Spring 6-4-2013

Climate Change Impact on the Spatio-Temporal Variability of Hydro-Climate Extremes

Mohammad Reza Najafi
Portland State University

Let us know how access to this document benefits you.

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Part of the Climate Commons, Environmental Indicators and Impact Assessment Commons, and the Other Civil and Environmental Engineering Commons

Recommended Citation


10.15760/etd.1114

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.
Climate Change Impact on the Spatio-Temporal Variability of Hydro-Climate Extremes

by

Mohammad Reza Najafi

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Civil and Environmental Engineering

Dissertation Committee:
Hamid Moradkhani, Chair
Scott Wells
David Jay
Aslam Khalil
Alan Yeakley

Portland State University
2013
Abstract

The rising temperature of the earth due to climate change has shown to alter the variations of hydro-climate variables, including their intensities, frequencies and durations. Extreme events such as floods are, in particular, susceptible to any disturbances in climate cycles. As such it is important to provide policymakers with sufficient knowledge about the probable impacts of climate change on hydrologic extremes and most importantly on floods, which have the highest impacts on the societies. For this reason analysis of hydro-climate extremes is commonly performed using data at each site (or grid cell), however due to the limited number of extreme events, these analyses are not robust. Current methods, such as the regional frequency analysis, which combine data from different locations are incapable of incorporating the spatial structure of the data as well as other explanatory variables, and do not explicitly, assess the uncertainties. In this thesis the spatial hierarchical Bayesian model is proposed for hydro-climate extreme analyses using data recorded at each site or grid. This method combines limited number of data from different locations, estimates the uncertainties in different stages of the hierarchy, incorporates additional explanatory variables (covariates), and can be used to estimate extreme events at un-gaged sites. The first project develops a spatial hierarchical Bayesian method to model the extreme runoffs over two spatial domains in the Columbia River Basin, U.S. The model is also employed to
estimate floods with different return levels within time slices of fifteen years in order to detect possible trends in runoff extremes.

Continuing on the extreme analysis, the impact of climate change on runoff extremes is investigated over the whole Pacific Northwest (PNW). This study aims to address the question of how the runoff extremes will change in the future compared to the historical time period, investigate the different behaviors of the regional climate models (RCMs) regarding the runoff extremes, and assess the seasonal variations of runoff extremes.

Given the increasing number of climate model simulations the goal of the third project is to provide a multi-model ensemble average of hydro-climate extremes and characterize the inherent uncertainties. Outputs from several regional climate models provided by NARCCAP are considered for the analysis in all seasons. Three combination scenarios are defined and compared for multi-modeling of extreme runoffs. The biases of each scenario are calculated and the scenario with the least bias is selected for projecting seasonal runoff extremes.

The aim of the fourth project is to quantify and compare the uncertainties regarding global climate models to the ones from the hydrologic model structures in climate change impact studies.
Various methods have been proposed to downscale the coarse resolution General Circulation Model (GCM) climatological variables to the fine scale regional variables; however fewer studies have been focused on the selection of GCM predictors. Additionally, the results obtained from one downscaling technique may not be robust and the uncertainties related to the downscaling scheme are not realized. To address these issues, in the fifth study we employed Independent Component Analysis (ICA) for predictor selection which determines spatially independent GCM variables (as discussed in Appendix A). Cross validation of the independent components is employed to find the predictor combination that describes the regional precipitation over the upper Willamette basin with minimum error. These climate variables along with the observed precipitation are used to calibrate three downscaling models: Multi Linear Regression (MLR), Support Vector Machine (SVM) and Adaptive-Network-Based Fuzzy Inference System (ANFIS).
Dedication

“In The Name of GOD, the Most Compassionate and Merciful”

I dedicate this dissertation to my wife and my best friend, Zahra, who has always been by my side, with unconditional love, help and support, encouraging me to follow what I am passionate about. To my daughter, Ava, whose presence makes me the happiest person and the proudest father.

And to my parents and sisters, Banafsheh and Mina, who have always been my role models and helped me start my journey and showered me with their love and support throughout the way.
Acknowledgements

I would like to acknowledge and thank my professor Dr. Hamid Moradkhani who supported my thesis, directed me through different research projects, motivated me to come up with new ideas and always kept a friendly relationship. I also thank my committee members for their valuable time and support. I thank all my instructors and professors especially Dr. Scott Wells, Dr. Aslam Khalil, Dr. Robert Fountain, and Kevin Martin for their guidance and help in and after class. I also thank Tom Bennet for supporting the Hydraulics and Fluid Mechanics laboratory classes.

I would like to thank all the post-doctoral, graduate and under-graduate students I had the pleasure to work with especially my group members: Matthew Meirs, Wen Yang, Marc Leisenring, Susan Wherry, Ilwon Jung, Caleb Dechant, and Shahrbanou Madadgar for providing a friendly environment in our lab.

Special thanks to Andy Wood, and Kevin Werner from NOAA/NWS River Forecast Center, David Garen from NRCS, and Daniel Cooley from Colorado State University for their help and support during the course of my research.

My gratitude is to the Department of Civil and Environmental Engineering’s staff, including Megan Niermeyer, Ariel Lewis, Elizabeth Alarid, and K.C. Hall, who have always been available to help and answer questions. Also special thanks to Joshua Davis from the Office of International Affairs who has been a great source of help from the beginning of my career in PSU.
Thanks to the National Science Foundation, Water Sustainability and Climate (WSC) program for providing partial financial support. I would like to thank the NARCCAP for providing the data used in this thesis. NARCCAP is funded by the National Science Foundation, the US Department of Energy, the National Oceanic and Atmospheric Administration and the US Environmental Protection Agency Office of Research and Development.

I deeply thank my lovely wife ‘Zahra’, who always encouraged me to continue on my path, and my sweet daughter ‘Ava’ for her presence. My sincere thanks to my lovely parents ‘AliAsghar Najafi’ and ‘Vajieh Poornaser’, who raised me with love and care, and always supported and guided me through my life. Also I deeply thank my sisters, Banafsheh and Mina for always being beside me and for their supports. I thank the rest of my family members and friends, whose support has been a reason for my confidence and efforts. Last but not least greatest thanks to Almighty God, the Most Compassionate and Merciful to whom I entirely trust and rely on.
Table of Contents

Abstract ............................................................................................................................. i
Dedication ........................................................................................................................ iv
Acknowledgements ........................................................................................................... v
Table of Contents ............................................................................................................. vii
List of Tables .................................................................................................................... x
1. Chapter 1. Introduction ................................................................................................. 1
   1.1. Climate Change ........................................................................................................ 1
   1.2. General Circulation Models (GCMs) ...................................................................... 1
   1.3. Downscaling ........................................................................................................... 2
   1.4. Challenges in Climate Modeling ............................................................................ 4
   1.5. Ensemble Multi-modeling ...................................................................................... 5
   1.6. Extreme Analysis ................................................................................................... 7
   1.7. Example of At-Site Extreme Assessment: .............................................................. 10
       1.7.1. Data Bias-correction ..................................................................................... 13
       1.7.2. Extremes ....................................................................................................... 14
       1.7.3. Summary ....................................................................................................... 21
   1.8. Outline .................................................................................................................... 22
2. Chapter 2. Analysis of Runoff Extremes using Spatial Hierarchical Bayesian
   Modeling ......................................................................................................................... 25
   2.1. Background ............................................................................................................ 25
   2.2. Theory and Methodology ....................................................................................... 28
       2.2.1. Extreme Value Theory ................................................................................ 29
       2.2.2. Spatial Hierarchical Bayesian Modeling ......................................................... 31
       2.2.3. Parameter Estimation .................................................................................... 36
   2.3. Case Study ............................................................................................................. 37
       2.3.1. Model Setup .................................................................................................. 41
       2.3.2. Model Structure ........................................................................................... 44
       2.3.3. Model Test ..................................................................................................... 51
       2.3.4. Prior Sensitivity .............................................................................................. 53
       2.3.5. Parameter Estimates ..................................................................................... 56
       2.3.6. Analysis of the Temporal Trend .................................................................... 60
   2.4. Conclusions .......................................................................................................... 64
3. Chapter 3. A Hierarchical Bayesian Approach for the Study of the Climate Change
   Impact on Runoff Extremes ........................................................................................... 67
3.1. Background ............................................................................................................. 67
3.2. Methodology ......................................................................................................... 75
  3.2.1. Hydrologic Modeling ...................................................................................... 75
  3.2.2. Spatial Hierarchical Bayesian Model .............................................................. 77
3.3. Application to the Runoff Extremes over the Pacific Northwest ....................... 82
  3.3.1. Data .................................................................................................................. 83
  3.3.2. Downscaling the NARCCAP Climate Variables ............................................ 84
  3.3.3. Hydrologic Modeling ...................................................................................... 86
3.4. Results and Discussion ......................................................................................... 87
  3.4.1. Analysis of GEV Parameters ........................................................................ 90
  3.4.2. Spatial and Temporal Variations of Runoff Extremes .................................. 105
  3.4.3. Hydro-Climate Analysis of the Pacific Northwest (PNW) ......................... 125
3.5. Conclusions ........................................................................................................ 128

