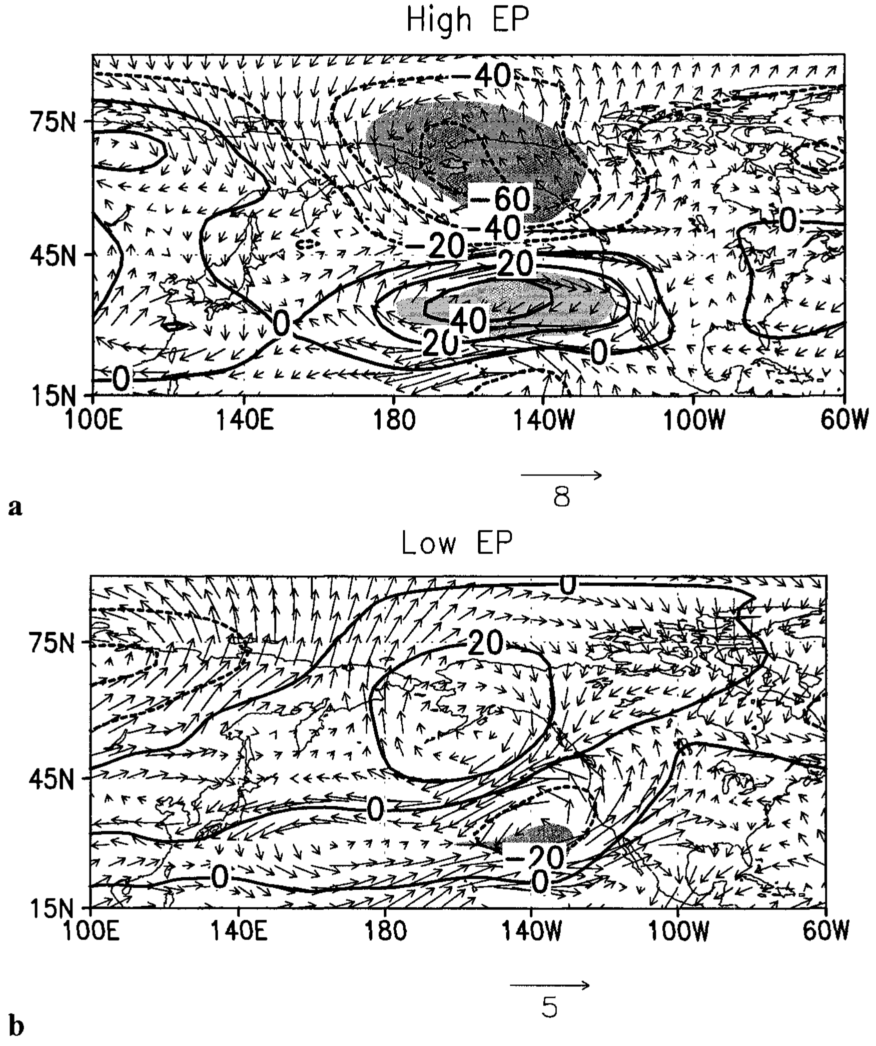


Figure 2.11. Composite DJF 300-mbar geopotential and wind vector anomalies associated with (a) high EP winters, and (b) low EP winters. Geopotential height anomalies significant at the 1% significance level are shaded. The units are m for geopotential height and ms^{-1} for wind speed.

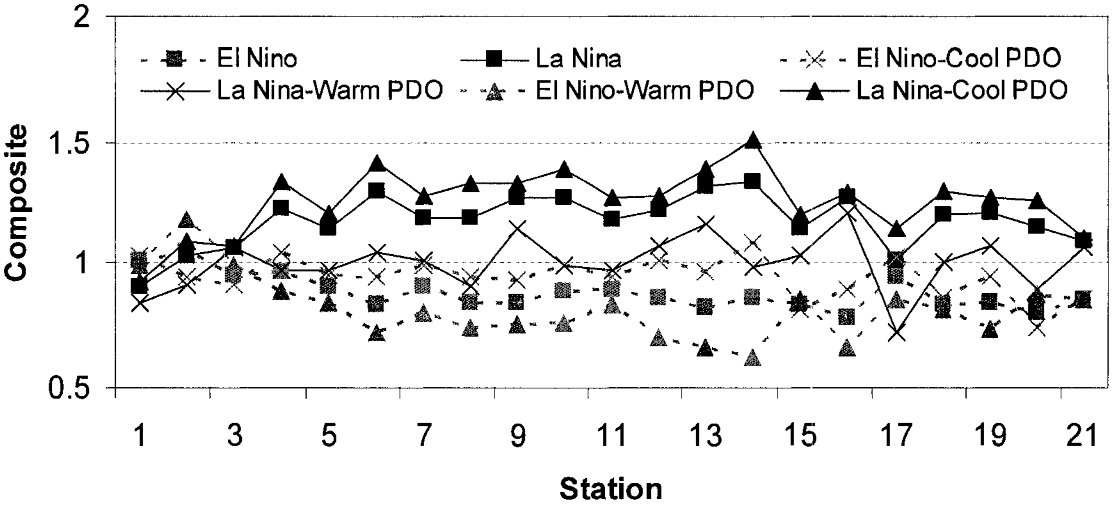


Figure 2.12. Western Canadian winter precipitation responses to ENSO stratified by PDO phases. Station numbers are same as in column 1 of Table 2.1.

Chapter 3 Links between Low Frequency Variability in Western Canadian Streamflow and Large-scale Climate Anomalies

3.1. Introduction

From hydropower generation and fisheries in British Columbia to agricultural production in the Prairie Provinces of Alberta, Saskatchewan and Manitoba, water resources play a prominent role in Western Canada's economic activities. Past studies and experience show that the Prairies' agricultural sector is highly susceptible to drought (Gan, 2000). The region's water resources managers are under increasing pressure from competing users such as hydropower production, municipality, recreation and others. For instance, Alberta has already reached the full allocation potential for which it has rights in the South Saskatchewan River basin. The problem of water availability is even more exacerbated during drought years. The vulnerability of the region to climatic extremes calls for a thorough understanding of the region's hydroclimatic variability to help in developing seasonal to long-range water supply forecasting models.

Relationships between hydroclimatic variability and large-scale climate anomalies such as ENSO could provide predictive skills up to several months of lead time. The ENSO phenomenon is an interannual source of climate variability with its origin in the tropical Pacific but its impact extends into the mid-latitudes, particularly during its "mature" phase in winter (Horel and Wallace, 1981). Several studies have documented precipitation patterns during ENSO events over the US (e.g. Yarnal and Diaz, 1986; Ropelewski and Halpert, 1987; Kiladis and Diaz, 1989), and Canada (Shabbar *et al.*, 1997; Gan *et al.*, 2007).

While streamflow characteristics should broadly reflect changes in the characteristics of the precipitation regime, the non-linear precipitation-streamflow dynamics may give rise to a different spatial and/or temporal picture of climate related signals in streamflow from that of precipitation (Kahya and Dracup, 1993). Streamflow may be preferred over precipitation for studying the impacts of climate variability on regional hydrology because of two reasons: (1) The fact that streamflow temporally integrates climatic forcings may intrinsically emphasize the low-frequency components of the climate system (Rajagopalan *et al.*, 1998), and (2) Unlike precipitation data, streamflow is less prone to contamination by noise in space and time since it is a naturally filtered product of precipitation (Piechota *et al.*, 1997). Even though streamflow represents an integrated response to various climatic inputs and catchment hydrologic processes (Zhang *et al.*, 2001), Yue and Gan (2004) found that annual average streamflow of Canada generally exhibits simple spatial scaling properties.

The relationships between US streamflow and ENSO have been extensively studied (e.g. Cayan and Peterson, 1989; Kahya and Dracup, 1993; Dracup and Kahya, 1994, Piechota *et al.*, 1997) and possible ENSO related response seasons have been identified. The understanding achieved from such studies has subsequently led to improved long-range forecasting capabilities (e.g. Hamlet and Lettenmaier, 1999; Piechota and Dracup, 1999; Piechota *et al.*, 2001). A few studies (Woo and Thorne, 2003; Coulibaly and Burn, 2004) have also explored the relationships between ENSO indicators and Canadian streamflow.

Gan *et al.* (2007) showed that in addition to ENSO, interannual precipitation variability over Western Canada is affected by several large-scale climate patterns that originate in the extratropical Pacific, including the Pacific/North American (PNA), West Pacific (WP) and East Pacific (EP) patterns. The PNA pattern is a natural, internal mode of variability which appears in differing types of diagnostics of the Northern Hemisphere winter atmospheric data such as in the correlation analysis of Wallace and Gutzler (1981) and the rotated principal components analysis of Barnston and Livezey (1987). The pattern shows little sensitivity to the choice of which particular winters in the record are used for the analysis (Straus and Shukla, 2002). Although the high (positive) phase of PNA appears to be favored during El Nino events (Moore, 1996; Trenberth and Hurrell, 1994), extreme phases of PNA can also occur during years with no anomalous tropical sea surface temperature (SST) forcing. The dynamics relating the PNA pattern and the dominant mid-latitude response associated with the equatorial SST forcing is still a major research subject in climatology. While some researchers (e.g. Molteni *et al.*, 1993; Palmer, 1993; Bladé, 1999) argue that the effect of the ENSO-like forcing is to reorganize the preferred internal modes of variability such as the PNA pattern, others (e.g. Straus and Shukla, 2000; 2002) argue that the ENSO forcing can produce mid-latitude circulation responses that are distinctly different from the internal modes of variability. Both groups of researchers used observed and ensemble GCM simulations of atmospheric data to support their respective arguments. Relationships between the PNA pattern and Canadian streamflow have been investigated at the regional scale (e.g. Moore and McKendry, 1996; Woo and Thorne, 2003; Coulibaly and Burn, 2004) and the local scale (e.g. Moore, 1996).

The WP pattern consists of a north-south dipole of height anomalies centered over the Kamchatka Peninsula, portions of southeastern Asia and the lower part of western North Pacific, whereas the EP Pattern reflects a north-south dipole of height anomalies over the eastern North Pacific, with its northern center located around Alaska and the west coast of Canada (Wallace and Gutzler, 1981). The high (low) phase of WP has been associated with drier (wetter) than normal winter precipitation over Western Canada. On the other hand, the EP pattern was linked only to winter precipitation variability over southwestern BC, with its high (low) phase being associated with wetter (drier) than normal conditions (Gan *et al.*, 2007). In this study, we consider the impacts of ENSO, PNA and WP patterns on interannual streamflow variability.

At the interdecadal scale, variations in the climate of western North America have been related to interdecadal oscillations in the North Pacific climate, with its oceanic component depicted by the PDO (Mantua *et al.*, 1997). The PDO is indexed by the leading principal component mode of the North Pacific SST poleward of 20°N (Mantua *et al.*, 1997). Although the dynamics of the PDO are not yet fully understood, the mode shows remarkable persistence over a span of several years. Only 2 complete cycles of PDO occurred between 1890 and 1998, with alternating cool and warm PDO regimes during 1890--1924, 1925–1946, 1947–1976 and 1977–1998. Western North American climate responses to the different phases of PDO are broadly similar to that of ENSO, with warm PDO being associated with warm, dry conditions and cool PDO being related to cool, wet conditions (Mantua and Hare, 2001). The modulating effects of PDO on western US precipitation-ENSO relationships have been reported in McCabe and Dettinger (1999), and Brown and Comrie (2004).

While previous studies on Canadian streamflow variability have provided significant insights into various aspects of the streamflow-climate relationships (e.g. Moore, 1996; Moore and McKendry, 1996; Woo and Thorne, 2003; Coulibaly and Burn, 2004), none of these studies has explicitly investigated the impacts of the different phases of ENSO, which is necessary to objectively identify seasons during which large-scale climate anomalies may provide potential prediction skills. In addition, interdecadal streamflow variability and its (possible) interaction with the variability at interannual scales have not been explored in any of the previous studies. Because of the slow evolution of interdecadal modes of variability such as the PDO, possible relations with streamflow will have far reaching implications for long-range forecasting efforts. Therefore, using a number of well established statistical techniques, this study addresses the following three objectives: (1) To investigate ENSO teleconnections to Western Canadian streamflow anomalies during the lifetime of El Nino and La Nina events, and identify periods when ENSO indicators would likely provide skill for long-range forecasting; (2) To investigate the relative importance of the equatorial (ENSO) forcing and the extratropical climate variability (the PNA and WP patterns) on inter-annual streamflow variability; and (3) To identify and attribute interdecadal streamflow variability and investigate whether ENSO related responses are modulated by such interdecadal oscillations. Upper air wind data were used to provide physical explanation for anomalous streamflow behaviors observed in association with the different phases of ENSO and the PNA pattern. This study contributes to an understanding of the hydroclimatic variability of the Western Canadian region on a broad scale and the potential predictability of streamflow anomalies based on large-scale atmospheric patterns. Details of the data and analysis techniques used in the study are described in sections 3.2 and 3.3, followed by the results in sections 3.4 through 3.7. A discussion of the results is presented in section 3.8, followed by concluding remarks in section 3.9.

3.2. Data Description

A total of 60 basins distributed over Western Canada were used in the study (Figure 3.1 and Table 3.1). Monthly mean, minimum and maximum streamflow data for 33 basins were obtained from the Reference Hydrometric Basin Network (RHBN) dataset (Harvey *et al.*, 1999) and the remaining 27 were obtained from Environment Canada's hydrometric data CD-ROM (HYDAT, 2001). The RHBN consists of hydrometric data for a collection of 255 basins for Canada that are pristine or have stable land use conditions, and have record length of at least 20 years. The 27 basins taken from the HYDAT data have also been identified by Woo and Thorne (2003) as either unregulated or moderately regulated. As a compromise between data length and spatial coverage, only stations having records beginning prior to or by 1960 and with relatively few missing data were selected. The data for the RHBN basins were updated to 2001 with data from the HYDAT CD-ROM. As shown in Figure 3.1, the data coverage in the Prairies Provinces is minimal because of flow regulations. Therefore, interpretation of the results for the Prairies should be handled with caution.

As shown in Table 3.1, the basins selected for the study have drainage areas varying from below 200 km² to over 50,000 km² and span 4 climatic zones (as used in the RHBN database), namely the Pacific Maritime, Montane Cordillera, Boreal forests and Prairies. As such, the raw streamflow values exhibit considerable spatial and seasonal variation. For the analysis involving ENSO, PNA and WP, the streamflow data were first subjected to logarithmic transformation to reduce the disparity in the magnitudes, from which monthly standardized streamflow anomalies (hereafter abbreviated as SSA) were computed so as to remove the effect of differences in drainage size and the annual cycle in streamflow. The use of the SSA instead of the raw values also facilitates inter-comparisons between basins and/or regions. The monthly SSA was constructed by subtracting the mean and dividing by the standard deviation for each month separately. For the study of interdecadal variability, we used the average April-September streamflow, partly because a major proportion of the annual flow occurs during this period and smaller streamflow stations are generally shut down during late autumn to winter months because of freeze-up.

The El Nino and La Nina years used in this study are similar to those of Shabbar *et al.* (1997) and contain moderate to strong El Nino and La Nina episodes where the 5-month moving average of the SOI index remained in the lower (higher) 25% of the SOI distribution for a period of 5 months or longer. This definition leads to 8 El Nino events between 1960 and 2001 (with onset years in 1965, 1969, 1972, 1976, 1982, 1986, 1991 and 1997) and 8 La Nina events (with onset years in 1964, 1970, 1971, 1973, 1975, 1988, 1995 and 1998). The monthly SOI time series was obtained from the Climate Research Unit of the University of East Anglia, UK.

The definition of a high (low) PNA year was based on that of Hsieh and Tang (2001), who used a winter season PNA index of ±0.5 standard deviation as a threshold to define high (low) PNA years. With this definition, 11 high PNA years (1961, 1963, 1970, 1977, 1978, 1981, 1983, 1986, 1987, 1992 and 1998) and 11 low PNA years (1965, 1966, 1969, 1971, 1972, 1979, 1982, 1988, 1989, 1997 and 1999) were identified for the period of 1960-2001. Six of the 11 high PNA years (1970, 1977, 1983, 1987, 1992 and 1998) coincided with the mature phase of El Nino whereas 5 of the low PNA years (1965, 1971, 1972, 1989 and 1999) coincided with the mature phase of La Nina. The definition of high (low) WP was also based on the winter index value of ±0.5 standard deviation as the threshold, resulting in 11 high WP years (1963, 1965, 1974, 1978, 1982, 1986, 1987, 1988, 1991, 1997 and 2000) and 12 low WP years (1960, 1961, 1962, 1964, 1967, 1970, 1973, 1980, 1985, 1990, 1995 and 1997). The PNA and WP indices were obtained from the NOAA Climate Prediction Center website. The monthly mean upper air data on a 2.5°×2.5° grid were obtained from the NOAA-CIRES Climate Diagnostics Center website. The winter anomaly field was defined as the seasonal deviation from the seasonal climatology based on 42 years (1960–2001).

3.3. Analysis Techniques

An appraisal of ENSO impacts on the streamflow was first obtained by applying the nonparametric Mann-Whitney test (see Appendix) to examine if streamflow distributions during El Nino (La Nina) years are significantly different from that of non-ENSO years. Since the test uses the rank of the data instead of real data values, it is robust to outliers and the underlying distribution of the data is irrelevant (Wilks, 1995). This property makes it more suitable for our purpose than the traditional *t*-test because the normality of streamflow distribution cannot be guaranteed even after log-transformation. The statistical test was applied on a month-by-month basis for 36 months beginning with the January of the onset year (abbreviated as Jan(0)) and ending with the December of the 2nd year after onset (abbreviated as Dec(2)).

Regional patterns in the streamflow response to ENSO events were identified using cluster analysis following an approach similar to that of Piechota *et al.* (1997). For each ENSO event, a 24-month SSA, beginning with Jan(0) and ending with Dec(1) was extracted. The 24-month period was chosen because earlier studies (Shabbar and Khandekar, 1996; Shabbar *et al*, 1997) showed that anomalous precipitation and temperature regimes in Western Canada following ENSO events occur within that time period. Streamflow data for only 16 of the 24 months (Mar(0) to Oct(0) and Mar(1) to Oct(1)) were selected for the cluster analysis because preliminary results from the Mann-Whitney test showed that significant differences between ENSO and non-ENSO years are observed during those months. Thus, the time series used in the cluster analysis consisted of 128 data points corresponding to the 8 El Nino (La Nina) events. We applied several hierarchical clustering methods based on the algorithms of Murtagh (1985), and by visual inspection, we selected spatially coherent clusters obtained by using Ward's minimum variance method (see Appendix). The cluster analysis was applied to the time series corresponding to El Nino and La Nina episodes separately. Based on the regional patterns identified from the cluster analysis, aggregate composites for a 36-month period (beginning with Jan(0) and ending with Dec(2)) were used to identify possible ENSO response seasons for each region.

Confidence intervals for the aggregate composites were established using a bootstrap re-sampling procedure (Efron and Tibshirani, 1993; Ntale and Gan, 2004). The bootstrap procedure involves constructing the El Nino (La Nina) database by randomly selecting 8 years from the record period (1960-2001) and constructing the aggregate composite. The procedure was repeated 1000 times and the empirical distributions of the random samples were used to establish the 90% confidence limits. Following previous studies (e.g. Ntale and Gan, 2004), the ENSO related streamflow response periods were then identified as those months during which the SSA was in the lower (upper) 5% of the distribution. Because spring snowmelt in several basins in the study area may be completed in a relatively short period, no threshold was set for the length of the response season. The consistency of the ENSO related response for each region was evaluated using an index time series (ITS) constructed by temporally averaging the SSA for months which were identified as ENSO related response periods.

Regionally averaged streamflow anomaly composites for ENSO, PNA and WP years were used to discriminate streamflow responses associated with the ENSO forcing from that of the extratropical Pacific climate variability. The relative importance of each phase of ENSO, PNA and WP in forcing regional streamflow variability was assessed by comparing the respective composites. Finally, a combination of wavelet transforms and principal components analysis (PCA) was used to identify decadal to interdecadal scale ($\geq$ 8years) streamflow variability and relate it to the North Pacific low frequency oscillation. A composite analysis was used to investigate interactions between the interdecadal and interannual responses of streamflow. Partial correlation analysis provided further insight into the potential of using the large-scale climate anomalies for long-range streamflow forecasting.

3.4. Interannual Variability: ENSO Teleconnections with Streamflow

3.4.1. Significance Test for Differences in Distribution

ENSO events predominantly affect the climate of Western Canada during their mature phase in the winter and early spring (Shabbar and Khandekar, 1996; Shabbar *et al.*, 1997). Most of the winter precipitation in the study area occurs in the form of snow. Therefore, major ENSO related responses of streamflow are expected to occur during and/or after spring snowmelt. The Mann-Whitney test was performed on a monthly basis. However, interpretation of the results should take into account the variation in the timing of snowmelt among different environments in the region (Woo and Thorne, 2003). Table 3.2 shows the statistical test results for the monthly mean, minimum and maximum SSA during March to October for the onset year (left), the following year (middle) and the 2nd year following the onset (right). The results are based on a 2-sided test at a significance level of 10%. Assuming independent samples, at least 7 of the 60 basins should have significant differences for field significance in any given month.

For La Nina events, basins with significant above-normal SSA emerge in Apr(1) and extend into early autumn, with the number peaking in Aug(1) (Table 3.2). Figure 3.2 shows the spatial distribution of basins with significant above-normal SSA during Apr(1) to Sep(1) of La Nina events for the monthly mean flow. During Apr(1) to May(1), basins with statistically significant above-normal anomalies (Figure 3.2, solid circles) are found in the Prairies, Vancouver Island, Central BC Plains and Southern BC Coast. In the spring, basins in these regions receive runoff predominantly from seasonal snowmelt. Between Jun(1) and Aug(1), the region of statistical significance mainly shifts to basins in the Canadian Rockies and Southern BC Interior, which receive the major portion of their runoffs from a combination of summer rain/mountain snowmelt and glacier melt. Given that there are very few basins with above-normal SSA for all 3 flow variables during the onset year and the second year following the onset (Table 3.2), the effect of La Nina is clearly evident. It is worth noting that for all three flow variables, more basins exhibit significantly above-normal anomalies during Aug(1) than during Jun(1) or Jul(1) of La Nina events. This is in contrast to previous studies for the region (e.g. Woo and Thorne, 2003) that are mainly limited to the relationships between ENSO and spring snowmelt runoff covering the periods of April to June.

For the sake of comparison, the results from the standard *t* test for La Nina are shown in parenthesis in Table 3.2. The null hypothesis of no difference between the two methods was tested using the binomial distribution in the following manner (see Appendix for a description of the binomial distribution). First, one of the methods (e.g. Mann-Whitney test) was assumed to give the 'true' number of basins with significant differences between La Nina and non-ENSO years. Then the 95% confidence region on the other value (e.g. *t* test) was constructed using a

Gaussian approximation of the binomial distribution (Wilks, 1995). The null hypothesis of no difference was accepted for all months since the results from the Mann-Whitney test for each month fell within the bounds of the 95% confidence region of the *t* test results.

The numbers of basins with significantly below-normal anomalies during El Nino events are also given in Table 3.3. Although there is a tendency for more basins with below-normal SSA during the summer and fall months following the onset year, the impacts of El Nino are not as strong as that of La Nina events. These results were also confirmed by the standard *t*-test (not shown). Perhaps, the negative impacts of El Nino events may be appreciated by looking at the number of basins that transited from above-normal conditions during the onset year to normal or below-normal conditions during the following year (see Table 3.4).

3.4.2. ENSO Related Spatial Patterns and Response Periods

Using the Ward's minimum variance cluster analysis, five spatially coherent patterns were identified from both the La Nina (Figure 3.3) and El Nino time series (not shown). The number of clusters was subjectively determined using the inflection point on the plot of the dissimilarity measure versus the number of clusters. In 55 of the 60 basins, the clusters for El Nino events agreed with that of La Nina events. Exceptions are observed for 4 basins located in Northern BC and one basin in Northern Manitoba (see Table 3.1). The clusters may reflect differences in ENSO response periods due to local influences (e.g. topography) or spatial variations in the atmospheric conditions that prevailed over the region during individual ENSO episodes, particularly the relative location with respect to the western Canadian height anomaly center. Some of the clusters include basins from more than one climatic region, e.g. basins in the Montane Cordillera fall into 3 different clusters. For the purpose of this study, those regions identified from the La Nina time series (Figure 3.3) were used as a basis for further regional analysis of the streamflow response to a 'typical' ENSO event. The remainder of the discussion will be based on the monthly mean flow variable since the spatial patterns obtained using the maximum and minimum flows are similar to that of the mean flow.

1. Region 1 (Canadian Prairies): This region encloses basins in the central and southern Prairies, where the Mann-Whitney test showed significant differences between the SSA of ENSO and non-ENSO years during the Apr(1) to May(1) period (see Figure 3.2) . Figure 3.4a shows the regional aggregate composite for this region. The horizontal lines in Figure 3.4a correspond to the lower 5% (upper 95%) of the distribution based on the bootstrap confidence interval. From Figure 3.4a, La Nina events are associated with significant positive anomalies from Apr(1) to Jun(1). The result is consistent with the positive precipitation anomalies during the winter season following La Nina events (Shabbar *et al.*, 1997; Gan *et al.*, 2007), with the delay in the streamflow response due to the timing of spring snowmelt. Although only about a third of the precipitation in the Canadian Prairies occurs in winter, spring snowmelt contributes as much as 80% of the annual surface runoff in some areas (Granger and Gray, 1990). The index time series (ITS) associated with La Nina events for this region was constructed by averaging the Apr(1) to Jun(1) SSA (3-month period) and is shown in Figure 3.4b. Six of the 8 La Nina years recorded positive SSA. The horizontal lines in Figure 3.4b correspond to the 10th and 90th percentiles of the ITS distribution. For the purpose of this study, periods with ITS values in the lower (upper) 10% of the distribution will be designated as extremely dry (wet). Two of the 4 extremely wet years are La Nina years, with the wettest year in the record (1974) associated with the 1973/74 La Nina event.

The regional composite for El Nino events (Figure 3.4c) shows a tendency for negative anomalies during Apr(1) to Jun(1) but the anomalies are not significant and so no ITS is constructed for category. However, these results should be used with caution because of the sparse data coverage.

2. Region 2 (Southern BC Interior and Rockies): This region comprises basins in the Southern BC Interior and the Canadian Rockies, where the Mann-Whitney test showed significant differences in the distribution of SSA during La Nina and non-ENSO years during Jun(1)-Aug(1) (see Figures 3.2 and 3.3). Basins in this region receive runoff from summer rain, mountain snowmelt and glacier-melt. Mountain snow and glaciers contain a memory of prior precipitation and temperature regimes (Vaccaro, 2002). Thus, above-normal snowpack conditions during winter/spring would likely lead to above-normal SSA during the summer period. During La Nina events (Figure 3.5a) significant positive anomalies are observed from Jun(1) to Aug(1), with a pronounced peak in Aug(1). The ITS constructed from the Jun(1) to Aug(1) period (Figure 3.5b) also shows the positive effects of La Nina, with 7 of the 8 years experiencing above average anomalies. All 4 extremely wet years in the record (those anomalies in the upper 90 percentile of the ITS data) have occurred following La Nina events.

Significant negative responses during El Nino years emerge in Jun(1) and extend to Oct(1) for this region (Figure 3.5c). The El Nino response period is longer than that of La Nina but the anomalies are not as pronounced as that of the latter. The result basically agrees with that of Hsieh and Tang (2001) who suggested that La Nina events may have far more impact on the snow water equivalent anomalies of the Columbia River basin than El Nino events. For the Pacific Northwest region of the USA, which lies immediately below this region, Dracup and Kahya (1994) showed that the La Nina signal has larger magnitudes than the El Nino signal. However, the length of the response season identified in Region 2 is shorter than that for the Pacific Northwest. The ITS based on the Jun(1) to Oct(1) response season (Figure 3.5d) shows that 7 of the 8 El Nino years are associated with dry conditions. Three of the five extremely dry years (anomalies in the lower 10 percentile) are associated with El Nino events. In terms of the temporal distribution, the impact of El Nino seems to be evenly distributed over the record period whereas that of La Nina dominated during the early part of the record, especially in the 1970's.

3. Region 3 (Central Boreal Plains): This region includes basins in the Boreal plains of western central Alberta and central BC (see Figures 3.2 and 3.3). The regional composite shows that following La Nina events, significant positive anomalies are observed during the Apr(1) to Aug(1) period (Figure 3.6a). From Figure 3.6b, the ITS based on this season shows that 7 of the 8 La Nina years are associated with wet SSA. With 3 of the 5 extremely wet years in the record also being associated with La Nina years, the result seems to confirm the positive effect of La Nina in this region. On the other hand, the composite for El Nino (Figure 3.6c) shows a shorter response period when compared to that of La Nina, with Jun(1) to July(1) being in the lower 5% of the distribution. The ITS for El Nino (Figure 3.6d) shows that 6 of the 8 El Nino years are associated with dry anomalies and 3 of the 5 extremely dry SSA in the record are associated with El Nino events.

4. Region 4 (Vancouver Region): This region comprises basins in the Greater Vancouver area that experience the Pacific Maritime climate, with 5 basins from the Vancouver Island and one basin from the Southern BC Coast (see Figure 3.3). The composite for La Nina in this region shows that significant wet anomalies exist during Apr(1) and Jul(1) (Figure 3.7a). The La Nina response season for this region is selected to be Apr(1) to May(1) since the May(1) composite is also close to the significance level. As shown in Figure 3.7b, the ITS constructed from the Apr(1) to May(1) SSA shows that 7 of the 8 La Nina events are associated with above average anomalies, with 2 of the 5 extremely wet anomalies being associated with La Nina. For El Nino, significant dry conditions are observed during Apr(1) to Jun(1) and Aug(1) (Figure 3.7c). The ITS based on the Apr(1) to Jun(1) SSA shows that 7 of the 8 El Nino events are associated with dry anomalies, with 3 of the 5 extremely dry years being associated with El Nino (Figure 3.7d). There appears to be intensification in the streamflow response to El Nino in the latter part of the record while La Nina responses were dominant in the early 1970's. The interdecadal variation in the strength of ENSO related responses appears to be well organized across Western Canada (see section 3.6).

5. Region 5 (Northern BC): This region that comprises basins in the Boreal Cordillera of Northern BC, shows a tendency of below-normal anomalies during the early periods of La Nina events, with significant negative anomalies being observed during Oct(0) to Dec(0) and then reverses its sign in Jun(1) but there are no significant positive anomalies observed (Figure 3.8a). During El Nino events, no significant response exists during the onset and the following year, with significant negative anomalies being observed only during Dec(1) to Jan(2) (Figure 3.8b). Since no significant response period exists during the spring/summer following the onset of both La Nina and El Nino events, the impacts of ENSO on the streamflow anomalies of this region are deemed insignificant and no ITS is constructed. This result is mostly in agreement with the Mann-Whitney test result for basins in this region except at three stations (Skeena, Morice, Bulkley) that show significant differences in the distribution of SSA during La Nina and non-ENSO years.

3.5. Effect of ENSO Forcing Versus Extratropical (PNA and WP) Variability

In this section, composite analysis is used to distinguish between Western Canadian streamflow anomaly responses to ENSO and PNA. A similar analysis for ENSO and the WP pattern was also performed. In order to capture the overall picture of the similarity/difference between responses related to ENSO and PNA over the study area, regionally averaged composite streamflow anomalies were constructed excluding basins in Region 5 (Northern BC) since no ENSO related response was identified for that region (see Section 3.4). Two types of regionally averaged composites were formed for the PNA pattern: (1) all high (low) PNA years irrespective of whether El Nino (La Nina) occurred or not, abbreviated as High PNA (Low PNA); and (2) high (low) PNA years conditioned on non-El Nino (non-La Nina) years, abbreviated as High PNA/non-El Nino (low PNA/non-La Nina). The exclusion of ENSO years from High PNA/non-El Nino and low PNA/non-La Nina composites ensures that any anomalous streamflow responses during those years are primarily due to internal atmospheric variability over the North Pacific/North American sector (e.g. see Straus and Shukla, 2002). Note that only 6 (5) of the 11 high (low) PNA winters during the period of 1960-2001 coincide with El Nino (La Nina) years.

The composite anomalies for El Nino, High PNA, and High PNA/non-El Nino years are shown in Figure 3.9a. Both El Nino and High PNA years are associated with below normal SSA over Western Canada during the spring and summer following the occurrence of the events. There is apparently no significant difference between the El Nino, High PNA, and High PNA/non-El Nino composites during most of the response season. However, the impacts of High PNA/non-El Nino tend to be of shorter duration than those of El Nino and High PNA years.

The regional composite anomalies for La Nina, Low PNA, and Low PNA/non-La Nina years are shown in Figure 3.9b. The Low PNA, and Low PNA/non-La Nina composites show a modest above-normal SSA during spring. On the other hand, La Nina events produce stronger positive anomalies that last for a longer period extending from the spring to summer. Application of the *t*-test showed that the spring and summer La Nina composites are significantly different from both the Low PNA, and Low PNA/non-La Nina composites at the 5% level. These results suggest that at the regional scale, the high phase of the PNA pattern may be associated with below-normal streamflow responses that are distinctly different from that of the ENSO forcing. On the other hand, streamflow responses associated with La Nina events dominate above normal anomalies and thus the role of the Low PNA years is not clear. This result agrees with the atmospheric analysis of Straus and Shukla (2002) in that both ENSO and PNA can produce anomalous responses independently of each other, particularly following their warm phases. The consequence of this result is that the presence of significant impacts of the internal mode of variability limits our ability to make successful long-range streamflow forecasting because of the unpredictability of the internal variability. However, more work using other hydroclimatic variables and preferably longer datasets is needed to confirm these findings.

To understand the atmospheric flow patterns that prevailed over Western Canada during the different ENSO and PNA categories considered above, a composite analysis of the 300-mbar wind anomalies and the vertically integrated moisture flux (VIMF) was performed for the El Nino, High PNA/Non-El Nino, La Nina, and Low PNA/non-La Nina winters. The VIMF is a measure of the flux of precipitable water in a column of unit area and is described by the magnitude and direction computed from vertically integrated zonal and meridional moisture fluxes:

$$Vx = \frac{1}{g}\int_{P_1}^{P_2} uq\,dp; \qquad Vy = \frac{1}{g}\int_{P_1}^{P_2} vq\,dp$$

where g is the acceleration due to gravity (ms^{-2}), u is the zonal wind speed (ms^{-1}), v is the meridional wind speed (ms^{-1}), q is the specific humidity (kg/kg), p is the pressure (Pa), and Vx and Vy are the vertically integrated zonal and meridional moisture fluxes ($Kgm^{-1}s^{-1}$), respectively. For our study, the VIMF composite was evaluated using 8 discrete pressure levels extending from the surface to the 300-mbar level. To show the relative strength of one field with respect to the other, the differences between El Nino and High PNA/Non-El Nino, and between La Nina and Low PNA/Non La Nina, are presented in Figures 3.10 and 3.11.

Figure 3.10a shows the wind anomaly difference composite between El Nino and High PNA/non-El Nino winters. Similarly, the VIMF difference composite between El Nino and High PNA/non-El Nino winters is shown in Figure 3.10b. It is well known that El Nino events can lead to a split in the jet stream and a subsequent southward displacement of the subtropical branch (e.g. Shabbar *et al.*, 1997). The strength of the net northwesterly flow over the eastern Pacific region in Figure 3.10a is consistent with this observation. Figure 3.10a suggests that the dominant flow pattern over Western Canada and the Pacific Northwest regions would be directed in the northwest-southeast direction. The southeasterly flow would bring to the region warm dry continental air that has already lost its moisture content over the southwestern USA whereas the northwesterly flow would bring to the region a cool dry Arctic air. There is also little difference between the VIMF fields of the two categories over Western Canada (Figure 3.10b). Thus, both El Nino and High PNA/non-El Nino years would lead to below-normal winter snowpack over Western Canada. This explains the similarities of the streamflow responses observed from Figure 3.9a for these two categories.

Figure 3.11a shows the vector wind anomaly difference composite between La Nina and Low PNA/non-La Nina winters whereas Figure 3.11b shows the VIMF difference composite between La Nina and Low PNA/non-La Nina winters. La Nina winters are associated with an erosion of the western Canadian ridge and strengthening of the Pacific westerlies. The net westerly flow in Figure 3.11a suggests that the westerlies over the eastern North Pacific and western Canada are stronger for La Nina winters than for Low PNA/non-La Nina winters. The associated VIMF difference field also shows a net eastward flux of moisture over the eastern North Pacific and Western Canada (Figure 3.11b). Thus, the greater moisture supply during La Nina winters would result in better winter snowpack and spring/summer streamflow than during Low PNA/non-La Nina winters (Figure 3.9b).

The streamflow composite analysis was repeated for all high (low) WP years and high (low) WP years conditioned on non-El Nino (non-La Nina) years (Figure 3.12). A comparison of Figures 3.9 and 3.12 suggests that unlike for ENSO and PNA, the effect of WP is mainly felt during winter and early spring months. However, most of the streamflow in Western Canada occurs during spring and summer months and as such the WP pattern does not seem to contribute much to the variability during the major flow season. Hence, the effects of WP are not discussed further.

3.6. Interdecadal Streamflow Variability

Most of the basins used in the foregoing sections have short records that were not suitable for identifying low frequency variability at interdecadal scales without ambiguity. As a first step, only 13 basins with more than 60 years of data were selected to study the general characteristics of the decadal or higher scale variability in streamflow of the region. The April-September mean flow of each basin was subjected to a continuous (Morlet) wavelet transform to determine if there are significant fluctuations at the interdecadal scale. Then the signal corresponding to the interdecadal ($\geq$ 8 years) scale was lowpass filtered using Equation 2.7 (Torrence and Compo, 1998). The lowpass signals were then subjected to PCA to retain the dominant low frequency mode of streamflow variability across the region. The leading PC mode (explaining 48.1% of the variance) was then compared to three lowpass filtered North Pacific climate indices, namely the PDO, NPO (North Pacific oscillation) and CNP (Central North Pacific) indices (Figure 3.13). Although both CNP and NPO indices also depict the variability in the streamflow at decadal scales ($\rho = 0.61$ for CNP and $\rho = 0.73$ for NPO), the streamflow-PDO relationship appears to be stronger than with the two indices (Figure 3.13). There is an inverse relationship between the filtered PDO and streamflow signals ($\rho = -0.75$), with warm (cool) PDO being associated with below- (above-) normal streamflow. It is interesting to note that the short-lived sign reversal of the PDO index between 1959 and 1961 was associated with a similar sign reversal in the streamflow PC. The sign reversal is due to the presence of significant power at a near 20-year scale for some of the basins used in the analysis as well as for the PDO index. This scale happens to be one of the two periodicities at which PDO has tremendous concentration of power, with the other being at 50-70 years (Mantua and Hare, 2001).

The strength of ENSO related streamflow responses should be related to the decadal PDO variability since we expect similar (opposite) phases of PDO and ENSO to complement (negate) each other. One way to investigate the interactions between the PDO and ENSO related responses is to stratify the streamflow into climate categories defined by a combination of the ENSO and PDO states (e.g. Bonsal *et al.*, 2001). Following this approach, 8 climate categories were defined: warm PDO, cool PDO, El Nino, La Nina, El Nino/warm PDO, La Nina/warm PDO, El Nino/cool PDO, and La Nina/cool PDO. For each basin, the April-September mean flow data in each category was composited and the ratio of each composite with respect to the 1947–1998 mean streamflow was computed. The composite analysis included 49 basins with data starting prior to or by 1957.

Figure 3.14 summarizes the ratios for the 8 categories, where the ratios for pairs of ENSO/PDO categories are plotted against each other for easy comparison (e.g., La Nina vs. El Nino, cool PDO vs. warm PDO, La Nina/warm PDO vs. El Nino/warm PDO, and La Nina/cool PDO vs. El Nino/cool PDO). The 10th and 90th percentiles of the ratios for each climate category are also given in Table 3.5. The horizontal and vertical lines on Figure 3.14 correspond to a ratio of unity. A shift towards any of the four quadrants is assumed to be due to the climate anomalies of the associated category. In agreement with the results obtained in Section 3.4.2, the majority of the ratios for La Nina vs. El Nino fall in the lower right quadrant of Figure 3.14. Similarly, most of the ratios for cool PDO vs. warm PDO lie in the lower right quadrant. This confirms that over Western Canada, both El Nino and warm PDO years are generally associated with below normal streamflow (ratio less than 1.0) whereas both La Nina and cool PDO years are associated with the opposite condition (ratio greater than 1.0).

When the El Nino and La Nina years are stratified according to the PDO phase, considerable changes occur in the strength of the streamflow response, especially during the warm phase of PDO. For several basins, the ratios for La Nina/warm PDO vs. El Nino/ warm PDO fall in the lower left quadrant of Figure 3.14. This means that during a warm PDO regime, not only El Nino but also La Nina events can, on average, be associated with below normal streamflow. Figure 3.14 also suggests that drier than normal responses associated with the El Nino/warm PDO category appear to be enhanced as compared to the typical El Nino related response. On the other hand, during cool PDO regimes, El Nino related responses are near neutral at several basins as the ratio is close to unity while wetter than normal responses in the La Nina/cool PDO category are on average almost identical to the typical La Nina response. As given in Table 3.4, the 10th and 90th percentiles of the ratios for the La Nina/warm PDO and El Nino/cool PDO categories are almost equally spread around unity, showing that the wetter (drier) than normal effect of La Nina (El Nino) is offset by the drier (wetter) than normal effect of warm (cool) PDO.

The interdecadal variations in the ENSO related responses observed in Figures 3.4 to 3.7 can now be explained in the context of the PDO modulation. The near-normal/positive streamflow responses during the 1965/66 and 1969/70 El Nino events occurred during the 1947–1976 cool PDO regime that suppressed the impact of the interannual El Nino signal. The enhanced La Nina impacts during the early 1970s are due to the same cool PDO regime that amplified the interannual La Nina signal. Similarly, the enhanced (reduced) El Nino (La Nina) related responses of the 1980s and 1990s appear to be due to the constructive (destructive) interference from the 1977–1998 warm PDO regime.

3.7. Correlations with Climate Indices

In this section we use correlation analysis to assess the potential of using information from the three climate anomalies (indexed by the SOI, PNA and PDO indices, respectively) for use in long-range streamflow forecasting. Since the three indices are correlated to one another, we computed the partial correlation between the streamflow and a climate index because partial correlations can help identify if a climate index contributes independent information at a site or not (Pizarro and

Lall, 2002). Our analysis focuses only on the periods identified as La Nina-related response seasons in Section 3.4.2, i.e. April-June for Region 1, June-August for Region 2, April-August for Region 3 and April-May for Region 4. We will not consider basins in Region 5 since no ENSO-related response season was identified. We will use standardized anomalies constructed from the untransformed streamflow data and the non-parametric Spearman rank correlation since it is resistant to outliers (Wilks, 1995). The partial correlation (r) between two variables x and y given variable z is computed from $r_{xy|z} = (r_{xy} - r_{xz} \times r_{yz}) / [(1 - r_{xz}^2)(1 - r_{yz}^2)]^{1/2}$.

For each climate index, the average of a 3-month window was considered as the candidate predictor, beginning with March-May of the year preceding the streamflow season (MAM(0)) and ending with January-March of the streamflow season (JFM(1)). Thus a total of 11 predictor windows were considered. For a streamflow season beginning in Apr(1), this provides lead times of up to 10 months. Detailed results are presented for Regions 2 and 4 in Figures 3.15-3.17 since the correlations observed for several basins in these regions appear to be promising. The correlations were computed for a common period of 40 years (1960-1999). With 37 degrees of freedom, partial correlation coefficients of 0.31 and 0.40 are statistically significant at the 5% and 1% level, respectively. The 1% and 5% significance levels are shown in Figure 3.15-3.17 by the solid and broken horizontal lines, respectively.

In a statistical forecasting exercise, one would choose a predictor that maximizes the amount of variance explained in the predictand that is attributable to the selected predictor. This could be achieved by choosing predictors from a time window where the correlation between the predictor and predictand is the highest. In Figure 3.18, we show the spatial distribution of the statistical significance levels for the highest partial correlation (in the absolute sense) of the 11 values for each basin computed using the 11 predictor windows considered in the analysis.

Only positive correlations with SOI and negative correlations with PNA and PDO are included in the maps.

Figure 3.15 shows the temporal evolution of the correlations between streamflow and SOI|PDO and PDO|SOI. The patterns in Figure 3.15a suggest that three types of basins are lumped together in this region: (1) Four of the 23 basins, namely Lillooet, Mistaya, Chilko and Oldman, show no significant correlations with SOI|PDO (the first two have glacierized flow regimes); (2) Six basins, namely Elk, Crowsnest, Waterton, Belly, Squamish and Salmo, show significant correlation with the summer to autumn SOI|PDO; and (3) The remaining thirteen basins show significant correlation with SOI|PDO for an extended period from spring to winter. The ENSO phenomenon commences in the equatorial Pacific during the summer to autumn period and reaches its mature phase in winter. Since previous studies (e.g. Shabbar *et al.*, 1997; Gan *et al.*, 2007) have showed ENSO teleconnections to occur during the boreal winter, it seems that only basins in the third group are directly affected by the mature phase of ENSO. Thus, winter SOI would be a more appropriate predictor choice for basins where snowmelt dominates the seasonal streamflow whereas for some other basins summer or autumn SOI would give better skill.

For Region 2, correlations with PDO|SOI are significant during the autumn to winter period except at two glacierized basins (Lillooet and Mistaya) (Figure 3.15b and 3.18b). Relatively strong correlations with the November-January PDO|SOI (significant at the 1% level) are observed for basins originating in the Rocky Mountains (e.g. Bow, Belly, Waterton, Crowsnest, Kootenay, Columbia and Elk), which also show significant correlations with the summer-autumn SOI|PDO. The fact that three of these basins (Columbia, Kootenay and Bow) receive some of their flow from glaciermelt in addition to the contribution from spring-summer snowmelt suggests that decadal scale climate oscillations (i.e. PDO) that affect the glacier balance also play a more important role than interannual climate variability (i.e. ENSO). Regarding Region 4, only SOI appears to play a significant role in the streamflow variability (Figure 3.15c-d and 3.18a-b). The spatial pattern of the partial correlations with SOI|PDO (Figure 3.18a) for the Mountains (Region 2) and Plains (Region 3) is in general agreement with that reported by Woo and Thorne (2003) who computed correlations between the October-March SOI and annual streamflow. However, Woo and Thorne's analysis did not show statistically significant correlations for basins in Region 4.

The temporal evolution of the partial correlations between streamflow and PDO|PNA and PNA|PDO are shown in Figure 3.16. For Region 2, PDO|PNA shows significant correlation with all but three basins, with the highest correlations observed again for basins in the Rocky Mountains area (Figure 3.16a and 3.18c). On the other hand, correlations with PNA|PDO turn out to be not statistically significant (Figure 3.16b and 3.18d). The reverse is true for Region 4 where the correlations with PNA|PDO are significant during winter (Figure 3.16d and 3.18c) but not with PDO|PNA (Figure 3.16c and 3.18d). In fact, the correlations with PDO|PNA show a tendency to move in the opposite direction (Figure 3.16d).

Figure 3.17 shows the evolution of the partial correlations between streamflow and SOI|PNA, and between streamflow and PNA|SOI. The partial correlations with SOI|PNA for both Regions 2 and 4 (Figure 3.17a, 3.17c and 3.18e) essentially show a similar pattern to that of SOI|PDO (cf. Figure 3.15a, 3.15c and 3.18a). The PNA pattern (Figure 3.17c-d and 3.18e-f) appears to exert statistically significant influence on several basins in all four regions (see Figure 3.18) but at different scales (see next paragraph). However, the connection with PNA appears to be weaker than that with SOI and PDO.

The correlation analysis can also be interpreted in terms of the dominant time scales at which the three climate anomalies operate. Since the dominant mode of variability of ENSO is at interannual time scales, these results confirm that interannual variability in streamflow in several Western Canadian basins is teleconnected to ENSO (Figure 3.18a and 318e). On the other hand, the dominant mode of variability of PDO is at decadal to interdecadal scale and so decadal fluctuations for basins in Region 2 are teleconnected to the PDO regime (Figure 3.18b and 3.18d). The PNA pattern exhibits both interannual and inter-decadal scale fluctuations (e.g. Gan *et al.*, 2007). Thus, PNA exerts significant influence on the interannual variability in several basins across the study area (Figure 3.18f). The fact that the partial correlations with PNA|PDO are statistically significant for basins in Region 4 (Figure 3.18c) suggests that decadal fluctuations of streamflow in this region are likely connected to the PNA pattern but not to the PDO regime. The correlation analysis also confirms our earlier assertion that the effects of ENSO are modulated at the decadal scale by the PDO and at interannual scale by the PNA.