4. Chapter 4. Multi-Model Ensemble Analysis of the Runoff Extremes for Climate
Change Assessments ................................................................................................... 131
  4.1. Background ......................................................................................................... 131
  4.2. Regional Climate Models ................................................................................... 134
  4.3. Methodology ..................................................................................................... 137
    4.3.1. Runoff Extreme Assessment ....................................................................... 137
    4.3.2. Bayesian Model Average ............................................................................ 140
    4.3.3. Multi-Modeling of Extreme Runoff ............................................................. 142
  4.4. Results ................................................................................................................ 144
  4.5. Conclusions ....................................................................................................... 157

5. Chapter 5. Assessing the Uncertainties of Hydrologic Model Selection in Climate
Change Impact Studies ............................................................................................... 159
  5.1. Background ......................................................................................................... 159
  5.2. Study Area and Data .......................................................................................... 161
  5.3. Methodology ..................................................................................................... 163
    5.3.1. Hydrologic models ...................................................................................... 164
    5.3.2. GCM Derived Climate Data ....................................................................... 167
    5.3.3. Hydrologic Modeling and Calibration ........................................................ 168
  5.4. Results and Discussion ...................................................................................... 171
    5.4.1. Precipitation and Temperature ................................................................. 171
    5.4.2. Hydrologic Impact Assessment ................................................................. 175
    5.4.3. Bayesian Model Averaging ...................................................................... 181
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.5. Conclusions</td>
<td>189</td>
</tr>
<tr>
<td>6.1. Future Research:</td>
<td>196</td>
</tr>
<tr>
<td>References</td>
<td>199</td>
</tr>
<tr>
<td>Appendices</td>
<td>220</td>
</tr>
<tr>
<td>Appendix A: Statistical Downscaling of Precipitation using Machine Learning with Optimal Predictor Selection</td>
<td>220</td>
</tr>
<tr>
<td>Study area and Data</td>
<td>223</td>
</tr>
<tr>
<td>Nonlinear Models</td>
<td>226</td>
</tr>
<tr>
<td>Independent Component Analysis for Predictor Selection</td>
<td>233</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>241</td>
</tr>
<tr>
<td>Summary and Conclusion</td>
<td>248</td>
</tr>
<tr>
<td>Appendix B: Metropolis Hastings within Gibbs Sampler</td>
<td>251</td>
</tr>
</tbody>
</table>
List of Tables

Table 1-1: Regional Climate Model Information ................................................................. 11
Table 1-2: General Circulation Model (Drivers for RCMs) Information ....................... 12
Table 2-1: Linear correlation between the scale parameter of the fitted generalized Pareto distribution at each gage (ψ) and the corresponding drainage area as well as the gage elevation. .................................................................................................................. 46
Table 2-2: The median and 95% confidence intervals of the estimated latent parameters corresponding to the scale parameter; the model has two covariates. ......................... 56
Table 3-1: Model selection based on the deviance information criterion...................... 90
Table 3-2: Overall bias (obs-sim) of the regional climate models compared with the observation. ........................................................................................................................................ 109
Table 4-1: Comparison between the mean absolute errors of the estimated 100-yr return level runoffs (mm) from each model and from the Bayesian model averaging. .......... 153
Table 5-1: Climate and streamflow stations used in this study ..................................... 162
Table 5-2: Description of the Global Climate Models used in this study (Randall et al., 2007)) ............................................................................................................................................... 164
Table 5-3: a) Lumped hydrologic models parameter description .................................. 165
Table 5-4: Hydrologic model performance in the calibration and evaluation periods. .. 170
Table 5-5: Seasonal precipitation change relative to the reference period ..................... 174
Table 5-6: Annual relative change of temperature and precipitation of the average GCMs for the two emission scenarios .................................................................................. 174
Table 5-7: Scenario analysis for BMA Nash Sutcliffe performance for summer and all seasons. Note that for each model included, three objective functions were considered to create plausible models. .................................................................................. 188
Table 1: Potential predictors with their respective elevations. Elevations listed as pressure correspond to the height at which the air pressure occurs. ............................................ 225
Table 2: Optimization process of the cluster radius in ANFIS initialized by subtractive clustering (e.g. winter predictors) ......................................................................................... 232
Table 3: Number of clusters used for the final predictor combination (ANFIS initialized by fuzzy c-means clustering).................................................................................................................. 233

Table 4: Determined combinations of predictors for each month and season based on MLR-IC.......................................................................................................................................................... 239

Table 5: Comparison of NSE values for training, checking and testing datasets for sigma=4 and C=50, 100, 1000........................................................................................................................................... 242

Table 6: Models’ performances for each season................................................................................................. 245
List of Figures

Figure 1-1: Example scenario depicting observed and simulated data CDFs over the historic time period from 1980-1998. .................................................................................................................. 14

Figure 1-2: Simulated historic (1980-1998) and future (2040-2069) 2-year return levels for all RCMs. ................................................................................................................................. 16

Figure 1-3: Simulated historic (1980-1998) and future (2040-2069) 25-year return level magnitudes for all RCMs. ....................................................................................................................... 17

Figure 1-4: Representative GPD distribution shape parameter ($\xi$) for historic (1980-1998) and future (2040-2069) RCM simulations. .................................................................................................. 18

Figure 1-5: Results from bias-correction procedure. Taking the values from the WRFG-CCSM simulation and the UW observation datasets for August and January, then calculating mean monthly precipitation values [mm] from 1980-1998. ........................................ 20

Figure 2-1: Geographical locations of the two study areas of CRBN and CRBS. ............ 38

Figure 2-2: Variogram plots for the two study areas with the bounding lines reflecting the limits of the prior uniform distributions chosen for the covariance parameters. .......... 39

Figure 2-3: Number of available gage sites for each fifteen year time period. ............. 40

Figure 2-4: Running mean plots to check mixing of the chains; results shown for the shape parameter and the scale parameter for one of the gages in CRB- North. M represents the mean of the samples up to each iteration. ......................................................... 42

Figure 2-5: Autocorrelation (i.e. the correlation between every draw and its $i^{th}$ lag) plots for the GPD shape parameter and the scale parameter for one of the gages in CRB- North, after thinning of 30 iterations; low values indicate satisfactory mixing of the MCMC samples. ........................................................................... 43

Figure 2-6: The histograms of exceedance rates for a) CRBN and b) CRBS. .......... 45

Figure 2-7: Spatial Hierarchical Bayesian model; the dark green density functions correspond to priors while the light green ones display the posterior distributions. Uninformative priors (with infinite integral under the distribution) are used for the mean latent parameters ($\lambda$) in $(-\infty, \infty)$ and for the covariance latent parameters ($\chi$) in the ranges shown in this figure. ............................................................................................................. 48
Figure 2-8: Use of the estimated scale and shape parameters along with the exceedance rate values (not shown) to calculate the 100-year return level flood; results shown for USGS gage#14220500, located in the CRB_N. ................................................................. 49

Figure 2-9: Analysis for CRB_N (a) and CRB_S (b) regions: Comparing the 95% confidence interval of 100-year flood by fitting the GPD distribution to each gage separately (left), and the one from the Spatial Hierarchical Bayesian Model (right). ........................................... 50

Figure 2-10: Evaluating the performance of the Spatial Hierarchical Bayesian model in gages not being used during the model setup; comparison is made based on separate GPD fit for each gage station (shown by triangles). The USGS gage numbers are shown on top of each panel followed by N and S representing CRB_N and CRB_S; the vertical lines represent the 95% confidence intervals of the hierarchical model estimations. ............. 52

Figure 2-11: Sensitivity analysis of the covariance latent parameter (\(\chi\)) and the scale parameter (\(\psi\)) to changes in prior distribution in a) CRB-North and b) CRB-South. ...... 55

Figure 2-12: 95% confidence intervals and the median of the GPD scale parameter for the CRB- North (top panels) and CRB- South (bottom panels) based on the results of the Spatial Hierarchical Bayesian Model. Triangle marks show the location of the streamflow gages. ........................................................................................................................................ 57

Figure 2-13: 95% confidence intervals for 25, 50 and 100- Year return level floods for the CRB- North based on the results of the Spatial Bayesian Hierarchical Model, the size of the circles represent the magnitude of the return levels.......................................................... 58

Figure 2-14: 95% confidence intervals and the median of the 25, 50 and 100- Year return level floods for the CRB- South based on the results of the Spatial Bayesian Hierarchical Model; , the size of the circles represent the magnitude of the return levels......................... 59

Figure 2-15: Change of the GPD scale parameter (top) and 100- Year return level (bottom) for three periods of 1965-1980, 1980-1995 and 1995-2011 in CRB-South. ..... 61

Figure 2-16: Change of the GPD scale parameter (top) and 100- Year return level (bottom) for three periods of 1965-1980, 1980-1995 and 1995-2011 in CRB-North. ..... 62

Figure 2-17: Changes of the 100- Year return level floods and the shape parameters over time periods of fifteen years starting from 1905 (CRB- North) and 1935 (CRB- South); Box- plots show the medians of the MCMC samples for the stations........................................ 64

Figure 3-1: The Pacific Northwest Region. ........................................................................................................... 72
Figure 3-2: Flowchart of the Modeling Procedure. ................................................................. 83
Figure 3-3: Posterior estimate mean of the location parameter for each grid shown for the current period of 1971-2000 on left and future period of 2041-2070 on right for (a) CRCM_ccsm (b) CRCM_cgcm3 (c) HRM3_gfdl (d) HRM3_hadcm3 (e) RCM3_cgcm3 (f) RCM3_gfdl (g) WRFG_ccsm and (h) WRFG_cgcm3 ................................................................. 95
Figure 3-4: Posterior estimate mean of the scale parameter for each grid shown for the current period of 1971-2000 on left and future period of 2041-2070 on right for (a) CRCM_ccsm (b) CRCM_cgcm3 (c) HRM3_gfdl (d) HRM3_hadcm3 (e) RCM3_cgcm3 (f) RCM3_gfdl (g) WRFG_ccsm and (h) WRFG_cgcm3 ................................................................. 100
Figure 3-5: Posterior estimate mean of the shape parameter for each grid shown for the current period of 1971-2000 on left and future period of 2041-2070 on right for (a) CRCM_ccsm (b) CRCM_cgcm3 (c) HRM3_gfdl (d) HRM3_hadcm3 (e) RCM3_cgcm3 (f) RCM3_gfdl (g) WRFG_ccsm and (h) WRFG_cgcm3 ................................................................. 105
Figure 3-6: Posterior mean of the 100-year return level runoffs (mm) for each grid cell estimated based on observation versus downscaled RCM data which were used as VIC model forcing (a) Winter (b) Spring (c) Summer (d) Fall ......................................................... 108
Figure 3-7: Difference between the posterior mean of the 100-year return level runoffs (mm) for each grid cell estimated based on observation versus downscaled RCM data which were used as VIC model forcing (a) Winter (b) Spring (c) Summer (d) Fall ...... 112
Figure 3-8: Change in the posterior mean of 100-year return level runoff (mm) estimated for each grid cell for the future period of 2041-2070 versus the current period of 1971-2000 (a) CRCM_ccsm (b) CRCM_cgcm3 (c) HRM3_gfdl (d) HRM3_hadcm3 (e) RCM3_cgcm3 (f) RCM3_gfdl (g) WRFG_ccsm and (h) WRFG_cgcm3 ......................................................... 114
Figure 3-9: Posterior standard deviation of the estimated 100-year return level runoff (mm) for each grid shown for the current period of 1971-2000 on left and future period of 2041-2070 on right for (a) CRCM_ccsm (b) CRCM_cgcm3 (c) HRM3_gfdl (d) HRM3_hadcm3 (e) RCM3_cgcm3 (f) RCM3_gfdl (g) WRFG_ccsm and (h) WRFG_cgcm3 ......................................................... 119
Figure 3-10: Posterior distribution of the precipitation coefficient for the mean (top), scale (middle) and shape (bottom) parameters for each season............................................. 120
Figure 3-11: Posterior distribution mean (Top) and standard deviation (Bottom) of the precipitation coefficient for all RCMs confronting current and future time periods; results shown for (a) mean, (b) scale and (c) shape parameters. 124

Figure 3-12 Aridity index (top) and mean precipitation (bottom) for historical period of 1971-2000 and future period of 2041-2070 as well as the changes over the PNW 127

Figure 3-13: Mean precipitation, 100-yr return level precipitation and runoff for winter and summer over the historical period of 1971-2000. 127

Figure 3-14: Change in the mean precipitation, 100-yr return level precipitation and runoff between the future period of 2041-2070 and 1971-2000. 128

Figure 4-1: Spatially distributed weights obtained for each RCM based on daily precipitation. 145

Figure 4-2: Spatially distributed weights obtained for each RCM based on daily maximum temperature. 146

Figure 4-3: Spatially distributed weights obtained for each RCM based on daily minimum temperature. 147

Figure 4-4: Spatially distributed weights obtained for each RCM based on daily wind speed. 147

Figure 4-5: Spatially distributed weights obtained for each RCM based on daily runoff from the VIC hydrologic model. 148

Figure 4-6: Ranges of the spatially averaged weights corresponding to eight RCMs based on each hydro-climate variable. 149

Figure 4-7: Comparing the 100-yr return level runoffs obtained from the observed forcing data (top) with scenarios of (from top to bottom respectively): assigning the BMA weights on estimated extremes, BMA of daily runoff estimates followed by extreme assessment, and BMA of RCM data (precip, Tmax, Tmin and wind) followed by hydrologic modeling and extreme analysis. 151

Figure 4-8: Biases of different multi-modeling approaches (sim-obs) including (top to bottom): assigning the BMA weights on estimated extremes, BMA of daily runoff estimates followed by extreme assessment, and BMA of RCM data (precip, Tmax, Tmin and wind) followed by extreme assessment, and BMA of RCM data (precip, Tmax, Tmin and wind) followed by hydrologic modeling and extreme analysis. 152
Figure 4-9: Between model standard deviation over 1971-2000 (red) and 2041-2070 (blue); results obtained from the estimated BMA weights on the daily runoff data and 100-yr return level runoffs; each circle (star) corresponds to a 1/8° cell.......................... 154

Figure 4-10: Multi-model average results showing the estimated 100-year return level runoff (mm) obtained for the current and future periods (top) and their corresponding differences (bottom)................................................................. 156

Figure 4-11: Bayesian model average results of the estimated 100-year return level runoff (mm) obtained from the hierarchical Bayesian model for the current versus future time periods; each circle shows one of the 6392 grid cells with 1/8 degree resolution in the Columbia River basin. .......................................................................................... 157

Figure 5-1: Tualatin River Basin, river network, climate stations, and the USGS streamflow gauging station. .......................................................................................................................... 162

Figure 5-2: Projected relative seasonal variation of the precipitation and temperature for three time-slices (2020s, 2050s, and 2080s) compared to the reference period of 1960-1989......................................................................................................................... 173

Figure 5-3: Runoff changes relative to the reference period showing the uncertainties in the GCMs for each hydrologic model for winter and summer seasons; RMSE, BIAS and HMLE reflect the hydrologic model parameter set. ............................................................................................................. 179

Figure 5-4: Probability density functions of runoff change in fall and summer seasons for 8 GCMs under A1B and B1 emission scenarios. The models were calibrated using the RMSE objective function........................................................................................................... 179

Figure 5-5: Hydrologic model uncertainty in high flow using the ccsm downscaled forcing data .................................................................................................................................................. 181

Figure 5-6: The weights obtained by the BMA for 12 hydrologic models.................. 185

Figure 5-7: Comparison of the BMA result with observation and uncertainty range obtained from 12 hydrological models for the period of 1983-1985............................................... 186

Figure 5-8: Relative changes of streamflow for the winter (wet season, upper panel) and the summer (dry season, lower panel) obtained from the BMA and the average GCMs; the bars show the extreme changes observed in the 12 hydrologic model results for all GCMs.................................................................................................................. 188
Figure 5-9: The range of total standard deviation of the 12 hydrologic models for each GCM in the reference and future time periods in A1B (upper panel) and B1 (lower panel) emission scenarios ................................................................. 189

Figure 1: Upper Willamette basin (left) in the Oregon state, U.S. along with the GCM data grids (red dots right) ......................................................................................................................... 224

Figure 2: a) “Sugeno” fuzzy model with two inputs (surface_specific_humidity(huss) and pressure at sea level (psl); b) The ANFIS model architecture .................................................................................. 229

Figure 3: The procedure for predictor selection ................................................................................................................................. 235

Figure 4: Root Mean Square Error (RMSE) of monthly and seasonal precipitation with respect to their corresponding MLR-IC predictors: a) ANFIS (grid partitioning) simulation; b) ANFIS (subtractive clustering) simulation; c) ANFIS (fuzzy c-means clustering) simulation; d) SVM simulation ...................................................................................... 241

Figure 5: Downscaled monthly precipitation using ANFIS (grid partitioned) and MLR-IC for the test data set .............................................................. 243

Figure 6: Downscaled seasonal precipitation using seasonal predictors ................................................................................................................. 244

Figure 7: Uncertainty range obtained from the ensemble of downscaled data using various techniques at the monthly scale compared with mean observed value shown by straight solid line ................................................................................................................................. 246

Figure 8: Uncertainty range in downscaled precipitation data at seasonal time scales compared with the mean observed shown by straight solid line .................................................................................. 247
1. Chapter 1. Introduction

1.1. Climate Change

Projected changes in climate show an increase in temperature, and the alterations in the spatio-temporal pattern and amount of precipitation over the globe (Solomon, 2007). The hydrologic impacts of these climatological changes include: reduction of snow cover and earlier onset of snow ablation, higher evaporation and transpiration, and spatial and temporal changes in the regional water resources (Brooks, 2009). Using a high resolution hydrologic model forced by General Circulation Models (GCMs), (Barnett et al., 2008) showed that up to 60% of the climate-related trends of river streamflow for 1950-1999 are human-induced. Though water resources will be affected by global climate change, the vulnerabilities to the resulting impacts vary between regions across the continent (Bates et al., 2008).