For Regions 1 and 3, the partial correlations between streamflow and SOI, and between streamflow and PNA are generally found to be stronger than those with PDO (e.g. cf. Figure 3.18a and 3.18b). However, the predictor window with the highest partial correlation for all three indices varied from basin to basin without exhibiting any clear spatial pattern. The lack of sufficient data coverage makes it difficult to make a definitive interpretation of these results.

3.8. Discussion of Results

Statistical significance test and composite analysis showed that ENSO related streamflow responses in certain regions of Western Canada occur during the spring and/or summer seasons following the onset of the ENSO events. While the length of the response seasons varies from one region to the other, basins in Regions 2, 3 and 4 typically experience wet (dry) conditions following La Nina (El Nino) events. Basins in Region 1 showed above average conditions following La Nina events but their response to El Nino events could not be clearly detected. The regional composites for Regions 2 and 3 show that El Nino onset years are associated with above-normal streamflow anomalies. This could be due to the biennial tendency associated with the Southern Oscillation, which has also been found in streamflow signals elsewhere (Dracup and Kahya, 1994).

The regional streamflow conditions during individual ENSO events are summarized in Table 3.5. The extremely wet (dry) years indicated in Table 3.5 generally agree with those during which the precipitation anomalies for the January-March (JFM) period following the onset of ENSO events were in the lower (upper) 10% of the distribution (Shabbar *et al.*, 1997). For instance, the precipitation for JFM 1973 (following the 1972 El Nino) was in the driest 10% of all JFMs whereas those of JFM 1972 and 1974 were among the wettest 10%. A one-on-one comparison with the results of Shabbar *et al.* was not possible because their study was based on a longer dataset covering 1910-1994 and most of the extreme precipitation cases occurred prior to 1960.

The relatively low number of ENSO related extreme streamflow conditions (Table 3.5) and modest correlations with SOI (Figure 3.15 and 3.17) suggest that while the ENSO forcing may have an impact on shifting the distribution of streamflow anomalies, it may not always dictate the occurrence of individual extreme events. This is also supported by the finding that the internal variability of the North Pacific climate (i.e. the PNA pattern) may produce anomalous streamflows that are distinctly different from that of El Nino events. In addition, ENSO related responses were found to be modulated by interdecadal oscillations in the North Pacific climate as depicted by the PDO. The interaction between ENSO and PDO was found to be constructive when the two are in phase (i.e. during El Nino/warm PDO, La Nina/cool PDO) and destructive when they are not in phase (i.e. during El Nino/cool PDO and La Nina/warm PDO). These interferences would lead to a nonlinear ENSO-streamflow relationship, and hence weak correlations with SOI. Investigation on the role of other internal modes of variability such as the Arctic Oscillation (AO) and North Atlantic Oscillation (NAO) may also reveal additional interactions and warrants further research.

Apparently, successful long-range forecasting of streamflow requires the state of the PDO and PNA in addition to that of ENSO. Whereas skillful predictions of ENSO have been achieved at lead times of several months (e.g. Tang *et al.*, 2000), the PNA pattern is a result of the nonlinear internal dynamics of the North Pacific and can not be skillfully predicted at present. Similarly, long-range forecasts for the PDO are currently not available. However, the slow evolution of the North Pacific climate system at decadal scales means that the future PDO state may be anticipated using analogs from past observations (McCabe and Dettinger, 1999). Regardless of the state of the PNA pattern, the consistency of the signs of ENSO related responses in certain regions and the persistent behavior of PDO means that categorical water supply outlooks may be issued for certain basins based on the state of these two climate regimes.

The identification of ENSO related spatial patterns and response seasons in this study assumed that there is a 'typical' ENSO event. However, it is known that the climate of western North America is sensitive to relatively small changes in the location of the height anomaly centers (Yarnal and Diaz, 1986). Thus, the associated variation in the storm track position may lead to significant variations in regional precipitation (and streamflow) response during individual ENSO events. For instance, during the 1988/89 La Nina event, there was a major amplification and eastward extension of the Pacific ridge that spanned over a large landmass of western North America (Molteni *et al.*, 1993). Western Canada was dominated by a northwesterly flow which would bring cool dry Arctic air into the region. Thus, several basins experienced below-normal streamflow instead of the usual wet conditions associated with La Nina years (see Table 3.5). Similarly, the 1965/66 El Nino event was associated with streamflow responses that were contrary to that of a typical El Nino event, with above-normal conditions observed in several basins (Table 3.5). As shown in Section 3.6, these variations result from decadal modulations of the strength of ENSO-streamflow relationships.

It should be noted that a different set of spatial patterns and ENSO response seasons may be identified if the analysis is carried out on ENSO events sorted by the PDO phase. This approach has not been explored in the current study due to the limitation in the length of the streamflow record. However, the stability of the results was examined for the extreme scenario by excluding the two ENSO events with the weakest streamflow responses. Accordingly, the exclusion of the 1965/66 El Nino event from the analysis did not affect the length of the response periods identified in Section 3.4.2. When the 1988/89 La Nina event was left out, the response season for Regions 1 and 2 extended to Aug(1) and Sep(1), respectively whereas that of Regions 3 and 4 remained unchanged. The boundaries between Regions 3 and 4 were also found to be sensitive to the exclusion of a particular ENSO year from the cluster analysis. The ENSO sample size of only 8 events is one major limitation of this study.

3.9. Summary and Conclusions

By applying an ensemble of statistical methods to 42 years of streamflow data for 60 basins in Western Canada, this study investigated the spatial and temporal characteristics of streamflow anomalies during the lifetime of ENSO events and their interactions with other climate anomalies such as PNA and PDO. The nonparametric Mann-Whitney test for distribution and composite analysis showed that basins in certain regions of Western Canada exhibit strong response to ENSO forcing. From cluster analysis on streamflow anomalies during the life cycle of ENSO events, five spatially coherent regions were identified. The response seasons for each region were objectively identified from aggregate composite streamflow anomalies. While the lengths of the response seasons exhibited regional variations, 3 of the 5 regions (Regions 2, 3 and 4) exhibited wet (dry) conditions with a consistency better than 75% during the spring and/or summer period following the onset of La Nina (El Nino) events. The response seasons correspond to the period when forecasts based on ENSO indicators are likely to have useful skill. Region 5 (Northern BC) does not show clear response during the year following the onset of both El Nino and La Nina events.

Streamflow variations associated with the PNA pattern and the PDO regime were investigated as possible sources of interference in ENSO-related responses. It was shown that high PNA years conditioned on non-El Nino years can also produce El Nino-like streamflow responses over Western Canada. On the other hand, there was no clear response during low PNA years, probably because over Western Canada the west-east air flow during low PNA years was weaker than that associated with La Nina years. At decadal scale, it was found that the strength of ENSO related responses varied as a function of the PDO phase. The interaction appears to be constructive when ENSO and PDO are in phase and destructive when they are out of phase. These interferences have important implications for long-range forecasting based on ENSO indicators because the current generations of climate prediction models do not offer skillful long-range predictions for PNA and PDO. The potential of using the three large-scale climate anomalies for long-range forecasting was assessed by computing partial correlation coefficients at lead times of up to 10 months between the La Nina ITS for each basin and SOI, PNA and PDO indices. The results showed that basins with flows dominated by spring snowmelt show better correlations with SOI than with PDO while basins dominated by spring-summer snow/glacier melt exhibit an opposite trend. Although partial correlations with SOI are statistically significant at several basins, ENSO was found to explain less than 30% of the variability in streamflow at most of the basins considered.

A number of techniques may be explored for seasonal forecasting of streamflow for certain basins that were found to be moderately correlated to one or more of the large-scale climate anomalies. Piechota and Dracup (1999) developed a methodology for probabilistic seasonal forecasts conditioned on streamflow persistence and two ENSO indicators. Piechota and Dracup stratified the streamflow and predictor (SST, SOI, or persistence) data into three categories and used a univariate kernel density estimation to fit the probability density function (PDF) for each of the three categories for a predictor. A Bayesian framework was then used to estimate the posterior probability of each streamflow category conditioned on the predictor variable. The probabilistic forecasts from the three models were then combined into a consensus forecast by linear regression. For cases where streamflow variability is affected by two or more climate regimes such as ENSO and PDO, such a method could be extended to incorporate information from multiple predictors using a multivariate framework that takes into account the interdependence between the climate anomalies. In addition, the ENSO and PDO information can be used to sample appropriate historical data as surrogates for future climate in the ensemble streamflow prediction (ESP) framework using deterministic hydrologic models (e.g. Hamlet and Lettenmaier, 1999). Finally, it is noteworthy that the major portion of streamflow in Western Canada comes from spring-summer snow/glacier melt, which lags the atmospheric circulation by about 2-3 months. Hence, for certain regions of Western Canada, information on the PNA pattern could still improve forecasting by up to one season lead time.

Acknowledgements

This work is partly supported by the Natural Science and Engineering Research Council (NSERC) of Canada and partly by the Canadian Water Networks (CWN). The first author is also supported by a graduate teaching assistantship of the University of Alberta. We thank Dave Harvey of Environment Canada for supplying the RHBN dataset. The PDO data were supplied by Dr. Nathan Mantua of the University of Washington, Seattle.

References

Barnston, A. G., and Livezey, R. E. (1987). Classification, seasonality and persistence of low-frequency circulation patterns. *Mon. Weather Rev.*, 115, 1083-1126.

Bladé, I. (1999). The influence of midlatitude ocean–atmosphere coupling on the low-frequency variability of a GCM. Part II, Interannual variability induced by tropical SST forcing. *J. Climate,* 12, 21–45.

Bonsal, B. R., Shabbar, A., and Higuchi, K. (2001). Impacts of low frequency variability modes on Canadian winter temperature. *Int. J. Climatol.,* 21, 95-108.

Brown, D. P., and Comrie A. C. (2004). A winter precipitation 'dipole' in the western United States associated with multi-decadal ENSO variability. *Geophys. Res. Lett.,* 31, L09203, doi:10.1029/2003GL018726.

Cayan, D. R., and Peterson, D. H. (1989). The influence of North Pacific atmospheric circulation on streamflow in the west. In D. H. Peterson (Ed.), *Aspects of Climate Variability in the Pacific and Western America* (pp. 375-397). Geophysical Monograph Series, 55. AGU, Washington, DC.

Coulibaly, P., and Burn D. H. (2004). Wavelet analysis of variability in annual Canadian streamflows. *Water Resour. Res.,* 40, W03105, doi:10.1029/2003WR002667.

Dracup, J. A., and Kahya, E. (1994). The relationships between U.S. streamflow and La Nina events. *Water Resour. Res.,* 30, 2133-2141.

Efron, B., and Tibshirani R. J. (1993). *An Introduction to the Bootstrap.* San Francisco, CA: Chapman & Hall.

Gan, T. Y. (2000). Reducing vulnerability of water resources of Canadian Prairies to potential droughts and possible climatic warming. *Water Resources Management,* 14, 111-135.

Gan, T. Y., Gobena, A. K., and Wang, Q. (2007). Precipitation of southwestern Canada – Wavelet, scaling, multifractal analysis, and teleconnection to climate anomalies. *J. Geophys. Res.*, 112, D10110, doi:1029/2006JD007157.

Granger, R. J., and Gray, D. M. (1990). A net radiation model for calibrating daily snowmelt in open environments. *Nordic Hydrology,* 21, 217-234.

Hamlet, A., and Lettenmaier, D. (1999). Columbia River streamflow forecasting based on ENSO and PDO climate signals. *J. Water Resour. Plann. Manage.*, 125, 333-341.

Harvey, K. D., Pilon, P. J., and Yuzyk, T. R. (1999). Canada's Reference Hydrometric Basin Network (RHBN). In *Partnerships in Water Resources Management, Proceedings of the CWRA 51st Annual Conference,* Nova Scotia.

Horel, J. D., and Wallace, J. M. (1981). Planetary-scale atmospheric phenomena associated with the southern Oscillation. *Mon. Weather Rev.*, 109, 813-829.

Hsieh, W., and Tang, B., (2001). Interannual variability of accumulated snow in the Columbia basin, British Columbia. *Water Resour. Res.,* 37, 1753-1759.

HYDAT (2001). National Surface Water Data (Water survey of Canada). Environment Canada, CD-ROM.

Kahya, E., and Dracup, J. A. (1993). U.S. Streamflow Patterns in Relation to the El Nino/Southern Oscillation. *Water Resour. Res.,* 29, 2491-2503.

Kiladis, G. N., and Diaz, H. F. (1989). Global climatic anomalies associated with extremes of the southern oscillation. *J. Climate,* 2, 1069-1090.

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., and Francis, R.C. (1997). A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin Amer. Meteorol. Soc.,* 78, 1069–1079.

Mantua, N., and Hare, S. (2001). The Pacific Decadal Oscillation. *J. Oceanography* 58, 35-44.

McCabe, G. J., and Dettinger, M. D. (1999). Decadal variations in the strength of ENSO teleconnections with precipitation in the western United States. *Int. J. Climatol.,* 19, 1399-1410.

Molteni, F., Ferranti, L., Palmer, T. N., and Viterbo, P. (1993). A dynamical interpretation of the global response to equatorial Pacific SST anomalies. *J. Climate,* 6, 777–795.

Moore, R. D. (1996). Snowpack and runoff responses to climatic variability, southern Coast Mountains, British Columbia. *Northwest Science,* 70, 321-333.

Moore, R. D., and McKendry, I. G. (1996). Spring snowpack anomaly patterns and winter climatic variability, British Columbia, Canada. *Water Resources Research,* 32, 623-632.

Murtagh, F. (1985). *Multivariate clustering algorithms*. Physica-Verlag.

Ntale, H. K., and Gan, T. Y. (2004). East African rainfall anomaly patterns in association with El Nino/Southern Oscillation. *J. Hydrol. Engrg., ASCE,* 9(4), 1-12.

Palmer, T. N. (1993). Extended-range atmospheric prediction and the Lorenz model. *Bulletin Amer. Meteorol. Soc.,* 74, 49-65.

Piechota, T. C., and Dracup, J. A. (1999). Long-range streamflow forecasting using El Nino-Southern Oscillation indicators. *J. Hydrol. Engrg, ASCE,* 4(2), 144-151.

Piechota, T. C., Dracup, J. A., and Chiew, F. H. S. (2001). Development of exceedance probability streamflow forecasts. *J. Hydrol. Engrg, ASCE,* 6(1), 20-28.

Piechota, T. C., Dracup, J. A., and Fovell, R. G. (1997). Western US streamflow and atmospheric circulation patterns during El Nino-Southern Oscillation. *J. Hydrol.,* 201, 249-271.

Pizarro, G., and Lall, U. (2002). El Nino and floods in the U.S.: What can we expect? *EOS, Transactions of the AGU,* 83(32), 349-352.

Rajagopalan, B., Mann, M. E., and Lall, U. (1998). A multivariate frequency-domain approach to long-lead climatic forecasting. *Weather and Forecasting,* 13, 58-74.

Ropelewski, C. F., and Halpert, M. S. (1987). North American temperature and precipitation patterns associated with the El Nino/Southern Oscillation (ENSO). *Mon. Weather Rev.,* 114, 2352-2362.

Shabbar, A., and Khandekar, M. (1996). The impact of El Nino-Southern Oscillation on the temperature field over Canada. *Atmosphere-Ocean,* 34, 401-416.

Shabbar, A., Bonsal, B., and Khandekar, M. (1997). Canadian precipitation patterns associated with the Southern Oscillation. *J. Climate,* 10, 3016-3027.

Straus, D. M., and Shukla, J. (2000). Distinguishing between the SST-forced variability and internal variability in mid latitudes: Analysis of observations and GCM simulations. *Quart. J.Royal Meteorol. Soc.,*126, 2323–2350.

Straus, D. M., and Shukla, J. (2002). Does ENSO force the PNA? *J. Climate,* 15, 2340-2358.

Tang, B., Hsieh, W., Monahan, A. H., and Tangang, F. T. (2000). Skill comparisons between neural networks and canonical correlation analysis in predicting the equatorial Pacific sea surface temperature. *J. Climate,* 13, 287-293.

Trenberth, K. E., and Hurrell, J. (1994). Decedal atmosphere-ocean variations in the pacific. *Climate Dynamics,* 9(6), 303-311.

Torrence, C., and Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin Amer. Meteorol. Soc.,* 79(1), 61-78.

Vaccaro, J. (2002). Interdecadal Changes in the Hydrometeorological Regime of the Pacific Northwest and in the Regional-to-Hemispheric Climate Regimes, and Their Linkages. U.S. Geological Survey, Water-Resources Investigations Report 02-4176, Tacoma, WA.

Wallace, J. M., and Gutzler, D. S. (1981). Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Weather Rev.,* 109, 784-812.

Wilks, D. S. (1995). *Statistical Methods in the Atmospheric Sciences: An Introduction*. San Diego, CA: Academic Press.

Woo, M.-K., and Thorne, R. (2003). Comment on 'Detection of hydrologic trends and variability'. *J. Hydrol.,* 277, 150-160.

Yarnal, B., and Diaz, H. F. (1986). Relationships between extremes of the Southern Oscillation and the winter climate of Anglo-American Pacific coast. *J. Climatology,* 6, 197-219.

Yue, S., and Gan, T. Y. (2004). Simple scaling properties of Canadian maximum annual streamflow. *Advances in Water Resources,* 27, 481-495.

Zhang, X. B., Harvey, K. D., Hogg, W. D., and Yuzyk, T. R. (2001). Trends in Canadian streamflow. *Water Resour. Res.,* 37, 987-998.

Table 3.1. List of basins used in the study, their drainage area, data length in years and ENSO region identified from cluster analysis.

Basin Name	WSC Station ID	Long, deg	Lat, deg	Area, km^2	ENSO Region	Data Length [c]	Basin Name	WSC Station ID	Long, deg	Lat, deg	Area, km^2	ENSO Region	Data Length [c]
Salmon	08HD006	125.45	50.10	1200	4	46	Pembina	07BB002	114.55	53.55	4420	3	47
Somass	08HB017	124.50	49.15	1280	4	44	Mcleod	07AF002	116.40	53.35	2560	3	47
Chemainus	08HA001	123.70	48.88	355	4	49	Smoky	07GJ001	118.10	54.55	50300	3	47
Koksilah	08HA003	123.67	48.73	209	4	48	Wapiti	07GE001	119.30	54.50	11300	3	41
Sproat	08HB008	124.91	49.29	347	4	61[a]	Mcgregor	08KB003	121.30	54.20	4770	3	42
Capilano	08GA010	123.14	49.40	172	4	61[a]	Salmon	08KC001	123.25	54.25	4300	3	49
Harrison	08MG013	121.45	49.20	7870	2	51	Stuart	08JE001	124.28	54.42	14600	3	61[a]
Nicola	08LG006	120.55	50.10	7280	2	44	Nautley	08JB003	125.35	54.10	6030	3	52
N. Thompson	08LB047	119.55	51.35	4450	2	42	Clearwater	07CD001	111.25	56.69	30800	1	45
Slocan	08NJ013	117.30	49.50	3320	2	61[a]	Waterhen	05LH005	99.55	51.85	55000	1	44
Salmo	08NE074	117.10	49.15	1230	2	53	Brokenhead	05SA002	96.43	50.09	1610	1	60
Elk	08NK016	114.55	49.45	1870	2	51	Beaver	06AD006	110.22	54.35	14500	1	46
Quesnel	08KH006	122.22	52.84	11500	2	61[a]	Moose Jaw	05JE006	104.55	50.10	9230	1	58
Clearwater	08LA001	120.07	51.66	10200	2	52	Wascana	05JF005	104.45	50.25	3850	1	57
Chilliwack	08MH016	121.46	49.08	329	2	55	Red Deer	05LC001	102.25	52.50	11000	1	48
Barnes Creek	08NE077	118.13	49.91	201	2	51	Sturgeon	05EA001	113.50	53.45	3350	1	61[a]
Kootenay	08NF001	116.04	50.89	420	2	57	Atlin [b]	09AA006	133.81	59.60	6810	5	52
Similkameen	08NL007	120.50	49.46	1850	2	61[a]	Swift	09AE003	131.30	59.50	3320	5	46
Crowsnest	05AA008	114.41	49.60	404	2	53	Stikine	08CE001	131.15	57.90	29300	5	48
Oldman	05AA023	114.18	49.81	1440	2	53	Iskut	08CG001	130.20	57.05	9350	5	43
Waterton	05AD003	113.84	49.11	614	2	54	Nass	08DB001	129.05	56.20	18500	5	61[a]
Belly	05AD005	113.70	49.10	319	2	61[a]	Skeena	08EF001	127.30	55.00	42200	5	61[a]
Columbia	08NB005	117.55	51.55	9710	2	57	Morice	08ED002	127.25	54.05	1910	5	40
Squamish	08GA002	123.25	50.10	2330	2	47	Bulkley	08EE004	127.05	54.50	7360	5	61[a]
Chilko	08MA002	124.14	51.63	2110	2	61[a]	Liard	10AA001	126.00	59.45	33400	5	42
Lillooet	08MG005	122.80	50.34	2160	2	61	Toad	10BE004	125.00	58.55	2570	5	41
Fraser	08KA007	122.35	50.50	1700	2	47	Beatton [b]	07FC001	120.55	56.55	15600	5	41
Bow	05BB001	115.57	51.18	2210	2	61[a]	Sikanni Chief [b]	10CB001	122.69	57.23	2160	5	58
Mistaya	05DA007	116.69	51.88	249	2	52	Muskwa [b]	10CD001	122.66	58.79	20300	5	58
Pine	07FB001	121.35	55.35	12100	3	41	Seal [b]	06GD001	96.28	58.89	48100	5	47

[a] The 13 basins used to obtain the interdecadal streamflow PC time series in Figure 3.13.

[b] The 5 basins for which the El Nino-based clusters were different from the La Nina-based clusters.

[c] All basins with streamflow records beginning before or by 1957 (data length $\geq$ 45 years) were used to obtain Figure 3.14.

Table 3.2. Number of basins with significantly above-normal SSA during March to October of La Nina events based on Mann-Whitney test at the 10% level. Values for the onset year (left), the following year (middle) and the second year following the onset (right) are separated by comas. The values in parenthesis are from the standard t test at the same confidence level.

	March	April	May	June	July	August	September	October
Minimum flow	2, 1, 2	4, 2, 0	2, 14, 0	3, 15, 6	2, 27, 10	1, 30, 4	1, 11, 3	2, 4, 3
	(0, 1, 2)	(0, 2, 1)	(1, 16, 2)	(5, 13, 7)	(3, 30, 10)	(2, 33, 3)	(0, 8, 3)	(0, 3, 2)
Maximum flow	0, 0, 1	2, 15, 1	3, 8, 3	7, 18, 5	0, 17, 5	1, 33, 4	2, 12, 2	2, 5, 3
	(0, 0, 1)	(2, 15, 1)	(2, 9, 3)	(5, 19, 7)	(0, 17, 7)	(1, 30, 5)	(1, 12, 2)	(1, 7, 4)
Mean flow	0, 1, 1	1, 11, 2	1, 14, 3	7, 23, 6	0, 24, 10	0, 29, 3	1, 11, 2	1, 7, 3
	(0, 1, 2)	(2, 11, 2)	(2, 16, 4)	(4, 20, 7)	(2, 24, 9)	(0, 33, 3)	(0, 10, 3)	(1, 6, 3)

Table 3.3. Number of basins with significantly below-normal SSA during March to October of El Nino events based on Mann-Whitney test at the 10% level. Values for the onset year (left), the following year (middle) and the second year following the onset (right) are separated by comas.

	March	April	May	June	July	August	September	October
Minimum flow	2, 0, 0	7, 4, 3	0, 2, 0	2, 6, 0	0, 3, 1	1, 5, 0	0, 10, 1	1, 8, 1
Maximum flow	5, 0, 3	0, 0, 0	0, 1, 3	0, 4, 0	0, 5, 1	0, 6, 0	0, 9, 3	0, 9, 1
Mean flow	3, 0, 2	2, 2, 0	0, 2, 0	1, 8, 0	0, 4, 1	0, 4, 0	0, 12, 2	1, 12, 1

Table 3.4. Number of basins with significant above-normal SSA during March to October of El Nino events based on Mann-Whitney test at the 10% level. Other attributes are the same as in Table 3.3.

Flow	March	April	May	June	July	August	September	October
Minimum flow	3, 2, 0	2, 0, 1	10, 2, 6	11, 0, 5	18, 0, 9	16, 1, 10	9, 1, 2	8, 1, 2
Maximum flow	3, 2, 1	7, 2, 2	15, 0, 8	19, 0, 8	16, 0, 7	14, 0, 9	8, 1, 6	10, 0, 2
Mean flow	4, 3, 0	6, 6, 2	18, 0, 6	23, 0, 9	18, 0, 10	16, 0, 10	8, 1, 3	12, 1, 3

Table 3.5. The 10th and 90th percentiles of the ratios of the April-September composite mean streamflow to long-term mean streamflow for each climate category.

Climate Category	Ratio (10%, 90%)	Climate Category	Ratio (10%, 90%)
La Nina	1.036, 1.328	El Nino	0.763, 1.004
Cool PDO	1.003, 1.141	Warm PDO	0.844, 0.996
La Nina/Cool PDO	1.027, 1.398	El Nino/Warm PDO	0.547, 0.983
La Nina/Warm PDO	0.849, 1.195	El Nino/Cool PDO	0.905, 1.137

Table 3.6. Summary of streamflow conditions associated with ENSO events. Dry (wet) conditions are relative to the regional ITS based on the ENSO response periods.

	El Nino and wet anomaly	El Nino and dry anomaly
Region 1	-	-
Region 2	1966	1970[a], 1973, 1977[a], 1983, 1987[a], 1992, 1998
Region 3	1966, 1977	1970, 1973, 1983, 1987[a], 1992[a], 1998[a]
Region 4	1966	1970, 1973[a], 1977, 1983, 1987, 1992[a], 1998[a]

	La Nina and wet anomaly	La Nina and dry anomaly
Region 1	1965, 1971, 1972, 1974[a], 1976, 1996[a]	1989, 1999
Region 2	1965, 1971, 1972[a], 1974[a], 1976[a], 1996, 1999[a]	1989
Region 3	1965[a], 1971, 1972[a], 1974, 1976[a], 1996, 1999	1989
Region 4	1971, 1972[a], 1974[a], 1976, 1989, 1996, 1999	1965

[a] ITS values in the lower (upper) 10 percentile of the distribution.

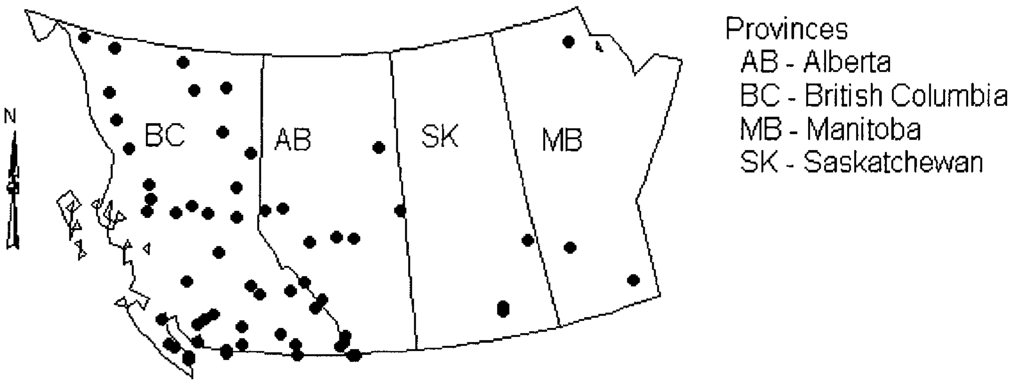


Figure 3.1. Spatial distribution of basins used in the study.

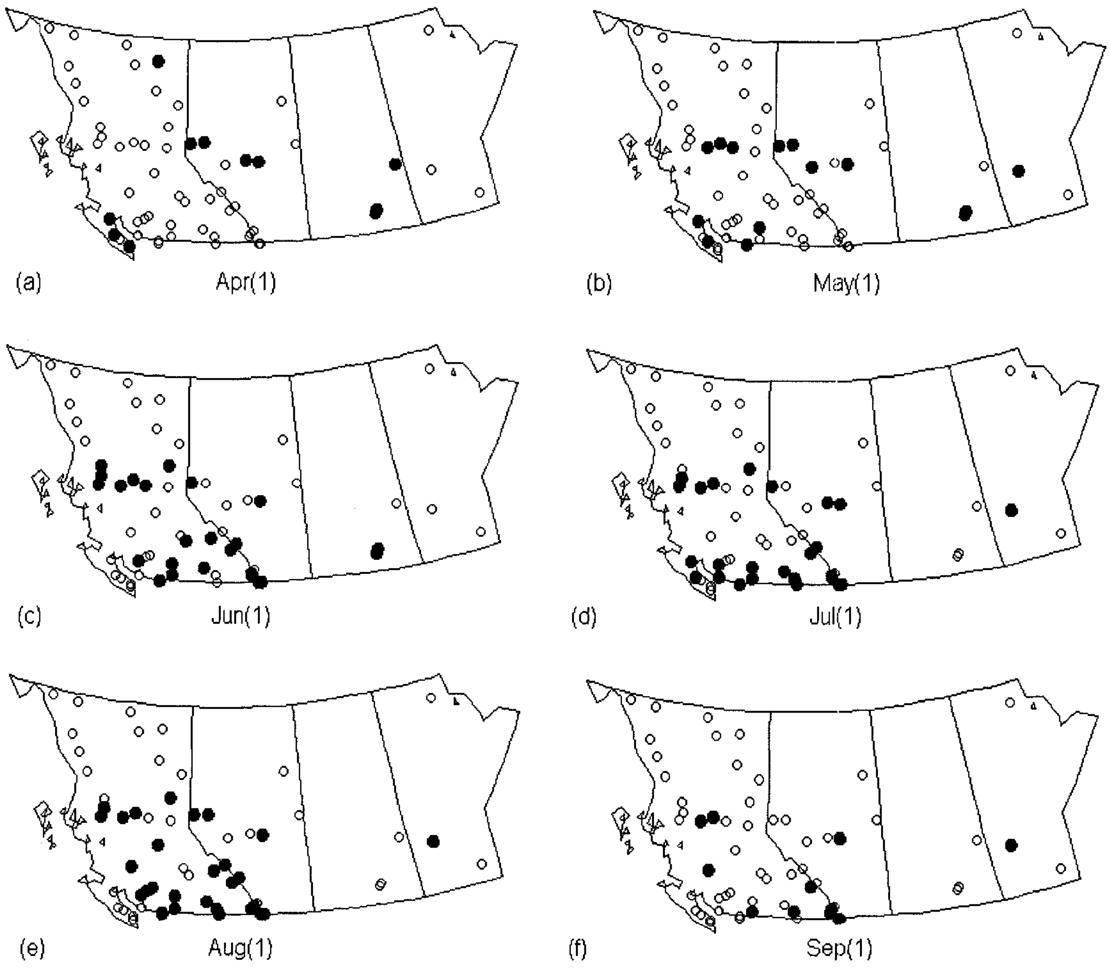


Figure 3.2. Spatial distribution of basins with statistically significant above-normal SSA (solid circles) following the onset of La Nina events based on a 2-sided Mann-Whitney test at the 10% significance level: (a) Apr(1), (b) May(1), (c) Jun(1), (d) Jul(1), (e) Aug(1) and (f) Sept(1).

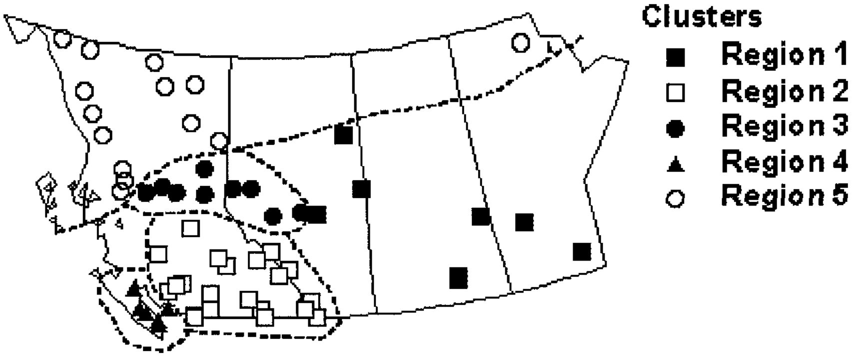


Figure 3.3. ENSO-related spatial patterns identified from cluster analysis of the La Nina time series.

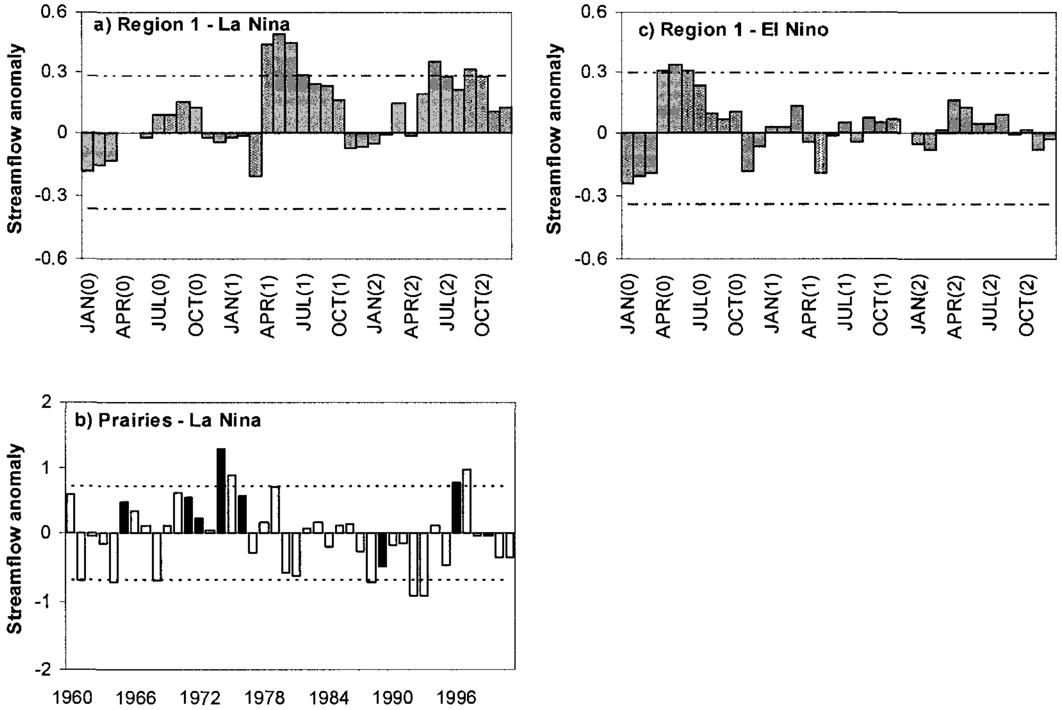


Figure 3.4. The aggregate composites and ITS for Region 1. (a) La Nina composite. (b) La Nina ITS. (c) El Nino composite. The horizontal lines in (a) and (c) are the 90% confidence intervals based on the empirical distribution of 1000 random samples obtained by bootstrap sampling. The solid blocks in (b) correspond to the La Nina years.

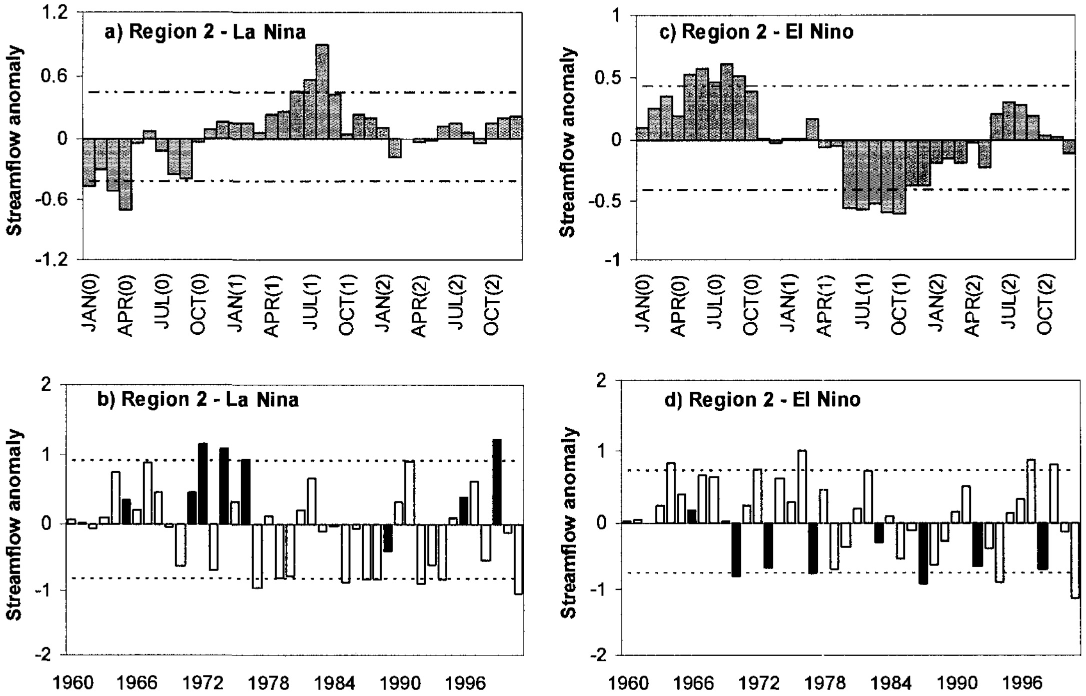


Figure 3.5. The aggregate composites and ITS for Region 2. (a) La Nina composite. (b) La Nina ITS. (c) El Nino composite. (d) El Nino ITS. The horizontal lines in (a) and (c) are the 90% confidence intervals based on the empirical distribution of 1000 random samples obtained by bootstrap sampling. The solid blocks in (b) and (d) correspond to the La Nina and El Nino years, respectively.

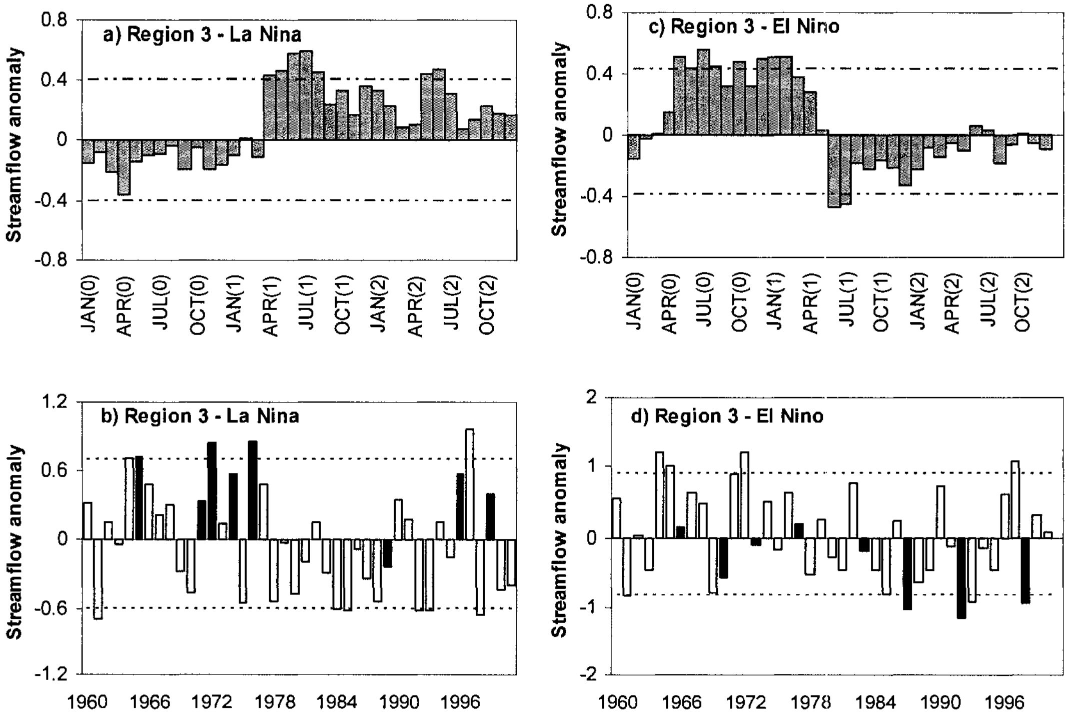


Figure 3.6. The aggregate composites and ITS for Region 3. All features are the same as in Figure 3.5.

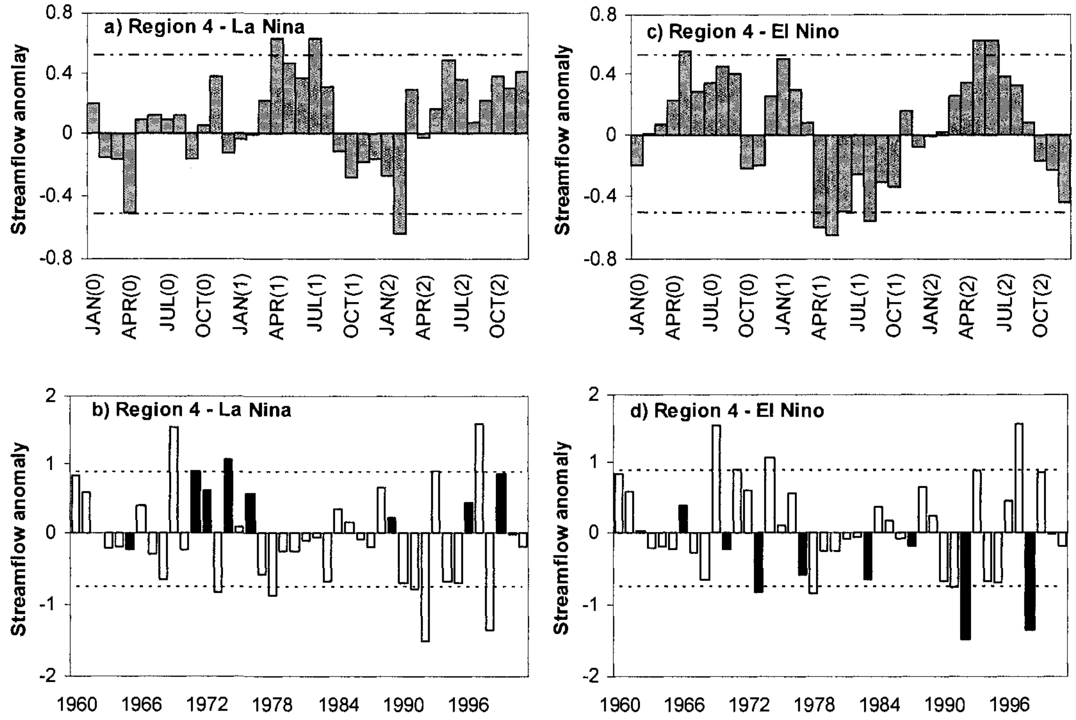


Figure 3.7. The aggregate composites and ITS for Region 4. All features are the same as in Figure 3.5.

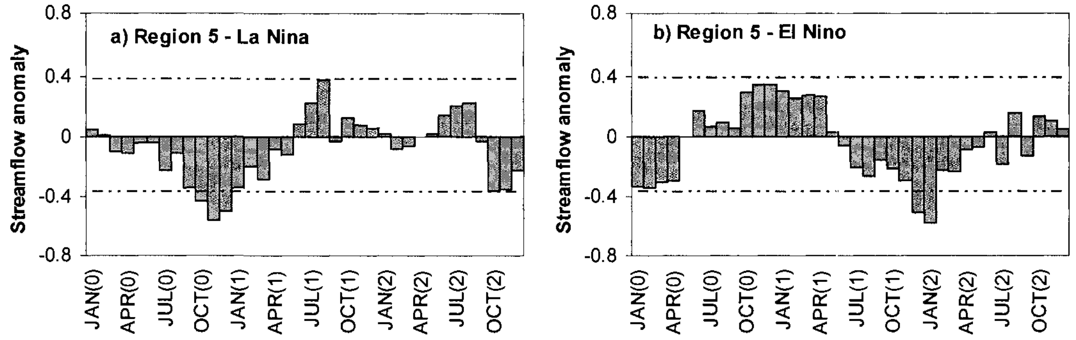


Figure 3.8. The aggregate composites for Region 5. (a) La Nina composite. (b) El Nino composite. The horizontal lines in (a) and (b) are the 90% confidence intervals based on the empirical distribution of 1000 random samples obtained by bootstrap sampling.

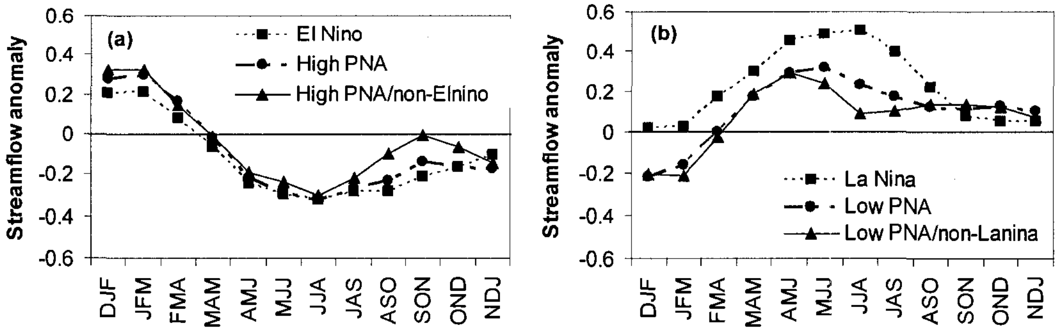


Figure 3.9. (a) Regionally averaged SSA response associated with El Nino, High PNA and High PNA/non-El Nino years. (b) Same as (a) but for La Nina, Low PNA and Low PNA/non-La Nina years. The composite anomalies are smoothed using a 3-month moving average filter.

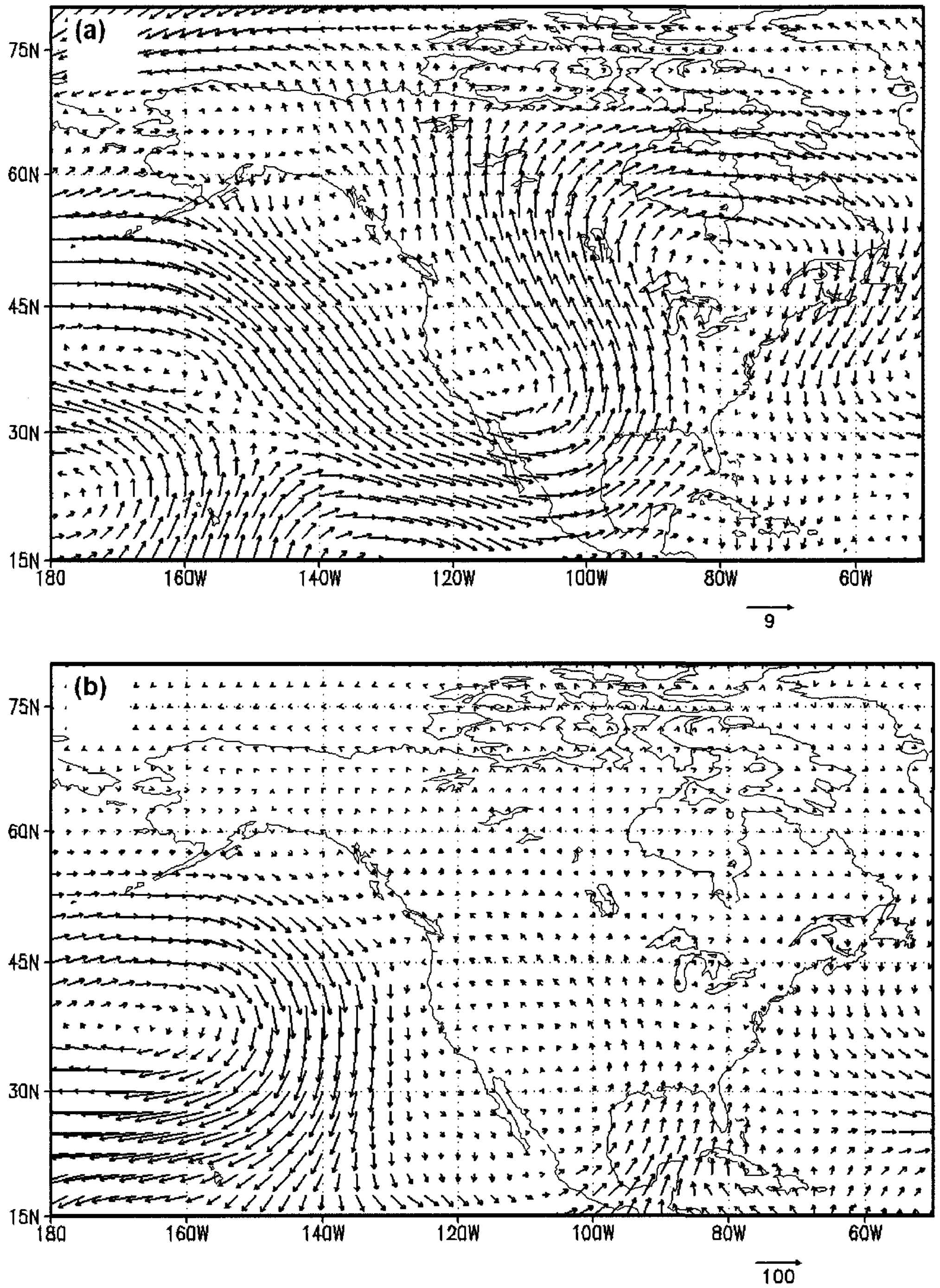


Figure 3.10. (a) The 300-mbar wind anomaly difference pattern obtained as the composite anomalies for El Nino winters minus High-PNA/non-El Nino winters. (b) The VIMF difference pattern obtained as the composites for El Nino winters minus High-PNA/non-El Nino winters.

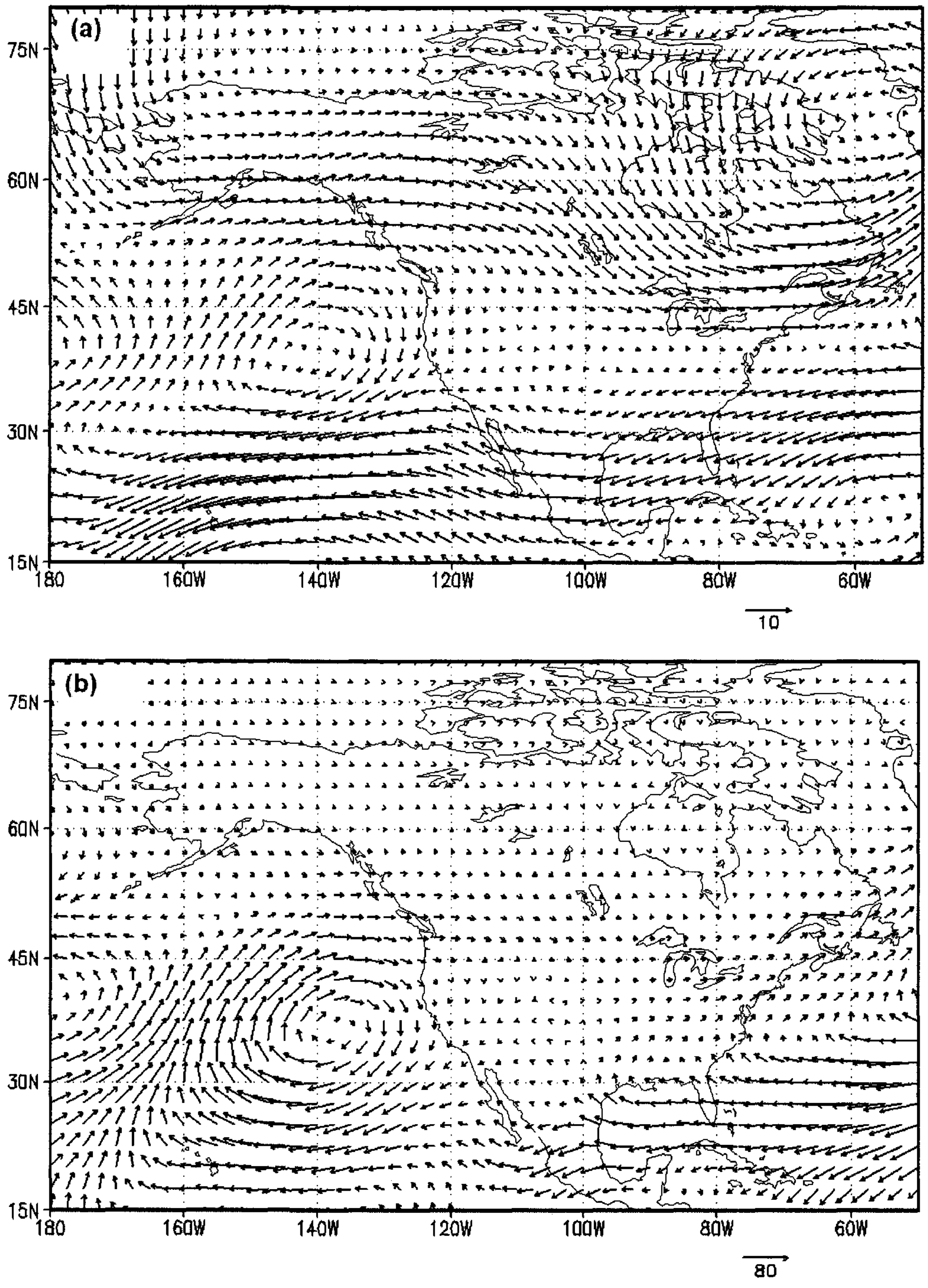


Figure 3.11. (a) The 300-mbar wind anomaly difference pattern obtained as the composite anomalies for La Nina winters minus Low-PNA/non-La Nina winters. (b) The VIMF difference pattern obtained as the composites for La Nina winters minus Low PNA/non-La Nina winters.

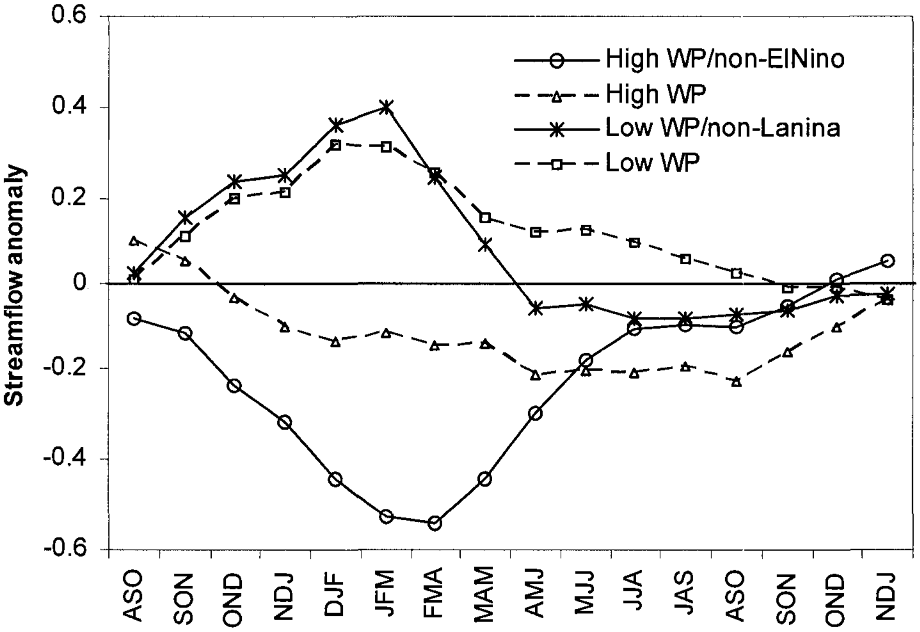


Figure 3.12. Regionally averaged SSA response associated with high WP, High WP/non-El Nino, low WP and low WP/non-La Nina years. The composite anomalies are smoothed using a 3-month moving average filter.

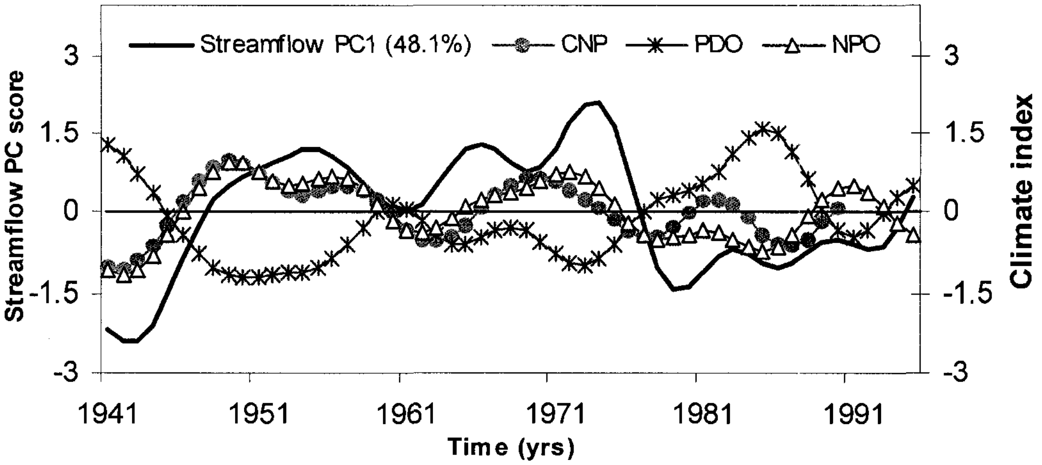


Figure 3.13. The leading PC time series of the lowpass filtered (≥ 8 years) streamflow signals in Western Canada. The lowpass filtered time series of the PDO, CNP and NPO indices are also plotted alongside the streamflow PC.

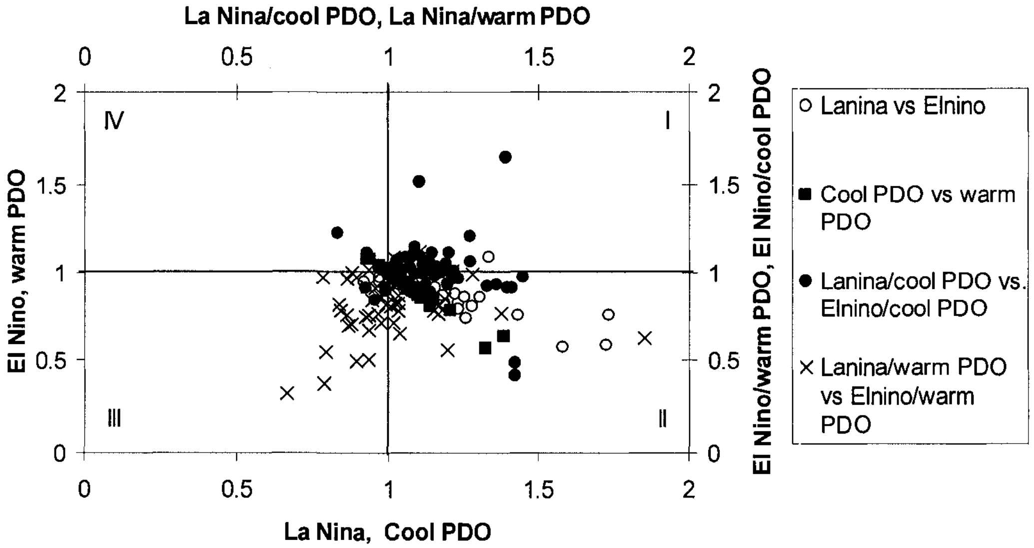


Figure 3.14. Scatterplots of the ratio of the April-September composite mean to the long-term mean streamflow for 49 basins. The ratios are plotted on a plane defined by ENSO/PDO phases, where the primary and secondary axes labels indicate the respective climate category.

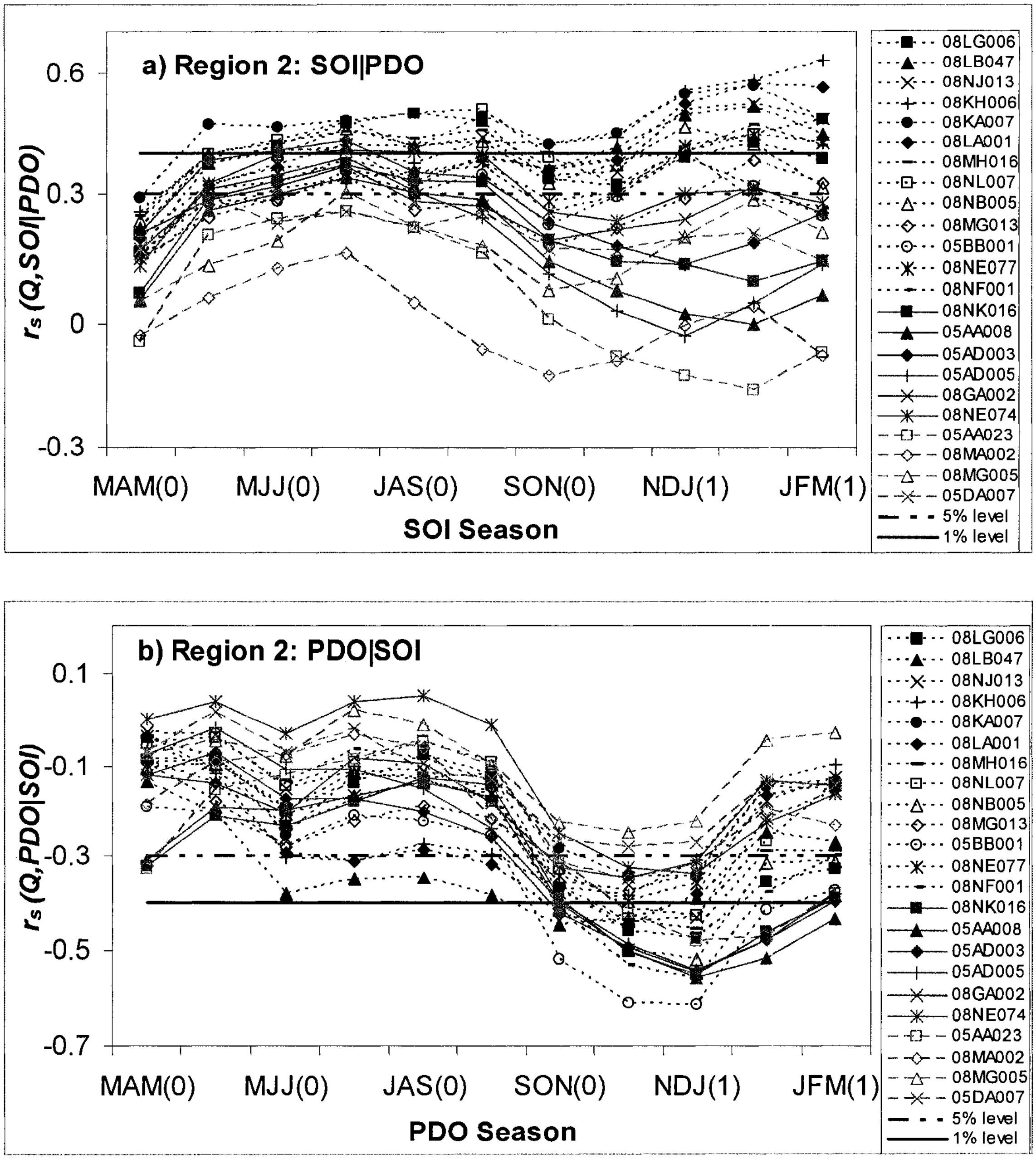
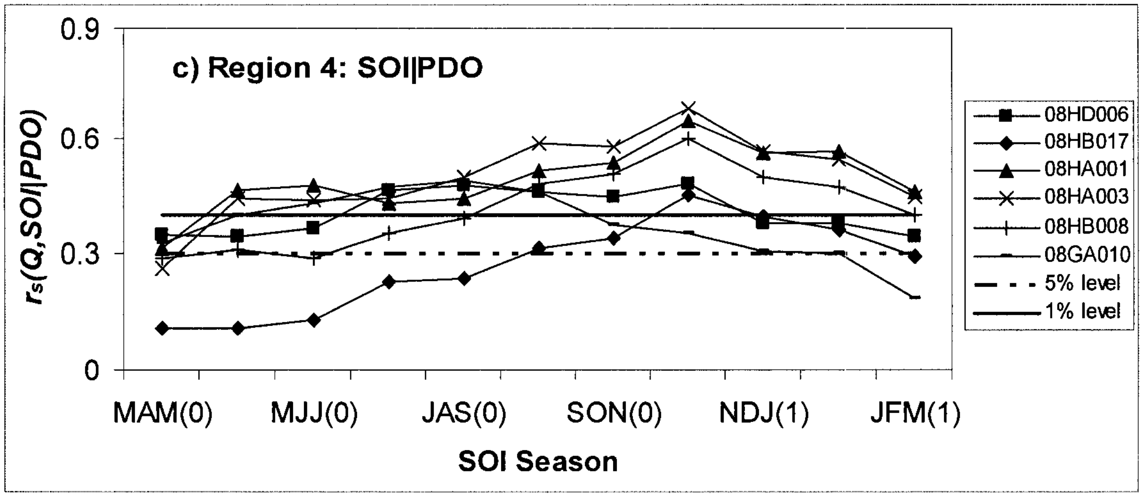


Figure 3.15. The temporal evolution of the partial correlations between streamflow and (a) SOI|PDO for Region 2, (b) PDO| SOI for Region 2, (c) SOI|PDO for Region 4, and (d) PDO|SOI for Region 4. The solid and dashed horizontal lines indicate the 1% and 5% significance levels, respectively.



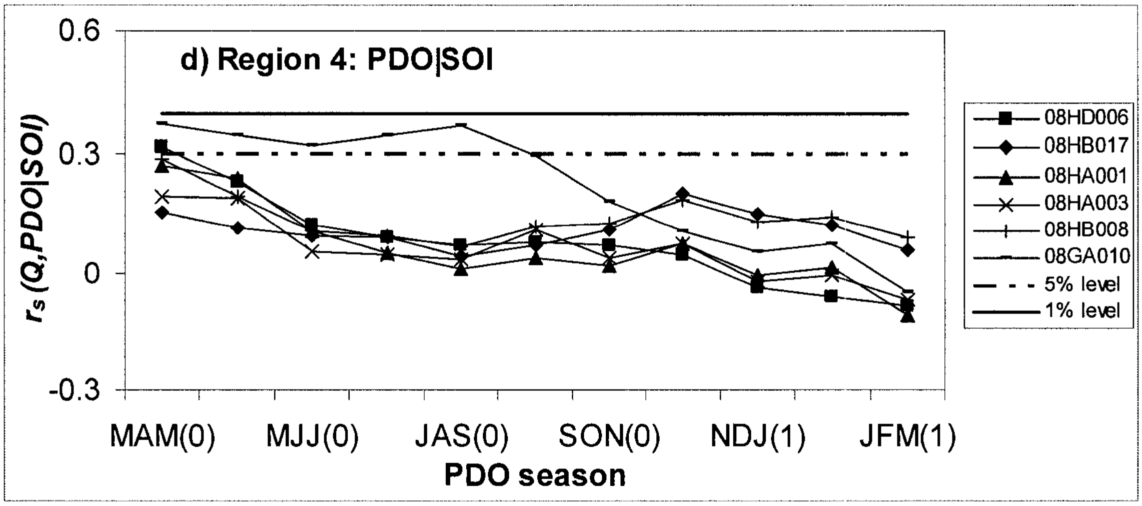


Figure 3.15. *Continued*

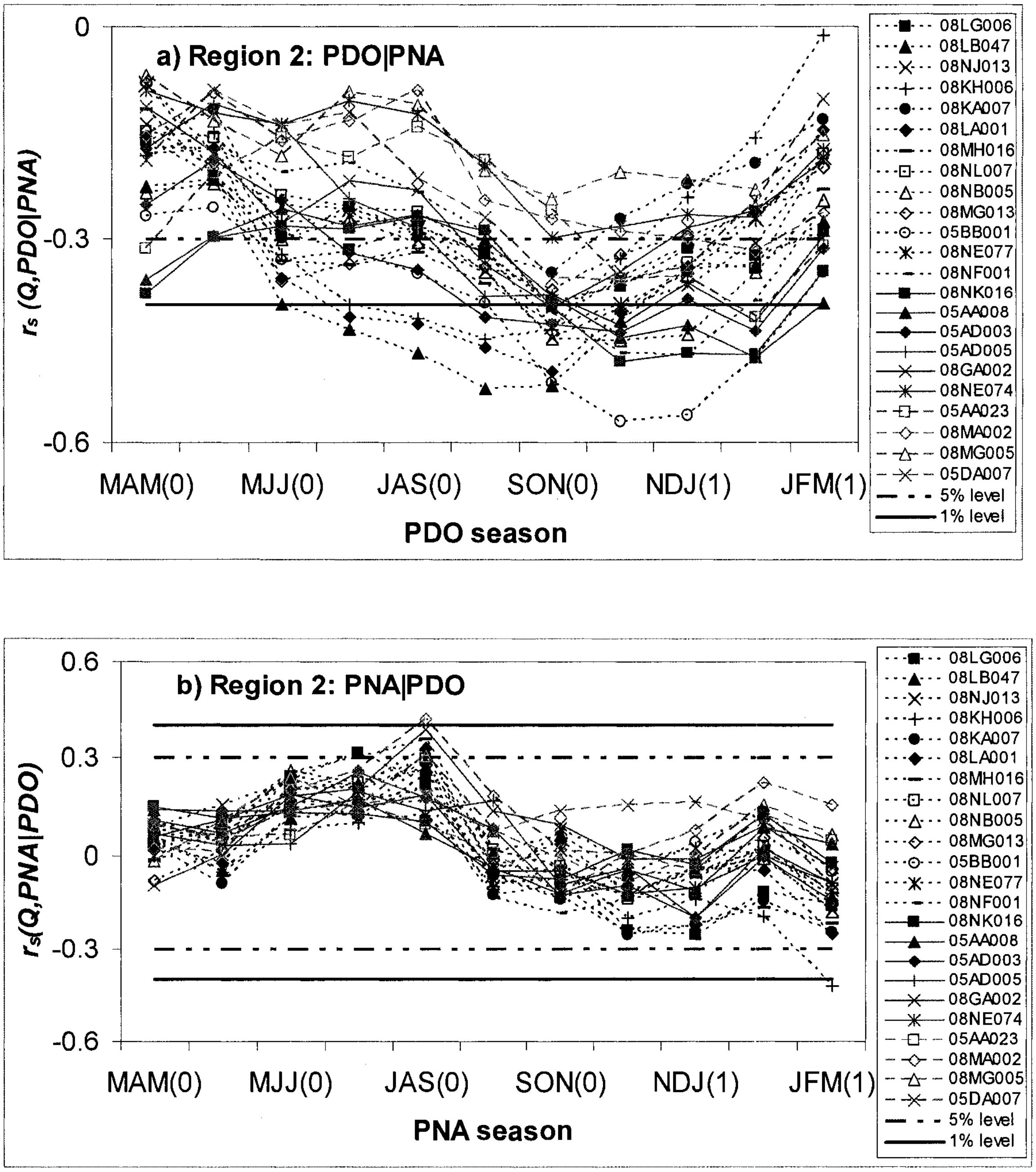
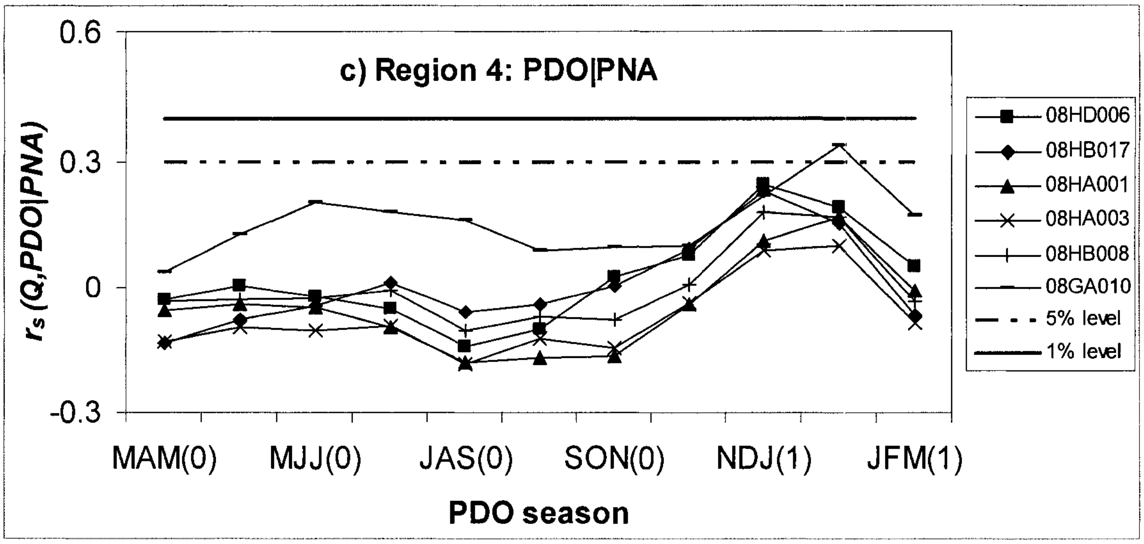


Figure 3.16. Same as in Figure 3.15 for (a) PDO|PNA for Region 2, (b) PNA|PDO for Region 2, (c) PDO|PNA for Region 4, and (d) PNA|PDO for Region 4.



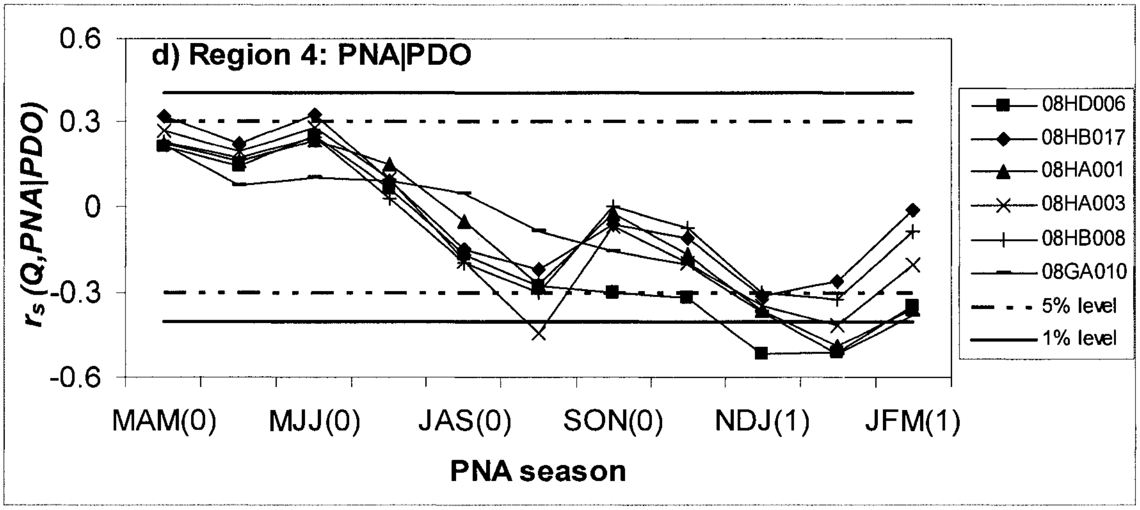


Figure 3.16. *Continued*

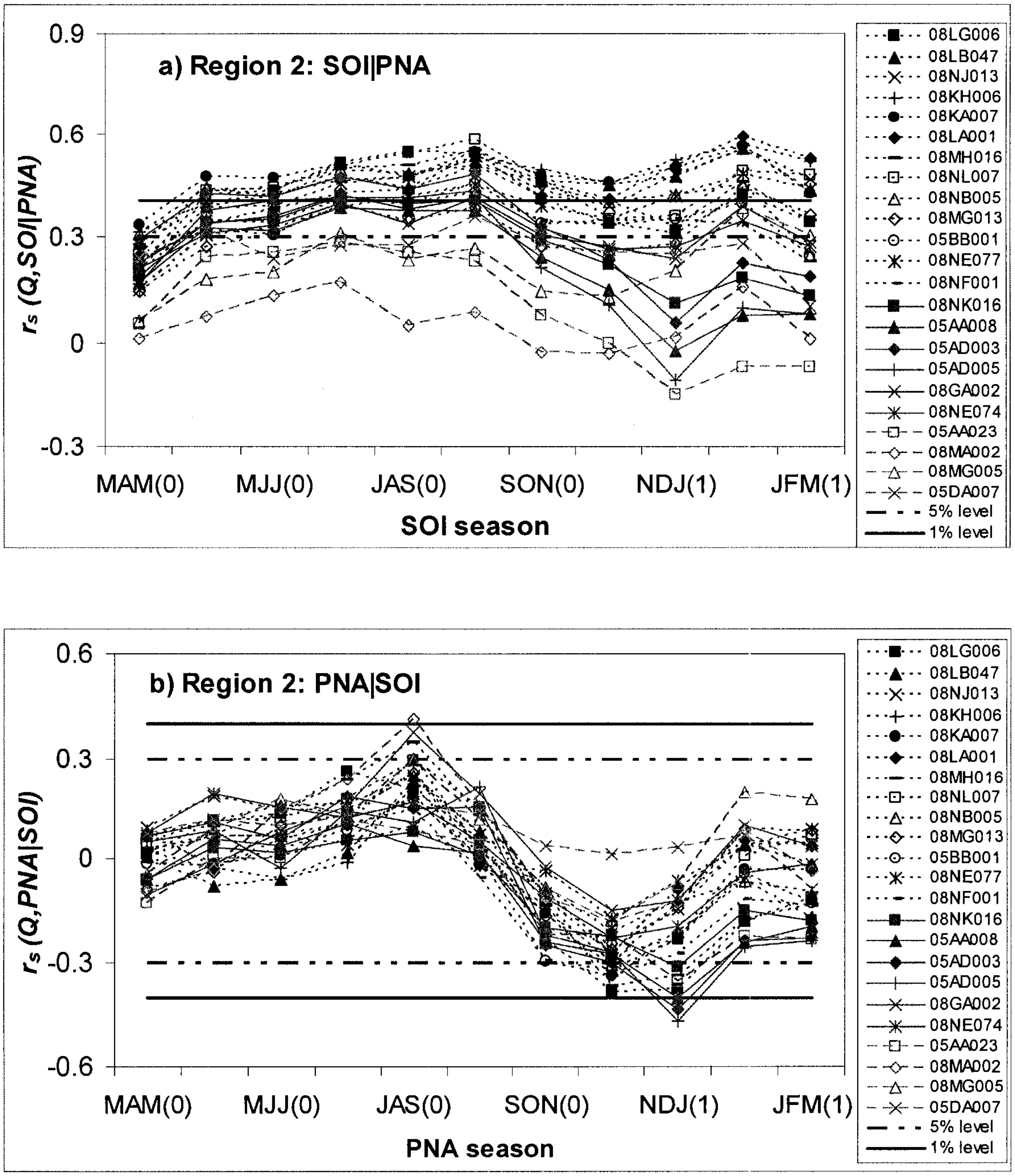
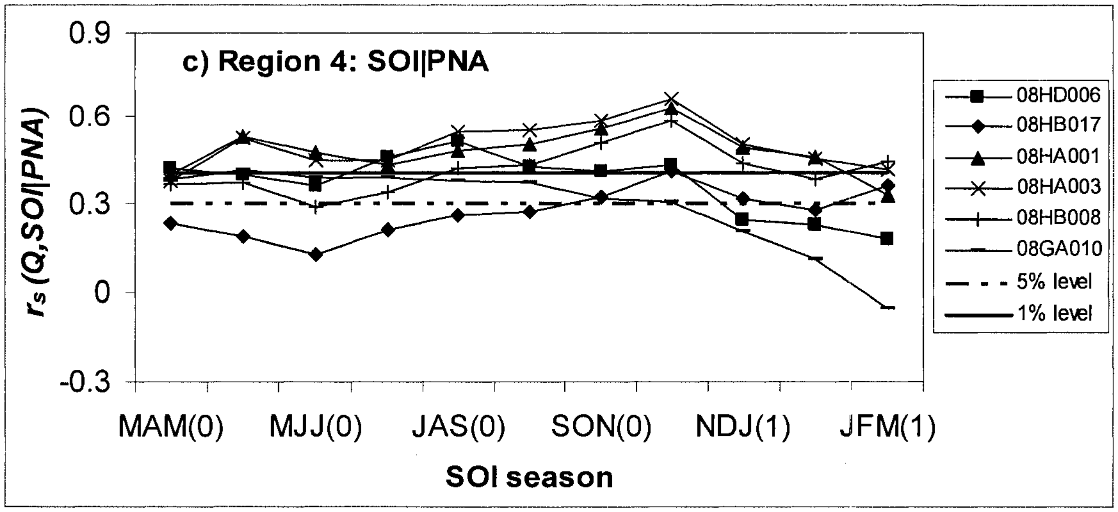


Figure 3.17. Same as in Figure 3.15 for (a) SOI|PNA for Region 2, (b) PNA|SOI for Region 2, (c) SOI|PNA for Region 4, and (d) PNA|SOI for Region 4.



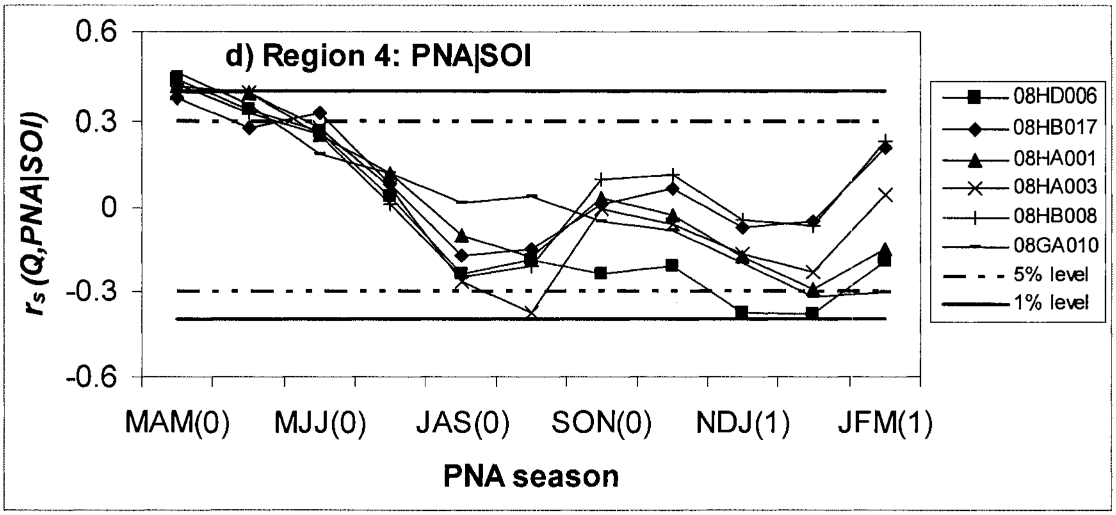


Figure 3.17. *Continued*

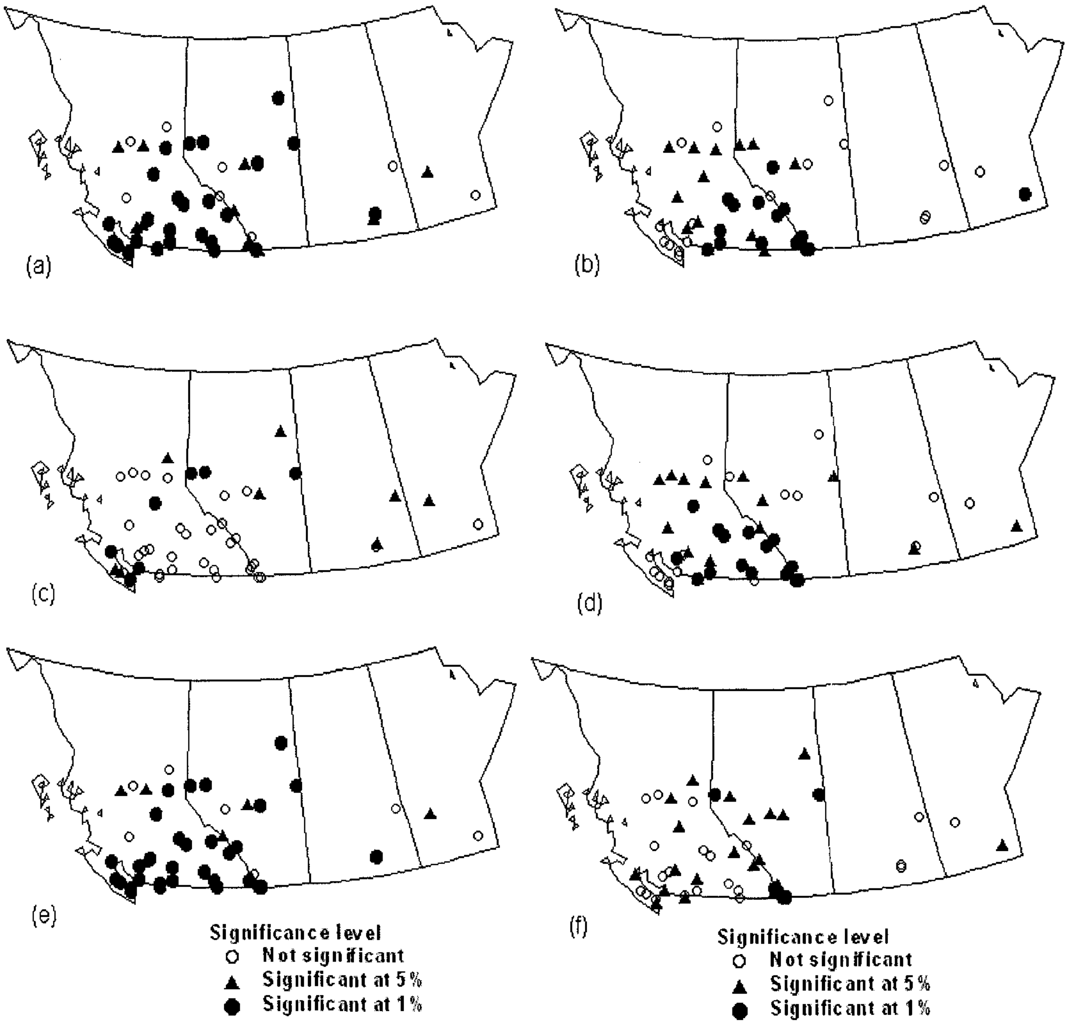


Figure 3.18. Statistical significance of the highest partial correlation out of 11 values for each basin computed using seasonal streamflow and lagged climate indices. Correlations are for (a) SOI|PDO, (b) PDO|SOI, (c) PNA|PDO, (d) PDO|PNA, (e) SOI|PNA, and (f) PNA|SOI. The streamflow seasons are defined by the La Nina-related response season identified from the composite analysis.

Chapter 4 Analysis of Hydroclimatic Variability and Predictability in the South Saskatchewan River Basin (Alberta)

4.1. Introduction

The South Saskatchewan River Basin (SSRB) provides over 57% of Alberta's allocated water even though it contains only about 6% of the water resources of the province (Alberta Environment, 2006a). A considerable portion of the annual flow in the basin comes from spring and summer snowmelt originating from the eastern slopes of the Canadian Rockies. As the Pacific Ocean is a primary moisture source for precipitation during the cold season (when most of the snow accumulation takes place), diagnostic analysis of low-frequency hydroclimatic variability and its relationship with the primary modes of Pacific climate variability such as the El Nino-Southern oscillation (ENSO), the Pacific/North American (PNA) pattern and the Pacific decadal oscillation (PDO) could provide vital information for Alberta's water resources managers, particularly with regard to long-range streamflow forecasting.

ENSO teleconnections to Western Canada's hydroclimatic variability are well documented (e.g. Shabbar *et al.*, 1997; Woo and Thorne, 2003; Coulibaly and Burn, 2004; Gobena and Gan, 2006; Gan *et al.*, 2007). Several studies have also investigated the relationships between hydroclimatic data and the PNA pattern (e.g. Moore, 1996; Woo and Thorne, 2003; Coulibaly and Burn, 2004; Gobena and Gan, 2006; Gan *et al.*, 2007). The ENSO and PNA climate anomalies exhibit most of their variance at interannual time scales. Apart from the work of Bonsal *et al.* (2001) who investigated the impacts of the PDO on Canadian surface temperature, little attention has been paid to decadal to interdecadal hydroclimatic variability in Western Canada and its association with the Pacific climate decadal variability. As described in Chapter 3, the strength of ENSO-streamflow relationships in Western Canada appears to be influenced by interdecadal variations in the PDO mode (Gobena and Gan, 2006). The present study focuses on the interannual to interdecadal precipitation and streamflow variability in the South Saskatchewan River basin (SSRB) of southern Alberta.

This study has two main objectives: (1) to identify the dominant inter-annual to inter-decadal modes of variability in the hydroclimatic data of the SSRB using wavelet analysis, and (2) to investigate the links between the low-frequency components of the basin's hydroclimate and Pacific climate variability using wavelet coherence and rank correlation. By analyzing streamflow and precipitation independently, our study provides a comprehensive picture of the low frequency features of the hydroclimate of the basin. We will also investigate the interactions between interannual and interdecadal modes of variability. The reader is referred to Section 2.3 for description of the wavelet transform methodology used in this chapter. A description of the study area and data is presented in section 4.2. Hydroclimatic variability in the SSRB and its relation to large-scale climate anomalies are discussed in sections 4.3 and 4.4, followed by summary and conclusions in section 4.5.

4.2. Study Area and Data Description

The study area covers that portion of the SSRB lying within the province of Alberta (Figure 4.1). The Alberta portion of the SSRB includes the Red Deer, Bow, Oldman and South Saskatchewan River sub-basins with a combined drainage area of 121,095 km^2, which is about a quarter of the surface area of the province. As per the 1996 census, the SSRB is home to 46.7% of the population of Alberta, of which 68% reside in the Bow River sub-basin, which constitutes only 21% of basin area (Alberta Environment, 2002). All four sub-basins originate in the Rocky Mountains and generally flow eastwards through the foothills and semi-arid prairies of southern Alberta and into Saskatchewan. The basin elevation ranges from about 600 m at the Alberta-Saskatchewan border to well over 3400 m in the Rocky Mountains.

The precipitation and streamflow data used in the study were extracted from the "South Saskatchewan River Basin Historical Weekly Natural Flows" CD-ROM obtained from Alberta Environment. The database contains weekly natural flows and weekly precipitation at a number of gauging stations. The natural flows were developed by adjusting recorded daily streamflow for known manmade modifications such as municipal use of water, historical irrigation diversions and return flows, and modifications of streamflow by hydroelectric power projects (Alberta Environment, 1998). The weekly precipitation data were derived from Environment Canada's climatic record, with missing values filled from nearby stations. Details of the procedures used to prepare the weekly data are available in the technical report provided with the CD-ROM (Alberta Environment, 1998). For our study, we selected 13 hydrometric stations (Figure 4.1 and Table 4.1) and 16 precipitation stations (Figure 4.1 and Table 4.2) with data covering the 1913 to 2001 period.

Annual precipitation in the SSRB increases from about 265 mm at Empress near the Alberta-Saskatchewan border to 658 mm at Lake Louise in the foothills of the Rockies. The total precipitation during the November to March period – when precipitation mostly occurs as snowfall – accounts for about 20% of the annual precipitation in the plains (e.g. at Calgary and Red Deer) to about 48% at Lake Louise. All of the 16 stations selected for our study are located either in the plains or at the foothills of the Rocky Mountains (Table 4.2) and so the data may not properly reflect the precipitation fields in the mountains.

Peak flows in the major rivers of the SSRB occur between mid June and early July. Although the eastern slopes of the Rocky Mountains and the adjacent foothills constitute only about 25% of the entire area of the basin (Environment Canada, 1974), annual peak flows in the major rivers are generated by a combination of mountain snowmelt and summer rainfall from higher altitudes. The April-September flow makes up over 75% of the annual flow for each of the 13 hydrometric stations, thus acting as an integrator of the winter to summer precipitation. In order to focus on those variables relevant to the major flow season, we considered the November-March (hereafter Nov-Mar) total precipitation (29 October – 2 April), the November-August (hereafter Nov-Aug) total precipitation (29 October – 27 August), and April-September (hereafter Apr-Sep) average streamflow (3 April – 1 October). Note that because the original data were at weekly time steps, the start (end) dates of the aggregation periods do not necessarily coincide with the first (last) day of the month in question.

To investigate the relationships between SSRB's hydroclimatic variables and the Pacific climate, we used the Nino3, PDO and PNA indices. Nino3, a time series of equatorial Pacific SST anomalies averaged over the window of 5°S–5°N and 150°W–90°W, is commonly used as a measure of the strength of ENSO. The PDO regime represents inter-decadal oscillations in the North Pacific climate system. Observational studies indicate that only 2 complete cycles of PDO have occurred between 1890 and 1998, with alternating cool and warm PDO phases during 1890–1924, 1925–1946, 1947–1976 and 1977–1998 (Mantua and Hare, 2001).The PDO index is a time series of the leading principal component of the North Pacific SST anomalies poleward of 20°N (Mantua and Hare, 2001). The PNA pattern represents a quadripole of 700 mbar geopotential height anomalies, with opposite anomalies centered over the Aleutian Low and western Canada, and the Hawaiian Islands and southeastern US (Wallace and Gutzler, 1981). In this study, we used the November-March average Nino3 and PDO indices, and the December-February average PNA index.

4.3. Hydroclimatic Variability of SSRB

For interdecadal scales, the time periods of locally significant streamflow and precipitation wavelet power are listed in Table 4.1 and Table 4.2, respectively.

For each period of significant wavelet power, the scale corresponding to the peak global wavelet spectrum is shown in parenthesis in Table 4.1 and Table 4.2. For streamflow, the dominant scales are consistently near 19–22, 41–42 and 62 years (Table 4.1 and Figure 4.2), while the dominant scales of precipitation exhibit a higher degree of variability than streamflow (Table 4.2). The dominant scales of precipitation are also found to be season dependent.

For interannual (2–8-year) scales, the scale-averaged wavelet power (SAWP) computed from the normalized wavelet power spectrum is shown in Figure 4.3. Figure 4.3a is the station-time diagram of the streamflow SAWP for the 13 stations. Figure 4.3b shows the time-averaged SAWP at individual stations whereas Figure 4.3c is the 'basin-averaged' SAWP. The peaks in the streamflow SAWP in Figure 4.3c correspond to periods during which statistically significant activities were observed at several hydrometric stations in the basin. Spatially coherent streamflow activities occurred with peaks in 1915, 1927, 1950, 1974 and 1993. Significant activities in the 1910s and 1970s occurred at a lesser number of stations than in the 1920s and 1950s, resulting in a suppressed basin-averaged SAWP for the former two time periods (Figure 4.3c). Although some activity in the 1990s is evident, the power was not significant at any of the 13 stations. Clearly, there was more streamflow activity before 1950s as the streamflow power is seen to decrease significantly during the second half of the twentieth century.

The 'basin-averaged' precipitation SAWP for the 2–8-year scale is also shown in Figure 4.3c. Similar to streamflow, coherent activities in the Nov-Mar and Nov-Aug precipitation were observed in the 1920s, 1950s, 1970s and 1990s. A comparison of the Nov-Mar and Nov-Aug precipitation SAWPs reveals that the former exhibited more variance in the 1920s and 1990s whereas the latter exhibited more variance in the 1920s and 1950s. The temporal fluctuation in the streamflow SAWP shows better agreement with the Nov-Aug precipitation SAWP (Pearson's correlation coefficient, ρ = 0.77) than with the Nov-Mar precipitation SAWP ($\rho = 0.55$). However, there are some mismatches between the temporal locations of significant precipitation and streamflow activities (e.g. in the 1950s, 1970s and 1990s).

The peaks of the streamflow and precipitation SAWP in Figure 4.3 appear to be organized on a background scale of approximately 20–25 years. These peaks correspond to years when streamflow records at the majority of gauging stations were in the upper 25% of the distribution. On the other hand, the low points on the SAWP correspond to periods with streamflow in the lower 25% of the distribution. Those years during which at least two-third of the stations recorded streamflow (precipitation) in the extreme 25% of the distribution (i.e. at least 9 stations for streamflow and 11 stations for precipitation) are listed in Table 4.3. The decrease in streamflow/ precipitation variance during the 1920s, 1930s, 1960s and 1980s was coincident with periods of extensive drought episodes in western Canada (Godwin, 1986; Gan, 2000). The gradual ascension and recession of the SAWP over the quasi 20–25 year cycle means that the SAWP may be used as an early indicator of whether an extended drought period is likely to occur several years in advance. For instance, observation of the continuous recession of the SAWP at the end of the time series since 1996 could be used as an indicator for the onset of extensive drought conditions that occurred in SSRB during 2001–2002.