1.2. General Circulation Models (GCMs)

The primary approach for evaluating potential changes to hydro-climatic variables is through the use of climate models (i.e. General Circulation Models or GCMs) that simulate aspects of the global climate cycle. Over the last three decades, the number and complexity of climate models have increased substantially, more physical processes are simulated and the coupling between individual sea, atmosphere and land-based processes has been improved (Mearns, 2007; Mearns,
Recent advancements in modeling spatial and temporal climate variables at finer scales allow for regional impact analysis studies. The ability to investigate the impact of climatic change at a regional scale has the potential to inform water resource managers and decision makers regarding the changes in the climate cycle that will influence extreme events, i.e. floods and droughts. In order to provide valuable information regarding potential climatic changes, the results from multiple climate model simulations can be investigated and compared to reduce the over-reliance on one model and quantify model uncertainty.

1.3. Downscaling

General Circulation Models (GCMs) are relatively coarse in resolution and therefore unable to resolve significant subgrid scale features including topography and land use as needed in hydrologic modeling and impact assessment. Within the climate modeling community, it has long been speculated that increasing the resolution of climate models is necessary to improve the estimates of regional-scale phenomena, such as precipitation (e.g. Giorgi, 1990; McGregor, 1997; Murphy, 1999; Di Luca et al., 2011; and Caldwell, 2010). The process of downscaling outputs from GCMs has been established as the primary approach for addressing the inadequacies of large scale resolution models. Downscaling is the process of translating large scale atmospheric data, called predictors, to finer spatial resolution data, called
predictands, to allow for the local analysis of climate change effects. The
downscaling process relies on the premise that there exists a linear or nonlinear
relationship between the predictors and the predictand. Air pressure, sea surface
temperature, moisture fluxes, geopotential height and wind field data outputs from
GCMs are commonly used as inputs, or predictors, to the downscaling process
(Crawford et al., 2007; Fowler et al., 2007c; Salathe et al., 2007). By means of the
future values of the GCM variables, downscaling methods can calculate the
climatologic values of temperature and precipitation, which are then used to assess
the hydrologic impacts due to climate change.

Generally two types of downscaling approaches are employed, which are known
as “dynamic” and “statistical” methods. The dynamic downscaling techniques such
as nested or regional climate modeling are performed by nesting a physically based,
small spatial resolution regional climate model within the grid of a GCM output.
GCM outputs are used as boundary conditions to drive the regional model (Wood et
al., 2004a; Wood et al., 2004b). These models are computationally expensive and
time consuming which limit their applicability (Salathe et al., 2007), however they
consider small scale topographical effects and nonstationarity in climate variables.
The second method, statistical downscaling, is performed by deriving empirical
mathematical relationships between GCM climate variables and regional scale
variables (Wood et al., 2004b). Statistical methods are sufficiently accurate in
representing the relationship between GCM data and mean observation (e.g. precipitations) over the regional basin (Fowler et al., 2007c), and are computationally inexpensive and flexible in their adjustment and movement to different regions (Anandhi et al., 2008; Ghosh and Mujumdar, 2008; Hewitson and Crane, 1996; Tripathi et al., 2006; Wilby et al., 2004). These methods, however, assume stationarity between relationships of large scale climate variables and small scale observed data. An analysis of predictor selection and uncertainty assessment in downscaling is provided in section 0.

1.4. Challenges in Climate Modeling

Model bias exists within climate models for multiple reasons. Commonly identified causes of bias are attributed to model structure and initial/forcing condition treatment. In order to more accurately compare historic and future climate model simulations, recent studies suggest the use of bias-correction techniques such as the quantile mapping approach e.g (Fowler et al., 2007a; Mote and Salathe, 2010; Shrestha et al., 2011). The need for bias-correction of climate model simulations over future periods is widely accepted for hydrologic impact studies (Wood et al., 2004a), however the relative strengths and weaknesses of each individual correction technique is still a focus of research (Johnson and Sharma, 2011; Najafi et al., 2011c).
Water resources decision makers are also confronted with complex problems attributed from the uncertainty of future hydrologic projection under climate change. Projecting changes in runoff is particularly difficult because it is subject to various uncertainties associated with the future Green House Gas (GHG) emission scenarios, GCM structure, downscaling method, hydrologic model structure and parameters (Najafi et al., 2011b; Wilby and Harris, 2006). These uncertainties confound regional water resources decision makers in sustainable water resources management and conservation (Wood et al., 2004b). Therefore, if the projection uncertainties are quantified properly, they can potentially be reduced or removed through advanced modeling and research (Kay, 2009).

1.5. Ensemble Multi-modeling

Due to the expansion of climate modeling efforts resulting in an abundance of distinct climate models, there is a need to evaluate how these models perform relative to one another. Multiple model inter-comparison projects have been organized to meet this need. On a global scale, the Coupled Model Inter-comparison Project (CMIP) and Atmospheric Model Inter-comparison Project (AMIP) are the most notable collaborations undertaken with this goal in mind (Kreienkamp et al., 2011; Meehl et al., 2007b; Meehl et al., 2005). Beginning in the mid-1990s the World Climate Research Programme (WCRP) committee, now known as the WCRP/Climate Variability and Predictability (CLIVAR) Working Group on Coupled Models (WGCM),
set about to organize one of the first generations of inter-comparison projects (Meehl et al., 2007b). Their efforts have since resulted in multiple CMIP generations, recently culminating in an open-access dataset, the WCRP CMIP3 multi-model dataset which represents “an unprecedented, comprehensive coordinated set of global couple climate model experiments” (Meehl et al., 2007b).

Several regional programs have been conducted in the last decade focused on addressing the need for appropriate scale level assessment of climate change impacts. In Europe, the Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects (PRUDENCE) project described in (Christensen and Christensen, 2007), followed by the Ensembles-Based Predictions of Climate Changes and Their Impacts (ENSEMBLES) project (Van der Linden and Mitchell, 2009) provided an array of regional datasets for investigating future climate variation. The STAistical and Regional dynamical Downscaling of EXtremes for European regions (STARDEX) project focused on the frequency and intensity of twenty-first century extreme events over Europe (http://www.cru.uea.ac.uk/projects/stardex/).

In North America, the North American Regional Climate Change Assessment Program (NARCCAP) provides data from multiple GCM-RCM coupled simulations over the majority of the continent (Mearns et al., 2009). The RCM data used in this
study was provided by NARCCAP efforts. NARCCAP’s goal is the production of climate simulations at a resolution which allows for regional-scale investigation of future climate variation. The products are intended to be useful in generating and studying impact scenarios across much of North America. The program consists of multiple RCMs driven by multiple AOGCMs. Simulations of both future (2041-2070) and historic (1971-2000) periods were produced by the NARCCAP modelers at a spatial resolution of 50km and sub-daily temporal resolution. Future scenarios were forced for the twenty-first century using the Special Report on Emissions Scenarios (SRES) A2 emissions scenario.

1.6. Extreme Analysis

Due to the climate change impacts on hydro-climatic events, analyses of the ability of climate models to capture and simulate extreme events is of paramount importance. (Frei et al., 2006; Villarini et al., 2011) studied the variability in both observed and climate model simulated extreme event occurrence. The potential for change in the occurrence of extreme events in the future in conjunction with projected climate variability is also prevalent in recent studies (Mote and Salathe, 2010; Tryhorn and DeGaetano, 2011). In order to evaluate the characteristics of extreme events, statistical extreme value theory is commonly utilized in water resource and hydrology related studies in recent decades (Katz et al., 2002). Extreme Value Theory (EVT) is a statistical method used for analyzing the tails of
probability distributions of a random variable. The parameters of the extreme value distributions yield estimates of the intensity and frequency of extreme events. As such, the distributions can be used for estimating the magnitude of extreme event return values in the future and historical time periods to investigate the climate change impact on hydro-climate extremes (Katz, 2010). EVT has been widely applied in studies of precipitation (Acero et al., 2010; Katz et al., 2002; Kharin et al., 2010b; Wehner et al., 2010), temperature (Kharin et al., 2010a; Kharin et al., 2010b; Rusticucci and Tencer, 2008), streamflow (Hurkmans et al., 2010; Katz et al., 2002; Lima and Lall, 2010) and wind speed (Brabson and Palutikof, 2010; Caires and Sterl, 2010; Hundecha et al., 2008) among others.