4.4. Wavelet Analysis of Teleconnections

4.4.1. Scale Averaged Wavelet Power (SAWP)

The SAWP for the Nino3 and PNA indices is plotted alongside the streamflow and precipitation SAWP in Figures 4.3c. As has been described elsewhere (Torrence and Compo, 1998), Nino3 exhibited inter-annual oscillations of large amplitude during the pre-1920 and post-1960 periods, and a reduced level of activity in between. In spite of the weak amplitudes of Nino3 between 1920 and 1960, sixteen moderate to strong ENSO events (10 warm and 6 cool) have been recorded during the same period (e.g. Shabbar *et al.*, 1997). While the ENSO events during 1939–43 may be attributed to moderate Nino3 activities observed from the SAWP, the remaining episodes can not be explained by the Nino3 SAWP.

The temporal locations of the peaks of the streamflow/precipitation SAWP in the 1910s and 1970s generally agree with significant Nino3 activities (Figure 4.3c). Significant streamflow/precipitation activities centered on 1927 and 1950 coincided with ENSO episodes that occurred during 1925–31 and 1951–54. The power fluctuations of the PNA index since 1948 closely follow the intense ENSO activities of the 1970s and 1980s and do not seem to offer additional information to that of Nino3.

4.4.2. Wavelet Coherence and Phase Difference

To provide a more quantitative picture of the links between climate indices and SSRB's hydroclimate, the wavelet coherence between the leading principal component (PC1) of the April-September streamflow and each of the indices were computed (Figure 4.4). The contours in Figure 4.4 enclose periods of statistically significant coherence based on a red noise process as determined by a Monte Carlo experiment (Jevrejeva *et al.*, 2003). Note that the existence of significant wavelet power is not a necessity for the two signals to exhibit significant coherence. The phase differences between the two signals for coherences greater than 0.5 are plotted as vectors in Figure 4.4, where a right pointing arrow indicates that the two signals are in-phase while a left pointing arrow indicates an anti-phase relationship. In spite of the generally weak Nino3 activities between 1920s and 1960s, streamflow and Nino3 show high coherency in the 2-8 year scale prior to 1940s, in the 1950s, 1970s and 1980s (Figure 4.4a). The inconsistency in the relationship between streamflow and Nino3 is clearly evident from the phase distribution in the interannual scale, where the phase difference changes from near 210° prior to 1940s to near 0° in the 1950s and 180° in the 1970s and 1980s (Figure 4.4a). Albeit not as strong as for the interannual scale, Nino3 and streamflow also show significant covariance near the 20-year scale. The coherency between streamflow and PNA is shown in Figure 4.4b. Relatively strong covariance is observed around 1960 and 1970 near the 2-year scale, and in the 1980s near the 5-year scale. The phase distribution appears to be even more inconsistent than that of Nino3.

The coherency between PDO and streamflow is shown in Figure 4.4c. The strongest and most consistent covariance between streamflow and PDO occurs for scales greater than about 20 years. This is not surprising as the PDO regime is an interdecadal oscillatory mode with two dominant scales centered at 15–25 years and 60–75 years (Mantua and Hare, 2001). The phase difference for the interdecadal scale is rather stable near 180° (Figure 4.4c). There are also periods of significant coherence in the interannual scale but with less consistent phase distribution e.g. 1920s to 1940s near the 2-year scale. McCabe and Dettinger (2002) stated that the unfiltered PDO index reflects important ENSO episodes in addition to the interdecadal variability of the North Pacific climate. Indeed, some of the time periods of significant coherency in the interannual scale appear to coincide with ENSO events recorded in earlier studies.

4.4.3. Wavelet Filtered Time Series

We further investigated the PDO-streamflow/precipitation relationship using a low-pass filtered time series. To emphasize the inter-decadal components, the streamflow, precipitation and PDO time series were filtered by using a low-pass cutoff scale of 15 years in Equation 2.7. To facilitate comparison among different time series, the filtered signals were normalized by the standard deviation of the respective original time series. For streamflow and precipitation, the first principal components of the filtered signals were used as the interdecadal signals. The proportion of variance accounted for by PC1 is 87.5%, 36.7% and 52.5% respectively for the Apr-Sep streamflow, Nov-Mar precipitation and Nov-Aug precipitation.

As shown in Figure 4.5, there is a strong agreement between the PDO and streamflow PC1 since 1930, with an increase in PDO associated with a decrease in streamflow and vice versa (Pearson's $\rho = -0.93$). Between 1930 and 2001, the historical drought years of 1936, 1941, 1977, 1983–85 and 1988 (Table 4.3) occurred when the PDO signal was in its extreme warm phase. In contrast, the anomalously wet periods of 1948–54 and 1972 (Table 4.3) coincided with years when the PDO was in its extreme cool phase. It is also noteworthy that the negative interdecadal streamflow signal in 1961 (e.g., the historical drought of 1961) was associated with a temporary warming of the PDO signal in the midst of the cool PDO phase of 1947–1976. Using Equation 2.7 with a low-pass cutoff scale of 8 years, Gobena and Gan (2006) found a similar sign reversal in southwestern Canadian streamflow that closely followed a sign reversal in the inter-decadal PDO signal between 1959 and 1961. The PDO-precipitation relationship after 1930 is broadly similar to that of streamflow (Pearson's $\rho = -0.72$ for Nov-Aug and $\rho = -0.76$ for Nov-Mar).

El Nino (La Nina) episodes coinciding with extreme positive (negative) values of the interdecadal PDO signal seem to have an enhanced negative (positive) effect on the hydrology of SSRB (cf. Table 4.3 and Figure 4.5). For instance, the mature El Nino years of 1931, 1940, 1941, and 1983 coincided with periods when the interdecadal PDO signal was over 0.5 standard deviations above normal. Similarly, the mature La Nina years of 1951 and 1972 coincided with interdecadal PDO signal in excess of 0.5 standard deviations below normal. On the other hand, the impacts of mature El Nino (La Nina) years that occurred during a cool (warm) PDO phase appear to be either muted or are in opposition to the expected response. A good example for the latter case is the 1954 El Nino event which was associated with one of the wettest years in the basin. Hence, hydroclimatic responses to ENSO and PDO could interfere with one another constructively or destructively depending on whether the two climate modes are in phase or out of phase.

Figure 4.5 suggests that a quasi 22-year drought cycle is superimposed on a quasi 42-year cycle. Given that the late 1880s to early 1890s drought was described as the worst on record (Godwin, 1986), one may argue that multiple-year droughts similar to the 1930s and 1980s could be organized on a time scale of 40–50 years. Since the PDO phase shift occurs every 20–30 years, this means that the region could experience a severe inter-decadal drought during every warm PDO phase. Using diatom-inferred lake water salinity levels in the Chauvin Lake of eastern Alberta (52.69°N, 110.10°W) and the 1988–89 regional droughts as a benchmark measure, Leavitt and Chen (2001) predicted that the probability of occurrence of a drought as severe as the 1988 drought by 2030 was 45%, with a mean inter-arrival time of 60 years and an average duration of a decade. An extension of the oscillation in Figure 4.5 into the future also suggests that the region could face the next major multiple-year drought between 2020 and 2030. It is noteworthy that Leavitt and Chen characterized the 1930s droughts as among the mildest on record while our results show that the signature of the 1930s drought in the SSRB hydroclimatic variables is at least comparable to that of the 1980s (Figure 4.5).

4.5. Correlation with Climate Indices

In this section, we use the Spearman rank correlation to examine the value of ENSO and PDO in long-range forecasting. The Spearman correlation is chosen because it is robust to outliers as it finds the correlations between the ranks of the data instead of the values of the data (Wilks, 1995). Rank correlations are computed between the streamflow/winter precipitation at each station and the Nov-Mar Nino3 and PDO indices. Since the climate indices are correlated to one another, partial correlations were also computed so as to assess the relative influence of each climate index. Correlations were assessed for 1913–2001, 1930–2001 and 1950–2001 periods in an attempt to examine the temporal stability of any monotonic relationships. The correlations for the period of 1950–2001 are shown in Figure 4.6a. Table 4.4 shows the number of stations with statistically significant correlations at the 5% and 1% levels. Also shown in Table 4.4 are the lowest, median and highest correlations.

From Table 4.4, it is clear that the PDO shows better correlations with the streamflow of SSRB than Nino3, both in terms of the number of stations with statistically significant correlations and the correlation magnitudes. The median partial correlations between streamflow and PDO|Nino3 (read as PDO given Nino3) for the 1913–2001, 1930–2001 and 1950–2001 periods are –0.36, –0.46 and –0.46, respectively. On the other hand, the median partial correlations between streamflow and Nino3|PDO for all three data windows are close to zero (Table 4.4). These results are consistent with that of Gobena and Gan (2006) who showed that basins originating in the Rocky Mountains were more strongly correlated to PDO than to SOI. The lack of significant partial correlation between streamflow and Nino3|PDO means that tropical Pacific SST conditions during the mature phase of ENSO provide redundant information once the PDO-streamflow relationship is accounted for.

The correlations between the Nov-Mar precipitation and PDO are also stronger than those with Nino3 (Table 4.4). In general, the correlations between the two climate indices and SSRB's precipitation are weaker than with streamflow, underlining the influence of the spatial and temporal noise in precipitation as compared to streamflow. With respect to precipitation, the distribution of stations with significant partial correlations with both PDO|Nino3 and Nino3|PDO suggest that the winter season indices of both climate modes provide important information on the precipitation variability in the basin. However, it should be noted that for the Nov-Mar precipitation, the indices are from the same time window with the precipitation season and as such their value for prediction purposes is limited.

As mentioned in Section 4.2, the precipitation data used in this study come from stations located in the plains or at the foothills of the Rocky Mountains. Could the apparent weakness of the relationship between SSRB's streamflow and Nino3 as compared to the modest precipitation-Nino3 relationships be due to failure of the precipitation data to capture the mountain snowpack variability? A preliminary analysis of Alberta Environment snow course data from 14 stations (see Appendix B for list of stations) shows that the partial correlations between the April 1 SWE and PDO|Nino3 are stronger than those between SWE and Nino3|PDO (Figure 4.6b). The calculations in Figure 4.6b were based on snow data for the period of 1970–2004 and obtained at locations shown in Figure 4.1. From these results, it seems that mountain snowpack, which plays an important role in runoff generation in the major rivers of the SSRB, is primarily forced by the North Pacific winter SST than by the equatorial Pacific SST. These results concur with that of McCabe and Dettinger (2002) who indicated that PDO, not ENSO, is the primary driving force for the April 1 snowpack variability in the western United States. It should be noted that the lack of strong correlations with an ENSO index does not necessarily indicate the lack of ENSO impacts on hydroclimatic variability of the SSRB, but rather that the unfiltered PDO index could reflect the role of the North Pacific decadal climate variability and that of important ENSO episodes. Newman *et al.* (2003) indicate that to first order, PDO can be considered as a reddened response to both ENSO and atmospheric noise, resulting in more decadal variability than either phenomenon.

4.6. Summary and Conclusions

This study has used wavelet transforms and rank correlation analysis to investigate low-frequency hydroclimatic variability in the SSRB of southern Alberta and its dynamical links with the Pacific climate variability. The results of the study are summarized as follows:

1) Dominant modes of streamflow variability occur at inter-decadal scales oscillating near 19–22, 41–42 and 62 years. Clusters of significant streamflow activities at inter-annual scales were also observed at intervals

of approximately 20–25 years. The intensity of the inter-annual oscillations has been on the decline since 1950. The spatially averaged SAWP of streamflow revealed that the clusters of significant activities at the inter-annual scales coincide with years when streamflow was in the upper quartile of the distribution and in between were years when streamflow was in the lower quartile of the distribution. The precipitation variability was also dominated by inter-decadal oscillations although the dominant oscillatory modes show more spatial variability than that of streamflow.

2) As expected, the strongest coherence between streamflow and Nino3 occurs in the interannual scale. In spite of the generally weak Nino3 activities between 1920s and 1960s, SSRB's streamflow and Nino3 showed high coherency in the 2-8 year scale prior to 1940s, in the 1950s, 1970s and 1980s. However, the relationship is highly inconsistent as observed from the shifting phase distribution in the interannual scale, where the phase difference changed from near 210° prior to 1940s to near 0° in the 1950s, and close to 180° in the 1970s and 1980s. On the other hand, PDO and streamflow exhibited consistently strong covariance with a rather stable phase difference of 180° for scales greater than about 20 years.

3) Streamflows in the lower and upper quartiles showed inter-decadal variations that were synchronized with the PDO phase. There were more years with flows in the lower (upper) 25% of the distribution during the warm (cool) PDO phase than during the cool (warm) phase. To a lesser extent, this was also true for winter precipitation. A comparison of the inter-decadal components of basin streamflow and precipitation to that of PDO also revealed that there is a strong agreement between the leading precipitation/streamflow PC and the PDO time series. Since 1930, droughts in the region have occurred at quasi 22-year and 42-year modes.

The multiple-year droughts of the 1930s and 1980s were part of the quasi 42-year mode that occurred when the inter-decadal PDO signal was in its extreme warm phase persistently for several years. The role of PDO in the initiation and/or maintenance of drought in this region deserves further investigation.

4) The Spearman rank correlation analysis with climate indices averaged over the November-March season showed that SSRB's streamflow is highly influenced by the PDO regime, with all 13 stations having significant correlations at the 1% level. Even though streamflow correlations with Nino3 were also significant at 5 of the 13 stations, partial correlation analysis showed that tropical Pacific SST during the mature phase of ENSO did not add substantial new information to that already contained in the PDO index. Preliminary analysis indicates that mountain snowpack, which plays a major role in runoff generation in SSRB, is more strongly correlated to PDO than to Nino3, thus partly explaining the weak ENSO-streamflow relationship. Even though correlations between SSRB's winter precipitation and Nino3 (PDO) were generally weaker than with streamflow, there are still statistically significant correlations observed at some precipitation stations.

Alberta Environment uses the current antecedent soil moisture, snow water equivalent, precipitation, snow pillow information and temperature to determine seasonal runoff by statistical techniques (Alberta Environment, 2006b). Currently, water supply outlooks for the March-September runoff volume are issued in the first week of each month beginning in February. The improved understanding of the relationships between Pacific climate variability and basin streamflow/precipitation could be used as a basis to objectively incorporate large-scale climate dynamics into long-range stream flow forecasting in the SSRB. The inclusion of this information into flow forecasting frameworks may lead to improved forecast lead times and/or skills.

One approach to include large-scale climate information in long-range flow forecasting is to predict statistically precipitation and temperature at the desired lead time and then use the predicted values to drive a hydrologic model (e.g. see Mwale *et al.*, 2004; Mwale and Gan, 2005). With regard to the SSRB, the applicability of this approach seems limited because of a number of reasons. First, significant wavelet power fluctuations at the inter-annual scale were mainly observed at intervals of about 20–25 years. In the absence of strong interannual persistence of precipitation power, the climate-precipitation relationship may not be strong enough to produce precipitation forecasts that are accurate enough for driving a hydrologic model. Second, modest correlations between precipitation and large-scale climate anomalies are observed only during the winter season (e.g. Shabbar *et al.*, 1997; Gan *et al.*, 2007) whereas SSRB's streamflow integrates precipitation over the winter to summer period and thus the effect of the winter precipitation alone, even if accurately predicted, may not yield a marked improvement in streamflow forecasts for the major flow season. Thirdly, the precipitation data from the foothill stations – where sufficiently long and reliable records are available – do not seem to accurately represent the mountain snowpack variability.

A second and more straightforward approach is to use indices of climate patterns as predictors in statistical regression models provided that the indices have a lag relationship with hydroclimatic data. For the SSRB, the use of the winter season PDO, plus PNA and Nino3/SOI from earlier seasons (beginning with the initiation of ENSO) as predictors may yield some improvement in the skill and/or lead time of the current regression models. Although tropical Pacific SST conditions during mature ENSO periods do not add new information to that contained in the PDO index, the influence of the early stages of ENSO on SSRB's streamflow variability cannot be ruled out and warrants further consideration. One limitation of regression-based forecasts is that they do not contain information about forecast uncertainty. In Chapter 5, a robust regression model based on M-estimators is used as a basis to develop ensemble forecasts using a modified nearest neighbors resampling algorithm.

A third approach could be to use the large-scale climate information as a conditioning variable for resampling appropriate historical precipitation and temperature data as surrogates for future climate and then use the re-sampled historical data as multiple scenarios for input to a hydrologic model in an ensemble streamflow prediction (ESP) framework (e.g. Hamlet and Lettenmaier, 1999). One attractive feature of the ESP approach is that probabilistic statements can be attached to streamflow forecasts. However, the mismatches in the temporal locations of the streamflow and precipitation powers as revealed by the SAWP at the interannual scale may pose problems with regard to the validity of this approach. A case by case analysis of the power fluctuations in the streamflow and precipitation in each sub-basin is required to establish its applicability to SSRB. The ESP approach will be explored in Chapter 6.

Acknowledgement

The snow course data used in this study was generously provided by Mr. Chacko Abraham of Alberta Environment.

References

Alberta Environment. (1998). South Saskatchewan River basin historical weekly natural flows – 1912 to 1995, Main Report, Alberta Environment Protection, Edmonton, Canada.

Alberta Environment. (2002). South Saskatchewan River Basin – Non-irrigation water use forecasts, Publ. No.: T/656, Edmonton, Canada.

Alberta Environment. (2006a). *Water in Alberta.* Alberta Environment website: http://www3.gov.ab.ca/env/water/gwsw/quantity/waterinalberta/index.html (accessed August 2006).

Alberta Environment. (2006b). *Water Supply Outlook for Alberta.* Alberta Environment website: http://www3.gov.ab.ca/env/water/ws/WaterSupply/Index.html (accessed August 2006).

Bonsal, B. R., Shabbar, A., and Higuchi, K. (2001). Impacts of low frequency variability modes on Canadian winter temperature. *Int. J. Climatol.,* 21, 95–108.

Coulibaly, P., and Burn, D. H. (2004). Wavelet analysis of variability in annual Canadian streamflows. *Water Resour. Res.,* 40, W03105, doi:10.1029/2003WR002667.

Environment Canada. (1974). Streamflow forecasting: South Saskatchewan River below Red Deer River, Technical report to the Prairie Provinces Water Board Committee on Hydrology, Water Survey of Canada, Ontario.

Gan, T. Y. (2000). Reducing vulnerability of water resources of Canadian Prairies to potential droughts and possible climatic warming. *Water Resources Management,* 14, 111–135.

Gan, T. Y., Gobena, A. K., and Wang, Q. (2007). Precipitation of southwestern Canada – Wavelet, scaling, multifractal analysis, and teleconnection to climate anomalies. *J. Geophys. Res.*, 112, D10110, doi:1029/2006JD007157.

Gobena, A. K., and Gan, T. Y. (2006). Low-frequency variability in southwestern Canadian streamflow: links to large-scale climate anomalies. *Int. J. Climatol.,* 26, 1843-1869.

Godwin, R. B. (1986). Drought: a surface water perspective. In Drought: The Impending Crisis? *Proc. Canadian Hydrology Symposium No 16*, Regina, Saskatchewan, Canada.

Hamlet, A., and Lettenmaier, D. (1999). Columbia River streamflow forecasting based on ENSO and PDO climate signals. *J. Water Resour. Plann. Manage.,* 125(6), 333–341.

Jevrejeva, S., Moore, J. C., and Grinsted, A. (2003), Influence of the Arctic Oscillation and El Nino-Southern Oscillation on ice conditions in the Baltic sea, The wavelet approach. *J. Geophys. Res.,* 108(D21), 4677, doi:10.1029/2003JD003417.

Leavitt, P., and Chen, G. (2001). Prairie Drought Project Website http://www.uregina.ca/drought (Accessed August 2005).

Mantua, N., and Hare, S. (2001). The Pacific Decadal Oscillation. *J. Oceanography,* 58, 35–44.

McCabe, G. J., and Dettinger, M. D. (2002). Primary modes and predictability of year-to-year snowpack variations in the western United States from teleconnections with Pacific Ocean climate. *J. Hydrometeorology,* 3, 13-25.

Moore, R. D. (1996). Snowpack and runoff responses to climatic variability, southern Coast Mountains, British Columbia. *Northwest Science,* 70, 321–333.

Mwale, D., Gan, T. Y., and Shen S. S. P. (2004). A new analysis of variability and predictability of seasonal rainfall of central southern Africa for 1950–94. *Int. J. Climatol.,* 24, 1509–1530.

Mwale, D., and Gan, T. Y. (2005). Wavelet analysis on variability, teleconnectivity and predictability of the September-November East African Rainfall. *J. Appl. Meteorol.,* 44, 256-269.

Newman, M., Compo, G. P., and Alexander, M. A. (2003). ENSO-forced variability of the Pacific Decadal Oscillation. *J. Climate*, 16(23), 3853-3857.

Shabbar, A., Bonsal, B., and Khandekar, M. (1997). Canadian precipitation patterns associated with the Southern Oscillation. *J. Climate,* 10, 3016–3027.

Torrence, C., and Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin Amer. Meteorol. Soc.,* 79(1), 61-78.

Wallace J. M., and Gutzler D. S. (1981). Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Weather Rev.,* 109, 784–812.

Wilks, D. S. (1995). *Statistical Methods in the Atmospheric Sciences: An Introduction.* San Diego, CA: Academic Press.

Woo M.-K., and Thorne, R. (2003). Comment on 'Detection of hydrologic trends and variability'. *J. Hydrol.,* 277, 150–160.

Table 4.1. Summary of 13 stations of streamflow data used in the study, showing periods of statistically significant wavelet power at the 10% level with respect to a red noise background spectrum. The corresponding scale in years is shown in parenthesis.

WSC Index[a]	Station Name	Lat – Lon, deg	Drainage area, km^2	Time period of significant wavelet power[b]
05BB001	Bow River at Banff	51.17, 115.57	2210	1925-1950 (19), 1913-1998 (41)[c]
05BC001	Spray River at Banff	51.16, 115.55	749	1917-63 (19)[c], 1921-65 (42)
05BJ001	Elbow River below Glenmore Dam	51.02, 114.09	1230	1924-50 (41)
05BH004	Bow River at Calgary	51.05, 114.05	7820	1913-90 (41)[c]
05BL009	Highwood River near Aldersyde	50.70, 113.86	2340	1943-81 (22)
05AD003	Waterton River near Waterton Park	49.11, 113.84	614	1939-75 (21), 1928-83 (41)[c], 1913-2001 (62)[c]
05CC002	Red Deer River at Red Deer	52.28, 113.82	11100	-
05AD005	Belly River near Mountain View	49.10, 113.70	319	1939-87 (21)[c], 1927-80 (42)[c], 1913-2001 (62)[c]
05AE027	St. Mary River at Int'l Boundary	49.00, 113.31	1170	1935-80 (21)[c], 1935-70(42), 1913-2001 (62)[c]
05AE002	Lee Creek at Cardston	49.20, 113.30	307	1933-82 (21)[c]
05AD007	Oldman River near Lethbridge	49.71, 112.87	15800	1937-93 (22)[c], 1927-81 (42)[c]
05AE006	St. Mary River near Lethbridge	49.57, 112.84	3340	1938-79 (20)[c], 1938-67 (42), 1913-2001 (62)[c]
05AJ001	S. Saskatchewan River at Medicine Hat	50.04, 110.68	42100	1941-86 (21)[c], 1913-82 (42)[c], 1921-94 (62)[c]

[a] WSC: Water Survey of Canada.

[b] The corresponding scale in years is given in parenthesis.

[c] The global wavelet power at that scale is also statistically significant at the 10% level.

Table 4.2. Summary of 16 stations of precipitation data used in the study, showing periods of statistically significant wavelet power at the 10% level with respect to a red noise background spectrum. The corresponding scale in years is shown in parenthesis.

	Lat – Lon,	Elevation,	Time period of significant wavelet power[b]	
Station Name	deg	m[a]	Nov-March precipitation	Nov-August precipitation.
Banff	51.18, 115.57	1397	1920-85 (20)	1983-97 (8), 1953-70 (21), 1913-2001 (41)[c]
Coleman	49.63, 114.58	1341	1918-80 (17)[c]	1960-2001 (20.5)[c], 1913-2001 (41)[c]
Pekisko	50.37, 114.42	1439	1924-87 (19)[c]	1956-61 (20.5), 1913-2001 (62)[c]
High River	50.48, 114.17	1219	1913-75 (22)[c]	1913-2001 (13)[c]
Beaver Mines	49.47, 114.17	1286	1930-80 (14.7, 20)	1967-2001 (21.5)
Calgary	51.12, 114.02	1084	1935-62 (24), 1913-2001 (62)[c]	1913-32 (13), 1922-55 (22) 1913-2001 (62)[c]
Pincher Creek	49.50, 113.95	1145	1928-81 (25)[c]	1927-2001 (22.5)[c]
Red Deer	52.18, 113.90	860	1915-45 (11), 1953-77 (32)	1913-43 (11)[c], 1921-38 (22)
Claresholm	49.93, 113.75	1052	1913-2001 (62)[c]	1913-53 (41), 1913-2001 (62)[c]
Meadow Creek				
Lacombe	52.47, 113.75	847	1940-2001 (62)[c]	1950-70 (9), 1953-77 (22), 1913-2001 (62)[c]
Mountain View	49.13, 113.63	1318	1920-68 (22)[c]	1947-70 (13.5)[c], 1927-79 (21.5)[c]
Gleichen	50.88, 113.05	905	-	1930-45 (22), 1913-2001 (62)[c]
Lethbridge CDA	49.70, 112.78	921	1913-2001 (62)[c]	1971-2001 (13)[c]
Drumheller	51.47, 112.72	687	1913-2001 (62)[c]	-
Vauxhall CDA	50.05, 112.13	779	1945-79 (14)[c], 1919-45 (31), 1913-2001 (62)[c]	1937-77 (13)[c]
Medicine Hat	50.02, 110.72	717	1919-59 (17.5)[c], 1913-2001 (62)[c]	1920-40, 1990-2001 (6)[c], 1937-70 (13)

[c] Elevation data taken from Environment Canada's climate CD-ROM.

[b] The corresponding scale in years is given in parenthesis.

[c] The global wavelet power at that scale is also statistically significant at the 10% level.

Table 4.3. List of years with streamflow and precipitation in the lower or upper 25% of the distribution at least at two-third of the stations considered in the study. The years are arranged according to the PDO phase.

Lower 25%	Upper 25%	PDO phase
April-September streamflow		
1919^{e}, 2001	1916, 1948, 1951^{l}, 1953, 1954^{e}, 1965^{l}, 1972^{l}	Cool PDO (1890–1924, 1947–1976, 1999– present)
1926^{e}, 1931^{e}, 1936, 1940^{e}, 1941^{e}, 1944, 1977^{e}, 1983^{e}, 1984, 1985, 1988	1927^{e} 1928, 1995	Warm PDO (1925–1946, 1977–1998)
November-March precipitation		
1917, 2001	1947, 1948, 1951^{l}, 1972^{l}, 1974^{l}	Cool PDO
1926^{e}, 1929^{l}, 1931^{e}, 1942^{e}, 1988, 1992^{e}, 1995	1925^{l}	Warm PDO
November-August precipitation		
1918^{l}, 1919^{e}, 2000, 2001	1915^{e}, 1948, 1951^{l}, 1953, 1954^{e}, 1965^{l}	Cool PDO
1929^{l} 1931^{e} 1936	1927^{e}, 1978, 1993, 1998^{e}	Warm PDO

^{e} Mature El Nino year; ^{l} Mature La Nina year.

Table 4.4. Number of streamflow and precipitation stations that exhibit statistically significant Spearman rank correlations with climate indices. The lowest, median and highest correlations with each index are also shown in parenthesis. Correlations significant at the 5% (1%) level are indicated in *italics* (boldface).

Data period	Climate index	Streamflow		Precipitation	
		5% level	1% level	5% level	1% level
1950–2001	Nino3	5	1	7	3
		(0.12, –18, **–0.38**)		(–0.06, –0.24, **–0.50**)	
	PDO	12	12	12	7
		(*–0.25*, **–0.50**, **–0.57**)		(–0.08, *–0.31*, **–0.63**)	
	Nino3|PDO	0	0	6	2
		(–0.14, 0.07, *0.33*)		(0.05, –0.14, **–0.30**)	
	PDO|Nino3	13	13	3	0
		(*–0.35*, **–0.46**, **–0.53**)		(0.06, –0.25, **–0.51**)	
1930–2001	Nino3	4	3	10	5
		(0.03, –0.18, **–0.36**)		(–0.03, *–0.25*, **–0.50**)	
	PDO	13	12	11	6
		(*–0.29*, **–0.50**, **–0.57**)		(–0.13, *–0.27*, **–0.56**)	
	Nino3|PDO	0	0	4	2
		(–0.15, 0.04, *0.24*)		(–0.05, –0.15, **–0.34**)	
	PDO|Nino3	13	13	6	2
		(**–0.34**, **–0.46**, **–0.51**)		(-0.02, –0.18, **–0.44**)	
1913–2001	Nino3	5	4	8	4
		(0.03, –0.18, **–0.38**)		(–0.06, *–0.21*, **–0.41**)	
	PDO	13	12	11	4
		(*–0.23*, **–0.43**, **–0.49**)		(–0.09, *–0.25*, **–0.50**)	
	Nino3|PDO	1	0	3	1
		(0.19, 0.00, *–0.21*)		(0.04, –0.13, **–0.34**)	
	PDO|Nino3	13	13	5	1
		(*–0.27*, **–0.36**, **–0.42**)		(0.01, –0.16, **–0.40**)	

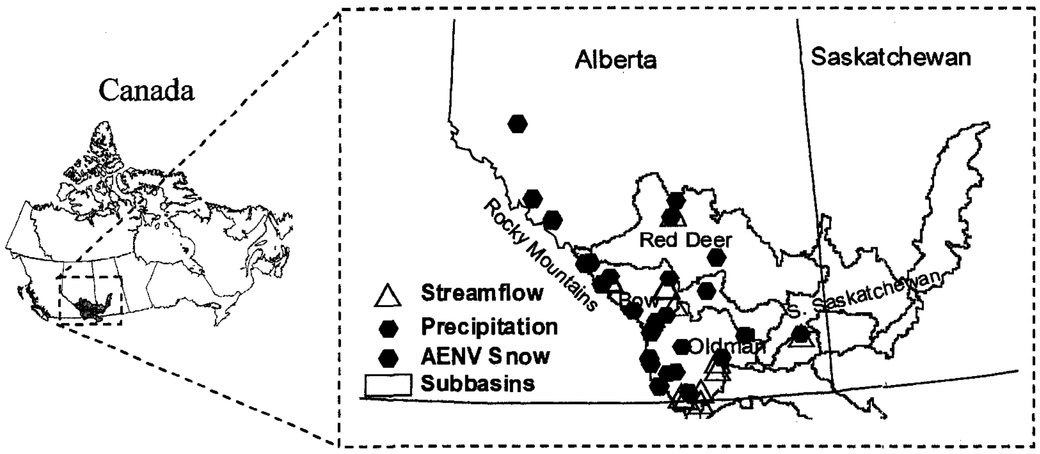


Figure 4.1. Location of precipitation, streamflow and snow course stations used in the study.

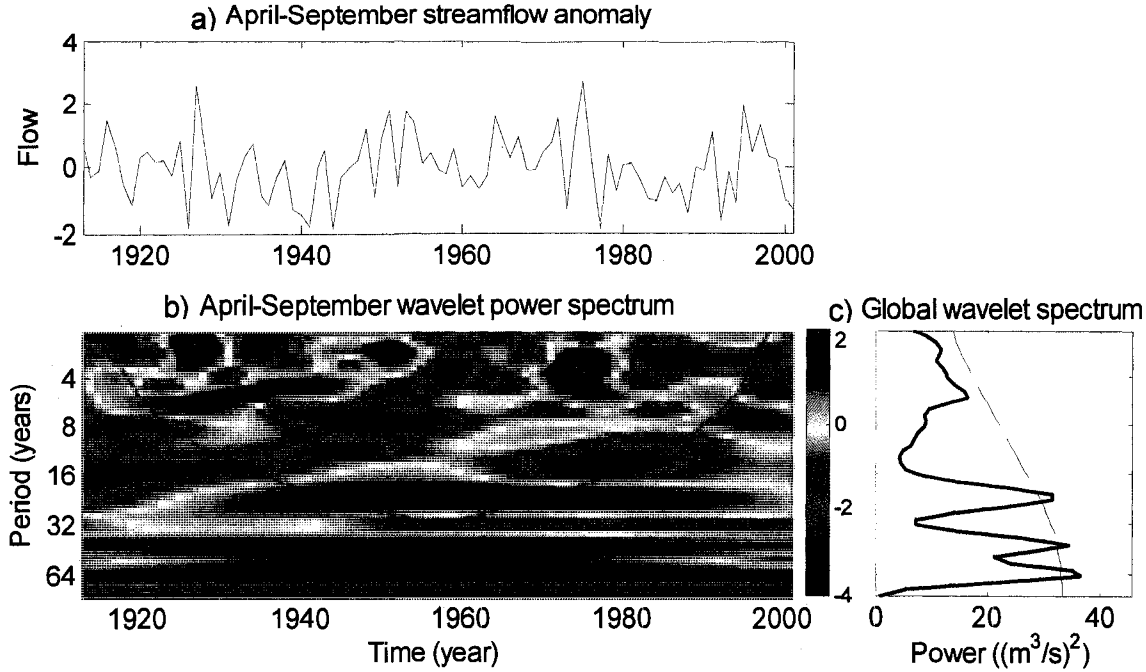


Figure 4.2. Wavelet decomposition of the April-September average streamflow of Belly River near Mountain View. (a) Streamflow anomaly time series. (b) Wavelet power spectrum. (c) Global wavelet power spectrum. The solid lines in (b) enclose regions in the time-frequency domain where the streamflow power was statistically significant against a red-noise spectrum at the 10% level. The dashed line in (b) is the cone of influence outside which the effect of zero-padding may suppress the wavelet power. The dashed line in (c) is the 90% confidence level for the global wavelet power spectrum.

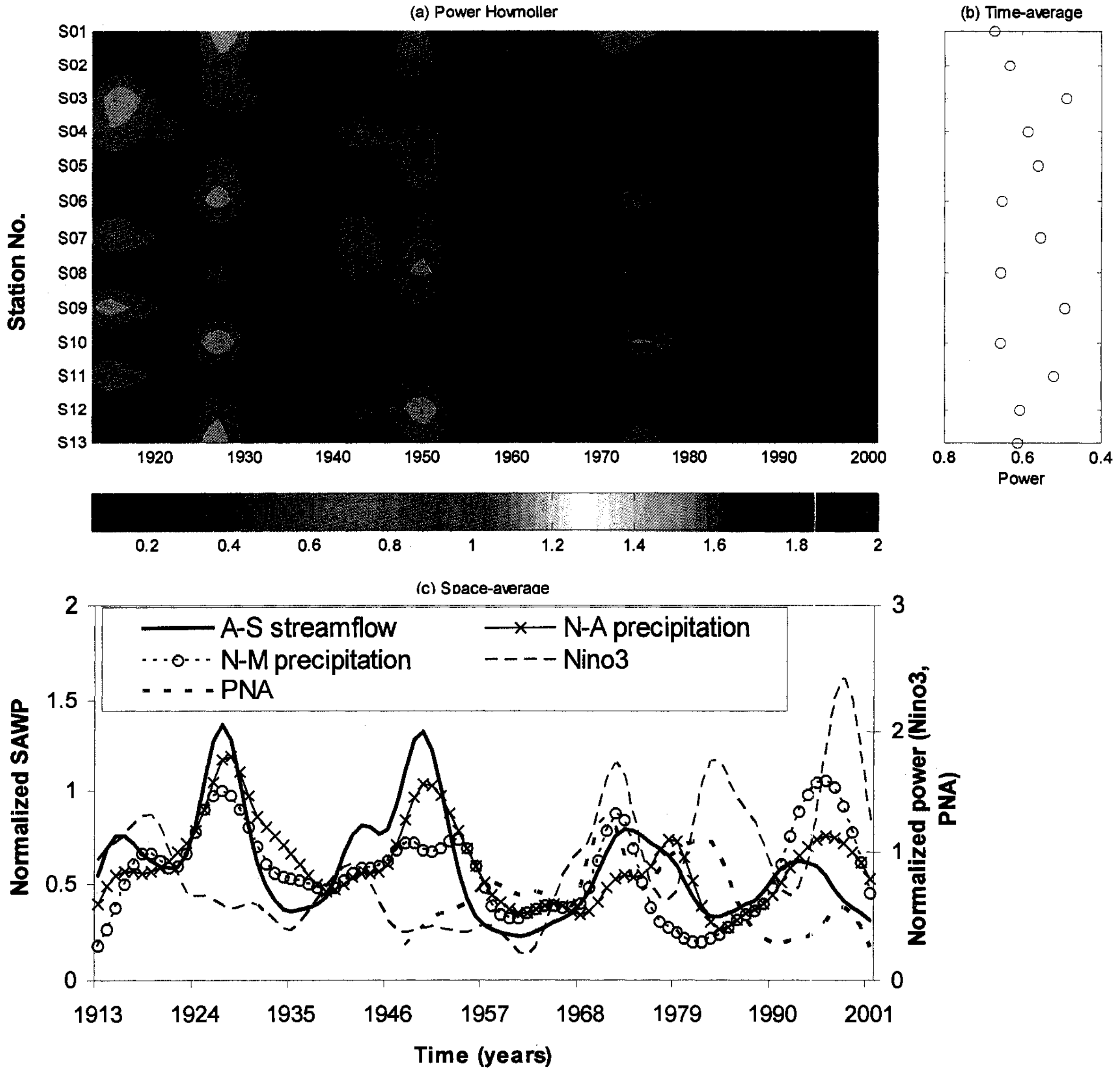


Figure 4.3. (a) Station-time diagram of the Apr-Sep streamflow SAWP for 13 stations in the SSRB. The station numbers follow the listing in Table 4.1. (b) Time-averaged Apr-Sep streamflow SAWP. (c) Basin-averaged SAWP of the Apr-Sep streamflow, Nov-Mar precipitation and Nov-Aug precipitation for the SSRB. The wavelet power was averaged over the 2–8-year scale.

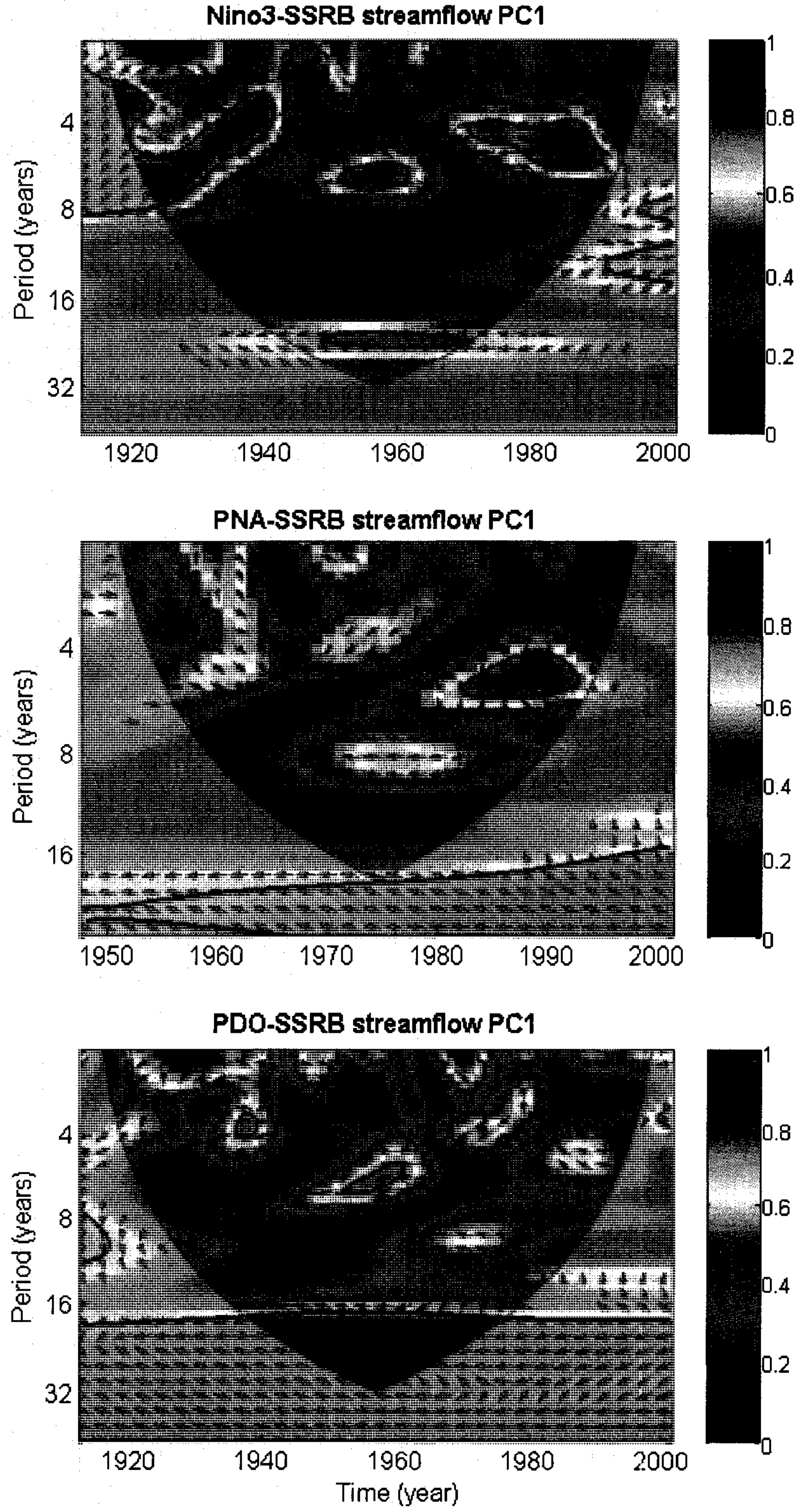


Figure 4.4. Wavelet coherence between streamflow PC1 and (a) Nino3, (b) PNA, (c) PDO. The contours indicate the 95% confidence level. The vectors show the phase difference between the two signals where phase difference is shown for coherence greater than 0.5.

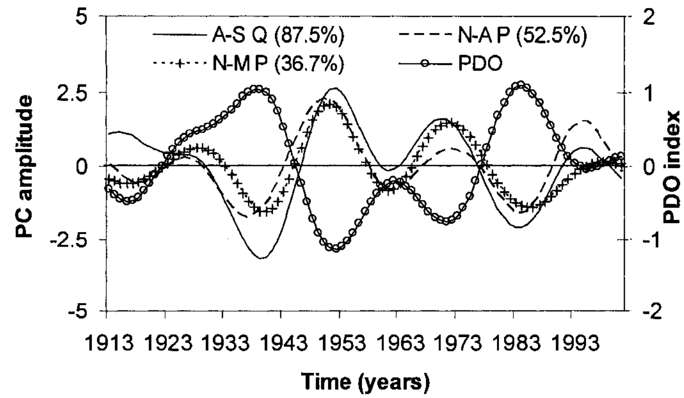


Figure 4.5. The leading PC of inter-decadal Apr-Sep streamflow (A-S Q), Nov-Aug precipitation (N-A P) and Nov-Mar precipitation (N-M P) and the inter-decadal component of the PDO index. A low-pass wavelet filter with a cutoff scale of 15 years was used to reconstruct the signals.

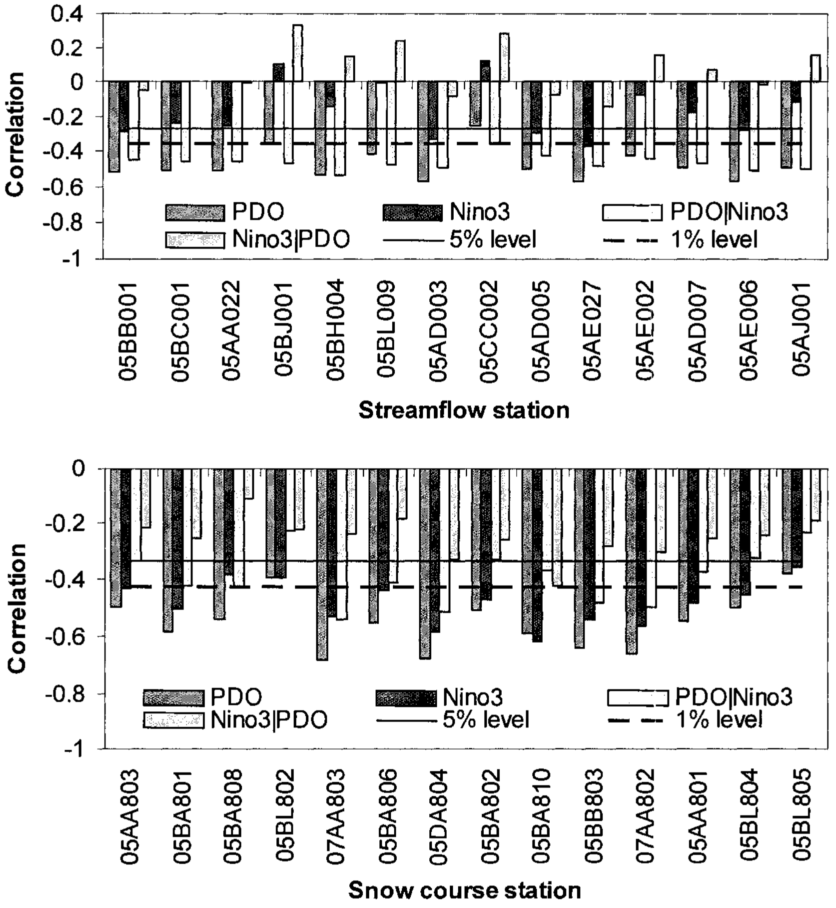


Figure 4.6. Spearman rank correlations and partial correlations between (a) Apr-Sep streamflow and Nov-Mar PDO and Nino3 indices (1950-2001); (b) April 1 SWE and Nov-Mar PDO and Nino3 indices (1970-2004). The horizontal lines are the 5% and 1% significance levels.

Chapter 5 Statistical Ensemble Streamflow Forecasting in the South Saskatchewan River Basin with Robust M-Regression and Modified Nearest Neighbors Resampling

5.1. Introduction

In Alberta, a Canadian Prairie province, rivers have served as the primary source of water supply for most users since the dawning of the 20th century. By the year 2005, the South Saskatchewan River Basin (SSRB) alone provided over 57% of the water allocated in the province (Alberta Environment, 2006a). The SSRB, which is home to about 47% of the provincial population, occupies about a quarter of the surface area of the province, but it contains less than 6% of the province's water resources. As stipulated in the *Master Agreement on Apportionment*, Alberta can use only up to half of the natural flow of the basin and pass the remaining flow to Saskatchewan (Alberta Environment, 2006a).

There is clear evidence that the water resource of the SSRB is under stress from demands exerted by fast population increase and an expanding economy. For instance, the Bow, Oldman and South Saskatchewan sub-basins have already reached their allocation potentials and thus the new Water Management Plan for the SSRB contains recommendations to suspend allocation of new water rights in these sub-basins until the Minister of Environment specifies how water not currently allocated is to be used (Alberta Environment, 2006b). In the face of increased uncertainty due to climate variability, sustaining existing water rights and meeting new water demands in southern Alberta will undoubtedly become a formidable challenge in the immediate future. Long-range streamflow forecasts that take into account the effects of climatic variability will be instrumental for proper planning and management of water resources in the basin.

Seasonal streamflow forecasts in Alberta are currently produced by statistical techniques using antecedent conditions (e.g. soil moisture, snow course/snow pillow data, precipitation and temperature) and assumed future scenarios (Alberta Environment, 2006c). Water supply outlooks for the March-September runoff volume of major rivers in the SSRB are usually issued within the first ten days of each month beginning in February. The probable forecast is produced by assuming normal weather in the future. Three additional forecasts are made by assuming that precipitation will be in the lower 10th, 25th and 75th percentiles (Alberta Environment, 2006c).

The links between western Canadian hydroclimatic variables and large-scale climate anomalies such as the El Nino-Southern Oscillation (ENSO), Pacific/North America pattern (PNA) and Pacific Decadal Oscillation (PDO) have been a subject of several previous studies (Moore, 1996; Moore and McKendry, 1996; Shabbar and Kandekar, 1996; Shabbar *et al.*, 1997; Bonsal *et al.*, 2001; Hsieh and Tang, 2001; Woo and Thorne, 2003; Gan *et al.*, 2007; Gobena and Gan, 2006a, 2006b). The incorporation of large-scale climate information into forecasting models for basins with strong climate signals should provide improvement in forecast skill and/or lead times. However, such information has rarely been used for long-lead streamflow forecasts for western Canadian basins. A notable exception is the study of Hsieh *et al.* (2003), who used a linear regression model based on local precipitation signal and indices of ENSO, PNA and PDO climate anomalies up to the end of November to forecast the April-August streamflow of the Columbia River at Donald, BC.

Regardless of the type of predictors used, deterministic statistical forecasts have a limited value since they do not provide information about the error distribution of the forecasts (Day, 1985). However, deterministic forecasts can be used as a basis to generate ensemble forecasts by way of local estimators such as nonparametric kernel density estimators and K-nearest neighbor (K-NN) bootstrap methods. In the original K-NN bootstrap method developed by Lall and Sharma (1996), the *K* nearest neighbors of the point of interest (based on some distance metric) are selected from historical record and then the ensemble members are generated by resampling the neighbors via a kernel that assigns large weight to the nearest and small weight to the furthest neighbors. While this approach can easily be extended to include large-scale climate information for conditionally selecting the K-NN, ensemble members not seen in historical record can not be simulated (Prairie *et al.*, 2006). To overcome this limitation, Grantz *et al.* (2005) used a modified K-NN bootstrap method to generate ensemble streamflow forecasts from deterministic local polynomial regression forecasts. They generated ensembles by resampling from the residuals of the K-NN and adding them to the "mean" forecast obtained from the local regression. Prairie *et al.* (2006) combined the modified K-NN bootstrap method with a lag-1 local polynomial regression for stochastic streamflow simulation.

The objective of this study is to develop a statistical ensemble streamflow forecasting model for two watersheds in the SSRB. The model uses both antecedent conditions and large-scale climate information as predictors as well as conditioning vectors for generating ensembles by a modified K-NN resampling approach. Our study differs from previous studies on streamflow forecasting (e.g., Grantz *et al.*, 2005; Prairie *et al.*, 2006) in the following aspects: (1) We introduce the recently developed robust M-regression model for obtaining the mean forecast, (2) We use a kernel function that is entirely driven by the predictor data using the Mahalanobis distance metric for resampling from the K-NN, and (3) the number of nearest neighbors (*K*) and the weights attached to each neighbor are optimized. The paper is organized as follows. Section 5.2 presents a description of the study area and data. In section 5.3, the development of the forecast model and criteria used to evaluate the performance of the forecasts are described. The results are discussed in section 5.4 followed by conclusions in section 5.5.

5.2. Data Description

The Bow River at Banff (WSC Station 05BB001) and Castle River near Beaver Mines (WSC Station 05AA022), both located in the headwaters of the SSRB were selected for this study (Figure 5.1). These watersheds are representative of the upper parts of the basin where roughly 90% of the annual flow is generated. The Bow River at Banff (hereafter referred to as "Bow River") is located in the headwaters of the Bow River sub-basin and has a drainage area of about 2210 km^2, with elevation ranging from around 1200 m at the outlet near Banff to over 3000 m in the mountains. About 36% of the drainage area of the Bow River lies below an elevation of 2000 m whereas about 17% of its drainage area has elevation greater than 2500 m. There are two meteorological stations with long historical records located in the watershed, namely Banff (Elevation: 1397m; Location: 51°11'N, 115°34'W) and Lake Louise (Elevation: 1534 m, Location: 51°26'N, 116°13'W). The mean annual flow of Bow River (05BB001) is about 562 mm whereas mean annual precipitations at Banff and Lake Louise are about 473 mm and 658 mm, respectively. The small difference between annual precipitation and streamflow does not necessarily imply negligible evapotranspiration losses but rather emphasizes the effect of topography in enhancing precipitation. The monthly climatologies of precipitation at Banff and Lake Louise stations and streamflow of the Bow River at Banff are shown in Figure 5.1. Peak flows in major rivers of the SSRB typically occur between mid June and early July.

The Castle River near Beaver Mines (hereafter referred to as "Castle River") is located in the headwaters of the Oldman River sub-basin and has a drainage area of 823 km^2. The watershed elevation ranges from 1200 m at the outlet to over 2500 m. About 85% of the drainage area of the watershed lies below an elevation of 2000 m. The only meteorological station in the watershed with long record is found at Beaver Mines near the basin outlet (Elevation: 1286m, Location: 49°28'N, 114°10'W). The mean annual flow of Castle River (05AA022) is about

600 mm whereas the mean annual precipitation at Beaver Mines is about 617 mm. The monthly precipitation and streamflow for the Castle River at Beaver Mines are also shown in Figure 5.1.

Natural weekly streamflow data for the two watersheds were extracted from the "South Saskatchewan River Basin Historical Weekly Natural Flows" CD-ROM obtained from Alberta Environment (Alberta Environment, 1998). The seasonal average flow was computed from the weekly streamflow data. Snow water equivalent (SWE) data were extracted from snow course data, also obtained from Alberta Environment (Chacko Abraham, Personal communication, 2004; See Appendix B for list of stations). The earliest available snow course record dates back to 1937 but most of the stations have data only since the 1960's. The highest snow course station is located at an elevation of 2380 m.

5.3. Forecast Model Development

5.3.1. Model for Mean Forecast

Statistical forecast models describe the empirical relationships between the predictor and response variables by a function of the form $y = f(X) + \varepsilon$ where y is an $N \times 1$ vector of observations, X is an $N \times M$ matrix of predictors, and ε is the error vector. The function f is fitted to a set of training examples by minimizing some cost function. Because of its computational efficiency and simplicity, multiple linear regression (MLR) has been widely used for forecasting hydroclimatic time series. For optimal performance of MLR models, the error terms should be independent, normally distributed and have a constant variance. However, such requirements may not always be met in hydroclimatic time series where outliers are more likely to be observed. The class of M-estimators on the other hand, have a good efficiency over a wide range of distributions (Huber, 1981) and can be made robust against outliers. In M-estimators, the squared residuals in the linear regression estimators are replaced by a general cost function of the form,

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{N} \rho(r_i) \tag{5.1}$$

where $r_i = y_i - X_i\beta$ is the i$^{\text{th}}$ residual, and ρ is a symmetric and non-decreasing function (Rousseeuw and Leroy, 1987). A weighting scheme that gives less importance to outliers can be achieved by choosing a bounded (e.g. convex) loss function (Serneels *et al.*, 2005). The weight attached to observation *i* can be written as,

$$w_i^r = \rho(r_i)/r_i^2 \Rightarrow \rho(r_i) = w_i^r r_i^2 \ . \tag{5.2}$$

Substituting Equation 5.2 into Equation 5.1, we have

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{N} w_i^r (y_i - X_i\beta)^2 \ . \tag{5.3}$$

Thus, the M-estimator is a weighted least squares estimator but with weights depending on β. For the special case of $\rho(r) = r^2$, Equation 5.3 reduces to the traditional least squares estimator.

Equation 5.3 provides robustness only against outliers in the residual terms (or vertical outliers). Serneels *et al.* (2005) have introduced additional algorithm to provide protection against outliers in the predictor space (also called leverage points). The idea is to multiply the loss function in Equation 5.3 by a second weight w_i^x:

$$\hat{\beta} = \operatorname{argmin}_{\beta} \sum_{i=1}^{N} w_i^r w_i^x (y_i - X_i\beta)^2 \tag{5.4}$$

For observations close to the center of the data cloud, the leverage weights w_i^x take on a value close to one whereas for outliers, they receive a value close to zero. Since Equation 5.4 protects against all kinds of outliers, it is called robust M-estimator (Serneels *et al.*, 2005). The weights w_i^r and w_i^x are estimated by an iterative reweighted least squares scheme, with the starting values computed from the residuals using the following equations,

$$w_i^r = f\left(\frac{r_i}{\hat{\sigma}}, c\right)$$

$$f(z,c) = \frac{1}{\left(1+\left|\frac{z}{c}\right|\right)^2}, \quad z = \frac{r_i}{\hat{\sigma}} \tag{5.5}$$

$$w_i^x = f\left(\frac{\|X_i - med_{L1}(X)\|}{median_i \|X_i - med_{L1}(X)\|}, c\right) \tag{5.6}$$

where *f* is called the "Fair" function, $\hat{\sigma}$ is an estimate of the residual scale (taken as the median absolute deviation), *c* is a tuning constant (usually set to 4), $\|.\|$ is the Euclidean norm, and med_{L1} is the L_1-median computed from the pool of predictor variables. The L_1-median minimizes the sum of Euclidean distances to all points in the predictor set. For more details and a MATLAB script on robust M-estimators, the reader is referred to Serneels *et al.* (2005).

In this study, we used the traditional MLR and robust M-regression models to obtain the mean forecast of the seasonal flow. We found that the robust M-regression model performs at least as well as MLR, if not better. Hence, we will only consider mean forecasts produced by the former model in subsequent discussions.

5.3.2. Selection of Predictors

In addition to the already established links between standard climate indices and streamflow, correlations between the April-September streamflow and gridded global ocean/atmosphere (SST, SLP, geopotential height and wind) data were evaluated at different lead times. Regions of high correlations with the global climate data generally collapsed onto the standard ENSO (equatorial central Pacific) and PDO/PNA (North Pacific/North American) regions (not shown). Use of new indices constructed from spatial averages from regions of high correlations offered no obvious advantage over standard indices. Therefore, the more readily available standard climate indices, and antecedent precipitation, temperature and a sub-basin now water equivalent (SWE) index (SWEI) are chosen as candidate predictor variables for use in the forecasting exercise presented here.

In this study the robust M-estimator is combined with a generalized cross-validation (GCV) algorithm (Loader, 1997) to automatically select the best subset of predictors that provide the least GCV score. The GCV function is defined as

$$GCV(M) = \frac{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{(1 - M/N)^2} \tag{5.7}$$

where M is the number of predictors in the model, such that $M << N$. The GCV function introduces a penalty for each predictor added to the model, thus minimizing the risk of using too many predictors that could artificially inflate the model skill at development stage. In a predictive environment, the GCV criterion is more appropriate for model selection than the least squares criterion because the former estimates the average squared prediction error (Loader, 1997) whereas the later estimates the average squared estimation error.

The climate indices presented to the GCV algorithm consist of rolling 3-month averages of PDO, SOI/Nino3 and PNA indices. Although PDO is the primary climate mode of importance to streamflow variability in the SSRB (see Chapter 4), correlation analysis indicates that there are modest relationships between the April-September streamflow and Nino3/SOI indices starting from late spring of the previous year. Thus, the 3-month averages for Nino3/SOI start with May-June-July (MJJ) of the previous year and end with January-February-March (JFM) of the runoff year. For PNA and PDO indices, the 3-month averages start with August-September-October (ASO) of the previous year and end with JFM of the runoff year. Forecasts are issued on the first day of each month beginning in November and ending in August. Thus, the candidate predictors for November include 6 time series (MJJ SOI, JJA SOI, JAS SOI, ASO SOI, ASO PNA, and

ASO PDO) and so on, where the season is abbreviated using the first letter of each of the 3 months forming the average. Climate indices beyond JFM of the runoff year were not entered as candidate predictors.

For temperature and precipitation, different averaging windows, beginning from as early as October and ending with the month immediately prior to the forecast issue date, were tried using data from Lake Louise for the Bow River and from Beaver Mines for the Castle River. By trial and error, a two-month averaging window immediately prior to the forecast issue date was generally found to give better results.

For each sub-basin, the SWEI for a desired date (e.g. April 1) was derived as follows. The SWE data from each snow course station was first expressed as a fraction of the long-term mean SWE at that station. These fractions were averaged to form the sub-basin SWEI. In an attempt to utilize as many data points as possible, all snow course stations with 20 years of data or more were included in sub-basin averaging irrespective of their record starting dates. Thus, the number of sites used to derive the SWEI for each watershed is generally higher for recent years. We opted for SWEI because it slightly improved skill over the raw average SWE. Where available, the SWEI on the first day of the forecast season is used. For forecasts issued after April 1, the April 1 SWEI was entered as a candidate predictor. SWE data are not available for forecasts issued before February 1.

5.3.3. Ensemble Forecast Method

The ensemble forecasting framework used in this study is similar to K-NN methods used in Grantz *et al.* (2005), Regonda *et al* (2005) and Prairie *et al* (2006). The main difference from those methods is the use of a kernel function that is entirely driven by the selected predictors. In addition, we determine the optimum number of nearest neighbors that maximize the ensemble forecast skill. The methodology is described as follows.

1) For the best combination of predictors selected by the GCV algorithm, generate robust M-regression ("conditional mean") forecasts for each data point in a leave-one-out cross-validation mode. Consider a station for which data are available for *N* years. In the leave-one-out cross-validation, the model is fitted using $N-1$ points excluding the data for a single year *t*. The fitted model is then used to generate the streamflow forecast for the holdout year *t*. Compute the model residuals for each forecast and save them.

2) Compute the Mahalanobis distances between the predictors of the forecast year *t* and the predictors of year *j*, where j = 1, ..., $N-1$. The Mahalanobis distance metric is defined as,

$$d_{tj} = [(X_t - X_j)C^{-1}(X_t - X_j)^T]^{1/2} \tag{5.8}$$

where C^{-1} is the inverse of the covariance matrix of the predictors and T stands for the 'transpose' operation. This distance metric has the advantage that it accounts for the redundancy in information content in the case of correlated variables (Wilks, 2006), which is not the case with the Euclidean distance metric. In addition, it does not require pre-standardization of variables.

3) Sort the $N-1$ Mahalanobis distances for forecast year *t* in ascending order and store them as $d_{t(j)}$.

4) Assign weights to the first *K* nearest neighbors of forecast year *t* in step 4 according to one of the following kernels,

$$w_i = \frac{1/i}{\sum_{j=1}^{K} 1/j}; \; K = \sqrt{N} \tag{5.9}$$

$$w_i = \left(1 - \frac{d_{t(i)}}{d_{t(K)}}\right)^{\eta}, \eta \geq 0; \; K = \text{int}\left(\frac{N-1}{\alpha}\right), \quad \alpha \geq 1. \tag{5.10}$$

The kernel in Equation 5.9 was originally proposed by Lall and Sharma (1996) and has been implemented in Grantz *et al.* (2005) and others. Werner *et al.* (2004) used Equation 5.10 to assign probabilities to ensemble traces where the distance metric was taken as the absolute difference between the November-January Nino-3.4 index of the forecast year and all other years. The use of the Mahalanobis distance metric provides a simple way to extend the method to the case of multiple conditioning vectors. In this study, α and η are determined by maximizing the ranked probability skill score (RPSS) using the Simulated Annealing algorithm. The weights w_i are normalized so as to sum up to unity,

$$w_i = \frac{w_i}{\sum_{j=1}^{K} w_j}. \tag{5.11}$$

For each of the K nearest neighbors, compute the cumulative probability metric p_i,

$$p_i = \sum_{j=1}^{i} w_j \ , \ i = 1,...,K. \tag{5.12}$$

5) The probability metric p_i is used to generate ensemble members for forecast year t from the model residuals of the K nearest neighbors as follows. Generate a uniform random number $u \subset (0,1)$. If $u \le p_1$, the residual corresponding to p_1 is selected and added to the conditional mean forecast of year t. If $p_1 < u < p_K$, the residual corresponding to the minimum of the absolute value of the difference between p_i and u is selected and added to the conditional mean forecast for year t. If $u \ge p_K$, the residual corresponding to p_K is selected and added to the conditional mean forecast of year t. Repeat this step as many times as desired to obtain a stable probability density function (PDF). In this study, we generate 100 ensemble members for each forecast year.

6) The ranked probability skill score (RPSS) is used as the objective function to optimize α and η (and hence K) using the simulated annealing algorithm. Steps 4 and 5 are executed with a trial value of α and η, and the RPSS is computed for the 100-member ensemble. The algorithm is repeatedly executed until the RPSS is maximized.

7) For short datasets, repeated residual resampling from the K nearest neighbors (step 5) leads to limited variety in the ensembles and may not model the error distribution properly. Grantz *et al.* (2005) and Singhrattna *et al.* (2005) simulated the perturbation terms by generating random normal deviates using the standard errors of a local polynomial regression model. In this study, we first fit a nonparametric kernel density estimator to the residuals of the K nearest neighbors and then simulate the perturbation terms from this density function as follows. A random variate is sampled from the kernel slice with variance $\lambda^2 s_K$ as

$$\tilde{y}_{t,e} = \hat{y}_t + \lambda S_K{}^{1/2} z_{t,e} \tag{5.13}$$

where $\hat{y}_t$ is the conditional mean forecast for year t, λ is the reference bandwidth of the Gaussian kernel, S_K is the sample variance of the residuals of the K nearest neighbors, $z_{t,e}$ is a random normal variate and $\tilde{y}_{t,e}$ is the ensemble member e for forecast year t. One hundred ensemble members are generated for each forecast year. The Gaussian reference bandwidth for an l-dimensional feature vector is computed as (Silverman, 1986)

$$\lambda = \left(\frac{4}{l+2}\right)^{\left(\frac{l}{l+4}\right)} N^{\left(-\frac{l}{l+4}\right)} \tag{5.14}$$

5.3.4. Forecast Verification Criteria

For the best combination of predictors selected by the GCV algorithm, the model performance was evaluated in a leave-one-out cross-validation mode for the period 1964-2001. Although streamflow data are available for a longer period, the length of data used in the model development is constrained by the availability of SWE data. Retrospective forecasts (hereafter simply called "forecasts") are issued at the beginning of each month starting in November prior to the runoff year and ending in August of the runoff year. For the April-September target season, forecasts begin on November 1 (at 5-month lead) and are updated at the beginning of the month until April 1 (0-month lead). For all other target seasons, forecasts are issued at 0-month lead, e.g. forecasts for the May-September target season are issued on May 1; forecasts for the June-September flow are issued on June 1, and so on.

Forecast performance was evaluated in terms of the correlation coefficient, ranked probability skill score (RPSS), relative operating characteristic (ROC) and potential economic value. For correlation coefficient, the ensembles have to be reduced to a single value (which is the median in our case). Thus, we use the correlation coefficient as a measure of the correspondence between the median forecast and observations. The remaining three criteria directly quantify the skill of the ensemble forecast. To do this, the ensemble forecasts are first converted into forecast probabilities as follows. For each forecast year, the ensemble members were grouped into "below normal", "near-normal" and "above normal" categories based on terciles determined from the distribution of the observed flows. The forecast probability (*f*) in each category is equal to the fraction of ensemble members falling in that category.

The ranked probability score (RPS) for a forecast year is given by (Wilks, 1995)

$$RPS = \sum_{m=1}^{J}\left(\sum_{j=1}^{m} f_j - \sum_{j=1}^{m} o_j\right)^2 \qquad (5.15)$$

where J is the number of categories (here $J = 3$), and o is the observation probability (o equals 1 if the event occurs in the jth category and zero otherwise). The RPSS for a collection of RPS values is computed as (Wilks, 1995)

$$RPSS = 1 - \frac{\overline{RPS}}{\overline{RPS}_{\text{ref}}} \tag{5.16}$$

where "ref" refers to the reference forecast (usually taken as the climatological probability) and the overbar represents averaging of RPS values for each forecast/event pair. For the 3-category forecast, the climatological forecast probability in the j^{th} category is 1/3. Thus, the RPS for the climatological forecast is computed using $f_j = 1/3$ in Equation 5.15. For a perfect forecast with respect to climatology, RPS is equal to zero and RPSS is equal to 1. Forecasts that are worse than the climatology receive negative RPSS values. Hence, positive RPSS values indicate improvement over climatological forecasts.

From management perspective, the forecaster's ability to forewarn the right category of an event (e.g. below normal, above normal) is of paramount importance. The RPSS only measures the overall performance of the ensemble forecast system, not its performance for specific conditions. Here we use the relative operating characteristic (ROC) to assess the performance of the median and ensemble forecasts for the "below normal" and "above normal" categories. Even though ROC has been used to evaluate the performance of meteorological forecasts (Harvey *et al.*, 1992; Mason and Graham, 1999), it has rarely been used for evaluation of hydrologic forecasts. Consider the "below normal" category. A warning is issued if the forecasts indicate that streamflow will be below-normal. The forecasts are converted into a binary format (e.g. "yes" or "no") depending on whether a warning has been issued or not issued. For the median (deterministic) forecast a warning is issued if the forecast anticipates an event to occur in that category. For ensemble (probabilistic) forecasts, a warning can be issued when the forecast probability for an event exceeds a certain cutoff probability (Mason and Graham, 1999). The four possible outcomes are given in Table 5.1. Two of these are correct outcomes: a hit (*h*: a warning was issued for a subsequent event) and a correct rejection (*r*: a warning was not issued for a subsequent nonevent). The remaining two are errors: a false-alarm (*f*: a warning was issued for a subsequent nonevent) and a miss (*m*: a warning was not issued for a subsequent event).

The ROC is a graphical display of the hit rate (the proportion of events for which a warning was provided correctly) against the false-alarm rate (the proportion of nonevents for which a warning was provided incorrectly) (Mason and Graham, 1999). Hit rate (H) and false-alarm rate (F) are computed as follows:

$$H = h / (h + m) \tag{5.17a}$$

$$F = f / (f + r) \tag{5.17b}$$

For probabilistic forecasts, a set of hit rate and false-alarm rate can be computed by using different probability thresholds. A plot of the false-alarm rate against hit rate generates the ROC curve. For a random forecast, hit rate and false-alarm rate are equal and the ROC curve falls on the 45° line. For skillful forecasts, hit rate exceeds false-alarm rate and the ROC curve falls above the 45° line; for a forecast with negative skill false-alarm rate exceeds hit rate the ROC curve falls below the 45° line. The area beneath the curve is often taken as the ROC score.

For an end-user of the forecast system, the ultimate measure of the utility of a forecast is its economic value and other benefits associated with its use in day-to-day decision making (Zhu *et al.*, 2002). Suppose that a hypothetical user of a forecast decides to protect or not protect against the possibility of a potential loss based on the warnings issued by the forecaster. The potential economic value associated with the use of a forecast can be estimated using a simplified decision model called the cost-loss ratio problem (Murphy, 1977; Richardson, 2000; Zhu *et al.*, 2002). If the event occurs and the user is not protected (a miss), the user incurs a loss L. If the user takes protective action against this potential loss (assuming that the protection is completely effective), the user will incur a cost C irrespective of whether the event occurs subsequently (a hit) or not (a false alarm). In the forecast correctly identified the non-occurrence of an event, the user doesn't incur any cost since no action is taken.

The optimal course of action for an end-user is to take protective action whenever the cost of protection is less than the loss incurred by lack of protection. Thus, a meaningful range for the cost-loss ratio (C/L) is 0 to 1. If only climatological information is available, i.e. the event occurs with a relative frequency $\bar{o}$, the optimal course of action is to protect if $C < \bar{o}L$ and not to protect otherwise. The expected expense for the climatological forecast, rescaled by *L*, is given by

$$E_{\text{climatology}} = \min(\frac{C}{L},\ \bar{o}) \tag{5.18}$$

The minimum possible expense is obtained for a perfect forecast, in which case the end-user would protect only on those occasions when the event occurred.

$$E_{perfect} = \bar{o}\frac{C}{L} \tag{5.19}$$

For a user who acts every time a forecast is issued, the expected expense is obtained from the 2×2 expense table (Table 5.1) as follows:

$$E_{forecast} = \frac{h*C + f*C + m*L}{L} = (h+f)\frac{C}{L} + m \tag{5.20}$$

where h, f and m are now expressed as relative frequencies. Noting that $h + m = \bar{o}$ and $f + r = 1 - \bar{o}$, Equation 5.20 can be expressed in terms of hit rate and false-alarm rate as follows:

$$E_{forecast} = F \times \frac{C}{L}(1-\bar{o}) - H \times \bar{o}(1-\frac{C}{L}) + \bar{o} \tag{5.21}$$

The economic value (*V*) relative to the climatology can then be expressed as,

$$V = \frac{E_{\text{forecast}} - E_{\text{climatology}}}{E_{\text{perfect}} - E_{\text{climatology}}} \tag{5.22}$$

For a perfect forecast, V is equal to 1 whereas for the climatological forecast, V is equal to zero. Thus, positive values of V indicate an improvement over the climatological forecast. For perfectly reliable forecasts (where the conditional relative frequency of occurrence of an event is equal to the forecast probability), V has a lower bound of zero (Wilks, 2001). If the forecast is taken at face value, V may be negative, which means that the user would be better off by discarding the forecast and adopt the climatological forecast. This happens because the conditional event relative frequency is seldom equal to the forecast probability due to conditional or unconditional biases in the model. Ideally, this requirement is satisfied by recalibrating the forecast probabilities using data from independent verification periods. Since we do not have independent verification data, the forecasts in our study are taken at face value.

For a deterministic forecast system, the computation of V is straight forward since there is only one set of hit rate and false-alarm rate for a dichotomous forecast. By varying C/L over the allowed range, an economic value curve can be generated. For ensemble forecasts, the hit rates and false-alarm rates corresponding to different warning thresholds are used to generate a set of economic value curves. The envelope of these curves, obtained by choosing the optimal warning threshold for each C/L, is an indicator of the overall value of the ensemble forecast system (Richardson, 2000). Murphy (1977) states that the best decision that minimizes the forecast user's expected expenses for perfectly reliable forecasts is to protect when the probability (p) of an event occurring is $p \geq C/L$. However, for miscalibrated forecasts such a decision criterion may give suboptimal estimates of the economic value. The envelope curves presented in our study are based on the decision threshold p that maximizes the economic value irrespective of whether $p \geq C/L$ or not. Thus, calibration is implicitly built into the evaluation of the optimum economic value.

5.4. Results and Discussion

The predictors selected using the GCV criteria are given in Table 5.2 for all lead times considered in this study. Among climate indices, the PDO seems to provide better information about the spring to summer season streamflow. Thus, the predictor selection corroborates our results on climate diagnostics discussed in section 3. For the Bow River, antecedent conditions other than SWE enter the model only for forecasts issued after June. For the Castle River, the same occurs only for forecasts issued after April. This means that forecasts issued prior to February, when there are not enough observed snow course data, are entirely based on climate indices alone.

Figure 5.2 provides a comparison of observations to median forecast of the standardized April-September streamflow issued on Apr 1 for both the Bow and Castle Rivers. Also shown in Figure 5.2 are the 5th and 95th percentiles of the ensemble forecast. Visually, there appears to be a good agreement between observations and forecasts, particularly for the Bow River. The overall skill of the forecasts as measured by the correlation coefficient at all lead times is shown in Figure 5.3. For the Bow River, the median forecast issued on Nov 1 has a correlation coefficient significant at the 10% level (Figure 5.3a). Median forecasts of the Bow River issued on Dec 1 or later all have correlation coefficient significant at less than the 1% level. The forecast skills for the Castle River are generally not as good as those for the Bow River. For Castle River, only median forecasts issued on Jan 1 or later do have correlation coefficients significant at less than 1% level.

Also shown in Figure 5.3 are the correlation coefficients for median forecasts issued on Feb 1, Mar 1 and Apr 1 using SWE data alone. These correlations indicate that the addition of large-scale climate information leads to better forecast skill for all three lead times for both rivers (the remaining skill measures were not shown since they lead to similar conclusion).

The effect of the choice of kernel on ensemble forecast skill was evaluated by comparing the RPSS based on weights assigned to the nearest neighbors according to Equations 5.9 and 5.10. With the exception of the forecast issued on Nov 1 based on Equation 5.9, the RPSS for the Bow River were all above zero. For the Castle River, the RPSS based on Equation 5.10 were all positive but those based on Equation 5.9 were negative for forecasts issued on Nov 1 and Dec 1. Using Equation 5.10 improves the RPSS for the Bow River by 7% (Apr 1) to 18% (Dec 1). Similarly, the improvement in the RPSS for the Castle River ranges from 19% (Apr 1) to 73% (Jan 1). It is noted that the importance of the choice of kernel decreases with a decrease in the forecast lead time. All subsequent presentations on ensemble forecasts are based on the kernel in Equation 5.10. One disadvantage is that the optimum number of nearest neighbors (*K*) selected for generating ensemble members using the kernel in Equation 5.10 was generally found to vary from one basin to the other as well as from one forecast issue date to another. For the Bow River, *K* varied from 7 (e.g. for Nov 1) to 18 (e.g. for Mar 1) whereas for Castle River, *K* varied from 6 (e.g. for Jan 1) to 8 (e.g. for Apr 1).

In agreement with the median forecast skill, the ensemble forecast of the Bow River issued on Nov 1 has RPSS close to zero, showing negligible improvement over the climatology. The Bow River ensemble forecasts for all remaining forecast issue dates have positive RPSS, indicating a 33% (Dec 1) to 59% (Apr 1) improvement over the climatology (Figure 5.3a). This means that water resources managers can derive some useful information as early as December about the coming spring to summer streamflow of the river from large-scale climate information alone. Since issuance of water supply outlooks in the SSRB begins in February, the inclusion of large-scale climate information can increase the lead time by two months. Ensemble forecasts for Castle River issued prior to Jan 1 show little overall skill with respect to climatology. In terms of RPSS, the Jan 1 ensemble forecast skill is about 22% better than climatology. The maximum RPSS value of 48.6% is obtained for the forecast issued on May 1, after which there appears to be a slight loss of skill (Figure 5.3b).

The areas beneath the ROC curve for median forecasts issued on Dec 1, Apr 1 and Jun 1 are given in Table 5.3. Areas above 0.5 are indicative of the skillfulness of the forecasts for that specific flow category when compared to climatology. Unlike the correlation measure which showed no skill for the Dec 1 median forecast of the Castle River, the ROC shows that there is some useful skill when the forecast is stratified into three categories. Figure 5.4 shows the ROC curves for "below normal" and "above normal" categories for ensemble forecasts issued on Dec 1 and Apr 1 for both the Bow and Castle rivers. The hit rate and false-alarm rate used to generate the curves in Figure 5.4 are based on warnings issued when the forecast probability of an event occurring in a category exceeds 10%, 20%, ..., 90% (for clarity only a few of these cutoff probabilities are shown along the curves). The top right corner of the ROC curve corresponds to perpetual warnings whereas the bottom left corner corresponds to perpetual no warnings. Between these two extremes, warnings are issued at successively increasing cutoff probabilities towards the bottom left corner of the curve. The more the ROC curve bends towards the top left corner, the higher the skill.

For ensemble forecasts of the Bow River issued on Apr 1 (Figure 5.4c), all events were successfully forewarned when as low as 10% of the ensemble members anticipated below-normal conditions but false alarms were also issued 23.1% of the time that conditions were not below normal. The percentage of successful forewarnings drops to 83.3% (66.7%) when 30% (50%) of the ensemble members anticipate below-normal conditions, with the false alarm rate also dropping to 7.7% (3.8%). In contrast, when 30% (50%) of the ensemble members anticipate above-normal conditions, 84.6% (84.6%) of the events were forewarned correctly but false alarms were also issued 16% (12%) of the time that conditions were not above-normal. This shows that the ensemble forecast system is slightly better at simulating dry conditions than wet conditions for this watershed. The Apr 1 forecasts for Castle River (Figure 5.4d) have a similar trend to that of the Bow River except that there are fewer hits and more false-alarms at each cutoff probability. For instance, when 30% (50%) of the ensemble members anticipate below-normal conditions, 58.3% (50.0%) of the events were forewarned correctly but a false alarm was issued 15.4% (7.7%) of the time that conditions were normal or above normal.

As expected, the forecasts issued on Dec 1 for both the Bow (Figure 5.4a) and Castle (Figure 5.4b) rivers generally have a lower rate of correctly anticipating an event in both below-normal and above-normal categories. For instance, for the Bow River, 58.3% (41.7%) of the events were forewarned correctly when 30% (50%) of the ensemble members anticipate below-normal conditions, and with false alarm rates of 15.4% (3.8%). The corresponding figures for above-normal forecasts are correct forewarnings of 76.9% (61.5%) and false alarm rates of 20% (16%). From Figure 5.3b, the overall skill of the Castle River forecasts issued on Dec 1 was not better than the climatology in terms of RPSS. The use of a single skill measure like the RPSS across the whole spectrum of forecasts does not reveal how the forecast system performs in different parts of the spectrum. In fact, Castle River forecasts issued on Dec 1 contain considerable probabilistic information for both above-normal and below-normal conditions (Figure 5.4b). For instance, 58.3% (25.0%) of the events were forewarned correctly when 30% (50%) of the ensemble members anticipate below normal conditions. Similarly, 61.5% (30.8%) of the events were forewarned correctly when 30% (50%) of the ensemble members anticipate above-normal conditions. However, the percentage of times false alarms were issued is higher when compared to forecasts issued on Apr 1.

Also shown in Figures 5.4e and 5.4f are the ROC curves for ensemble forecasts issued on Jun 1. For the Bow River, there is a decrease in ROC skill from that of Apr 1 for both below- and above-normal forecasts. On the other hand, the ROC skill shows an increase between Apr 1 and Jun 1 for the Castle River.

The potential economic value curves for below-normal and above-normal forecasts are presented in Figure 5.5 for median and ensemble forecasts at four forecast issue dates. For the Bow River ensemble forecasts, the economic value generally shows an increasing trend from Dec 1 (Figure 5.5a) to Apr 1 (Figure 5.5c) but Jun 1 (Figure 5.5d) shows slightly less value than that of Apr 1. For the Castle River ensemble forecasts, the value increases from Dec 1 (Figure 5.5e) through Apr 1 (Figure 5.5g) to Jun 1 (Figure 5.5h), thus corroborating the trends in the ROC skill (Figure 5.4).

A perusal of Figure 5.5 also reveals that the ensemble forecasts outperformed single-value forecasts for both below- and above-normal categories, in terms of the overall value and the range of users that can benefit from the respective forecast system. The value curves for ensemble forecasts lie above those of median forecasts in all but one case (Figure 5.5e). In addition, the maximum economic value, which is attained when the *C/L* ratio is equal to $\bar{o}$ (Richardson, 2000), is higher for ensemble forecasts except for the case of the Bow River Apr 1 forecasts for the above-normal category (Figure 5.5d), in which case it is the same for both ensemble and median forecasts.

In general, the range of *C/L* for which the economic value is positive, increases with decrease in the lead time. For longer lead times, the forecast lacks sharpness, i.e., forecasts with extreme probabilities are rarely issued. As a result, only users within a narrow band of *C/L* (e.g. approximately between 0.2 and 0.6 for the Castle River Dec 1 forecasts (Figure 5.5e)) can benefit from the ensemble forecast. As the lead time decreases, the forecast becomes sharper, leading to an increase in the range of *C/L* over which *V* is positive. While the trend is similar for median forecasts, the range of useful *C/L* is narrower. As an example, for the above-normal category of the Castle River Dec 1 forecasts (Figure 5.5e), the median forecast offers virtually no economic benefit while the ensemble forecast offers a benefit for users with *C/L* of 0.2 – 0.6. It is also interesting to note that even in those cases when the benefits attained from both types of forecasts are similar for *C/L* close to the climatological relative frequency (e.g. above-normal forecasts of the Bow River issued on Apr 1 (Figure 5.5c)), the range of useful *C/L* is still wider for ensemble forecasts. For example, users with *C/L* less than 0.1 (greater than 0.8) can not derive any benefits from below-normal (above-normal) forecasts of the Bow River Apr 1 median forecasts whereas ensemble forecasts still offer positive economic value for these *C/L* ratios (Figure 5.5c).

More often than not, there are more than one downstream users of a forecast. Ideally, each user will have a decision criterion based on an economically viable *C/L* for his investment. One advantage of probabilistic forecasts over single-value forecasts is that they naturally provide the basis for such multiple decision criteria (Zhu *et al.*, 2002). Consider the case of two users who own properties in a flood prone area. Consider also that the cost of protection required for flood mitigation is the same for both users but their potential losses are somehow different, resulting in *C/L* of 0.1 and 0.9, respectively. Assuming that both of them have access only to the Bow River Apr 1 median forecasts (Figure 5.5.c), the user with much to lose (*C/L* = 0.1) will benefit from protection but the user with less to lose (*C/L* = 0.9) will opt for the climatological forecast and will never protect. On the other hand, if probability forecasts are accessible to both users, the later will also realize benefits from protection when over 90% of the ensemble members anticipate above-normal conditions. Together, the ROC and economic value curves for ensemble forecasts provide a comprehensive tool for conveying the forecast information, which may be used as a basis for making risk-based management decisions.

The skill assessment presented above may also provide useful information regarding the optimum time for issuance of forecasts that can be used as a basis for water resources planning in the two watersheds considered in the study. For the Bow River, it appears that forecasts issued in the spring (e.g. April 1) are generally more appropriate than those issued later in the flow season. On the other hand, water resources planning for the Castle River can successively be refined with each subsequent forecast until June 1. The two watersheds are typical of flow conditions in the headwater regions of the Bow and Oldman sub-basins but whether these findings will apply to the mainstreams of the two sub-basins is yet to be determined.

5.5. Summary and Conclusions

A modified K-NN resampling approach has been used for ensemble seasonal streamflow forecasting of two rivers in the SSRB of southern Alberta based on large-scale climate information and antecedent conditions. The approach differs from earlier studies in that the weight function used for resampling from the K-NN is entirely driven by the predictor data using the Mahalanobis distance metric. In addition, the optimum number of the K nearest neighbors for each forecast lead time is selected by maximizing a probabilistic skill measure. The mean forecasts used to generate the ensemble members were produced by a robust M-regression model where the subset of predictors was selected by a GCV criterion to avoid over-fitting. In general, SWE and winter season PDO were found to be the most important predictors of the spring to summer streamflow in the SSRB. In the presence of SWE and climate index data, other antecedent conditions (e.g. temperature and precipitation) become insignificant predictors for forecasts issued in winter and spring months.

Forecast skills progressively increased with decrease in lead time until April 1 for the Bow River and until June 1 for the Castle River. Even though snow course data are not available for forecasts issued prior to Feb 1, ensemble forecasts based on large-scale climate information alone possess considerable skill for forecasts issued as early as Dec 1. The ensemble forecasting scheme presented here offers several advantages over the current forecasting practice used for the SSRB. First, our results indicate that forecast lead times can be increased by up to two months since forecasts are currently issued beginning in February. Second, ensemble forecasts increase the range of downstream users that can benefit from the probabilistic information compared to single-value forecasts. Third, ensemble members may be used to produce probabilistic forecasts for user-defined quantiles other than the four percentiles for which forecasts are made currently. In fact, ensemble forecasts provide a venue for running downstream planning models on each ensemble member and generate a PDF of management scenarios, from which end-users can quantify the risks and benefits associated with alternative scenarios.

References

Alberta Environment. (1998). South Saskatchewan River basin historical weekly natural flows – 1912 to 1995. *Main Report.* Alberta Environment Protection, Edmonton, Canada.

Alberta Environment. (2002). South Saskatchewan River Basin – Non-irrigation water use forecasts. Publ. No.: T/656, Edmonton, Canada.

Alberta Environment (2006a). *Water in Alberta.* Alberta Environment website: http://www3.gov.ab.ca/env/water/gwsw/quantity/waterinalberta/index.html (accessed August 2006).

Alberta Environment. (2006b). Approved Water Management Plan for the South Saskatchewan River Basin (Alberta), Publ. No. I/011, Edmonton, Canada (Also available online at http://www.environment.gov.ab.ca/).

Alberta Environment. (2006c). *Water Supply Outlook for Alberta.* Alberta Environment website: http://www3.gov.ab.ca/env/water/ws/WaterSupply/Index.html (accessed August 2006).

Day, G. N. (1985). Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plann. Manage.,* 111, 157-170.

Gan, T. Y., Gobena, A. K., and Wang, Q. (2007). Precipitation of southwestern Canada – Wavelet, scaling, multifractal analysis, and teleconnection to climate anomalies. *J. Geophys. Res.*, 112, D10110, doi:1029/2006JD007157.

Gobena A. K., and Gan, T. Y. (2006a). Low-frequency variability in southwestern Canadian streamflow: links to large-scale climate anomalies. *Int. J. Climatol.,* 26, 1843-1869.

Gobena A. K., and Gan, T. Y. (2006b). Links between pacific climate variability and low frequency hydroclimatic variability in Southern Alberta. *Proc. CSCE Annual Conference*, Canadian Society for Civil Engineering, Abstracts & CD-ROM, Calgary, Alberta, May 23-26, 2006.

Grantz, K., Rajagopalan, B., Clark, M., and Zagona E. (2005). A technique for incorporating large-scale climate information into basin-scale ensemble streamflow forecasts. *Water Resour. Res.,* 41, W10410, doi:10.1029/2004WR003467.

Hsieh, W., and Tang, B. (2001). Interannual variability of accumulated snow in the Columbia basin, British Columbia. *Water Resour. Res.,* 37, 1753-1759.

Hsieh, W., Yuval, Li, J., Shabbar, A., and Smith, S. (2003). Seasonal prediction with error estimation f Columbia River streamflow in British Columbia. *J. Water Resour. Plann. Manage.,* 129(2), 146-149.

Huber, P. J. (1981). *Robust Statistics*, New York: Wiley.

Lall, U., and Sharma, A. (1996). A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resour. Res.,* 32, 679-693.

Loader, C. (1997). LOCFIT: An introduction. *Statistical Computing and Graphics Newsletter,* 8(1), 11-17.

Mason, S. J., and Graham, N. E. (1999). Conditional probabilities, relative operating characteristics, and relative operating levels. *Weather and Forecasting,* 14, 713-725.

Moore, R. D. (1996). Snowpack and runoff responses to climatic variability, southern Coast Mountains, British Columbia. *Northwest Science,* 70, 321–333.

Moore, R. D., and McKendry, I. G. (1996). Spring snowpack anomaly patterns and winter climatic variability, British Columbia, Canada. *Water Resour. Res.,* 32, 623-632.

Murphy, A. H. (1977). The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation. *Mon. Weather Rev.,* 105, 803-816.

Prairie, J. R., Rajagopalan, B., Fulp, T. J., and Zagona, E. A. (2006). Modified K-NN model for stochastic streamflow simulation. *J. Hydrol. Engrg., ASCE,* 11(4), 371-378.

Regonda, S. K., Rajagopalan, B., Lall, U., Clark, M., and Moon, Y.-I. (2005). Local polynomial method for ensemble forecast of time series. *Nonlinear Processes in Geophysics,* 12, 397-406.

Richardson, D. S. (2001). Skill and economic value of the ECMWF ensemble prediction system. *Quart. J. Royal Meteorol. Soc.*, 126, 649-667.

Rousseeuw, P. J., and Leroy, A. M. (1987). *Robust regression and Outlier Detection*, New York: Wiley.

Serneels, S., Croux, C., Filzmoser, P., Van Espen, P. J. (2005). Partial robust M-regression. *Chemometr. Intell. Lab. Syst.,* 79, 55-64.

Shabbar, A., Bonsal, B., Khandekar, M. (1997). Canadian precipitation patterns associated with the Southern Oscillation. *J. Climate,* 10, 3016–3027.

Shabbar, A., Khandekar, M. (1996). The impact of El Nino-Southern Oscillation on the temperature field over Canada. *Atmos.-Ocean,* 34, 401–416.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis.* Monographs on Statistics and Applied Probability. New York: Chapman & Hall.

Singhrattna, N., Rajagopalan, B., Clark, M., and Krishna Kumar, K. (2005). Forecasting Thailand summer monsoon rainfall. *Int. J. Climatol.,* 25, 649-664.

Werner, K., Brandon, D., Clark, M., and Gangopadhyay, S. (2004). Climate index weighting schemes for NWS ESP-based seasonal volume forecasts. *J. Hydrometeorology,* 5, 1076-1090.

Wilks, D. S. (1995). *Statistical Methods in the Atmospheric Sciences: An Introduction.* San Diego: CA, Academic Press.

Wilks, D. S. (2001). A skill score based on economic value for probability forecasts. *Meteorol. Appl.*, 8, 209-219.

Woo, M.-K., and Thorne, R. (2003). Comment on 'Detection of hydrologic trends and variability'. *J. Hydrol.,* 277, 150–160.

Zhu, Y., Toth, Z., Wobus, R., Richardson, D., and Mylne, K. (2002). The economic value of ensemble-based weather forecasts. *Bull. Amer. Meteor. Soc.,* 83, 73-83.

Table 5.1. A 2×2 contingency table for verification of dichotomous forecasts and associated expected cost of protection (*C*) or loss (*L*) for lack thereof. The expected expense for a correct rejection is zero.

		Forecasts	
		Warning	No warning
Observations	Event	*h* *C*	*m* *L*
	Nonevent	*f* *C*	*r* 0

Table 5.2. Streamflow predictors selected using the GCV criteria. The forecast target season is indicated in parenthesis in column 1.

Forecast issue date	Bow River [a]	Castle River [a]
Nov 1 (Apr-Sep)	SOI (JJA)	SOI (JJA)
Dec 1 (Apr-Sep)	PDO (SON, ASO), SOI (JJA)	PDO (SON)
Jan 1 (Apr-Sep)	PDO (OND, ASO), SOI (JJA)	PDO (OND)
Feb 1 (Apr-Sep)	SWEI, PDO (NDJ, ASO)	SWEI, PDO (NDJ)
Mar 1 (Apr-Sep)	SWEI, PDO (NDJ, DJF)	SWEI, PDO (NDJ)
Apr 1 (Apr-Sep)	SWEI, PDO (NDJ, DJF)	SWEI, PDO (NDJ), SOI (OND)
May 1 (May-Sep)	SWEI[b], PDO (NDJ, DJF)	SWEI[b], TMP[c], PDO (NDJ)
Jun 1 (Jun-Sep)	SWEI[b], PDO (OND)	SWEI[1], PCP[c]
Jul1 (Jul-Sep)	SWEI[b], TMP[c], PDO (NDJ)	PDO (DJF), PCP[c], SWEI
Aug 1 (Aug-Sep)	TMP[c], SWEI	TMP[c], PCP[c], SOI (ASO)

[a] For climate indices, the selected predictor is abbreviated using the first letter of each of the 3 months forming the average.

[b] The April 1 SWEI is used for all forecasts issued after April 1.

[c] Precipitation (PCP) and temperature (TMP) are averages of the two months preceding the forecast issue date.

Table 5.3. Area beneath the ROC curve for median forecasts (BN = Below-normal, AN = Above-normal)

Forecast	Bow River			Castle River		
category	Dec 1	Apr 1	Jun 1	Dec 1	Apr 1	Jun 1
BN	0.69	0.81	0.71	0.61	0.67	0.61
AN	0.75	0.86	0.73	0.52	0.71	0.82

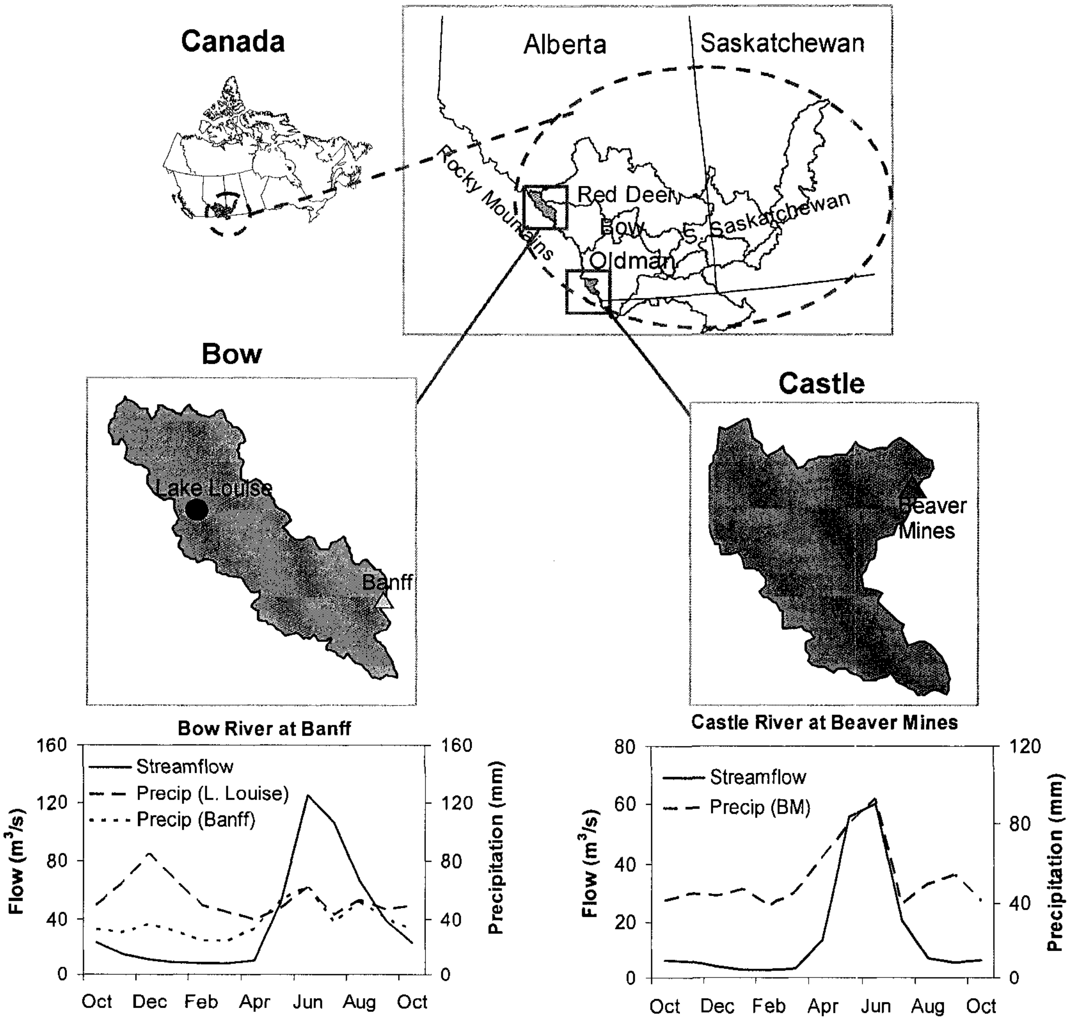


Figure 5.1. Location of the Bow and Castle rivers in the SSRB in southern Alberta. The lower panel shows Bow River (05BB001) annual hydrograph and precipitation climatology (at Lake Louise and Banff), and Castle River (05AA022) annual hydrograph and precipitation climatology (at Beaver Mines).