Limited records of hydro-climatic extremes in space and time (Fuentes et al., 2010) requires the collection of information from different locations in order to reduce the uncertainty of the simulations and provide more reliable results. Methods have been developed for regional frequency analysis (RFA) which are shown to be superior to the at site flood estimations (Burn, 1990; Chokmani and Ouarda, 2004; Dalrymple and Survey, 1960; Grehy, 1996; Gupta et al., 2007; Ouarda and El-Adlouni, 2011; Stedinger and Tasker, 1985). Index flood method (Dalrymple and Survey, 1960) is an approach to combine extreme data from different locations in order to improve the accuracy of the estimates and to predict flood at ungagged sites. This method was further improved by (Hosking and Wallis, 1997) who used
the L-moments approach to estimate the parameters of the extreme value distribution. They divided the method into three steps: outline a homogeneous region, divide the extreme data at each gage by the index flood and then fit a distribution to the combined data from all gages. RFA however, does not consider the spatial components of the point data (i.e. geographic coordinate, elevation) and cannot incorporate additional variables (i.e. covariates) into the analysis. Also it is not possible to explicitly estimate the uncertainties based on the L-moments approach.

Recently with the accessible records of spatially scattered or gridded data and high performance computing machines, there is growing interest in the analysis of spatially distributed extremes. Applications are found in the studies of wind (Fawcett and Walshaw, 2006), precipitation and temperature (Aryal et al., 2009; Cooley et al., 2007; Sang and Gelfand, 2010; Schliep et al., 2010) among others. For this purpose, it is possible to consider a univariate extreme value distribution at each point (or grid) generating a spatial model on its parameters (Cooley et al., 2007; Cooley and Sain, 2010; Lima and Lall, 2010; Renard, 2011; Sang and Gelfand, 2009; Schliep et al., 2010). The spatial dependence can also be modeled using the theory of max-stable process (Coles, 1993; De Haan and Pereira, 2006; Padoan et al., 2010). In addition Bayesian approach is a formal way to quantify the uncertainties (Majda and Gershgorin, 2010; Najafi et al., 2011b; Tebaldi et al., 2005) and is flexible.
in combining different sources of uncertainties (Tebaldi et al., 2005). By considering the parameters of the extreme distributions as random variables one would utilize the Bayesian method to find their corresponding distributions.

1.7. Example of At-Site Extreme Assessment:

This section provides an example of extreme precipitation analysis based on extreme value theory using at-site extreme data. The results are shown for the Willamette River basin. Oregon’s Willamette River basin (WRB), covers a drainage area of 29,728 square kilometers (11,478 square miles), roughly twelve percent of the entire state, and intersects or contains thirteen of the thirty-six counties in the state (Hulse et al., 2002). The temperate marine climate of the basin translates into cool wet winters, with 80 percent of annual precipitation occurring between October and May, and warmer mostly dry summers (Lee and Risley, 2002). Average annual temperatures in the region depend primarily on elevation and range from forty to sixty-five degrees Fahrenheit (F). Annual mean precipitation also varies with elevation, from about forty inches at the lowest elevations up to 175 inches at the highest elevations. Precipitation in the form of snow at the higher elevations within the basin is an influential component of the overall water cycle. Recent studies estimate that as much as 75 percent of precipitation falls as snow at or above 6500 feet (Chang and Jung, 2010). Above 4000 feet, thirty-five percent of precipitation falls as snow (Lee and Risley, 2002).
The NARCCAP project provides dynamically downscaled GCM outputs at a spatial resolution averaging 50km. Data from six RCM simulations were available at the time of this study; these RCMs are listed in Table 1-1.

Table 1-1: Regional Climate Model Information

<table>
<thead>
<tr>
<th>Model</th>
<th>Aliases</th>
<th>Modeling Group</th>
<th>Full Name</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECP2</td>
<td>ECPC, RSM</td>
<td>UC San Diego/Scripps</td>
<td>Experimental Climate Prediction Center Regional Spectral Model</td>
<td><a href="http://www.emc.ncep.noaa.gov/mmb/RSM/">http://www.emc.ncep.noaa.gov/mmb/RSM/</a></td>
</tr>
<tr>
<td>HRM3</td>
<td>PRECIS,HadRM3</td>
<td>Hadley Centre</td>
<td>Hadley Regional Model 3</td>
<td><a href="http://www.metoffice.gov.uk/precis/">http://www.metoffice.gov.uk/precis/</a></td>
</tr>
<tr>
<td>MM5I</td>
<td>MM5, MM5P</td>
<td>Iowa State University</td>
<td>MM5-PSU/NCAR mesoscale model</td>
<td><a href="http://www.mmm.ucar.edu/mm5/">http://www.mmm.ucar.edu/mm5/</a></td>
</tr>
<tr>
<td>RCM3</td>
<td>RegCM3</td>
<td>UC Santa Cruz</td>
<td>Regional Climate Model version 3</td>
<td><a href="http://users.ictp.it/~pubregcm/RegCM3/M3/">http://users.ictp.it/~pubregcm/RegCM3/M3/</a></td>
</tr>
</tbody>
</table>

The RCMs were driven by boundary condition datasets provided by the AOGCMs.

Four distinct AOGCMs were selected by NARCCAP in order to provide boundary conditions required as inputs to the RCMs. These AOGCMs are listed in Table 1-2 along with the group name, aliases, and other distinct information regarding the model differences.
Table 1-2: General Circulation Model (Drivers for RCMs) Information

<table>
<thead>
<tr>
<th>Model</th>
<th>Full Name</th>
<th>Modeling Group</th>
<th>Ensemble Member Used</th>
<th>More Information (References)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCSM</td>
<td>Community Climate System Model</td>
<td>National Center for Atmospheric Research (NCAR),</td>
<td>b30.030e (ctl), b30.042e (fut)</td>
<td><a href="http://www.cesm.ucar.edu/">http://www.cesm.ucar.edu/</a></td>
</tr>
<tr>
<td>CGCM3</td>
<td>Third Generation Coupled Global Climate Model</td>
<td>Canadian Centre for Climate Modeling and Analysis</td>
<td>CGCM #4</td>
<td><a href="http://www.ec.gc.ca/ccmac-ccma/default.asp?lang=En&amp;n=1299529F-1">http://www.ec.gc.ca/ccmac-ccma/default.asp?lang=En&amp;n=1299529F-1</a></td>
</tr>
<tr>
<td>GFDL</td>
<td>Geophysical Fluid Dynamics Laboratory GCM</td>
<td>GFDL/NOAA</td>
<td>20C3M, run2; sresa2, run1</td>
<td></td>
</tr>
</tbody>
</table>

Precipitation rate data [kgm⁻²s⁻¹], at a temporal resolution of three hours, was obtained over both a historical period (1979-2004 for the NCEP reanalysis driven data and 1976-2000 for the GCM driven data) as well as a future period (2038-2069).

To bias-correct the climate simulations, the University of Washington (UW) gridded dataset, described by (Maurer et al., 2002) was used. This dataset covers the time period 1950-2000 and provides surface level information regarding numerous climatic variables at three-hourly time intervals. Specifically for this study, the UW dataset provides values of total daily precipitation over the continental United States obtained from the National Oceanic and Atmospheric
Administration’s (NOAA) Cooperative Observer (Co-op) stations (Maurer et al., 2002). The precipitation data over the WRB used in this study was obtained at 1/8th degree resolution and served as an observational benchmark upon which the dynamically downscaled NARCCAP datasets were compared.

1.7.1. Data Bias-correction

For this study, the quantile mapping approach was implemented on the NARCCAP datasets. In this procedure, the observed and simulated datasets are each characterized in terms of their full distribution of daily values for each month, a so-called non-parametric approach (Johnson and Sharma, 2011) since it does not rely on adjusting the mean, standard deviation, or other standard statistical parameters. As in most bias-correction approaches, a scaling factor is developed between the observed and simulated datasets over a historic period. For both observed and simulated datasets, the cumulative distribution functions (CDFs) are computed on a monthly basis. Figure 1-1 shows the CDFs for an example case. After computing the CDFs, the scaling factor determined based on the respective quantile values during the observed period are then applied for the projected (future) period.
1.7.2. Extremes