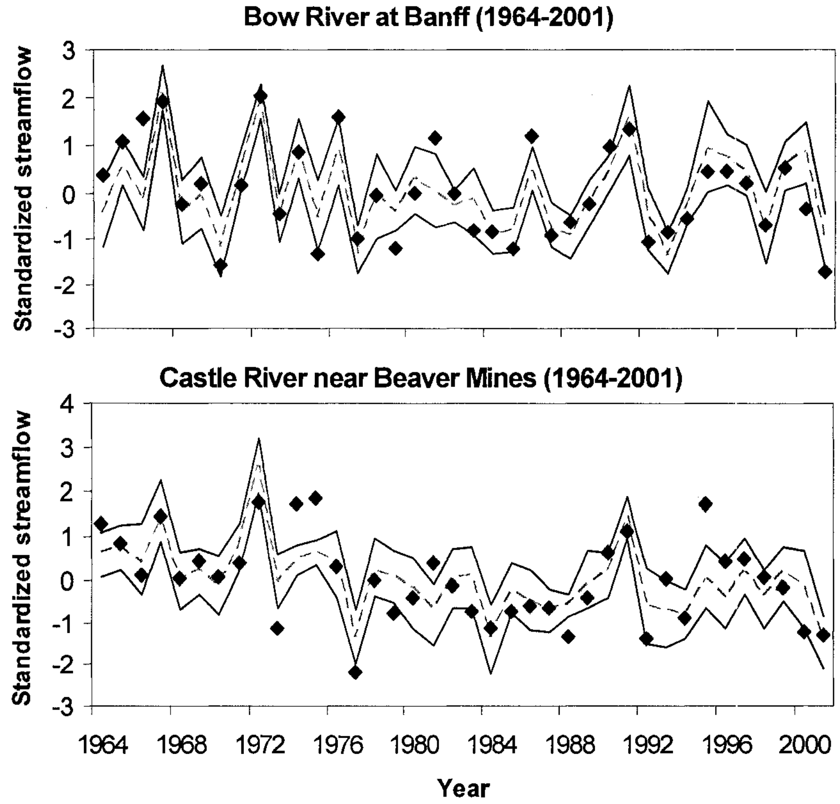
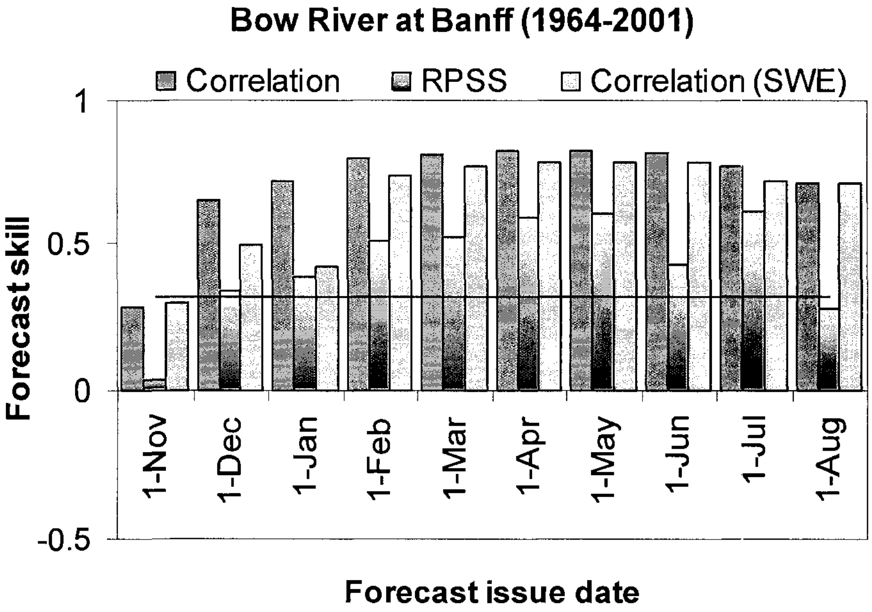


Figure 5.2. A comparison of observed and forecast streamflow for forecasts issued on April 1. The solid diamonds represent historical flow. The median forecast is shown by the broken lines. The 5th and 95th percentiles of the ensembles are shown by the solid lines.



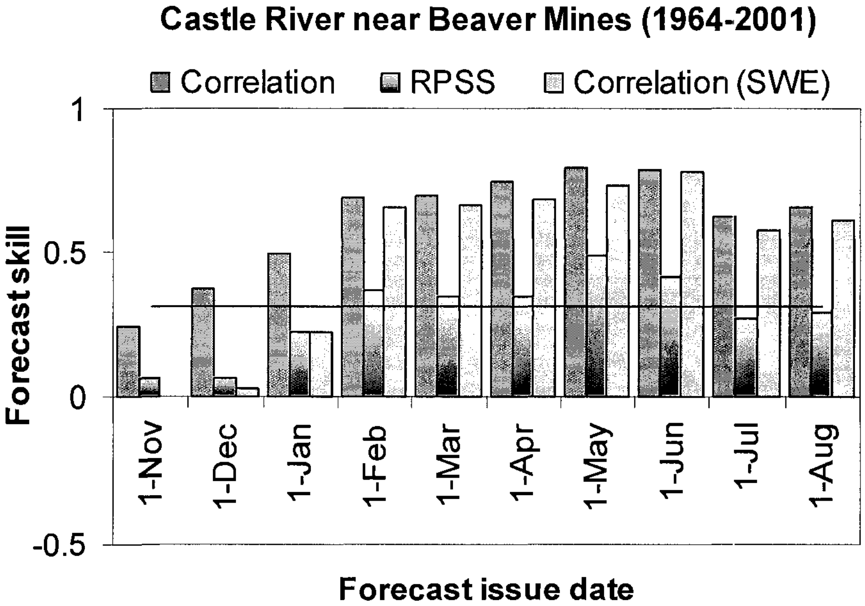


Figure 5.3. Forecast skill measures for (a) Bow River, and (b) Castle River.

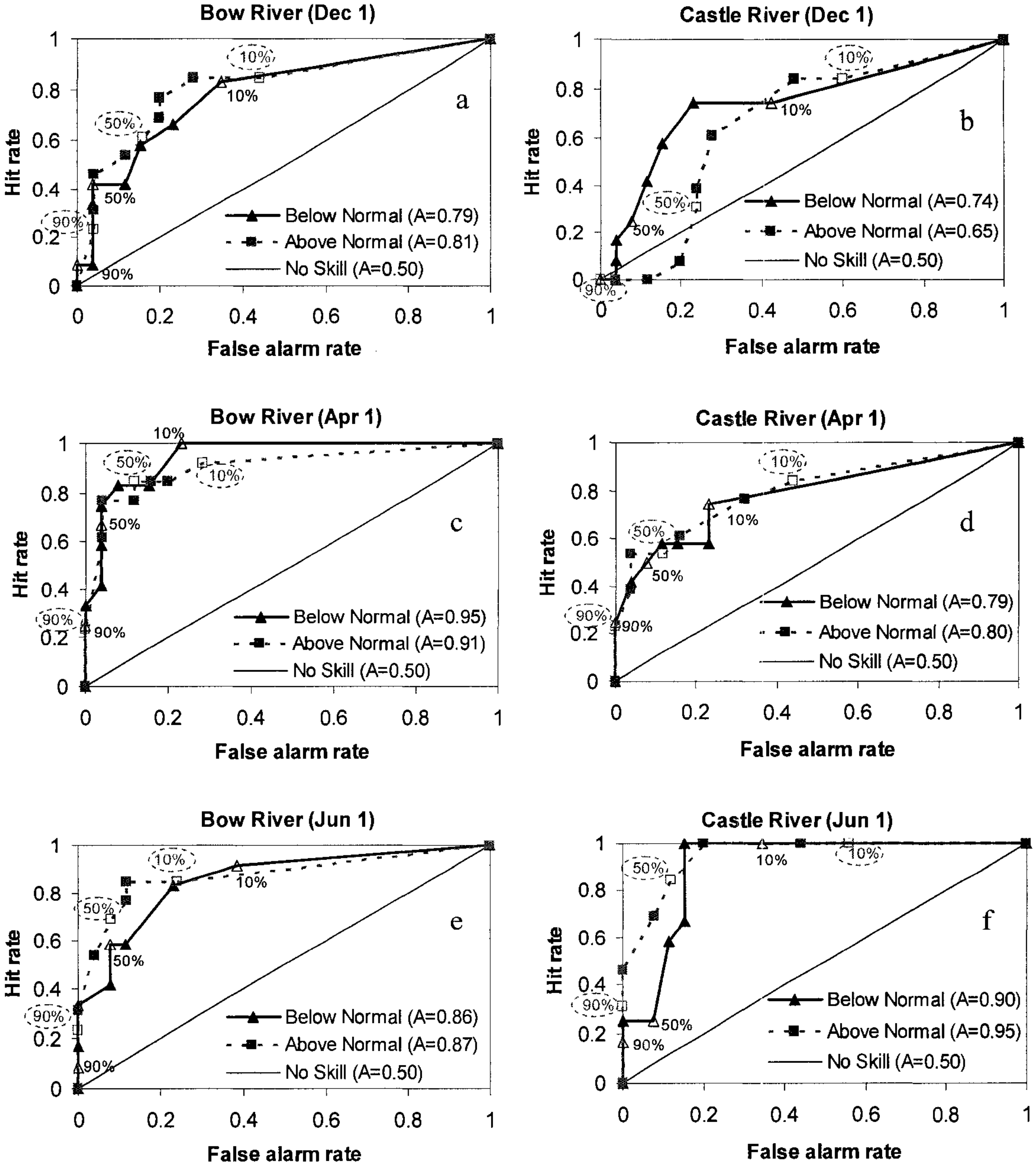


Figure 5.4. ROC curves for ensemble streamflow forecasts of the Bow and Castle rivers issued on December 1, April 1 and June 1. Results are shown for streamflow forecasts in the lower (solid line) and upper (broken line) terciles. The open markers on the curves indicate the 10%, 50% and 90% warning thresholds, which are circled for above-normal forecasts. The areas beneath the curves, *A*, are given the legend.

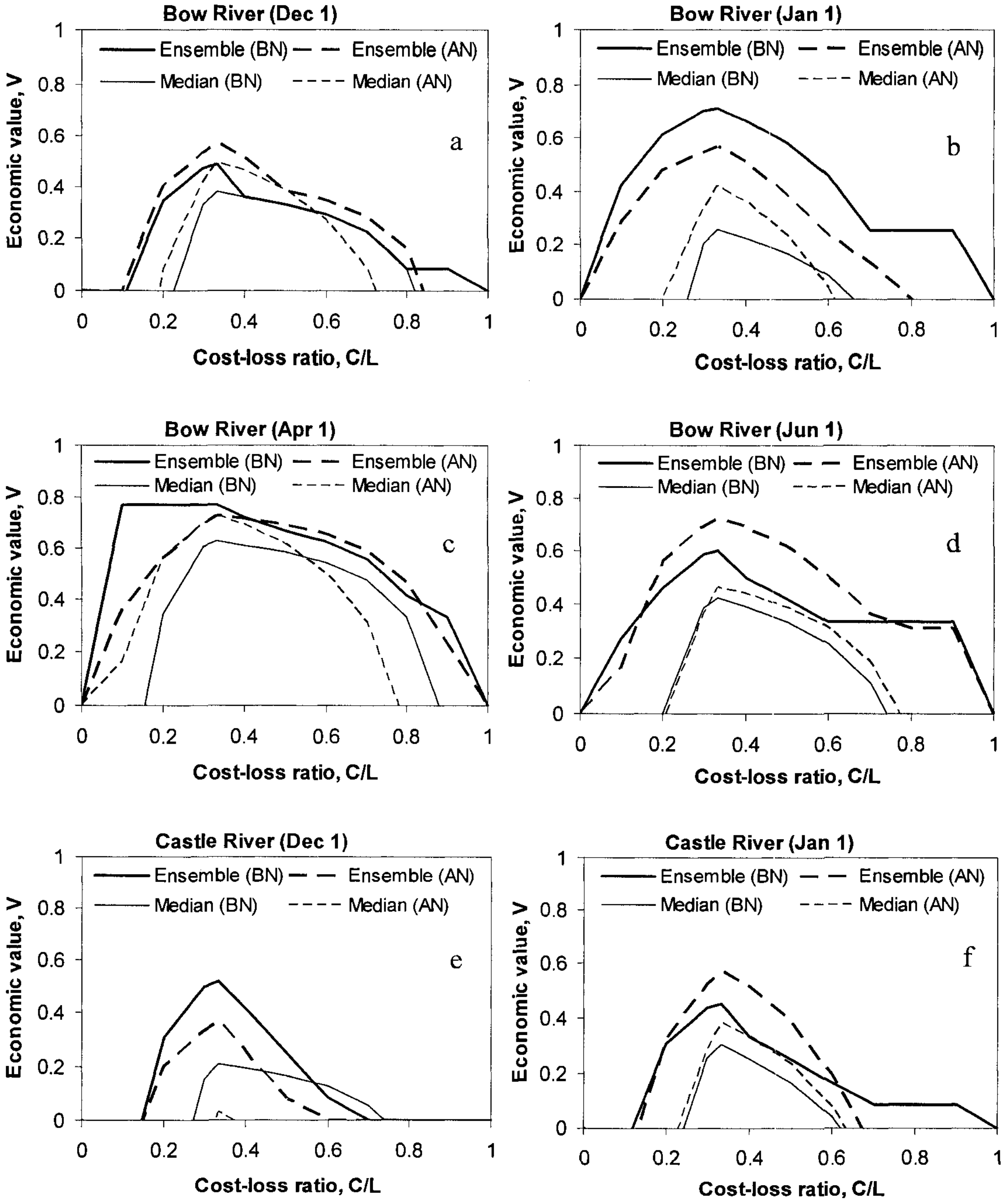


Figure 5.5. Optimal economic values of the deterministic (light lines) and ensemble (heavy lines) forecast systems for below-normal (BN, solid lines) and above-normal (AN, dashed lines) forecasts issued on December 1, January 1, April 1 and June 1 for the Bow River (a to d) and Castle River (e to h). The curves for ensemble forecasts are obtained by choosing the warning threshold that maximizes *V* for each *C/L*. Economic values less than zero are not plotted.

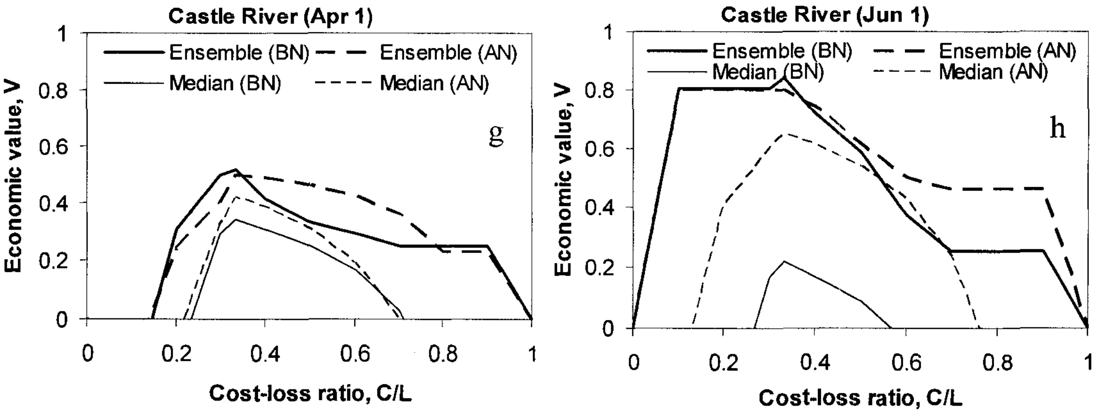


Figure 5.5. *Continued*

Chapter 6 A Technique for Incorporating Seasonal Climate Forecasts into the Ensemble Streamflow Prediction System

6.1. Introduction

Seasonal streamflow forecasts in western Canada are mostly produced by regression techniques that use current and/or antecedent conditions as predictors. For instance, Alberta Environment uses information on antecedent soil moisture, snow course/snow pillow data, precipitation and temperature to determine seasonal runoff volumes for rivers in the SSRB by statistical techniques (Alberta Environment, 2006). In British Columbia (BC), the Columbia River Treaty stipulates that BC Hydro produce seasonal runoff forecasts of the Columbia River inflow to the Mica Reservoir using a linear regression model (Druce, 2001). Recognizing the influence of the El Nino-Southern oscillation (ENSO) and Pacific/North American pattern (PNA) climate anomalies on the interannual variability of snowpack in the Columbia Basin in BC (Hsieh and Tang, 2001), Hsieh *et al.* (2003) used a linear regression model based on local precipitation and indices of ENSO, PNA and the Pacific Decadal Oscillation (PDO) to forecast streamflow of the Columbia River at Donald for the April-August target season.

Regardless of which predictors are included in forecast models, statistical forecasting is often criticized for it ignores basin rainfall-runoff dynamics that are important in controlling runoff generation. Moreover, Day (1985) states that deterministic statistical forecasts do not provide information about the uncertainty of predictions. Day (1985) introduced the ensemble streamflow prediction (ESP) method as a means to objectively incorporate uncertainties into deterministic hydrologic model forecasts. In the climatic ESP method, streamflow forecasts are produced by forcing a hydrologic model with current initial conditions and weather data from past observations. Given *N* years of observed weather data, the method produces *N* possible runoff forecasts (also called 'ensemble traces') for the current forecast year. This means that the ensemble traces are produced by assuming that the weather sequence of each historical year in the record has an equal likelihood of occurrence in the current forecast year (e.g. Day, 1985; Druce, 2001; Franz *et al.*, 2003). The uncertainty about the forecast is then estimated from the probability density function (PDF) constructed from the ensemble traces based on the assumption that each ensemble trace is a sample from the PDF. BC Hydro uses an approach similar to the climatic ESP method to generate forecasts of the Columbia River inflow to Mica Reservoir for the operation planning of non-Treaty storage (Druce, 2001).

There are some limitations in climatic ESP approach. First, for relatively small *N*, the range of variations captured by the PDF may be too small. Secondly, basins with highly variable flow regimes may exhibit flat PDFs as a result of too much dispersion in the ensemble traces (e.g. Franz *et al.*, 2003). In addition, the assumption that each historical year has an equal likelihood of occurrence during the forecast year does not take into account the forecaster's knowledge about the climate system from other sources such as the present state of large-scale climate anomalies and seasonal climate outlooks.

A number of approaches have been proposed to include large-scale climate information into the ESP method. Based on the observed relationships between streamflow and ENSO, Werner *et al.* (2004) examined the performance of different schemes for weighting the climatic ESP flow traces (i.e. post-ESP adjustment schemes). Hamlet and Lettenmaier (1999) restricted ensemble traces to years that are analogous to the forecast year in terms of ENSO and PDO phases. While ENSO-PDO climate categories may lead to improved ensemble spread, statistical problems may arise in the interpretation of the PDF when the number of years in a particular ENSO-PDO category is too small.

Efforts to include seasonal climate outlooks from numerical weather prediction (NWP) models into the ESP system have also been reported in the literature. For instance, in the U.S. National Weather Service (NWS) ESP system an adjustment is made to temperature and precipitation inputs to the hydrologic model based on climate forecasts (Werner *et al.*, 2004). Wood *et al.* (2002) describe a strategy for using monthly ensemble climate forecasts produced by the Global Spectral Model (GSM) of the National Centers for Environmental Prediction (NCEP) as inputs to a macroscale hydrology model for long-range hydrologic forecasting in the eastern U.S. Their strategy involves bias-correcting the forecasts, followed by downscaling to spatial and temporal scales suitable for running the hydrologic model. To disaggregate monthly forecasts into daily time steps, a year was randomly selected from the climatology period and then the daily data of the forecast month in the selected year were adjusted in such a way that the magnitude of the monthly forecast is maintained. Clark and Hay (2004) demonstrated that for forecast lead times of less than 2 weeks, using downscaled forecasts from the NCEP Medium Range Forecast (MRF) outputs significantly improved forecast skill when compared to the climatic ESP approach.

6.2. Research Objective

The primary objective of this study is to develop a technique for incorporating 0-3 month lead ensemble climate forecasts from Canadian NWP models into the ESP system for seasonal streamflow forecasting. Our study differs from that of Wood *et al.* (2002) in two aspects: (1) the raw NWP outputs are used to produce temperature and precipitation forecasts at station locations through downscaling by the model output statistics (MOS) approach; and (2) temporal disaggregation of the MOS-based forecasts from monthly to daily time steps is performed using the K-nearest neighbors (K-NN) resampling algorithm, where a year is selected among years in the climatology period that are closest to the forecast year in terms of the PDO index. Unlike the climatic ESP method, our technique does not suffer from statistical problems for stations with relatively short records since as many different ensemble members as desired can be generated using the K-NN algorithm.

Our study also explores the applicability of input data generated from daily historical observations by the K-NN algorithm (i.e. pre-ESP). In the pre-ESP scheme, daily weather data resampled from historical observations are used to drive the hydrologic model without any forecast guidance. By comparing the performance of the MOS-based scheme to that of the pre-ESP and post-ESP schemes, our study addresses the question of whether the added computational burden of the MOS-based scheme could lead to an improvement in forecast performance over those using historical data alone. The three schemes are applied to two snowmelt dominated sub-basins in the SSRB of southern Alberta. The rest of the paper is organized as follows. Section 6.2 describes the hydrometeorological data for the selected watersheds and the technique used to produce monthly climate forecasts at station locations. The re-sampling algorithms used to generate daily input weather data for the hydrologic model are presented in Section 6.3. This is followed by a brief description, set up and calibration of the hydrologic model in section 6.4. Section 6.5 discusses the criteria used to evaluate forecast performance. The discussion of the results in section 6.6 is followed by a comparison with results obtained from a statistical forecast model in Section 6.7. Finally, the summary and conclusions are presented in section 6.8.

6.3. Data

6.3.1. Hydrometeorological Data

The Bow River at Banff and Castle River near Beaver Mines, both located in the headwaters of the SSRB were selected for this study (Figure 5.1). These watersheds were selected due to low level of development above the respective gauging stations (e.g., minimal regulation effect on streamflow) and the availability of long-term hydrometeorological data. Natural daily streamflow data for both the Bow and Castle rivers were extracted from the Canadian hydrometric data database (HYDAT, 2001). Daily precipitation and mean temperature data were extracted from the Canadian Daily Climate Data CD-ROM of Environment Canada. Snow information used to calibrate the snowmelt component of the hydrologic model was obtained from Alberta Environment (Chacko Abraham, Personal communication, 2004). The earliest available snow course record dates back to 1937 but most of the stations have data only since the 1960's.

6.3.2. Seasonal Climate Forecasts

The Canadian Meteorological Centre produces 0-3 month lead climate forecasts for Canada using two NWP models, namely the Global Environmental Multiscale model (GEM) and the Canadian Centre for Climate Modeling and Analysis (CCCma) second generation atmospheric general circulation model (AGCM2). The operational seasonal forecasts for temperature and precipitation are based on a 12 member ensemble, with 6 members produced by each model. The model climatology comes from a 26 year hindcast (1969-1994) of the seasonal Historical Forecast Project (HFP). The HFP hindcast data are provided on a Gaussian grid of approximately 3.75° lat ×3.75° lon resolution and are generated four times a year for standard seasons only (i.e., DJF, MAM, JJA, SON).

Our study utilizes more comprehensive hindcast data from the Historical Forecast Project 2 (HFP2) obtained from the CCCma website. The HFP2 hindcast data are available on a 128×64 Gaussian grid of approximately 2.81° lat × 2.81° lon resolution. The database contains hindcasts produced by four NWP models, with each model producing 10-member ensembles of 4-month hindcasts for 12 rolling seasons (JFMA, FMAM, MAMJ, etc). The integrations are initialized from the NCEP/NCAR reanalysis spaced at 12-hour intervals prior to the forecast period. This means that the first member is initialized at 12 hours before the forecast period, while the 10th member is initialized at 5 days prior to the forecast period. In our study we utilize a 35 year (1969-2003) hindcast from the GEM and AGCM3 (third generation CCCma climate model). We opted for AGCM3 instead of the operational AGCM2 because the former contains some improved parameterizations (e.g. its land surface scheme) over the latter. Besides being of higher resolution than the HFP hindcasts, the HFP2 dataset is suitable for continually updating forecasts at the beginning of each month.

The temperature and precipitation data to be used for hydrologic modeling in our ESP system are taken from observations at Lake Louise and Beaver Mines stations. Thus, the NWP model grid point nearest to each of the two stations is used to forecast the temperature and precipitation at that station. As a first step, the following bias-correction and downscaling approach was used to translate the temperature and precipitation forecasts from the nearest grid to the station location. Let the precipitation and temperature climatologies for month m at station S be represented by $\overline{P}_{s,m}$ and $\overline{T}_{s,m}$. The bias-correction, to be applied to each NWP model grid separately, proceeds as follows:

1) For each grid g and month m, the NWP model climatologies for forecasts of precipitation ($\overline{P}_{g,m}$) and temperature ($\overline{T}_{g,m}$) are computed from the entire hindcast period.

2) For each ensemble member e, the precipitation forecast at each grid, $P_{g,m,e}$ is expressed as a ratio relative to $\overline{P}_{g,m}$, the model precipitation climatology (i.e. it is scaled) while the temperature forecast, $T_{g,m,e}$ is expressed as an anomaly with respect to $\overline{T}_{g,m}$, the model temperature climatology (i.e. it is shifted):

$$T'_{g,m,e} = T_{g,m,e} - \overline{T}_{g,m} \tag{6.1a}$$

$$P'_{g,m,e} = P_{g,m,e} / \overline{P}_{g,m} \tag{6.1b}$$

3) These anomalies are translated to monthly forecasts at station S through shifting (scaling) the monthly observed temperature (precipitation) climatology at the station as:

$$T_{s,m,e} = \overline{T}_{s,m} + T'_{g,m,e} \tag{6.2a}$$

$$P_{s,m,e} = \overline{P}_{s,m} * P'_{g,m,e} \tag{6.2b}$$

where $T_{s,m,e}$ and $P_{s,m,e}$ are respectively, the monthly average temperature and total precipitation forecasts for ensemble member e. In other words, the station forecasts are produced by replacing the monthly climatology of the grid by that of the station.

The skill of the station forecasts produced by this simple downscaling approach was assessed using the ranked probability skill score (RPSS) in Figure 6.1. As described in Chapter 5, RPSS is a skill measure for multi-category probability forecasts. The results in Figure 6.1 are based on equiprobable terciles (below normal, near normal and above normal categories), with the climatological forecast being used as the reference forecast. Station forecasts from both NWP models mostly exhibited RPSS less than zero, indicating that for these two stations the simple bias-correction and downscaling scheme employed above does not lead to improvement over the climatological forecast. Thus, a more elaborate downscaling approach using the model output statistics (MOS) technique is attempted in the next section to further improve the NWP forecasts before using them in a hydrologic prediction environment.

6.3.3. Downscaling NWP Forecasts Using MOS

Forecasts from NWP models are generally more skillful for upper air variables than for surface weather variables such as precipitation and temperature (e.g. see Clark and Hay, 2004) partly because the NWP model grids are too coarse to resolve small-scale effects such as topography that are important to local weather (Wilks, 2006). In operational weather forecasting, statistical relationships are developed between NWP model outputs and observational data to produce forecasts for variables and/or locations not explicitly represented by the NWP model. One of the most commonly used statistical post-processing techniques is the MOS, in which empirical relationships are developed between observational data (i.e. the predictand) and the NWP forecast outputs (i.e. the predictors) (Wilks, 2006). Development of MOS requires sufficiently long historical data for the predictand, and an archive of NWP forecast outputs for the same period as that of the predictand.

The HFP2 archive contains 35 years of hindcasts for 23 forecast variables for each ensemble member. There are three surface variables (near-surface daily mean air temperature, precipitation flux and surface temperature) and 4 atmospheric variables at 5 pressure levels (each of geopotential height, air temperature, eastward wind and northward wind at the 200-, 500-, 700-, 850- and 1000-hPa levels). In this study, we employ MOS to downscale NWP model forecasts to a station location from the nearest grid point. The objective is to forecast monthly near-surface air temperature and precipitation at the desired station using a subset of predictors selected from the pool of 24 variables including the 1000-850 hPa thickness.

Our study employs a two-step procedure to arrive at the forecasts. First, preliminary forecasts for each lead time and ensemble member are produced by the robust M-regression model described in Chapter 5. To minimize the risk of over-fitting, the subset of predictors is automatically selected among the 24 candidate variables through minimization of the generalized cross-validation (GCV) score (see Equation 5.7).

In cases where the predictor-predictand relationship is weak, the ensemble members may be regressed towards the climatology, leading to shrinking of the ensemble spread, which is not desirable. To avoid this pitfall, the empirical cumulative probabilities of the preliminary forecasts (based on ranks in the forecast time series) are matched with empirical cumulative probabilities of the observed time series. The preliminary forecast is then replaced with the observation with the same cumulative probability. This method also preserves both the mean and variance of the distribution. In practice, observations are not available for the forecast year and thus the distributions for the observations are defined by less than or equal to $N-1$ data points. Hence, a linear interpolation is generally required since the cumulative probability of a preliminary forecast lies between that of two observations. In cases where the cumulative probabilities of the forecasts fall outside the range of the cumulative probabilities of observations, the forecasts are estimated with theoretical distributions fitted to observations. In this study, the two-parameter gamma distribution is used for precipitation whereas the Gaussian distribution is used for temperature. To assess the improvement in skill over the bias-corrected NWP forecasts, the RPSS for the MOS-based forecasts is computed using the bias-corrected NWP forecasts as the reference forecast in Equation 5.16. A perusal of Figure 6.2 reveals that the MOS-based forecasts lead to improvement of up to 40% in terms of RPSS. However, there is no clear increasing or decreasing trend in skill improvement with respect to lead time.

6.4. Generation of Input Weather Data

6.4.1. K-NN re-sampling

Parametric models have traditionally been employed as a convenient method for the generation of synthetic weather sequences for assessing risk in water resources planning and management. Parametric models require a priori assumption of the underlying PDF of the variable to be modeled and as such the accuracy of the method depends on how well the sample statistics characterize the assumed PDF. Moreover, there is often a need to fit separate models for each season and thus the technique needs too many parameters to be specified. The use of seasonal models also means that proper representation of the low-frequency components of the simulated data becomes a problem. Recently, interest has shifted towards nonparametric approaches as a better alternative for synthetic weather generation. A nonparametric approach that has found wide application in hydrology is the nearest neighbor resampling algorithm in which data are simulated by bootstrap resampling from historical observations. The method makes no assumption about the PDF of the variables, can reproduce any arbitrary function, and can easily be extended to condition the simulation on exogenous variables (e.g. ENSO, PDO, etc) (Yates *et al.*, 2003).

In the K-NN re-sampling algorithm, first the K-NN of the point (e.g. day, year, etc) of interest are selected from historical observations and then one of these neighbors is re-sampled (with replacement) via a kernel function with the property that the nearest neighbors receive the largest weight while the furthest ones receive the least weight. Originally developed by Lall and Sharma (1996) for the simulation of monthly streamflow with dependence structure, the K-NN re-sampling technique has in recent years received wide application in stochastic simulation and downscaling of geophysical time series, and forecasting. Rajagopalan and Lall (1999) used a multivariate K-NN re-sampling scheme with a lag-1 dependence for the simulation of daily weather data for six variables. Gangopadhyay *et al* (2005) used a similar model for downscaling of the National Centers for Environmental Prediction (NCEP) medium range forecast model outputs for four basins in the USA. Yates *et al.* (2003) used a modified K-NN method to simulate regional daily weather sequences that are suitable for assessing a variety of climate change scenarios. Their modification involves an algorithm to introduce bias to the re-samples in order to simulate hypothetical climate scenarios, e.g. drier-warmer winters, wetter-colder springs and so on. The K-NN method is essentially a bootstrap technique and so values not seen in historical data can not be simulated. To overcome this limitation, Sharif and Burn (2006) improved up on the algorithm of Yates *et al.* (2003) by introducing a perturbation term based on the statistics of the nearest neighbors.

The use of the K-NN algorithm to date in the forecasting literature has mainly focused on generating ensembles from single-value forecasts produced by statistical models. Singhrattna *et al.* (2005) used a K-NN re-sampling scheme to generate ensemble forecasts of the Thailand summer monsoon from locally weighted polynomial (LWP) regression estimates and their residuals. Their method involves re-sampling one of the residuals of the regression model fit and adding it to the LWP estimate to generate an ensemble member. Grantz *et al.* (2005) used the same approach for ensemble streamflow forecasting while Prairie *et al.* (2006) used a lag-1 LWP regression for stochastic simulation of streamflow.

6.4.2. Selection of Conditioning Vectors

In regions where the climate is significantly affected by large-scale circulation patterns such as ENSO, PDO, etc, the selection of the K-NN should be conditioned on relevant climate indices. Previous studies have shown that the Pacific Decadal Oscillation (PDO) exerts more influence on the streamflow variability of the SSRB than the El Nino-Southern Oscillation (ENSO) and the Pacific/North America (PNA) climate anomalies (e.g., Gobena and Gan, 2006). In fact, as shown in Chapter 4, partial correlations between the April-September runoff at 13 hydrometric stations in the SSRB and the November-March Nino3|PDO, were found to be not statistically significant at the 5% level.

Figure 6.3 shows lagged Spearman rank correlations between the April-September runoff and rolling 3-month averages of selected climate indices for the two selected basins, based on the period 1964-2001. Correlations are shown for the year of the runoff (e.g., JFM) and the previous year (e.g., JFM(-1)). The 5% and 1% significance levels based on the standard significance test for correlation are indicated by the thick and thin horizontal lines, respectively. Correlations with PDO are consistently above the 1% significance level during SON(-1) to JFM for the Bow River and during SON(-1) to MAM for the Castle River. Thus, the average of the PDO index beginning with November and ending with the month immediately prior to the forecast issue date (e.g. ND for a January 1, NDJ for February 1, NDJF for March 1, and NDJFM for April 1 forecast dates) is used for determining neighborhoods.

6.4.3. Synthesis of Daily Input Data from MOS-based Forecasts

Consider a streamflow target season of April to September inclusive. For this target, a forecast issued at the beginning of January involves generation of nine months of input data for running the hydrologic model. Forecasts of monthly temperature and precipitation are available from the MOS procedure for 0-3 month lead (i.e. for January, February, March and April). Our intention is to assimilate these forecasts for the first four months of the hydrologic simulation period, and use resampled historical data for the remaining part of the simulation period. Suppose that historical daily weather data are available for N years. First, the input data for each month of the simulation period (January to September) are populated using a block K-NN resampling from $K \in N$ neighbors of the forecast year, where the PDO index is used as the criterion for selection of neighborhoods. Then, the daily data for each of the first four months are scaled so that their monthly average temperature (total precipitation) is equal to the MOS-based monthly forecast values. In our case, K includes all N-1 years excluding the forecast year. The complete algorithm is as follows:

1) One of the K years is resampled with replacement according to the following weight function proposed by Lall and Sharma (1996):

$$w_j = \frac{1/j}{\sum_{i=1}^{K} 1/i} \qquad (6.3)$$

This function assigns more weight to the closest neighbor and least weight to the furthest neighbor. The weights are first converted to cumulative probabilities p_j:

$$p_j = \sum_{i=1}^{j} w_i, \qquad j = 1, \ldots K \qquad (6.4)$$

Then a uniform random number $u \subset (0,1)$ is generated and compared to p_j. If $u \leq p_1$, the year corresponding to p_1 is selected. Let us refer to the selected year as the "base" year. If $p_1 < u < p_K$, the year which minimizes the absolute value of the difference between p_j and u is selected. If $u \geq p_K$, the year corresponding to p_K is selected.

2) Let the hydrologic simulation period start with January of the forecast year. The daily data of the January of the selected "base" year are used to populate the data for January of the forecast year.

3) Steps 1 and 2 are executed for each month in the hydrologic simulation period (e.g., January to September).

4) For each of the first four months (e.g., January to April), the populated daily temperature (precipitation) values are shifted (scaled) so that the monthly average temperature (total precipitation) is equal to the MOS-based forecast monthly average temperature (total precipitation). For the remaining part of the simulation period (e.g., May to September), the daily data synthesized in steps 1-3 are adopted without modification.

Since the MOS-based forecasts contain 10-member ensembles, a single execution of steps 1 through 4 produces only 10 streamflow traces, which may be too small to obtain a stable PDF. In order to circumvent this problem, steps 1 though 4 were repeated 5 times for each member of the climate forecast, thus producing a total of 50 streamflow traces from the 10-member climate forecasts.

6.4.4. Synthesis of Daily Input Data from Historical Observations (Pre-ESP Scheme)

Synthesizing daily data in the Pre-ESP scheme involves applying the K-NN algorithm twice – first to select years that are similar to the forecast year in terms of the conditioning vector(s), and then to obtain days that are similar to the day of interest in terms of the weather variables of interest. A "base" year for a given month is selected following the procedure explained in Section 6.4.3. The rest of the algorithm is similar to the one used in Sharif and Burn (2006) and is described below:

1) Let the weather generation start on day i of the forecast year. All days of the $K \in N$ years within a temporal window of width w and centered on day i are potential candidates for the weather of day i. For instance, if i is 1 January, K is 34 and w is 14, all days within the temporal window of 25 December to 8 January (i.e. $K \times (w+1) - 1 = 509$ days) are potential candidates for the weather of 1 January.

2) The weather for day i of the forecast year is initialized as the weather of the same day of the "base" year. The algorithm proceeds as follows to generate the weather for day i+1. First, Mahalanobis distances between the initial weather of day i and the remaining $(w+1) \times K - 1$ potential candidates are computed using Equation 5.8. The distances are sorted in ascending order and the first $K' = \sqrt{(w+1) \times K - 1}$ nearest neighbors are retained using the heuristic rule proposed by Lall and Sharma (1996). In our case, this produces K' = 23 days. Then Equations 6.3 and 6.4 are used to assign weights and cumulative probabilities to each of the K' neighbors and also to select one of these neighbors. Once the neighbor is selected, the weather for day i+1 is adopted as the observed weather for the day subsequent to the selected neighbor.

3) The re-sampling procedure described above produces a *sequence* of weather data mostly not seen in the historical record even though the *values* are not new. However, the procedure may also produce an exact replica of the weather sequence of the reference year under rare circumstances. Another limitation is that values not seen in the historical record cannot be simulated. In order to overcome these limitations, the re-sampled data are perturbed by innovations derived from the statistics of the neighborhood as follows (Sharif and Burn, 2006). First, a

nonparametric kernel density distribution is fitted to the *K'* nearest neighbors of day *i*. Then, a random variate is sampled from the kernel slice with variance $\lambda^2 S_{K'}$ as

$$\tilde{y}_{i+1} = x_{i+1} + \lambda S_{K'}^{1/2} z_{i+1} \tag{6.5}$$

where x_{i+1} is the re-sampled weather for day i+1 (from step 3), λ is the reference bandwidth of the Gaussian kernel (Equation 5.14), $S_{K'}$ is the sample variance of the *K'* nearest neighbors of day i, z_{i+1} is a random normal variate and $\tilde{y}_{i+1}$ is the final weather for day $i+1$.

4) Since z_{t+1} is unbounded, Equation 6.5 may produce a negative (and hence unacceptable) value when x_{t+1} lies on or close to the zero boundary. Sharma and Lall (1997) propose generating a new z_{i+1} until $\tilde{y}_{i+1}$ becomes positive. However, such a procedure may not reproduce the statistics of wet and dry days properly. Sharma and O'Neill (2002) used a "variable kernel" to minimize the effect of this problem. The idea is to reduce the bandwidth of the kernel depending on the distance of x_{i+1} from the zero boundary. Let the threshold probability for the precipitation being less than or equal to zero be α. The largest value of λ corresponding to generating a negative value with a probability of exactly α is given by $\lambda_\alpha = x_{i+1} / z_\alpha S_{K'}^{1/2}$ where z_α is the normal variate corresponding to α. Following Sharma and O'Neill (2002), we use $\alpha = 0.06$ ($z_\alpha = -1.55$). If the value of λ is greater than λ_α, then λ_α is used in Equation 6.5. Note that if x_{i+1} is zero, λ_α will also be zero and thus a dry day will be simulated for day $i+1$. If the precipitation generated for day i+1 is still negative, a new value of z_{i+1} is generated until the precipitation becomes positive.

Steps 1 through 4 are repeated for each day of the forecast year to generate a single sequence of input data to the hydrologic model. The algorithm is then reinitialized with the selection of a new "base" year to generate another sequence of inputs. In our case, input sequences were generated for 50 members. Since data are resampled at daily time steps, the pre-ESP scheme is convenient for generating a sufficient number of ensemble members to get a stable PDF even from stations with relatively short records.

6.4.5. The Post-ESP Scheme

The post-ESP scheme involves some form of post-model adjustment to the flow traces produced by the climatic ESP method. In this study, this adjustment was effected by using the K-NN resampling algorithm where the final ensemble members were assembled by resampling from years closest to the forecast year in terms of the PDO index. In order to avoid biases due to differences in ensemble size among the three schemes considered in the study, it was necessary to have the same number of ensemble traces for each method. Since we only have 35 years of data (i.e. 34 traces for each forecast year if we use the climatic ESP system), resampling with replacement was used to produce 50 ensemble members from the 34 traces for each forecast year. This means that some of the flow traces appear more than once in the ensemble members.

6.5. Setup and Calibration of the Hydrologic Model

In this study, we use the Sacramento Soil Moisture Accounting (SAC-SMA) model for simulating daily runoff, mainly because only precipitation and temperature data are available and partly because SAC-SMA is one of the most studied conceptual rainfall-runoff models (CRR) in the past several decades. SAC-SMA is a deterministic, lumped parameter, CRR model developed by Burnash *et al.* (1973). The model requires mean areal precipitation (MAP) and potential evapotranspiration (PET) inputs at daily or sub-daily time step. As a lumped conceptual model, SAC-SMA only accounts for water fluxes for evaporation, runoff and soil moisture in conceptual storages of "free moisture" and "tension moisture" types in two soil layers (Gan and Burges, 1990).

The soil moisture accounting framework the SAC-SMA model is shown in Figure 6.4. Rain falling on an impervious area produces direct runoff whereas rain falling on pervious areas goes on to fill up the tension water storage in the upper zone (UZTWM). Water in excess of the upper zone tension storage requirement fills up the upper zone free water storage (UZFWM), which in turn will feed lateral inflow (through UZK) or vertical percolation to the lower zone (through PERC and REXP). When the rainfall rate is greater than the sum of upper zone free water storage capacity, lateral inflow and vertical percolation, excess surface runoff is generated from the upper layer. Water reaching the lower zone first satisfies the lower zone tension requirement (LZTWM) and then fills up the primary (LZFPM) and secondary (LZFSM) free water storage reservoirs. The primary and secondary free storage reservoirs generate baseflow, the withdrawal rates of which respectively depend on the lower zone recession parameters LZPK and LZSK. Evapotranspiration extracts water from both types of storage in the upper and lower zones. The parameters controlling the partition of moisture between the various conceptual storages and fluxes must be calibrated to reproduce observed hydrographs.

The SMA process described above assumes that precipitation occurs in liquid form, which will end up in one of the conceptual storages or fluxes without much delay. For snow covered basins, precipitation is stored in the basin in the form of snow for an extended period of time. Thus, the input to the SMA process requires estimation of liquid precipitation using a snow accumulation and ablation model. In this study, the SNOW17 model is used to handle the snowmelt processes. The SNOW17 model is also a lumped conceptual model that uses air temperature as an index to estimate energy exchange across the air-snow interface (Anderson, 1973). For our purpose, SNOW17 was modified to handle snow accumulation and ablation in a semi-distributed mode, where the watershed is subdivided into a desired number of elevation bands (3 in our case). The primary input data to SNOW17 are raw daily precipitation and daily mean temperature.

A simple water balance calculation for the Bow River at Banff shows that precipitation recorded at both Banff and Lake Louise underestimates the sub-basin precipitation. Thus, a correction for elevation bias was necessary for proper representation of the precipitation distribution across the watershed. A comparison of long-term monthly mean precipitation at Banff and Lake Louise shows that most of the difference in precipitation between the two stations occurs during the cold season (Figure 5.1, inset), suggesting that orographic effects are more enhanced during the cold season than during the warm season. However, preliminary analysis showed that using precipitation gradients (i.e. of precipitation with respect to a change in elevation) derived from these two stations (i.e. 1.35mm/m) would lead to unreasonably high precipitation at higher elevations. Initial estimates of monthly precipitation gradients used in this study were derived from monthly precipitations at Banff (elevation 1397m) and two mountain stations (Skoki and Sunshine Village). The later two stations are located at elevations above 2000 m but data at these stations are available only for the period 1997/98 to 2002. Initial estimates of monthly precipitation gradients for the Castle River watershed were derived from precipitation observations at Beaver Mines (elevation 1286m), Pincher Creek (elevation 1145m) and Coleman (elevation 1341m) stations. A correction value for each elevation band, which is established through calibration against SWE, was then applied in the SNOW17 model.

Area-elevation curves derived from a 9-arc seconds digital elevation model (DEM) are shown in Figure 6.5. For Bow River, the precipitation and temperature observed at Lake Louise and monthly precipitation gradients were used to estimate the SNOW17 input data at the hypsometric elevation of each band. Likewise, the SNOW17 input data at the hypsometric elevations of the Castle River were estimated from observed precipitation and temperature at Beaver Mines and the respective monthly precipitation gradients. The area-weighted rain-plus-melt outputs from the SNOW17 model constitute the MAP input to SAC-SMA.

The daily PET input to SAC-SMA was estimated using the Hamon PET model. Oudin *et al.* (2005) demonstrated that for CRR applications, simple temperature- and radiation-based models such as the Hamon and Thornthwaite models are more efficient than PET models based on the energy balance approach. The Hamon model requires daily temperature and length of day (which is a function of latitude, slope and aspect) (Hamon, 1961; Oudin *et al*, 2005). PET was computed for each elevation band and then area-weighted to produce the input to SAC-SMA.

SAC-SMA was calibrated with daily weather data for 1981-1995 and validated for 1966-1980. For Bow River, the coefficient of determination (R^2) and the Nash-Sutcliffe efficiency (E_f) for the calibration and validation periods are $R^2 = 0.944$, $E_f = 0.889$ and $R^2 = 0.933$, $E_f = 0.868$, respectively. For Castle River, $R^2 = 0.898$, $E_f = 0.801$ for the calibration period and $R^2 = 0.877$, $E_f = 0.768$ for the validation period. Figure 6.6 shows the observed and simulated monthly hydrographs of the Bow and Castle rivers for the validation period. With comparable calibration and validation results which both demonstrate good agreements with the observed data (high R^2 and E_f), we have the basis to use the calibrated SAC-SMA to generate ensemble runoff forecasts, which will be used for assessing the forecast skill in Section 6.7.

6.6. Forecast Verification Criteria

The forecast performance was evaluated using hindcasts produced by a leave-one-out cross-validation approach. In this approach, the data for a retrospective forecast year are excluded from the database that constitutes the candidate weather data for that forecast year. For the April-September target season, forecasts begin on January 1 (3-month lead) and are updated on February 1 (2-month lead), March 1 (1-month lead) and April 1 (0-month lead). For all other target seasons, forecasts are issued at 0-month lead. For instance, forecasts for the

May-September target season are issued on May 1 whereas forecasts for the June-September target season are issued on June 1, and so on.