The Generalized Pareto Distribution (GPD), from the Extreme Value Theory (EVT), models the probability distribution of exceedances over the threshold ‘u’ [S Coles, 2001]. 95% quantile of the data was selected as the GPD threshold. Results from the extreme analysis are displayed in Figure 1-2, Figure 1-3 and Figure 1-4. The return level analysis in Figure 1-2 and Figure 1-3 reveal noticeable differences between both RCM simulations as well as GCM boundary condition influence. Figure 1-2 displays the spatial distribution of the 2 year return level magnitudes over the WRB as simulated by each of the RCM-GCM datasets modeled via representative GPD distributions. The 2 year return level magnitudes vary substantially between the RCMs, with values ranging from below 75 mm/day to
over 135 mm/day. The upper limit of the 25 year return level magnitudes extends to values in excess of 150 mm/day. Some very clear differences between both RCMs and GCMs are evident in these figures. It is notable that only the HRM3-HADCM3 simulation dataset contains magnitudes over 110 mm/day for the 2 year return level. The majority of future simulations do not reveal a substantial increase in the magnitude of the 2 year return level. The 25 year return levels also do not change dramatically between the historic and future periods. As with the 2 year return level estimates, the HRM3-HADCM3 simulations exhibit the highest magnitudes, however other model combinations, such as WRF-CGCM3&CCSM and RCM3-GFDL, also exhibit high magnitudes. The spatial distribution of these magnitudes is also informative. The topography of the WRB, high elevations on the Eastern and Western edges and low valley floor in between, should influence the distribution of precipitation, theoretically with more precipitation falling at higher elevations. However in most of the model simulations the topographical influence is not discernible.
Figure 1-2: Simulated historic (1980-1998) and future (2040-2069) 2-year return levels for all RCMs.
The shape parameter of the estimated GPD is displayed in Figure 1-4 across the WRB. In terms of the simulated precipitation events over the WRB, a positive shape parameter indicates the presence of a heavy upper tail (higher likelihood of extreme magnitude events), a negative shape parameter indicates a bounded distribution (an identifiable upper limit to those extreme events), and when the shape parameter is equal to zero the distribution is unbounded but has a thin upper tail. However, when the distribution is wider, or unbounded, the uncertainty in the distribution increases. Areas in green represent regions where the shape parameter is negative,
yellow areas represent shape parameter values that are positive but nearly zero, and red and pink areas are indicative of areas where the estimated distribution's shape parameter is above 0.1. The shape parameter results also reveal differences due to RCM characteristics as well as the influence of GCM driving conditions on the behavior of the RCMs. From Figure 1-4 it is evident that the GFDL (a GCM) simulations are represented by unbounded heavy upper tails regardless of RCM and time period.

Results of the quantile mapping bias-correction technique are displayed in Figure 1-5, for the WRFG-CCSM simulation. Mean monthly precipitation values for
January and August over the historical period, 1980-1998, are displayed over the WRB both before and after the bias-correction procedure was applied. These two months were selected because they represent both heavy precipitation, January, and relatively low precipitation, August, months. The observed gridded dataset used to bias-correct the simulated data is also displayed for the selected months in Figure 1-5 row 1.
Figure 1-5: Results from bias-correction procedure. Taking the values from the WRFG-CCSM simulation and the UW observation datasets for August and January, then calculating mean monthly precipitation values [mm] from 1980-1998.
Rows 2 and 3 of Figure 1-5 display the simulated data before and after, respectively, the application of the bias-correction procedure. Rows 4 and 5 of Figure 1-5 show the difference, simulated-observed, between the datasets before and after bias-correction. The bias is stronger in the month of January due to the higher magnitude of precipitation during the month, whereas August, a relatively dry month over the WRB, reveals a smaller magnitude in the bias. During the month of January the simulation tends to over predict precipitation in the WRB. This positive bias is an attribute that has been documented before in dynamically downscaled datasets (Caldwell, 2010). The bias present in the simulation dataset during the month of August is also slightly positive, demonstrated by the range of bias values present prior to bias-correction. The bias-correction procedure effectively corrects for this positive bias and is able to reduce the overall magnitude of bias as well.

1.7.3. Summary

In this study, precipitation data from the multiple RCMs and GCMs was investigated over the WRB. A bias-correction scheme, and extreme value analysis was applied to the data. The results of this study demonstrate two key facts regarding the use of dynamically downscaled climate datasets. First, applying a bias-correction scheme to any downscaled dataset is a needed and important step yielding more accurate results. For the NARCCAP datasets, the quantile mapping
procedure was implemented and successfully reduced the difference between observed and simulated precipitation over the WRB by correcting for the positive bias that is present in RCM datasets and reducing overall bias magnitude. Second, implementing fundamentals of EV theory to climate datasets provides estimates of changes to variable values, such as precipitation, due to climate change. Using the GPD distribution, this study obtained estimates of changes to 2 and 25-year extreme precipitation event magnitudes over the WRB. The results indicate that these return level magnitudes will increase in the future period 2038-2069 compared to simulations over the historical period 1980-1998. As well, the shape parameter of the GPD distribution derived from the NARCCAP datasets indicates that the RCM models in the NARCCAP study provide different depictions of future changes over the WRB.

1.8. Outline

In this thesis the problems associated with the limited number of extreme events for climate change studies, uncertainty quantification and multi-modeling analyses are addressed. In chapter 2 the spatial hierarchical Bayesian method is presented as an approach to model extreme runoffs. In this study the scale parameter of the Generalized Pareto Distribution (GPD) is spatially modeled and a homogenous shape parameter is assumed over the region. The model is then employed to estimate historical changes in flood events. Chapter 3 also presents the spatial
hierarchical Bayesian model for the analysis of runoff extremes. This study differs from the previous one in that it uses grid based data in a large area (6392 grid cells with 1/8° resolutions); Generalized extreme value (GEV) distribution is used rather than GPD; all the parameters of the GEV distribution are spatially modeled; and the analyses are performed over all seasons using several regional climate model datasets. Hydrologic modeling is performed by the variable infiltration capacity (VIC) model. In chapter 4 Bayesian model averaging (BMA) approach is presented for a multi-model ensemble analysis of the extreme runoff. Three combination scenarios are defined and compared for multi-modeling of extreme runoffs. The biases of each scenario are calculated and the scenario with the least bias is selected for projecting seasonal runoff extremes. Multi-modeling uncertainties are also compared between the historical and future time periods. Chapter 5 deals with the uncertainties associated with GCMs and hydrologic models by means of multi-modeling. The models are calibrated based on three objective functions to create more plausible models for the study. The hydrologic model simulations are then combined using the Bayesian Model Averaging (BMA) method according to the performance of each model in the observed period, and the total variance of the models. Chapter 6 summarizes and concludes the main findings of the thesis and provides recommendations for future studies. Appendices are given in chapter 0. Appendix A, presents a procedure for predictor selection in statistical downscaling
of large scale climate model outputs including precipitation. The presented method incorporates several GCM grids in the downscaling process which allows for considering more predictors in the model calibration and removes the predictors’ correlation and dependence through independent component analysis. Also, the study utilizes several downscaling techniques in order to develop an ensemble of precipitation time series that can be used in hydrologic climate impact assessment. Appendix B illustrates the Markov Chain Monte Carlo technique used for sampling from the posterior distributions of the hierarchical Bayesian model presented in chapter 2.
2. **Chapter 2. Analysis of Runoff Extremes using Spatial Hierarchical Bayesian Modeling**

2.1. **Background**

There are several factors which could affect the mean and peak of flow in a basin. Changes in precipitation and temperature as a result of climate change can directly influence the streamflow trend (Arnell et al., 2001; Arnell, 2004; Gao et al., 2011; Kundzewicz et al., 2008). The time of snowmelt which is dependent on seasonal temperature may also affect the peak flows through shifts from spring to winter due to the long term increase in climate temperature [Jung et al., 2012]. Although analyzing the contributing factors in streamflow generation provides a general view of its behavior, it does not suffice as an accurate image of the characteristics of a flow regime in a basin. Studying the historical records of streamflow is a meaningful way to detect long term trends in the face of natural variations as a result of climate and land use change (Fu et al., 2010; Karl and Melillo, 2009; Piao et al., 2010).

The U.S. Geological Survey (USGS) maintains streamflow records at gage sites across the US as a valuable resource for flood estimation and water resource management. However, gage stations are not uniformly distributed across the river basins with fewer stations in mountainous regions (Durrans and Tomic, 1996). Also
the recently installed gages provide short records of data which makes it almost impossible to perform a robust flood assessment.

The analysis of runoff extremes would be possible through the extreme value theory (Beirlant, 2004; Coles, 2001) comprising of various extreme value distributions. Commonly the block maximum (Huerta and Sansó, 2007; Kharin and Zwiers, 2005; Sang and Gelfand, 2010) and the extreme over a specific threshold (Durman et al., 2001) are adopted in hydrologic applications. Several studies have performed at-site analysis of extreme events (Frei et al., 2006; Guttorp and Xu, 2011; Halmstad et al., 2012; Kharin et al., 2007; Towler et al., 2010; Villarini et al., 2011).

Attempts to develop models for characterizing the spatial dependencies in hydro-climate data and the underlying processes root back to 1950’s (Cressie, 1992). (Besag, 1974) introduced the concept of Markov random fields which consist of random variables with Markov properties. Hierarchical spatial models became popular after the introduction of Markov Chain Monte Carlo methods (Berliner, 1996; Wikle et al., 1998). (Arab et al., 2007; Banerjee et al., 2004) and Arab et al. [2007] provide detailed explanation of the origins of the hierarchical spatial models. (Casson and Coles, 1999) performed one of the earliest studies on hierarchical spatial modeling over the hurricane wind speed. Recently, there has been growing interest in the development and application of hierarchical spatial models for
climate variables (Fawcett and Walshaw, 2006; Sang and Gelfand, 2009; Sang and Gelfand, 2010; Schliep et al., 2010). (Cooley et al., 2007) analyzed the precipitation extremes over Colorado using 56 gage records. (Cooley and Sain, 2010) studied the changes in precipitation extremes using regional climate model data for historical and future periods. (Sang and Gelfand, 2009; Schliep et al., 2010) utilized a spatial autoregressive model for the annual maximum rainfall. (Renard, 2011) developed a procedure to account for spatial dependency using copulas in the analysis of rainfall extremes.

Although much attention has been directed towards climate variable extremes such as precipitation and temperature and their spatial variations, fewer attempts have been made to provide more reliable models for hydrologic extremes such as streamflow. (Lima and Lall, 2010) employed a hierarchical Bayesian model based on a block maxima distribution and incorporated the drainage area as an additional variable to illustrate the parameters of the distributions at each gage station. However, no spatial analysis was performed for the hierarchical model. Furthermore this approach disregards other extreme events which are lower than the block maximum. Considering the fact that extreme events are rare, the peaks over threshold approach would provide additional data for the analysis (Coles, 2001).
In this study a procedure is developed to model the runoff extremes recorded at USGS gage stations given their spatial variations (e.g. latitudes and longitudes), drainage areas and elevations. This is done by spatially modeling the parameters of an extreme distribution through a hierarchical Bayesian process with latent parameters considered as random variables and simulated using Markov Chain Monte Carlo techniques. Estimates of the return levels for gages not being used in model fitting process are compared to point fit model results in order to validate the procedure. Furthermore, the trend in runoff extreme is assessed using the estimated parameters for time windows of fifteen years, starting from 1906 until 2011.

Section 2.2 will start by the extreme value theory and the generalized Pareto distribution that builds the basis for the hierarchical Bayesian model. The spatial hierarchical Bayesian modeling is then illustrated along with the MCMC parameter estimation. In section 2.3 the case study is presented along with the model test and trend analysis of the runoff extremes followed by concluding remarks, a brief summary and future research in section 2.4.

2.2. Theory and Methodology

In this section a procedure is presented to model the runoff extremes recorded at gage stations, given their spatial variations (e.g. latitudes and longitudes), drainage areas and elevations. This is done by spatially modeling the parameters of an extreme distribution through a hierarchical Bayesian process with latent
parameters considered as random variables and simulated using Markov Chain Monte Carlo techniques.

2.2.1. Extreme Value Theory

The extreme value analysis is developed to illustrate the tail of a distribution (Coles, 2001). Based on asymptotic outcomes it provides different classes of distributions for characterizing extremes. In hydrologic applications the generalized extreme value (GEV) distribution is commonly used when data are considered in annual, seasonal or monthly blocks in which their block maxima are taken as independent and identically distributed random variables. A pitfall of the GEV distribution is that it only considers one extreme value (the maximum) in each block and ignores other useful data. In order to increase the number of data the peaks-over-threshold approach is preferred and used in this study. Having a marginal distribution of \( F \), the distribution of the threshold exceedances is given by:

\[
\Pr(Y > \nu + x | Y > \nu) = \frac{1 - F(\nu + x)}{1 - F(\nu)}
\]  

(2-1)

where \( x > 0 \), and \( Y \) represents a set of \( iid \) random variables of \( Y_i \) (e.g. daily runoff). \( \nu \) is the threshold to distinguish the extreme events of \( Y \). \( F \) is not known in practice, therefore this distribution is approximated for high values of \( \nu \). The asymptotic result (for sufficiently large \( \nu \)) suggests that the observations exceeding \( \nu \) converge to the generalized Pareto distribution (GPD) (Pickands, 1975):

\[
F_{GPD}(y) = \left(1 + \frac{\xi(y - \mu)}{\sigma}\right)^{-\frac{1}{\xi}}
\]
Pr\left(Y > y \mid Y > \nu\right) = \left[1 + \kappa \left(\frac{y - \nu}{\sigma}\right)\right]^{-1/\kappa} \tag{2-2}

with the probability density function of:

\begin{equation}
\begin{aligned}
p(y) &= \frac{1}{\sigma} \left[1 + \kappa \left(\frac{y - \nu}{\sigma}\right)\right]^{-1/\kappa-1} \\
\end{aligned}
\end{equation} \tag{2-3}

where \(\kappa\) and \(\sigma\) represent the shape and scale parameters respectively and \(y > \nu\).

The shape parameter characterizes the tail of the distribution resulting in exponential type (light tail), Pareto type (heavy tail) and beta type (bounded tail) distributions. GPD considers all data exceeding \(\nu\) which highlights the importance of the threshold that affects the GPD scale parameter. In general one may choose the flood level in a region as a threshold. However, the threshold is generally selected statistically with a trade-off between a higher value providing a better GPD approximation, and a lower one providing a more reliable estimation. There are several methods available for threshold selection including mean residual life, dispersion index, and threshold choice (Anagnostopoulou and Tolika, 2011; Lang et al., 1999).

Let’s denote \(\varepsilon = P(Y > \nu)\), then above equation can be written as:
The exceedance rate $\varepsilon$ is the proportion of runoff observations exceeding the threshold $\nu$.

### 2.2.2. Spatial Hierarchical Bayesian Modeling

The power of the hierarchical Bayesian model compared to other extreme analysis techniques relies on the spatial modeling in the process stage which enables it to borrow strength from data in different locations. The spatial hierarchical model is divided into the so called data, process and prior stages (Arab et al., 2007; Banerjee et al., 2004; Cooley et al., 2007; Sang and Gelfand, 2009). In the first stage of the hierarchy (i.e. the data stage) the extreme data is modeled. At location $i$ in time $t$, $Y_{i,t}$ exceeds the threshold $\nu_i$ and follows the generalized Pareto distribution with separate scale parameters ($\sigma_i$) and exceedance rates ($\varepsilon_i$) at each gage. A homogenous shape parameter ($\kappa$) over the study region was assumed in this study. $Y_i$’s are assumed to be independent and conditional on $\sigma_i$. The likelihood function in the first stage of the model is then the product of the GPD equation with respect to space and time.

$$p(\mathbf{y}|\psi_{\sigma=1}, \kappa) = \prod_{i=1}^{n} \prod_{t=1}^{T} \frac{1}{\exp(\psi_{i,t})} \left[ 1 + \kappa \left( \frac{y_{i,t} - \nu_i}{\sigma} \right) \right]^{-\frac{1}{\kappa}}$$  \hspace{1cm} (2-5)
with $\psi_i = \log(\sigma_i)$.

In the second stage of the model, the GPD parameter $\psi(i)$ is defined based on a latent process that affects the runoff extremes. The scale parameter $(\psi)$ is specified through a Gaussian process:

$$p_z(\psi_{a1}, \psi_{b1}) = N(\mu_{\psi_{a1}}, \Sigma_{\psi_{a1}}) = \frac{1}{(2\pi)^{\frac{1}{2}}|\Sigma_{\psi_{a1}}|^\frac{1}{2}} \exp\left[-\frac{1}{2}(\psi_{a1} - \mu_{\psi_{a1}})^T \Sigma_{\psi_{a1}}^{-1}(\psi_{a1} - \mu_{\psi_{a1}})\right]$$  \hspace{1cm} (2-6)

This corresponds to a multivariate normal distribution with the mean and covariance defined by latent parameters of $\lambda$ and $\chi$. $n$ denotes the total number of gage stations for each area under study, and $q$ the total number of covariates plus one (as the intercept variable). Consequently a distinct distribution is generated for the scale parameter at each gage station. The scale parameter can be modeled in several ways depending on the definition of its mean ($\mu_\psi$) and covariance ($\Sigma_\psi$).

$$\mu_{\psi_{a1}} = \lambda_{\psi_1} + \lambda_{\psi_2} \times (Cov1)_{a1} + \lambda_{\psi_3} \times (Cov2)_{a1}$$  \hspace{1cm} (2-7)

where $Cov$ is the covariate such as elevation, area or any other physical/geophysical characteristics of the basin. Spatially distributed hydrologic data are either recorded at gage sites (i.e. geostatistical) or at areal units of grids or pixels (i.e. lattice). For the spatial analysis of the gage data, variograms, and for grid data the conditional, intrinsic and simultaneous autoregressive (CAR, IAR, SAR) models are commonly used (Banerjee et al., 2004; Cressie, 1992). Having the daily runoff data at gage location $i$ exceeding the threshold $\nu$ (i.e. $Y(i)$), we utilize a
A geostatistical model based on variograms to spatially define the covariance of the scale parameter (i.e. $\sigma(i)$). This is based on the assumption that $E[\sigma(i + l) - \sigma(i)] = 0$ a so called intrinsically stationary process. The variogram $2\gamma(l)$ is then defined as:

$$E[\sigma(i + l) - \sigma(i)]^2 = \text{var}(\sigma(i + l) - \sigma(i)) = 2\gamma(l) \quad (2-8)$$

Assuming an isotropic process, that is $\gamma(l)$ depends on the distance between gages ($l$) (i.e. $\gamma(||l||)$) and not its direction, a parametric function is considered to define the variogram. There are several functions available for this purpose such as the linear, exponential, spherical, Gaussian, rational quadratic, Matérn. Equation below corresponds to an exponential covariogram with a nugget value of zero.

$$\Sigma_{\psi, \psi} = \chi_{\psi 1} \times \exp\left(-\frac{\psi}{\chi_{\psi 2}} d\right) \quad (2-9)$$

where $d$ represents the geographic distance between gage stations. $\chi_{\psi 1}$ and $\chi_{\psi 2}$ represent the “sill” and “decay” parameters of the covariogram respectively. Usually the “range” of a covariogram is used rather than “decay” which is defined by the distance at which the correlation between the extreme runoff data is less than 0.05. In exponential covariograms this distance is approximately $3/\chi_{\psi 2}$ (Banerjee et al., 2004).