Four criteria were used to provide a quantitative assessment of the forecast skills, namely the correlation coefficient, root mean square error (RMSE), ranked probability skill score (RPSS) and relative operative characteristic (ROC) diagram. The RMSE (expressed as a percentage relative to the long term observed mean) is obtained as,

$$RMSE = \left(\left(\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \right)^{1/2} \right) \frac{1}{\bar{y}} \times 100 \tag{6.6}$$

where y is the observed flow, $\bar{y}$ is the mean of the observed flow, and $\hat{y}$ is the median forecast. The RMSE expressed in this form allows direct comparison among forecasts issued on different dates as well as across basins. For correlation and RMSE analysis, the ensemble forecasts must be reduced to deterministic forecasts. In this study, the median of the ensembles is used to compute these skill measures. RPSS and ROC are probabilistic skill measures and are suitable for assessing ensemble forecast performance at different levels of probability. The reader is referred to Chapter 5 for the description of the other skill measures.

6.7. Discussion of Results

To evaluate the forecast skills of the MOS-based scheme for seasonal streamflow forecasting, 50-member ensembles of hindcasts were produced for each year in the period of 1969-2003. These hindcasts are simply referred to as forecasts in subsequent presentation. The skills of median and ensemble forecasts based on the two NWP models (AGCM3 and GEM) are compared to the pre-ESP and post-ESP schemes (the later two schemes use no forecast guidance from NWP models). The skills of the median forecast as measured by the correlation coefficient and RMSE are shown in Figure 6.7. For the Bow River, MOS-based median forecasts based on either AGCM3 or GEM showed better correlation than the other schemes in three of the eight forecast issue dates. For the Castle River, MOS-based median forecasts based on both AGCM3 and GEM have better correlations than the other schemes for all eight forecast issue dates. The correlation coefficient shows an increasing trend towards summer, with values peaking for forecasts issued on June 1 for Castle River and on July 1 for Bow River. On the other hand, the lowest RMSE (as % of the mean) occurs for forecasts issued on April 1 for both rivers.

Also shown in Figure 6.7 are the RPSS for ensemble forecasts. For the Bow River, forecasts based on AGCM3 and GEM produced positive RPSS for all issue dates, indicating improvement over climatological forecasts. In four of the eight issue dates, the RPSS for forecasts based on either AGCM3 or GEM is better than the other schemes. For the Castle River, all schemes produced negative RPSS for forecasts issued on July 1 and August 1, indicating that these forecasts would be worse than one that is based on climatology alone. While forecasts for the remaining issue dates offer positive RPSS values, the improvement is marginal, with RPSS peaking at 32.5%. Only in two out of the eight forecast issue dates is the RPSS for forecasts based on either AGCM3 or GEM better than the other schemes.

Figures 6.8 and 6.9 show ROC diagrams for ensemble forecasts of the Bow and Castle rivers. These curves are for "below normal" and "above normal" categories for forecasts issued on January 1, April 1 and June 1. The hit rate and false-alarm rate used to generate the curves in Figures 6.8 and 6.9 are based on warnings issued when the forecast probability of an event occurring in a category exceeds 10%, 20%, …, 90%. Warnings are issued at successively increasing cutoff probabilities from the top right to the bottom left corner of the curve. Near the top right corner of the curve, warnings are issued more frequently and hence the hit rate increases but at the expense of increased false-alarm rate. Forecasts with high skill have larger hit rates and smaller false-alarm rates and thus the curve bends towards the top left corner. Curves falling below the diagonal line indicate that the forecast is worse than a climatological forecast.

Consider the ROC curves for the below-normal forecast category of the Bow River (Figure 6.8). For ensemble forecasts issued on January 1 (Figure 6.8a-d), all schemes show marginal skill up to an exceedance probability of 90%. Above this probability, the ROC curves for AGCM3 (Figure 6.8a) and Pre-ESP (Figure 6.8c) coincide with the no-skill line. For forecasts issued on April 1 (Figure 6.8e-h) and June 1 (Figure 6.8i-l), the ROC curves lie above the diagonal line, indicating improvement over climatology. For all schemes, there is a marked increase in the skill of the below-normal forecast category between January 1 and April 1 but the change in skill between April 1 and June 1 is minimal. For the above-normal forecast category, the ROC curves for all three issue dates lie above the diagonal line (Figure 6.8a-l). Above-normal flows appear to be forecast more successfully than below-normal flows, particularly for the January 1 (Figure 6.8a-d) and June 1 (Figure 6.8i-l) issue dates.

For the Castle River, the ROC curves for below-normal forecasts issued on January 1 (Figure 6.9a-d) exhibit marginal skill up to an exceedance probability of 40%. Beyond this threshold, forecasts based on the Pre-ESP (Figure 6.9c) show no skill at all whereas GEM (Figure 6.9b) looses skill for exceedance probabilities above 60%. For below-normal forecasts issued on April 1 (Figure 6.9e-h), GEM has the best performance with a hit rate of 27% at an exccedance probability of 90% (Figure 6.9f). At the higher end of the warning threshold, forecasts issued on June 1 (Figure 6.9i-l) are only marginally better than those issued on April 1. For the above-normal forecast category, forecasts issued on January 1 (Figure 6.9a-d) possess marginal skill across all exceedance probabilities but they also exhibit a much more variable skill among the various schemes. Compared to below-normal forecasts, above-normal forecasts issued on April 1 (Figure 6.9e-h) are much more skillful at higher warning thresholds, with a hit rate greater than 40% at an exceedance probability of 90%. Above-normal forecasts issued on June 1 (Figure

6.9i-l) exhibit a similar characteristic to those issued on April 1 for higher exceedance probabilities.

Figures 6.10 and 6.11 show plots of the hit rates and false-alarm rates against forecast issue dates at four exceedance probabilities (10%, 30%, 70% and 90%) for the below-normal and above-normal forecasts of the Bow River. These plots are helpful to track the evolution of the ensemble forecast skill for extreme flows as a function of forecast season and the forecaster's confidence. For below-normal forecasts, the hit rates at the 10% warning threshold remain above 70% for all forecast issue dates (Figure 6.10a). However, as the forecaster's confidence increases (i.e. as the warning threshold is increased), a reduction in the hit rate across the forecast period is observed. This is not surprising because as the forecaster's confidence increases, beating the climatology becomes more and more difficult. In addition, a trend emerges where forecasts issued at later dates become more accurate than those issued at earlier dates (Figure 6.10b-d). At higher warning thresholds (e.g. Figure 6.10d), the hit rates for below-normal forecasts peak out in the summer months. The hit rates for above-normal forecasts generally show similar characteristics to that of below-normal forecasts (Figure 6.10e-h).

For below-normal forecasts of the Bow River, the false-alarm rates at the 10% warning threshold show a minimum between April 1 and June 1 forecasts (Figure 6.11a). As the warning threshold is increased, the forecast with the minimum false-alarm rate moves more and more towards forecasts issued in the spring (Figure 6.11b-d). For above-normal forecasts, the minimum false-alarm rate consistently occurred for forecasts issued in the summer (Figure 6.11e-h). For above-normal forecasts, the peaking of hit rate in the summer, combined with a minimum false-alarm rate leads to the high ROC skill for the June 1 forecasts as shown in Figure 6.8f. As far as the issuance of false alarms is concerned, a comparison of Figure 6.11a-d and Figure 6.11e-h indicates that all schemes exhibit more skill for above-normal forecasts than for below-normal forecasts across the range of exceedance probabilities considered. A diagnosis of the model bias for each forecast category revealed that the hydrologic model underestimated low flows much more than it overestimated high flows for this watershed.

Figures 6.12 and 6.13 show the evolution of hit rates and false-alarm rates at the 10%, 30%, 70% and 90% exceedance probabilities for Castle River forecasts. For below-normal forecasts issued on or before June 1, the hit rates change from a slightly negative trend for lower warning thresholds (e.g. Figure 6.12a) to a positive trend for higher warning thresholds (e.g. Figure 6.12c). In the case of above normal forecasts, the hit rates exhibit a local minimum for forecasts issued on May 1 at lower warning thresholds (Figure 6.12e-f). At the higher end of the warning threshold, hit rates show an increasing trend as the forecast season progresses (Figure 6.12h). The higher hit rates for the July 1 and August 1 forecasts are accompanied by higher false-alarm rates (Figure 6.13e-h), indicating that the forecast issues too many false warnings of wet conditions. This is a consequence of a wet bias in the hydrologic model for the late target seasons.

The general increase in forecast skill towards the summer months and its tapering off in July and August could be explained by the relative contributions of initial conditions and meteorological forecasts to the forecast uncertainty. At the beginning of the season, the future meteorology is the major source of uncertainty (Franz *et al.*, 2003 and references therein). As the season progresses, the contribution of meteorological uncertainty decreases because more and more snow gets accumulated in the basin, leading to more accurate specification of initial conditions. Since snowmelt runoff constitutes the major portion of the streamflow during the spring and early summer target seasons, the forecasts issued between April 1 and June 1 show better skills than the others. As the summer progresses, the contribution of snowmelt to the actual streamflow progressively decreases, and thus the accuracy of individual rain forecasts becomes the major source of uncertainty. The absence of the dampening effect of snowmelt runoff on individual rain events would lead to ensemble members with too much dispersion, thus resulting in reduced forecast skills for late forecast dates (Franz *et al.*, 2003).

While the use of ensemble of meteorological inputs to the hydrologic model is meant to mimic uncertainties in the future meteorology, hydrologic modeling system uncertainties (due to initial conditions, calibration data, model structure) are not resolved by the ESP system. Under perfect model assumptions (i.e., a perfect model structure and error-free calibration data), a grasp on the "potential" forecast skill of the ESP system could be obtained by evaluating the forecast skill with respect to the model simulated streamflow instead of the historical flow. The RPSS measure based on this approach is shown in Figure 6.14. Averaged over all eight forecast issue dates, the actual RPSS (Figure 6.7) is about 50% and 25% of the "potential" RPSS for the Bow and Castle rivers, respectively.

Given that the "potential" RPSS is, at the current level of meteorological forecast skill, generally higher than the actual RPSS, the poor performance of the ESP forecasting system is attributed to uncertainty in hydrologic modeling. In our case, the major part of this uncertainty could be attributed to model inputs, particularly problems with elevation-adjusted input weather data from foothill stations to reflect the actual spatio-temporal distributions of weather variables in the mountains from where the bulk of the streamflow originates.

Another input error of minor importance may be the contribution of glacier melt, which has not been taken into account in this study. For the Bow River at Banff, glacier melt contributions make up only about 1.8% of the long term annual mean flow (Hopkinson and Young, 1998). However, glacier contributions may be considerably more than the long-term average during low flow years as well as late in the summer. For instance, Hopkinson and Young (1998) noted that the glacier contribution during the low flow year of 1970 increased to 13%, and to around 56% for August of the same year. The Castle River watershed has no glacier cover. As SAC-SMA is a simplified conceptual representation of nature, limitations in the model structure also contribute to uncertainties in the forecasts. Therefore, further improvement in the actual ESP forecast skill requires improvements in local-scale precipitation and temperature forecast skills, better spatial distribution of inputs, and accurate streamflow simulation.

6.8. Skill Comparison with Statistical Forecasts

In Chapter 5, ensemble forecasts produced by the robust M-regression model and K-NN resampling algorithm were presented for the two sub-basins. The subsets of predictors used in the M-regression model were selected among a pool of large-scale climate indices, SWE, antecedent precipitation and temperature using the GCV criterion explained in Chapter 5. Here, the skills of the ESP median and ensemble forecasts are compared to those produced by the statistical model. For each forecast issue date, the ESP scheme with the best skill (among AGCM3, GEM, Pre-ESP and Post-ESP) was selected for the comparison.

For forecasts issued between January and August, the correlations for median forecasts, and RPSS for ensemble forecasts for both sub-basins are compared in Figure 6.15. For the Bow River, median forecasts by the statistical model show better correlations for forecasts issued between January and June whereas ESP median forecasts are marginally better for forecasts issued after June (Figure 6.15a). That the difference in correlation decreases with forecast issue date suggests that the median forecasts by the statistical model are more reliable than any of the ESP median forecasts at long lead times. For the Castle River, only the ESP median forecast issued on June 1 has marginally better correlation than statistical median forecasts (Figure 6.15b). In terms or RPSS, ESP ensemble forecasts for the Bow River are also less skillful than statistical ensemble forecasts for the entire forecasting season (Figure 6.15c). Differences in RPSS between the two methods are greatest for forecasts issued before June 1. While statistical ensemble forecasts for the Castle River are also more skillful than ESP ensemble forecasts (Figure 6.15d), the difference in skill is more modest compared to that of the Bow River.

6.9. Summary and Conclusions

A framework for incorporating 0-3 month lead temperature and precipitation forecasts from two Canadian NWP models into the ESP system was presented. The framework involves improving the NWP forecasts by the MOS technique and then modifying input weather data generated by a block K-NN resampling from historical observations so as to reproduce the magnitudes of monthly forecasts of the desired variable. The technique was applied to two snowmelt dominated headwater basins in the SSRB of southern Alberta. Seasonal ensemble streamflow forecasts based on the two NWP models (AGCM3 and GEM) were compared to forecasts produced by pre-ESP and post-ESP K-NN resampling schemes.

Forecast performance was assessed using hindcasts produced by a conceptual rainfall-runoff model for a 35-year period. Forecasts were evaluated in both deterministic and probabilistic settings. The median forecast was selected as the deterministic ESP forecast and its performance was evaluated in terms of correlation coefficient and RMSE. For the Bow River, the correlations of the median forecasts based on AGCM3 and GEM were generally comparable to those produced by the pre-ESP and post-ESP resampling schemes. For Castle River, both AGCM3 and GEM produced better correlations than the pre-ESP and post-ESP schemes. In terms of overall ensemble forecast skill, all schemes gave modest improvement over climatology as measured by RPSS.

The streamflow forecast skills reported in this study are lower than the potential skill that could be achieved with the same meteorological input data, primarily due to uncertainties in hydrologic modeling. One possible source of hydrologic uncertainty is that the calibration data for the hydrologic model were based on single station observations located close to the watershed outlet. We have attempted to take into account elevation effects through gradients estimated from station observations but the gradients themselves are not expected to be free of elevation bias as they were also derived from stations located at relatively lower elevations. This is particularly the case for the Castle River, where all stations used to estimate precipitation gradients with elevation were located at < 1400m while the basin elevation ranges from 1200 to 2500m.

Estimates of the temperature distribution should be fairly accurate because temperature varies with elevation at a more or less uniform rate (e.g. environmental lapse rate). On the other hand, the actual precipitation distribution (particularly in higher elevation parts of the watersheds) could be highly variable due to the complex topography and orographic effects. The forecasting system performed slightly better for the Bow River, where the two mountain stations, albeit with short data records, provide some information about the precipitation fields in the mountains. Provided that data issues are properly resolved, future research should evaluate the methods presented in this study at additional watersheds of less complex topography in the SSRB, and watersheds of different climatic regimes.

In addition to input errors, CRR models like SAC-SMA are expected to incur some degree of uncertainty in watersheds of complex topography. For instance, snowmelt computation by the degree-day approach could contribute some uncertainties given that air temperature doesn't quite accurately represent the energy budget associated with snowmelt. Even though a semi-distributed approach was adapted in computing snowmelt, using lumped parameters for the soil moisture accounting routine also contributes to model uncertainties. While moving from conceptual to more physically based approach, from lumped to semi-distributed or fully distributed approach may improve hydrologic simulation, it could also mean requiring more input data to drive the hydrologic model. Therefore a pertinent question is how to strike a balance between input data requirement and model structure.

The MOS-based scheme involves downscaling and disaggregating NWP forecasts to produce suitable data for ESP. Similarly, the Pre-ESP scheme has a considerable computational demand because it involves random resampling of weather data for each day during the forecast period. For instance, for a forecast to be issued on January 1, the simulation period contains 273 days (1 Jan – 30 Sep). The random resampling and all associated steps (see Section 6.3) are carried out 273 times for a single ensemble member, as compared to a single time for the Post-ESP scheme.

Generally speaking, the increased computational burdens of both the MOS-based and Pre-ESP schemes do not appear to produce a pronounced improvement in forecast skills over the Post-ESP scheme. However, the decision about which of the three schemes to implement should also take into account the potential for further improvement of the forecast skill. While all three schemes can benefit from more accurate streamflow simulation, only the MOS-based scheme can gain additional improvement from more accurate local temperature and precipitation forecasts. Future work should also consider using an ensemble of hydrologic models so as to reduce the dependence of the results on specific model structures.

References

Alberta Environment. (2006). Water Supply Outlook for Alberta, Alberta Environment: http://www.environment.gov.ab.ca/ (Accessed August 2006).

Anderson, E. A. (1973). National Weather Service river forecast system – Snow accumulation and ablation model. *NOAA Technical Memorandum NWS Hydro-17*, U. S. Department of Commerce, Silver Spring, MD.

Burnash, R. J. C., Ferral, R. L., and McGuire, R. A. (1973). *A generalized streamflow simulation system - conceptual modeling for digital computers.* U.S. Department of Commerce, National Weather Service and State of California, Department of Water Resources.

Clark, M. P., and Hay, L. E. (2004). Use of medium-range numerical weather prediction model output to produce forecasts of streamflow. *J. Hydrometeorology*, 5, 15-32.

Day, G. N. (1985). Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plann. Manage.*, 111, 157-170.

Druce, D. J. (2001). Insights from a history of seasonal inflow forecasting with a conceptual hydrologic model. *J. Hydrol.*, 249, 102-112.

Franz, K., Hartmann, H., Sorooshian, S., and Bales, R. (2003). Verification of National Weather Service ensemble streamflow predictions for water supply forecasting in the Colorado River basin. *J. Hydrometeorology*, 4, 1105-1118.

Gan, T. Y. (1988). Applications of scientific modeling of hydrologic responses from hypothetical small catchments to assess a complex conceptual rainfall-runoff model. *Water Resources Series Tech. Rep. No. 11*, Seattle, WA.

Gan, T.Y., and Burges, S.J. (1990). An assessment of a conceptual rainfall-runoff model's ability to represent the dynamics of small hypothetical catchments 1. Model properties and experimental design. *Water Resour. Res.*, 26(7), 1595-1604.

Gobena, A. K., and Gan, T. Y. (2006). Links between pacific climate variability and low frequency hydroclimatic variability in Southern Alberta. *Proc. CSCE Annual Conference*, Canadian Society for Civil Engineering, Abstracts & CD-ROM, Calgary, Alberta, May 23-26, 2006.

Hamlet, A., and Lettenmaier, D. (1999). Columbia River streamflow forecasting based on ENSO and PDO climate signals. *J. Water Resour. Plann. Manage.*, 125, 333-341.

Hamon, W. R. (1961). Estimating potential evaporation. *Journal of the Hydraulics Division*, Proceedings of the American Society of Civil Engineers, 87, 107–120.

Hopkinson, C., and Young, G. J. (1998). The effect of glacier wastage on the flow of the Bow River at Banff, Alberta, 1951-1993. *Hydrol. Process.,* 12, 1745-1762.

Hsieh, W., and Tang, B. (2001). Interannual variability of accumulated snow in the Columbia basin, British Columbia. *Water Resour. Res.*, 37, 1753-1759.

Hsieh, W., Yuval, Li, J., Shabbar, A., and Smith, S. (2003). Seasonal prediction with error estimation f Columbia River streamflow in British Columbia. *J. Water Resources Planning and Management* 129(2), 146-149.

HYDAT (2001). *National Surface Water Data (Water Survey of Canada)*, Environment Canada, CD-ROM.

Lall, U., and Sharma, A. (1996). A nearest neighbor bootstrap for re-sampling hydrologic time series. *Water Resour. Res.*, 32, 679-693.

Oudin, L, Hervieu, F., Michel, C., Perrin, C., Andreassian, V., Antil., F., and Loumange, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model? Part 2 - Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modeling. *J. Hydrol.*, 303, 290-306.

Prairie, J. R., Rajagopalan, B., Fulp, T. J., and Zagona, E. A. (2006). Modified K-NN model for stochastic streamflow simulation. *J. Hydrol. Engrg.*, 11(4), 371-378.

Sharif, M., and Burn, D. H. (2006). Simulating climate change scenarios using an improved K-nearest neighbor model. *J. Hydrol.*, 325, 179-196.

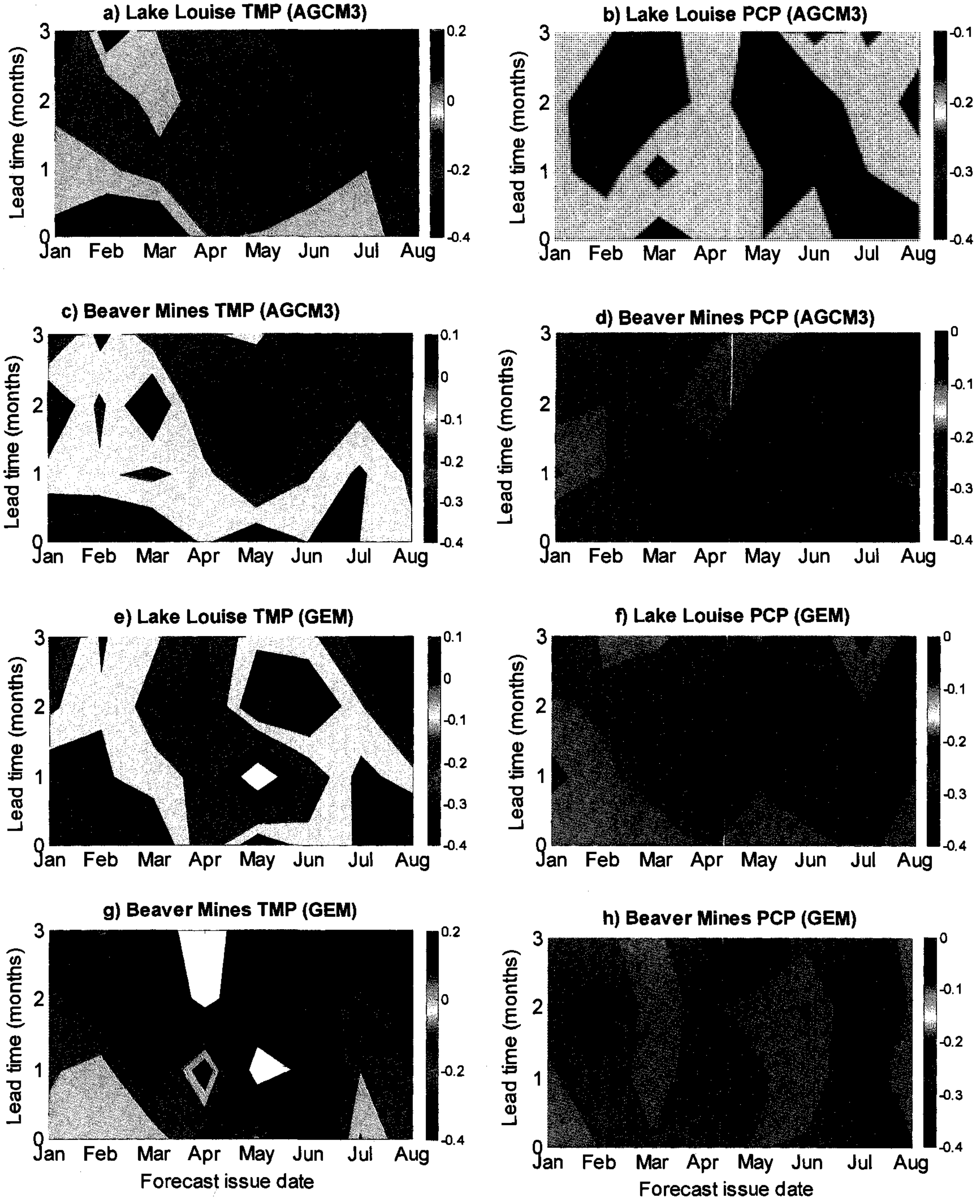
Singhrattna, N., Rajagopalan, B., Clark, M., and Krishna Kumar, K. (2005). Forecasting Thailand summer monsoon rainfall. *Int. J. Climatol.*, 25, 649-664.

Werner, K., Brandon, D., Clark, M., and Gangopadhyay, S. (2004). Climate index weighting schemes for NWS ESP-based seasonal volume forecasts. *J. Hydrometeorology,* 5, 1076-1090.

Wilks, D. S. (2006). *Statistical Methods in the Atmospheric Sciences* (2nd ed.) Amsterdam: Academic Press.

Wood, A. W., Maurer, E. P., Kumar, A., and Lettenmaier, D. (2002). Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.*, 107 (D20), 4429, doi:10.1029/2001JD000659.

Yates, D., Gangopadhyay, S., Rajagopalan, B., and Strzepek, K. (2003). A technique for generating regional climate scenarios using a nearest-neighbor algorithm. *Water Resour. Res.*, 39(7), 1199, doi:10.1029/2002WR001769.


a) Lake Louise TMP (AGCM3)

b) Lake Louise PCP (AGCM3)

c) Beaver Mines TMP (AGCM3)

d) Beaver Mines PCP (AGCM3)

e) Lake Louise TMP (GEM)

f) Lake Louise PCP (GEM)

g) Beaver Mines TMP (GEM)

h) Beaver Mines PCP (GEM)

Figure 6.1. RPSS for bias-corrected NWP (AGCM3 and GEM) monthly temperature and precipitation forecasts at Lake Louise and Beaver Mines stations. The reference forecast is the climatology.

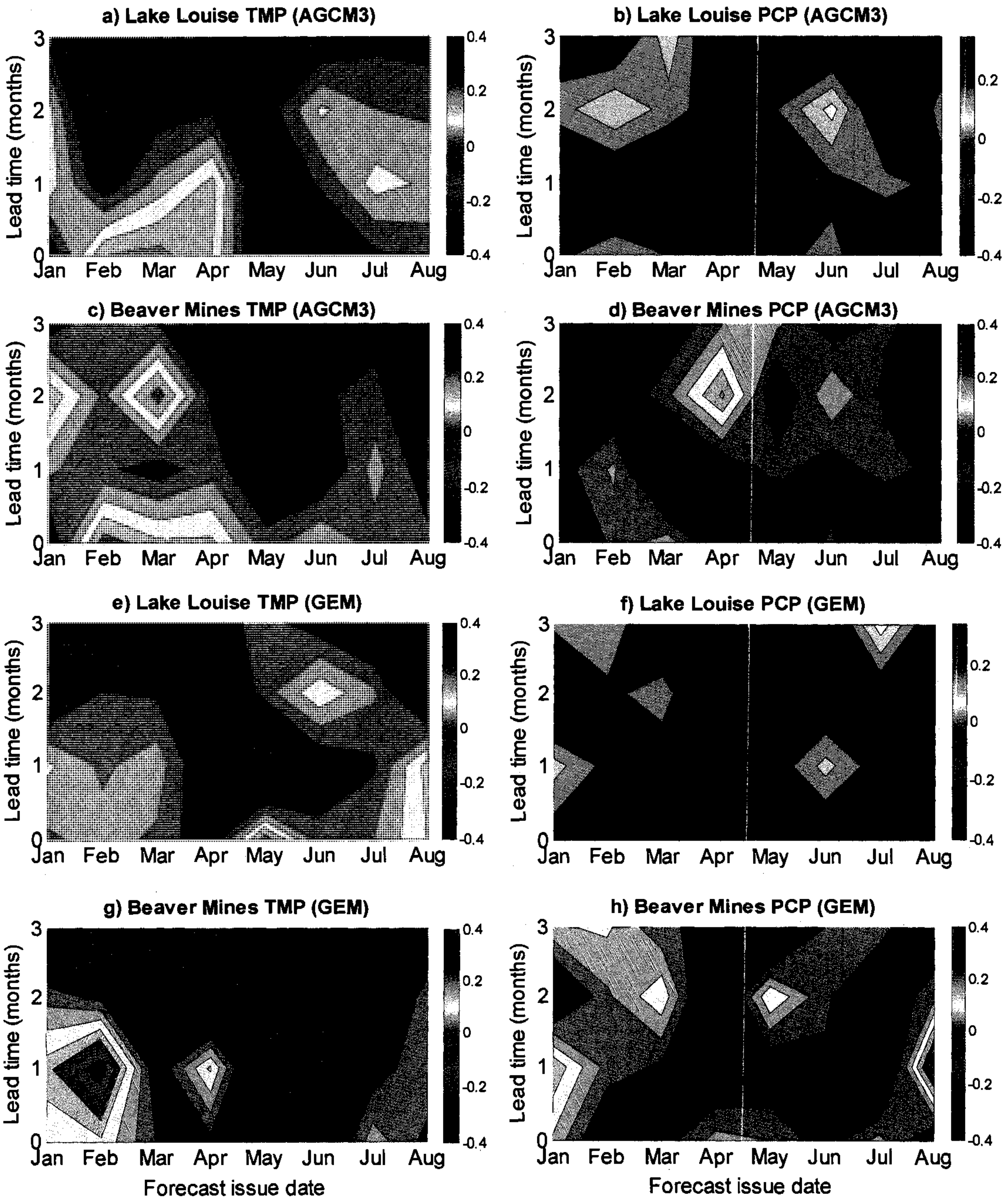


Figure 6.2. RPSS for MOS-based monthly temperature and precipitation forecasts at Lake Louise and Beaver Mines stations based on AGCM3 (a to d) and GEM (e to h). The reference forecast is the bias-corrected NWP forecast.

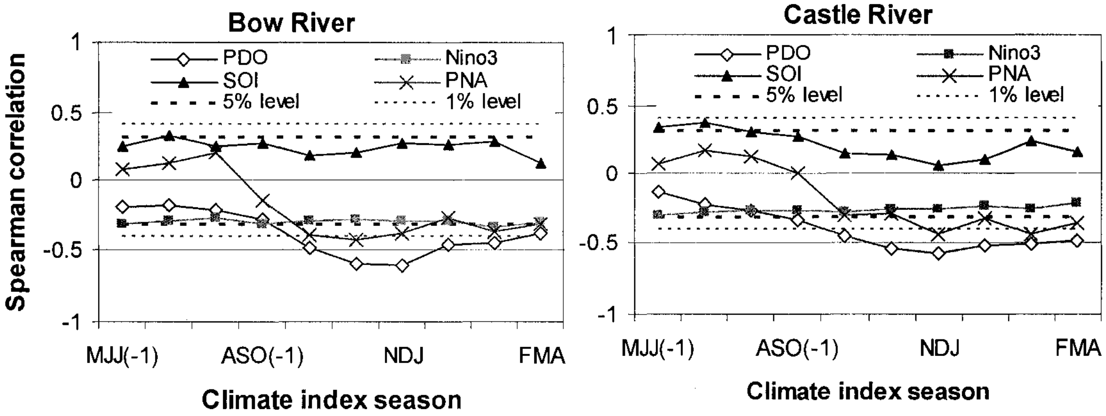


Figure 6.3. Spearman rank correlations between April-September average stream flow and four climate indices. Averages of 3-month moving windows starting with May-June-July of the previous year (MJJ(-1)) and ending with February-March-April (FMA) of the runoff year were used for the climate indices.

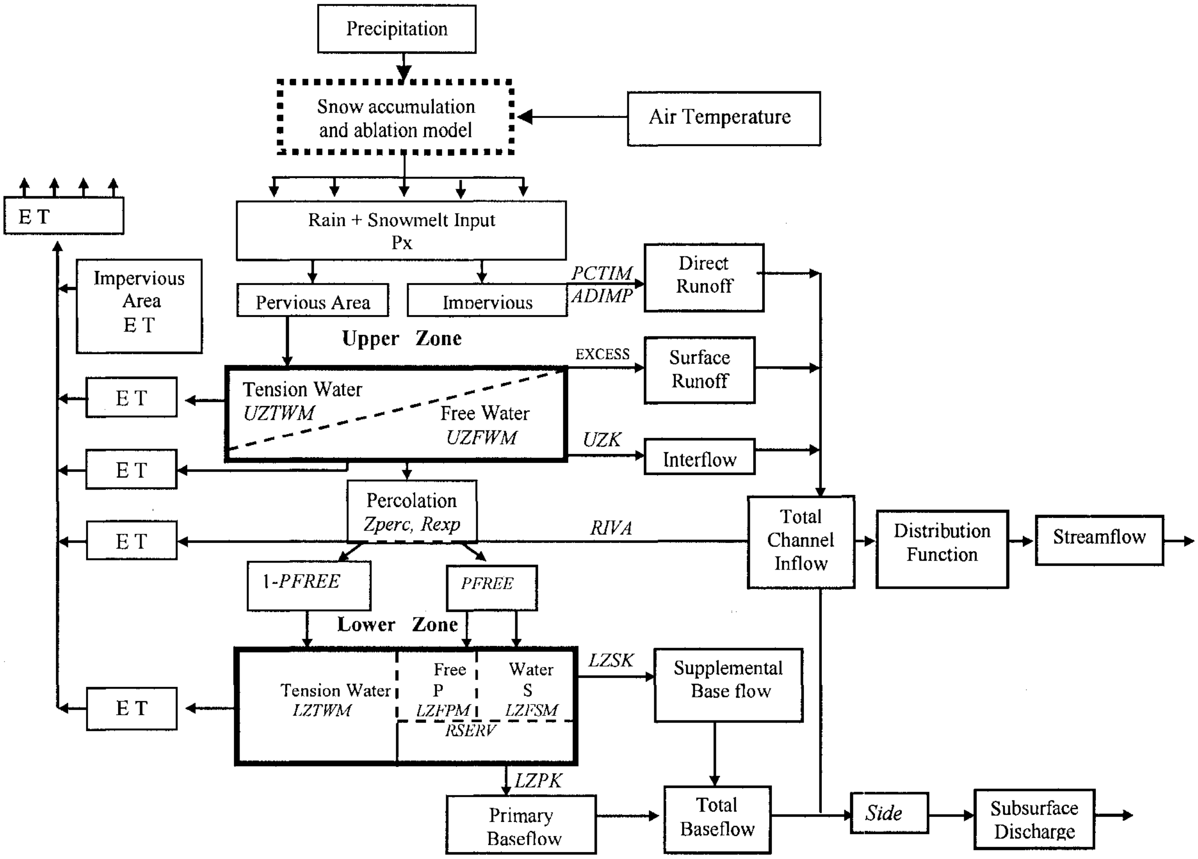


Figure 6.4. Structure of the Sacramento Soil Moisture Accounting model (adapted from Gan, 1988). Model parameters are shown in italics.

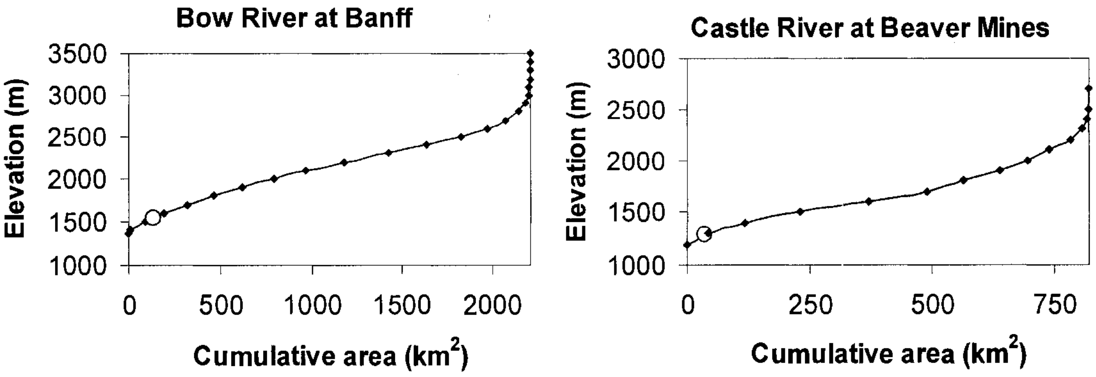


Figure 6.5. Area-elevation curves for the Bow and Castle rivers. The open circles on the curves indicate the location of the meteorological stations used to derive input weather data.

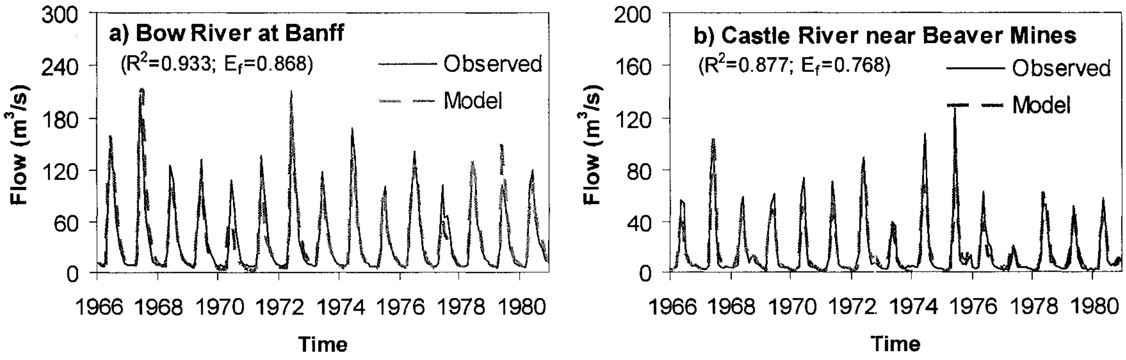


Figure 6.6. Observed and modeled monthly hydrographs for the validation period (1966-1980). (a) Bow River (05BB001). (b) Castle River (05AA022).

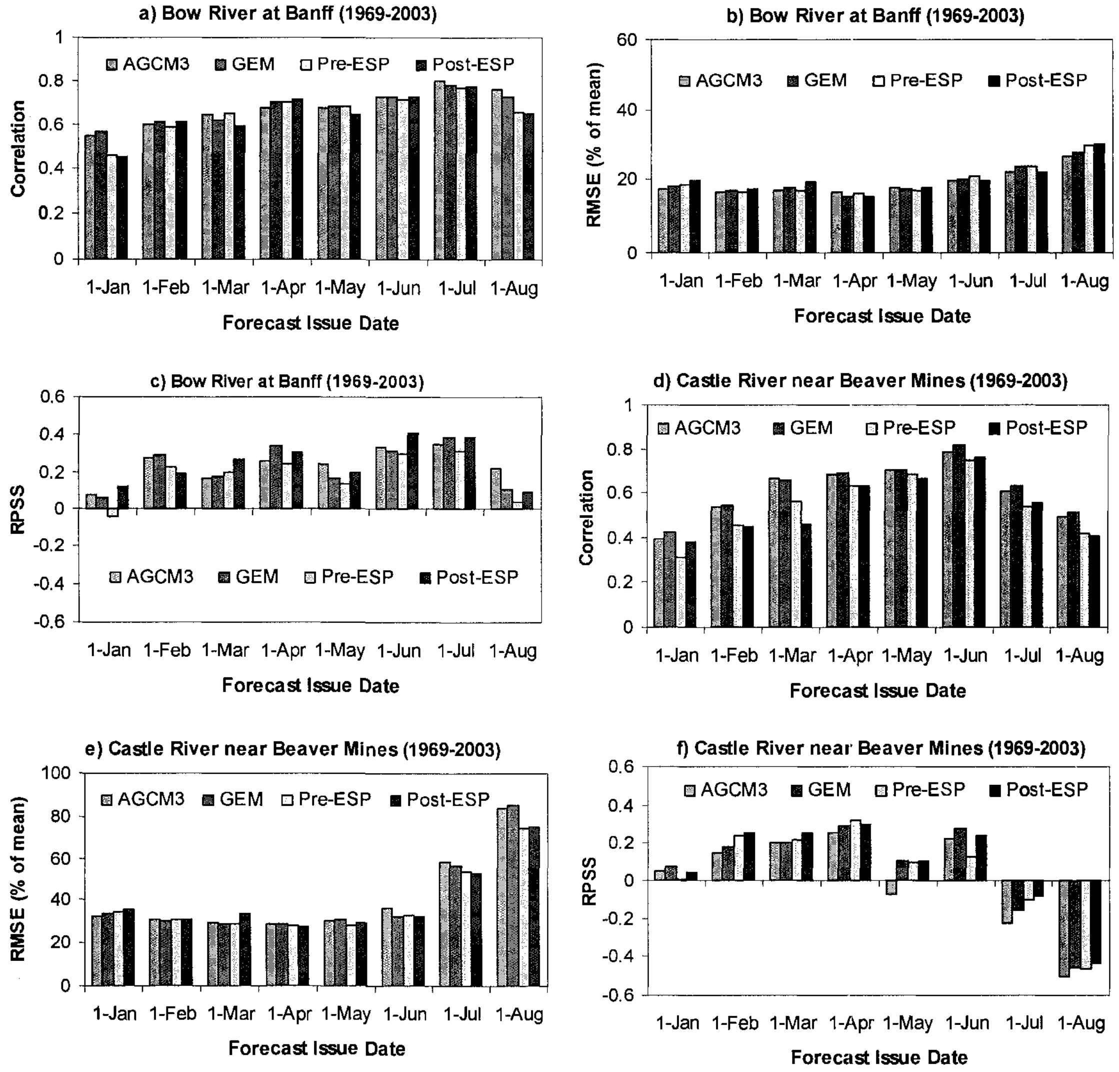


Figure 6.7. Correlation coefficient, RMSE and RPSS skill measures for forecasts of the Bow and Castle rivers

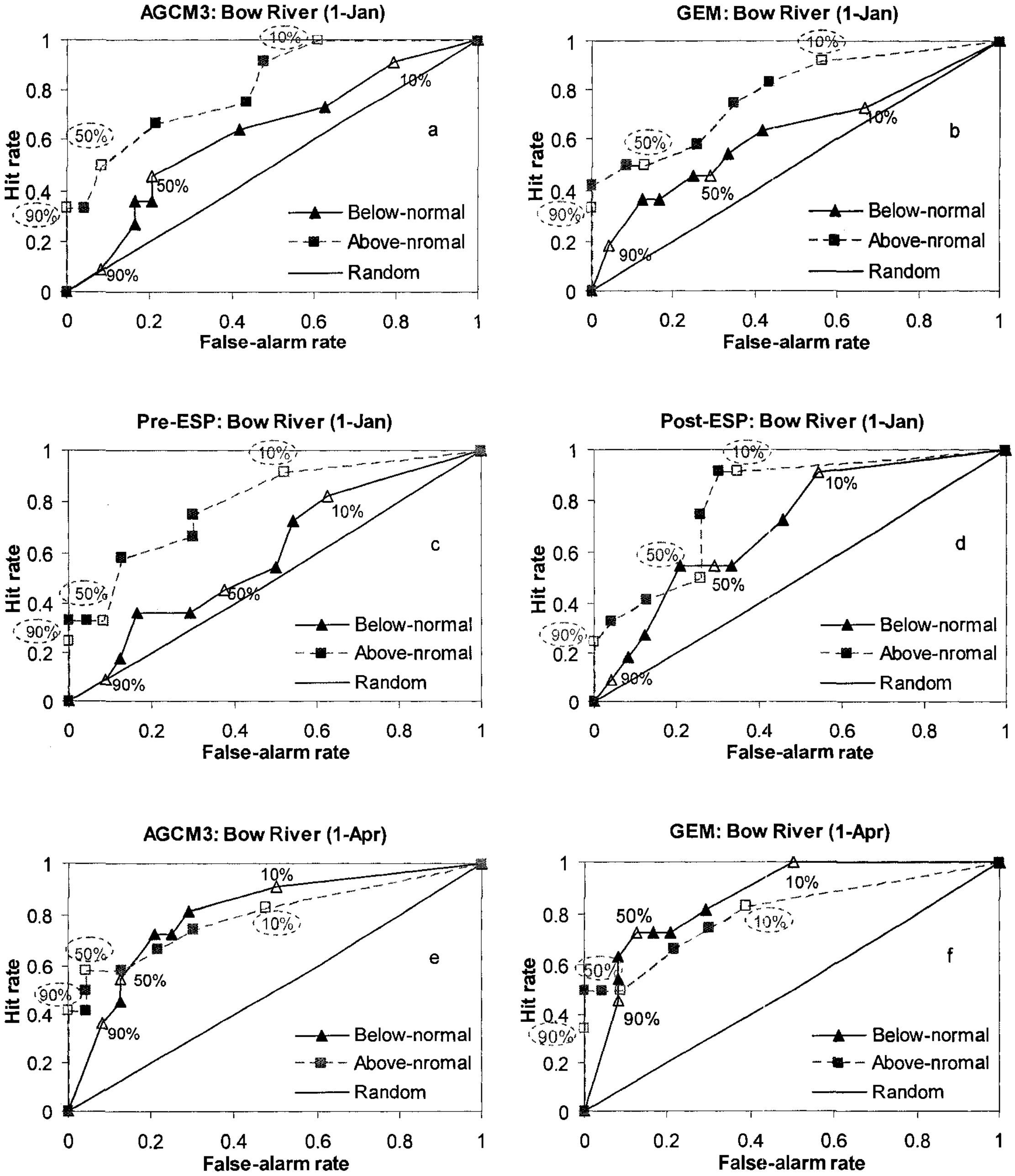
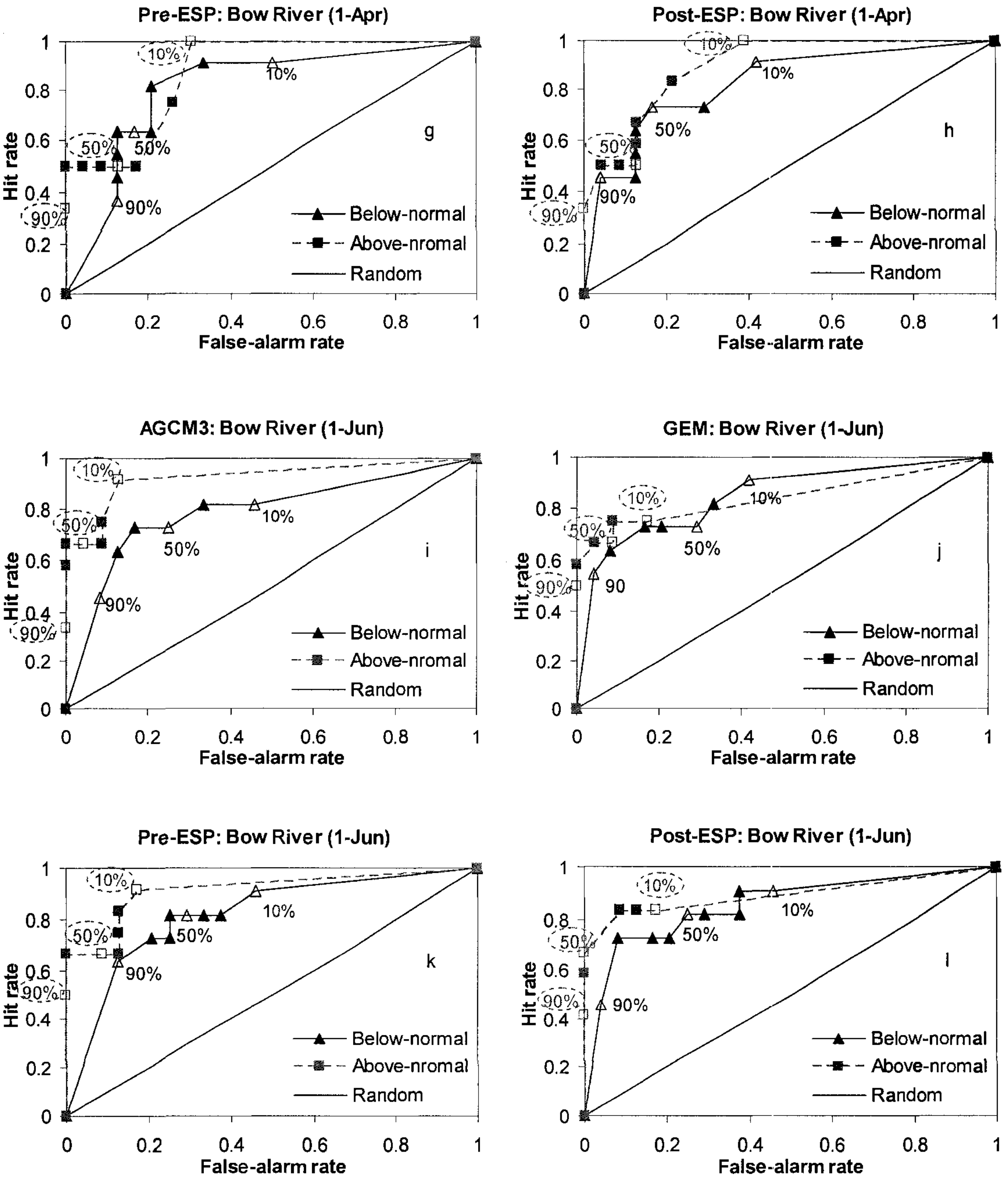


Figure 6.8. Bow River ROC diagrams for ensemble forecasts issued on 1-Jan, 1-Apr and 1-Jun for below-normal and above-normal flows based on AGCM3, GEM, Pre-ESP, and Post-ESP. The diagonal line indicates the no-skill line. The open markers on the curves show the 10%, 50% and 90% warning thresholds, which are also circled for above-normal forecasts.


Pre-ESP: Bow River (1-Apr)

Post-ESP: Bow River (1-Apr)

AGCM3: Bow River (1-Jun)

GEM: Bow River (1-Jun)

Pre-ESP: Bow River (1-Jun)

Post-ESP: Bow River (1-Jun)

Figure 6.8. *Continued*

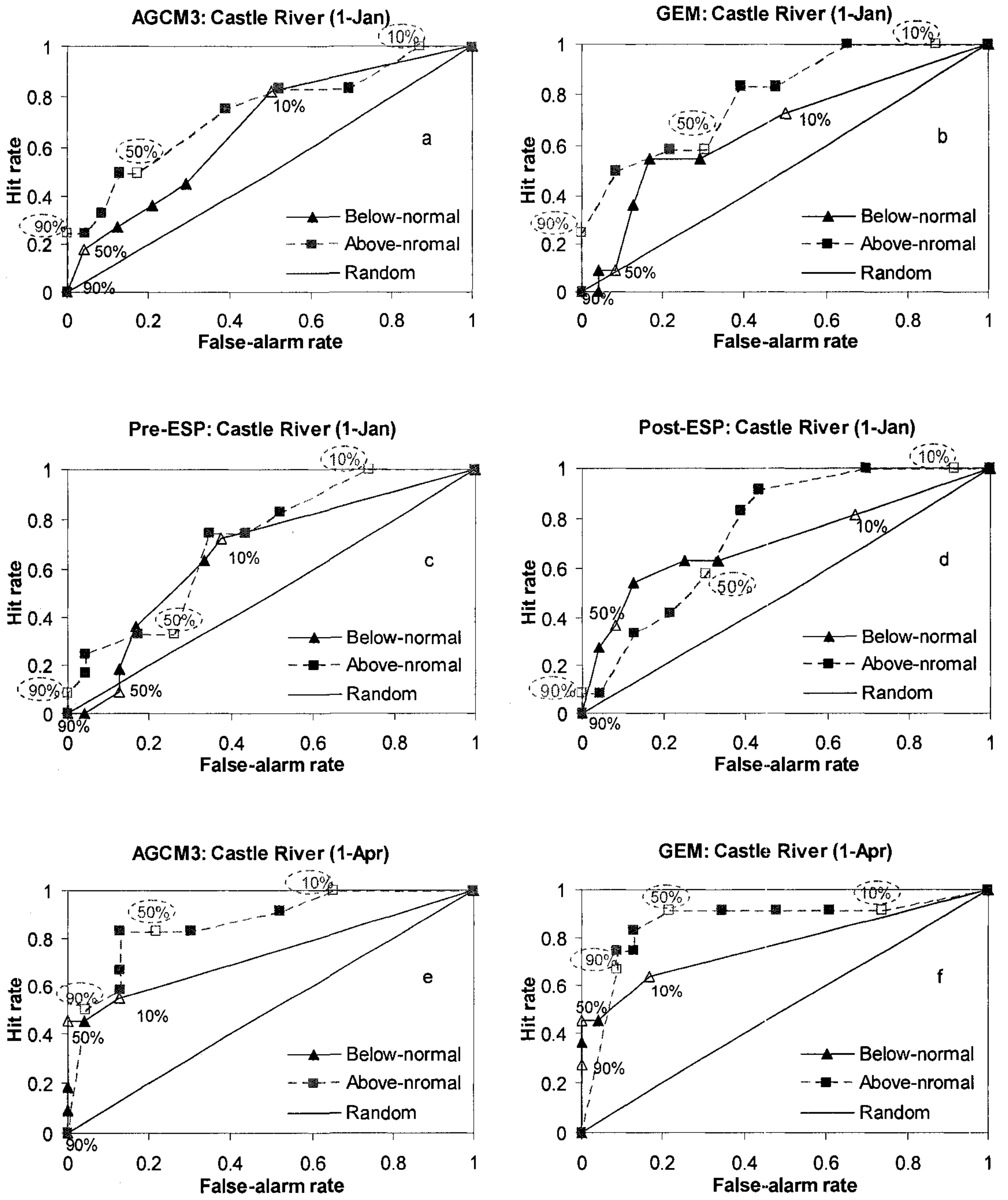


Figure 6.9. Castle River ROC diagrams for ensemble forecasts issued on 1-Jan, 1-Apr and 1-Jun for below-normal and above-normal flows based on AGCM3, GEM, Pre-ESP, and Post-ESP. The diagonal line indicates the no-skill line. The open markers on the curves show the 10%, 50% and 90% warning thresholds, which are also circled for above-normal forecasts.

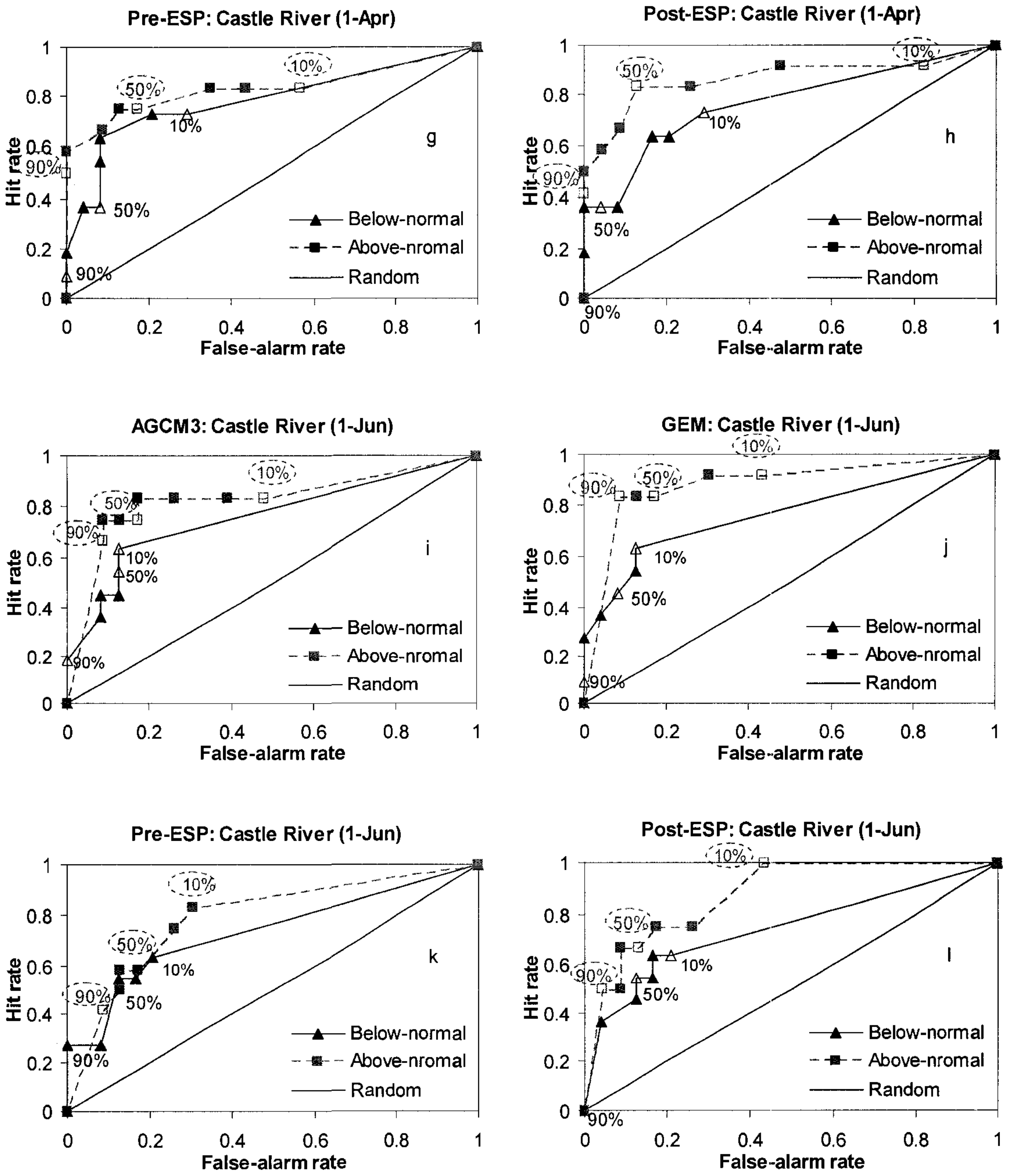
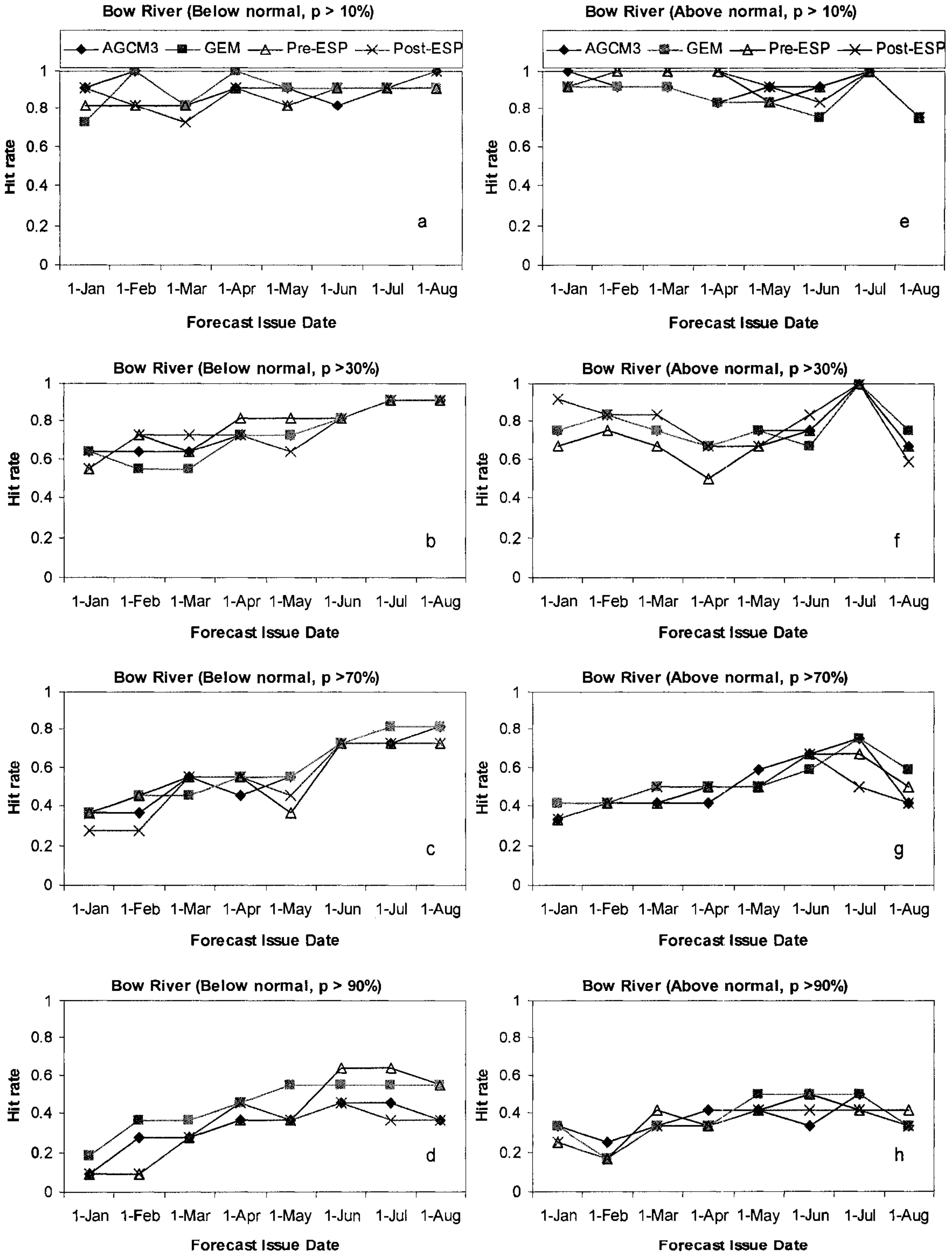


Figure 6.9. *Continued*


Bow River (Below normal, p > 10%)

Bow River (Above normal, p > 10%)

Bow River (Below normal, p >30%)

Bow River (Above normal, p >30%)

Bow River (Below normal, p >70%)

Bow River (Above normal, p >70%)

Bow River (Below normal, p > 90%)

Bow River (Above normal, p >90%)

Figure 6.10. Temporal variation of hit rates for below-normal (a to d) and above-normal (e to h) forecasts of the Bow River at exceedance probabilities of 10%, 30%, 70% and 90%.

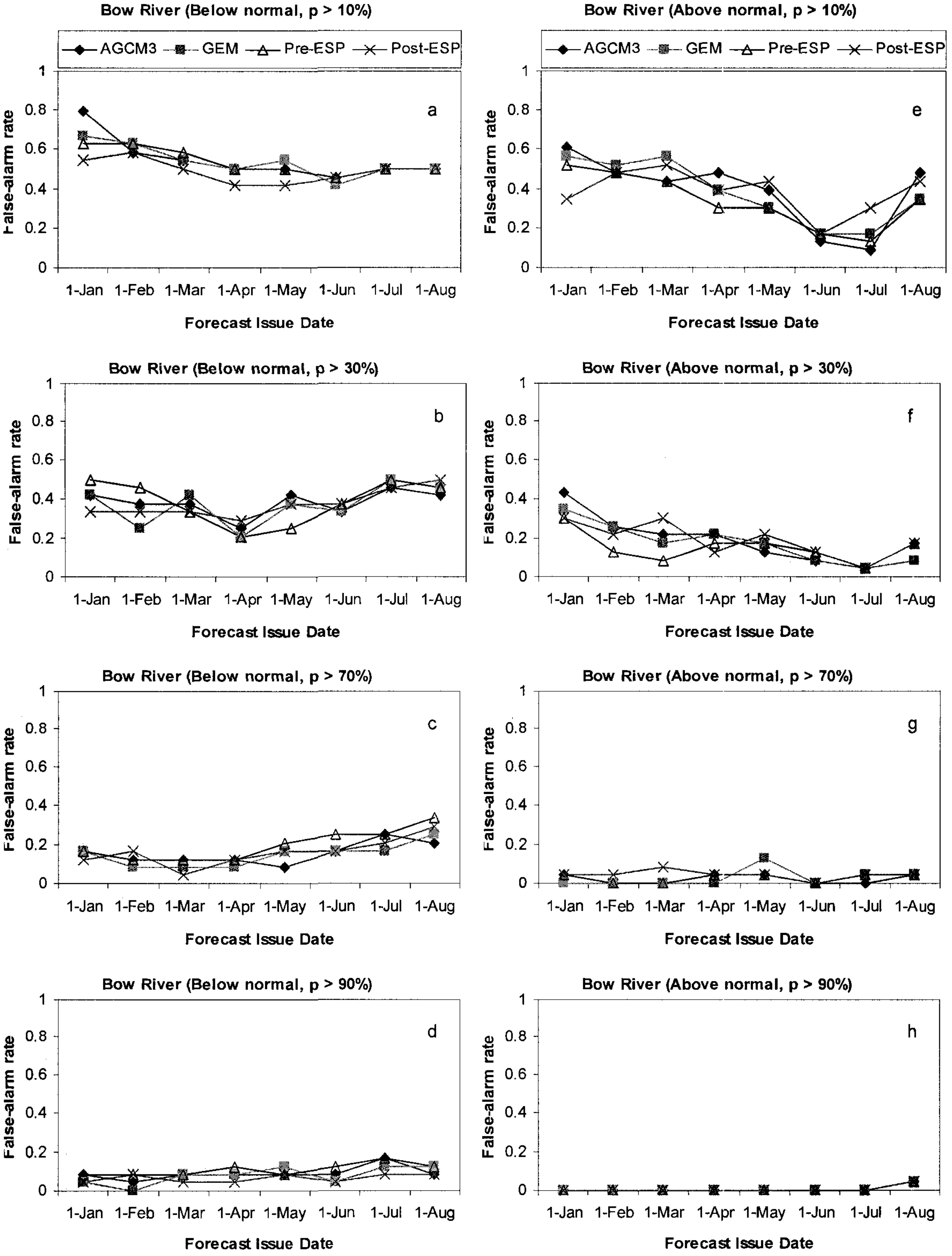


Figure 6.11. Temporal variation of false-alarm rates for below-normal (a to d) and above-normal (e to h) forecasts of the Bow River at exceedance probabilities of 10%, 30%, 70% and 90%.

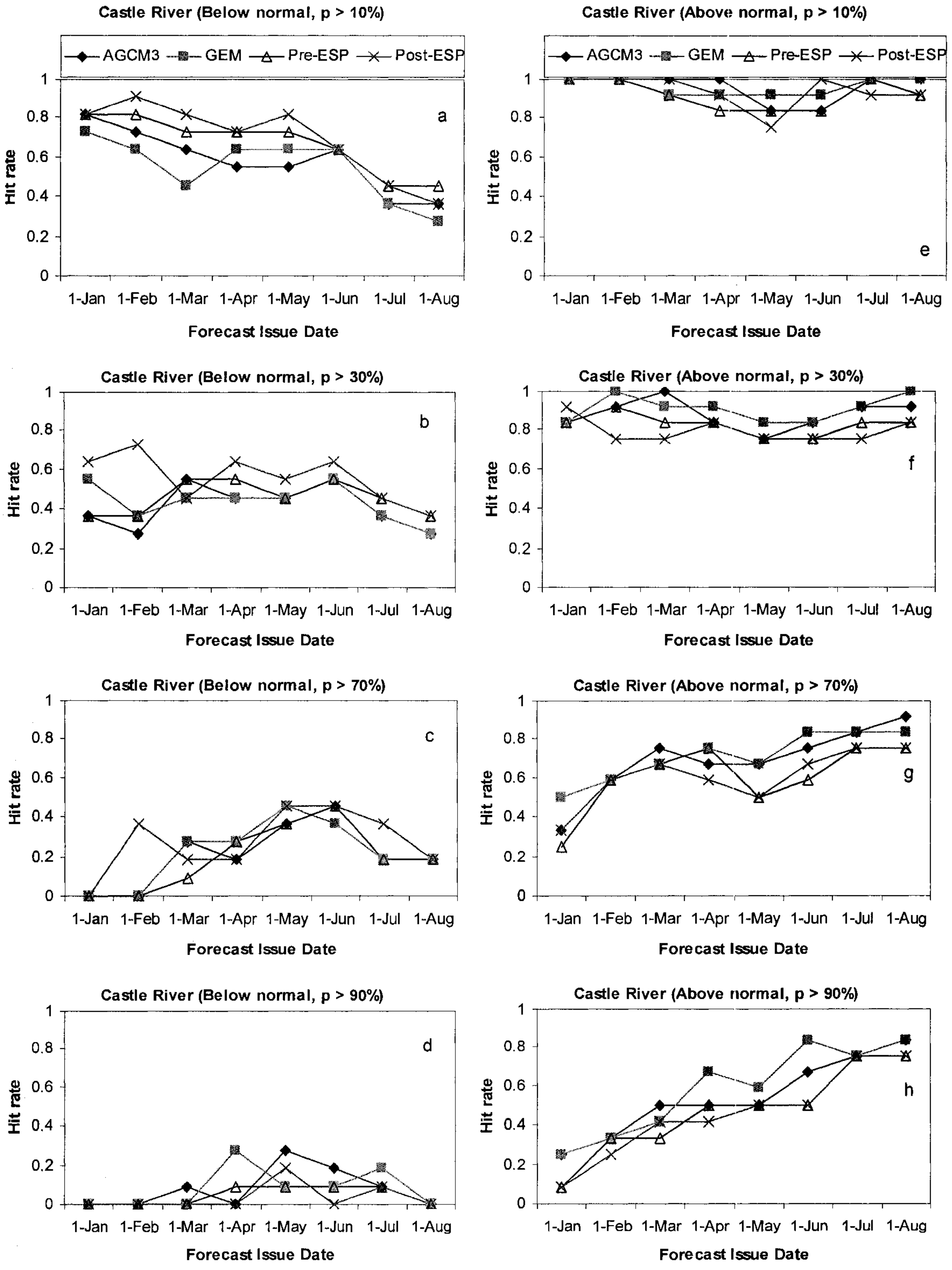


Figure 6.12. Same as Figure 6.10 but for Castle River.

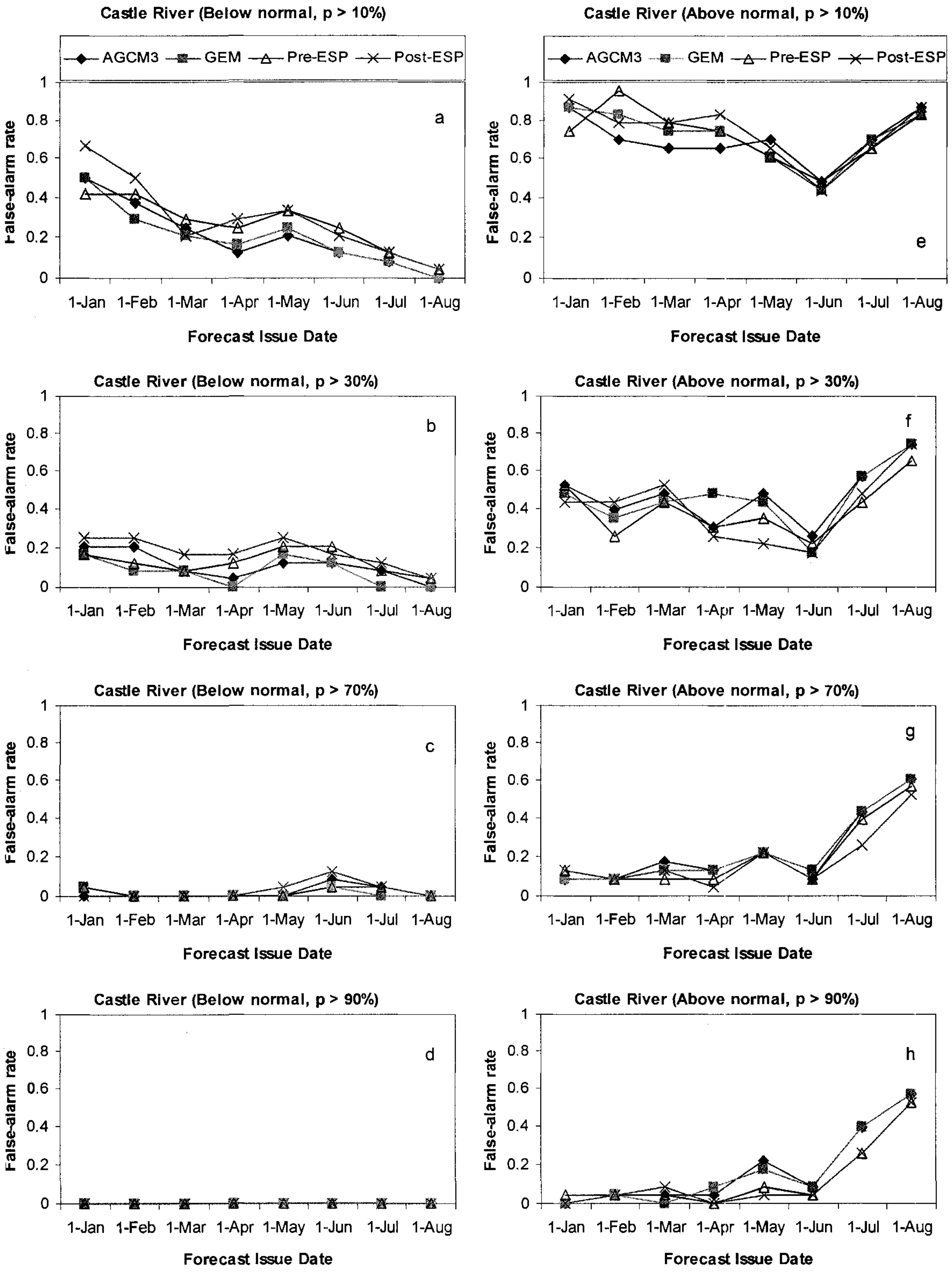


Figure 6.13. Same as Figure 6.11 but for Castle River.

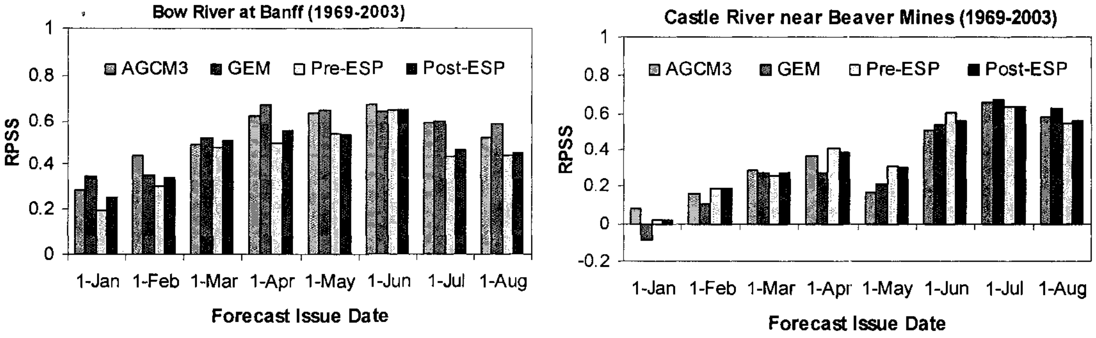


Figure 6.14. RPSS of ESP forecasts for an error-free hydrologic model. In the RPSS computation, the SAC-SMA simulated flow was used instead of the observed flow.

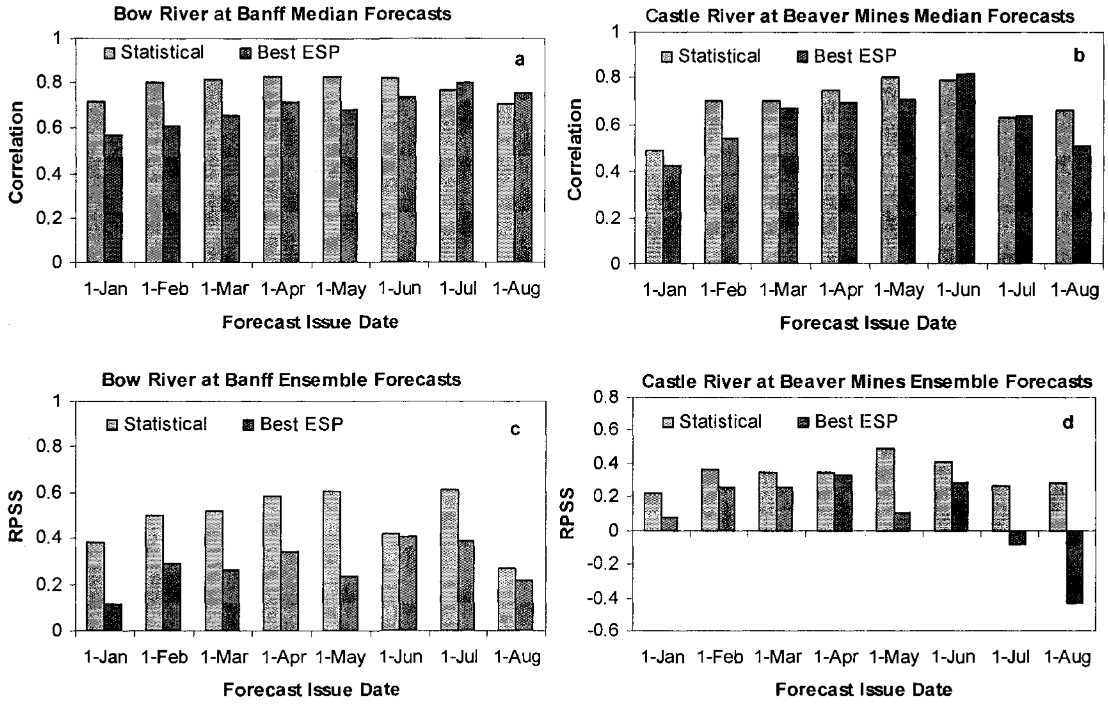


Figure 6.15. Skills comparison of ESP and statistical model forecasts. The statistical forecasts are produced by a robust regression model and K-NN resampling. 'Best ESP' stands for the ESP scheme (among AGCM3, GEM, Pre-ESP and Post-ESP) with the best skill for that forecast issue date.

Chapter 7 Summary, Conclusions and Recommendations

7.1. Summary and Conclusions

The theme of this research is the analysis of the role of large scale climate anomalies on Western Canadian hydroclimatic variability and its applicability for seasonal streamflow forecasting. The findings of the research are summarized in the following paragraphs.

In Chapter 2, wavelet and teleconnection analysis was performed on Western Canadian seasonal precipitation time series obtained from 21 stations with long historical record. It was found that statistically significant interannual oscillations in seasonal precipitation occurred haphazardly while interdecadal fluctuations, albeit detected only in a few stations, were found to be more consistent than interannual fluctuations. In order to facilitate regional analysis, Western Canada was qualitatively zoned into three regions based on the similarities in the precipitation wavelet power at interannual scales. These three zones – referred to as western, central and eastern regions – were used for teleconnection analysis with several large-scale oceanic/atmospheric circulation patterns (via climate indices). In particular, the precipitation signal for each of the three regions was represented by the leading principal components (PC1) of the precipitation data from each of the three regions.

Climate anomalies considered for the teleconnection analysis are the El Nino/Southern Oscillation (ENSO), Pacific/North America pattern (PNA), West Pacific pattern (WP), East Pacific pattern (EP), Central North Pacific (CNP) SLP index, and Pacific Decadal Oscillation (PDO). Teleconnections between regional precipitation signals and large-scale climate anomalies were investigated by using several statistical techniques including wavelet coherence and phase difference, wavelet filtering, compositing and correlation analysis. At interannual scales, the wavelet coherence and phase difference between all of the climate indices and the precipitation signal for each region was found to be highly inconsistent in both time and frequency. In addition, Pearson's correlations between the PC's of band-pass filtered precipitation and climate indices were generally found to be weak, indicating that the strength of teleconnections changes along both time and frequency dimensions. These observed inconsistencies corroborate the haphazard nature of interannual oscillations in the precipitation data. The consistency and strength of the relationships between regional precipitation and the PDO and CNP indices was much better at interdecadal scales.

Given that teleconnection effects are relatively stronger during boreal winter (Horel and Wallace, 1981; Shabbar *et al.*, 1997), composites of raw winter precipitation were analyzed for each of the six climate anomalies at all 21 stations. Considering the study area as a whole, it was found that ENSO exhibits relatively stronger influence on winter precipitation than the remaining climate anomalies. It was also observed that in terms of the sign of the anomaly, precipitation responses to opposite phases of ENSO, PNA and PDO are relatively more consistent across the study area than responses to opposite phases of WP, EP and CNP. The fact that the detected teleconnections could occur at interannual or inter-decadal levels (depending on which anomaly index is used), and their changing strength in time, frequency and space makes their applications for seasonal precipitation prediction by statistical techniques unreliable. Spatial variations in the multifractal properties of precipitation across the region were reported by Gan *et al.* (2007).

In Chapter 3, teleconnections between Western Canadian streamflow anomalies and ENSO, PNA and PDO were investigated using a suite of statistical tools on 42 years of monthly streamflow data (Gobena and Gan, 2006). Five coherent regions were identified based on cluster analysis of streamflow anomaly time series constructed from 8 El Nino and 8 La Nina years. Periods of significant streamflow response to El Nino and La Nina events exhibit spatial variations. The periods of significant response for rivers in the Prairies and the Greater Vancouver region correspond with spring snowmelt. On the other hand, the response period for rivers in the BC interior and Rocky Mountains predominantly occurs during late spring to summer months due to the fact that the dominant runoff regime for basins in those regions is mountain snowmelt/glacier melt and summer rain. Basins in Northern BC do not show significant response to ENSO. Streamflow Index Time Series, constructed by spatial averaging of the streamflow anomalies for each region, were used to study the consistency of streamflow response to El Nino and La Nina events. It was found that that while El Nino (La Nina) events are generally associated with negative (positive) streamflow anomalies, there were cases when opposite conditions were observed.

The role of extratropical Pacific interannual (PNA and WP) and interdecadal (PDO) variability modes as possible sources of interference in the ENSO-streamflow relationship was studied by stratifying ENSO events according to different phases of PNA and PDO. It was found that high PNA years conditioned on non-El Nino years could be associated with streamflow anomalies that are similar in magnitude and duration to streamflow anomalies that are observed during El Nino years. On the other hand, streamflow anomalies during Low PNA years, regardless of the La Nina state, appear to be much more suppressed than those observed during La Nina years. Composite analysis of the 300-mbar wind flow patterns over Western Canada provided evidence for the existence of favorable conditions that support the findings of this study. It was also found that the strength of ENSO-related streamflow response varied as a function of the PDO phase, with the relationship being synergetic when the two climate anomalies are in phase (i.e. El Nino/Warm PDO and La Nina/Cool PDO) and destructive when they are out of phase (i.e. El Nino/Cool PDO and La Nina/Warm PDO).

The current generations of climate models are capable of producing skillful forecasts of the SST in the equatorial Pacific where ENSO originates (Barnston *et al.*, 1994; Latif *et al.*, 1994). Although ENSO may provide useful long-range forecast skills for streamflow anomalies observed during the major flow season in

Western Canada, possible non-linear ENSO-streamflow relationships resulting from interferences by PNA and PDO will likely affect the skill of the forecast. Partial correlations computed at different lead times indicate that ENSO indices provide better predictability for basins that are dominated by spring snowmelt runoff regimes while the winter season PDO provides better predictability for basins originating in the Rocky Mountains, whose flow regimes are dominated by summer snowmelt/glacier melt. Since current climate prediction models do not offer skillful long-range predictions for PNA and PDO as they do for ENSO, the use of PNA and PDO for prediction purposes hinges on the time lag between atmospheric circulation and streamflow response due to snow accumulation.

In Chapter 4, wavelet analysis of streamflow and precipitation variability and teleconnectivity was performed for the SSRB using 89 years of streamflow and precipitation data. Besides having a sufficiently long dataset to resolve interdecadal oscillations, SSRB is of key interest to Alberta primarily because it provides over 57% of the water allocated in the province. The results of this analysis are consistent with those reported in the previous two chapters. Significant streamflow variance was observed at interdecadal scales with periods of 19–22, 41–42 and 62 years whereas significant variance in the 2–8-year scale was observed at intervals of 20–25 years. The intermittency in the interannual streamflow variance leads to inconsistent relationships with ENSO indices as observed from changes in the wavelet phase difference with time. On the other hand, PDO and streamflow exhibited consistently strong covariance with a rather stable phase difference of 180° for scales greater than about 20 years. Corroborating the results reported earlier for basins originating in the Rocky Mountains, the winter season PDO, not ENSO, provides the primary predictive information for the major streamflow season (April-September) in the SSRB. This is partly because mountain snow accumulation, which is a significant source of basin runoff during the major flow season, is more strongly correlated to PDO than to ENSO.

In Chapters 5 and 6, two ensemble seasonal streamflow forecasting models that incorporate large-scale climate information were developed and applied to two watersheds in the SSRB. The first used a statistical approach (Chapter 5), where single-value forecasts generated by a robust M-regression model are perturbed by ordinates resampled from a Gaussian kernel density function fitted to the K-NN residuals of the M-regression model. The method produces skillful forecasts of the April to September streamflow from as early as December prior to the runoff season. Given that forecasts are currently issued beginning in February, the inclusion of large-scale climate information enables us to extend the current forecast lead time by up to 2 months. The ultimate objective of a forecasting system is to help downstream users in making optimum decisions that will maximize their benefits. In terms of potential economic value, ensemble forecasts were found to provide better positive return for a wide range of hypothetical downstream users (i.e. users with a wide range of cost-loss ratios) as compared to single-value forecasts.

In Chapter 6, a technique for incorporating 0-3 month lead temperature and precipitation forecasts from two Canadian NWP models into the ESP system was presented. The skills of bias-corrected NWP forecasts, downscaled to Lake Louise and Beaver Mines stations, were generally found to be worse than the climatology. Therefore, the MOS technique was used to improve the NWP model forecasts before using them in a predictive environment. The robust M-regression model was used to forecast monthly precipitation and temperature, where the GCV criterion was used to select the subsets of predictors from a pool of forecasts of surface and atmospheric variables included in the HFP2 archive. The MOS forecasts were then replaced by observations having the same quantile values as the forecasts so as to preserve the ensemble spread, and the mean and variance of the distribution. The MOS-based forecasts showed up to 40% improvement in RPSS over the bias-corrected NWP forecasts. Since our hydrologic simulations in the ESP system are performed at a daily time step, a K-NN resampling algorithm was used to temporally disaggregate the sequence of MOS-based monthly temperature and precipitation forecasts to daily time steps for the first four months in the hydrologic simulation period.

In addition to the MOS-based scheme, two other schemes (Pre-ESP and Post-ESP) were considered where the weather data used to force the hydrologic model are entirely generated from historical observations (i.e. without any forecast guidance from NWP models). In the Pre-ESP scheme, ensemble members were obtained by forcing the hydrologic model with daily weather sequences generated by K-NN resampling from historical data.

Unlike the climatic ESP method, the number of ensemble members that can be generated by both the MOS and Pre-ESP schemes is not constrained by the length of historical records. Thus, these schemes are suitable for basins with relatively short historical records as well. In the Post-ESP scheme, ensemble members were obtained by K-NN resampling from the climatic ESP traces. Forecast skills computed from a 35-year hindcast indicate that the three schemes generally exhibit comparable performance. In spite of the improved temperature and precipitation forecasts with respect to bias-corrected NWP forecasts, in general the MOS-based scheme did not lead to noticeable improvement in forecast skill over the Pre-ESP and Post-ESP schemes. The only exception where the MOS-based scheme outperformed the other schemes is in the case of median forecasts for the Castle River.

We found that all three ESP schemes improved on climatological forecasts for most of the forecast issue dates, which is encouraging. However, compared to ensemble forecasts produced by the robust regression model, our ESP forecasts were found to be less skillful, both for median and ensemble forecasts. The fact that higher skills could be achieved under perfect model assumptions suggests that the relatively poor performance of ESP forecasts can partly be attributed to hydrologic modeling errors. Thus, any further improvement in the accuracy of local precipitation and temperature forecasts should be accompanied by a more comprehensive modeling of hydrologic processes if improvement in ESP forecast skill is to be pursued.

Both the regression and ESP methods require long historical observations of hydroclimatic data for model calibration and forecast verification. The regression model requires data to be available at monthly or coarser time steps. On the other hand, the ESP approach involves running a hydrologic model, which is usually executed at much finer time steps (e.g. daily frequency in our case). While the latter has an added complexity of understanding the basin rainfall-runoff dynamics, it has the advantage that streamflow forecasts can easily be partitioned into monthly or other desired time scales whenever such is required for operational purposes. When seasonal volume forecasts are the quantity of interest, ensemble forecasts produced by the more straightforward regression models can be used without loss of accuracy.

7.2. Suggestions for Future Work

Even though the systematic analysis carried out in this work has added substantial new insights to existing knowledge on Western Canadian hydroclimatic variability, there still remain a number of issues that require further investigation.

1) Our study addressed only the relationships between the hydroclimate and large-scale climate anomalies without regard to the influence of local factors. There is a need to investigate possible effects of Western Canada's complex topographic features on the regional and seasonal variations observed in hydrologic responses to ENSO and other climatic anomalies.

2) An investigation into the effect of other climatic anomalies (e.g. the Arctic oscillation and the North Atlantic oscillation) on regional climate variability will be helpful to provide a complete picture of climate systems at play in Western Canada.

3) Regarding statistical seasonal streamflow forecasting, a comparison of the robust regression model with other statistical forecasting methods such as local regression and kernel based approaches is warranted. A multi-model ensemble strategy could be pursued so as to exploit the strength of each modeling approach.

4) Further improvement in the accuracy of local temperature and precipitation forecasts and hydrologic simulation are necessary to improve ESP forecasts. The downscaling method used to improve the raw NWP forecasts should also consider other derived variables such as specific humidity and vertically integrated moisture flux as candidate predictors in the MOS technique. In addition, use of multiple hydrologic models in a multi-ensemble framework is recommended so as to reduce the effect of model structure uncertainty on the ESP forecast skill.

5) Before going operational with the seasonal streamflow forecast models proposed in this study, there is a need to evaluate the techniques in additional watersheds within the SSRB and other basins with different climatic regimes such as the North Saskatchewan and Athabasca river basins.

References

Barnston, A. G., Van den Dool, H. M., Zebiak, S. E., Barnett, T. P., Ji, M., Rodenhuis, D. R., Cane, M. A., Leetmaa, A., Graham, N. E., Ropelewski, C. R., Kousky, V. E., O'Lenic, E. A., and Livezey R. E. (1994). Long-lead seasonal forecasts – where do we stand?, *Bull. Amer. Meteor. Soc.,* 75, 2097-2114.

Gan, T. Y., Gobena, A. K., and Wang, Q. (2007). Precipitation of southwestern Canada – Wavelet, scaling, multifractal analysis, and teleconnection to climate anomalies. *J. Geophys. Res.*, 112, D10110, doi:1029/2006JD007157.

Gobena, A. K., and Gan, T. Y. (2006). Low-frequency variability in southwestern Canadian streamflow: links to large-scale climate anomalies. *Int. J. Climatol.,* 26, 1843-1869.

Horel, J. D., and Wallace, J. M. (1981). Planetary-scale atmospheric phenomena associated with the southern Oscillation. *Mon. Weather Rev.,* 109, 813-829.

Latif, M., Barnett, T. P., Cane, M. A., Flugel, M., Graham, N. E., von Storch, H., Xu, J.-S., and Zebiak, S. E. (1994). A review of ENSO prediction studies. *Climate Dynamics,* 9, 167-179.

Shabbar, A., Bonsal, B., and Khandekar, M. (1997). Canadian precipitation patterns associated with the Southern Oscillation. *J. Climate,* 10, 3016-3027.

Appendix A Statistical Methods

This appendix contains brief mathematical descriptions of some of the standard statistical methods used for diagnostic analysis in the study.

Binomial Distribution

Consider a random variable X that can take on one of two mutually exclusive and continuously exhaustive (MECE) outcomes in a success/failure experiment. The probability of realizing x success outcomes from a collection of N independent trials is given by (Wilks, 1995)

$$\Pr(X = x) = \binom{N}{x} p^x (1-p)^x, \qquad x = 0,1,2,\ldots.N \tag{A.1}$$

where p is the probability of occurrence of the event of interest (success) on any one of the N independent trials. For cases where $0 < p \pm [p(1-p)/N]^{1/2} < 1$, i.e. N is sufficiently large, the random variable X can be approximately characterized by a Gaussian distribution with

$$\mu = Np \tag{A.2a}$$

$$\sigma = [Np(1-p)]^{1/2} \tag{A.2b}$$

Equation A.2 can be used to estimate the confidence interval around the observed estimate of the binomial x. The $(1-\alpha)$ confidence interval around x is given by

$$\mu \pm \sigma z_{1-\alpha} \tag{A.3}$$

Mann-Whitney Test

Given two independent data samples, the non-parametric Mann-Whitney test uses the rank of the data to test for possible difference in location (e.g. mean) of the two samples (Wilks, 1995). Unlike the traditional t test, the Mann-Whitney test does not make any assumption about the underlying distribution of the data and is resistant to outliers. The null hypothesis is that the two data samples are drawn from the same distribution and the labeling of the individual values as belonging to one or the other sample is arbitrary. In order to compute the test statistic, the two samples are pooled together under this null hypothesis and ranked. Then the sums of the ranks held by the original data samples 1 and 2 are calculated separately as S_1 and S_2. The Mann-Whitney U-statistic is computed from

$$\begin{aligned} U_1 &= S_1 - n_1(n_1+1)/2 \\ U_2 &= S_2 - n_2(n_2+1)/2 \end{aligned} \tag{A.4}$$

For moderately large values of n_1 and n_2, the null distribution of the U statistic is approximately Gaussian and can be estimated from

$$Z = \frac{U - n_1 n_2 / 2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}} \tag{A.5}$$

where, for a two-sided test, U is taken as the minimum of U_1 and U_2 (Wilks, 1995).

Principal Component Analysis (PCA)

PCA is an orthogonal transformation method for reducing multidimensional data sets into lower dimensions through linear combinations of the original variables (Wilks, 1995). Let X be an $N \times p$ matrix of mean-centered data where N is the number of observations and p is the number of variables. The j^{th} principal component of the data is obtained as the projection of the data vector onto the j^{th} eigenvector e_j of the covariance matrix $\sum$,

$$Y_j = e_j^T X = \sum_{i=1}^{p} e_{ij} X_i , \quad j = 1, 2 \ldots\ldots k \tag{A.7}$$

where $k \leq p$. The transformation is subjected to two conditions: (1) the eigenvectors are orthogonal to one another and hence the principal components are mutually uncorrelated; (2) the principal component associated with an eigenvector accounts for the maximum joint variability in the original data set in the direction of that eigenvector (Wilks, 1995). The proportion of the total variance in the original data accounted for by the j^{th} principal component is given by

$$R_j^2 = \frac{\lambda_j}{\sum_{j=1}^{p} \lambda_j} \times 100\% \tag{A.8}$$

where λ_j is the eigenvalue associated with e_j, where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_j \geq \ldots \geq \lambda_p$.

Ward's Minimum Variance Clustering

Ward's method is an agglomerative hierarchical clustering algorithm which uses an ANOVA-type approach to evaluate the distance between clusters. Given an N×p data matrix, it begins with N single-member groups, and merges two groups at each step until all the data are in a single group after N-1 steps (Wilks, 1995). The pair to be merged at each step is chosen as the one that minimizes the sum of squared Euclidean distances between each point in a group and the centroids of their respective groups. Mathematically, the criterion to merge two of G + 1 groups to form G groups can be expressed as

$$Minimize \sum_{g=1}^{G} \sum_{i=1}^{n_g} \sum_{j=1}^{p} (x_{i,j} - \overline{x}_{g,j})^2 \tag{A.9}$$

where n_g is the number of members in a group, p is the number of variables, and $\overline{x}_{g,j}$ is the group centroid for the j^{th} variable. Ward's method minimizes within-group variances (Wilks, 1995).

Reference

Wilks, D. (1995). *Statistical Methods in the Atmospheric Science: An Introduction*. San Diego, CA: Academic Press.

Appendix B Snow Course Stations

This appendix includes a table showing the list of mountain snow course stations in the South Saskatchewan River basin from which snow data used in the study were derived. The data were obtained from Alberta Environment (Chacko Abraham, Personal communication, 2004).

Station	ID Code	Elevation (m)	Basin	Record starts	Lat (deg)	Lon (deg)
Akamina Pass [b]	05AD803	1800	Oldman	1980	49.03	-114.05
Allison Pass [a, b]	05AA803	1980	Oldman	1963	49.73	-114.60
Bow River [a, b]	05BA801	1580	Bow	1937	51.42	-116.18
Bow Summit [b]	05BA813	2080	Bow	1979	51.70	-116.47
Chatteau Lawn [a, b]	05BA808	1740	Bow	1940	51.42	-116.28
Cuthead [b]	15X07	2180	Bow	1979	51.45	-115.77
Gardiner [b]	05AA809	1970	Oldman	1984	49.35	-114.52
Highwood Summit [a]	05BL802	2210	Bow	1963	50.60	-114.98
Katherine Lake [b]	05BA814	2380	Bow	1980	51.68	-116.38
Larch Valley [b]	05BA812	2230	Bow	1979	51.32	-116.22
Marmot-Jasper [a]	07AA803	1830	Athabasca	1970	53.80	-118.08
Mirror Lake [a, b]	05BA806	2030	Bow	1940	51.42	-116.23
Nigel Creek[a]	05DA804	1920	N. Sask.	1968	52.20	-117.08
Pipestone Upper [a, b]	05BA802	1615	Bow	1937	51.43	-116.17
Ptarmigan Hut [a, b]	05BA810	2190	Bow	1967	51.47	-116.10
Skoki Mountain [b]	05CA805	2040	Bow	1979	51.53	-116.05
Sunshine Village [a, b]	05BB803	2230	Bow	1967	51.08	-115.78
Sunwapta Falls [a]	07AA802	1400	Athabasca	1968	52.55	-117.65
West Castle Bush [a, b]	05AA801	1520	Oldman	1967	49.28	-114.37
Wilkinson Bush [a]	05BL804	1980	Oldman	1963	50.20	-114.55
Wilkinson Open [a]	05BL805	1980	Oldman	1965	50.20	-114.55

[a] Snow course data used in Chapter 4

[b] Snow course data used in Chapter 5