The third stage of the hierarchical model defines the priors of the latent parameters. Independence between the parameters is assumed in all stages of the model including the prior stage:
\[ p_2(\lambda_{\psi=1q}, \chi_{\psi=12}) = p_{\lambda_{\psi=1q}} \times p_{\chi_{\psi=12}} \quad (2-10) \]

These stages are then combined in a Bayesian framework:

\[ p(\theta|y(x_i=\infty)) \propto p_1(y(x_i=\infty)|\theta_1) p_2(\theta_1|\theta_2) p_3(\theta_2) \quad (2-11) \]

where \( \theta_1 = \psi_{i=\infty}, \kappa \) and \( \theta_2 = \lambda_{\psi=1q}, \chi_{\psi=12} \) with \( \psi = \log(\sigma) \) which lets \( \psi \) take both positive and negative values. \( \sigma \) and \( \kappa \) are GPD scale and shape parameters. Equation above assumes an uninformative uniform prior in the range of \((-\infty, +\infty)\) for the shape (\(\kappa\)) parameter (i.e. \(\pi(\kappa)=1\) indicating the prior distribution function). \( \lambda \) and \( \chi \) are the mean and covariance latent parameters. \( n \) represents the number of USGS gage stations and \( q \) the number of covariates plus the intercept.

Extreme value distributions are developed based on the assumption that the data are independent and identically distributed. In some cases two consecutive days may experience flood events as a result of a single storm. (Fawcett and Walshaw, 2006) developed a first-order Markov chain model to account for the serial dependence but ignored lags greater than one in their dependence model. Declustering is another approach to deal with temporal dependency in which clusters of the extremes over a specified threshold are generated assuming that these extremes occur in consecutive days. A time lag of one day is used to separate the clusters. The maximum of each cluster is then taken as the single independent extreme event. In this study declustering is used to remove temporal dependence.
Furthermore conditional independence is assumed between the spatial data. The assumption of conditional independence implies that given the GPD parameters the extreme values in different gages are spatially independent.

The hydro-climate extremes can be expressed in a more descriptive manner through the return levels. The runoff rate with \( \tau \)-year return level \( (z_\tau) \) is the rate that is exceeded by an extreme event with a probability of \( 1/\tau \). Exceedance over the threshold once every \( N \) observation results in:

\[
\varepsilon \left[ 1 + \kappa \left( \frac{y_N - \nu}{\sigma} \right) \right]^{\frac{1}{\kappa}} = \frac{1}{N} \tag{2-12}
\]

Therefore:

\[
y_N = \nu + \frac{\sigma}{\kappa} \left( (N \varepsilon)^\kappa - 1 \right) \tag{2-13}
\]

And for \( \kappa = 0 \), \( y_N = \nu + \sigma \log(N \varepsilon) \)

In practice a \( \tau \)-year return level flood is expressed by:

\[
z_\tau = \nu + \frac{\sigma}{\kappa} \left( (\tau n_p \varepsilon)^\kappa - 1 \right) \tag{2-14}
\]

And for \( \kappa = 0 \), \( z_\tau = \nu + \sigma \log(\tau n_p \varepsilon) \)

where \( n_p \) is the number of observations per year (e.g. \( n_p = 120 \) for this study).
2.2.3. Parameter Estimation

Evidently the number of model parameters increases with the number of gages resulting in a complex hierarchical model. Because the analytical solution of the multi-dimensional integral over all the parameters is practically intractable, the Markov Chain Monte Carlo (MCMC) technique is utilized to estimate the parameters.

❖ Metropolis-Hastings within Gibbs Sampler

Gibbs sampling (Gelfand and Smith, 1990), along with the Metropolis-Hastings updating algorithm (Cooley et al., 2007; Gelman, 2004) is employed to infer the posterior estimates of the parameters in the hierarchical equation. As an initial step, the uniform priors ranging from \((-\infty, \infty)\) are considered for the mean latent parameters \(\lambda_{\psi}\) as well as the covariance latent parameters \(\chi_{\psi}\) with constrained ranges obtained from an exploratory analysis of the data as illustrated previously. The Metropolis-Hastings step with Gaussian proposal distribution is used to update the scale parameter, \(\log(\sigma)\). Assuming \(\kappa, \psi\) as the MLE estimates of the GPD parameters the latent parameters of \(\lambda_{\psi}, \chi_{\psi}\) as well as the GPD parameters of \(\kappa, \psi\) are initialized. The MCMC procedure shown in Appendix B is applied in order to estimate the posterior distributions of the model parameters.
2.3. Case Study

This study is performed using USGS daily streamflow data over two regions in the Columbia River basin (CRB) as shown in Figure 2-1. CRB is located in the western US covering parts of seven states along with the province of British Columbia in Canada. With a drainage area of 238,000 $mi^2$ it is the third largest basin in the US regarding the flow volume. The mountainous regions of CRB are snow dominated and receive most of the precipitation in winter (Matheussen et al., 2000), consequently the temperature fluctuations due to climate change have significant impact on the intensity, frequency and seasonality of the streamflow (Hamlet and Lettenmaier, 1999; Lettenmaier et al., 1992; Payne et al., 2004).

The northern region (CRB-N) includes 31 gage stations while the southern one (CRB-S) consists of 20 gage stations between 1905-2011. This number of gages was sufficient to produce the variogram plots (Figure 2-2) and provide the priors for the latent parameters.
Figure 2-1: Geographical locations of the two study areas of CRB$_N$ and CRB$_S$. 
Figure 2-2: Variogram plots for the two study areas with the bounding lines reflecting the limits of the prior uniform distributions chosen for the covariance parameters.

The number of gage sites available for time periods of fifteen years is shown for each region during 1905-2011 (Figure 2-3). It is evident that new gages are installed
in time for each region. The reduction of gages in some periods pertains to the missing data especially for CRB-S.

Figure 2-3: Number of available gage sites for each fifteen year time period.

As there is not a priori information about the parameters, we initially consider uninformative uniform priors. For the mean latent parameters $\lambda_i$, uninformative uniform priors in the range of $(-\infty, +\infty)$ are designated which provide proper posteriors as explained in (Banerjee et al., 2004) and shown in the following
sections. As recommended in previous studies (Banerjee et al., 2004; Cooley et al., 2007) bounded priors are assigned to the covariance latent parameters $\chi_{vi}$. The prior distribution of sill parameter ($\chi_{v1}$) for CRB-N and CRB-S are taken to be uniform in the range of [0.001, 6] and [0.001, 3] respectively using the variogram plots. Similarly, the decay parameter ($\chi_{v2}$) is uniformly distributed for both CRB-N and CRB-S in the range of [0.12, 3] and [0.15, 3]. This represents a minimum range of 1mi for both areas, a maximum range of 25mi for CRB-N and a maximum range of 20mi for CRB-S.

2.3.1. Model Setup

To ensure stationary posterior distributions the simulations are performed for 150,000 iterations with a burn in period of 30,000 iterations. To break the dependence between draws and improve the mixing of the posterior samples in the Markov Chain, we perform the thinning by choosing to keep every 30th draw in the chains. Three parallel chains are generated each time with different initial values, and then the chains are merged to produce the posterior distributions.

In an MCMC algorithm when samples are sufficiently drawn from the posterior distribution, adding new samples may not alter the mean of the draws. A plot of the sample means of a parameter distribution versus iteration, the so-called running mean plot (Figure 2-4), is an effective way to evaluate the convergence of a chain.
The results show that the means of parameter samples do not change after 30,000 iterations indicating convergence of the model. In addition, one approach to examine the MCMC efficiency is to plot the autocorrelation of the samples (Figure 2-5). Slow attenuation in the autocorrelation plot reflects slow mixing of samples resulting in slow convergence. The autocorrelation values for all the parameters are low and they attenuate quickly explaining the satisfactory mixing of MCMC samples in obtaining the posterior distributions.

*Figure 2-4: Running mean plots to check mixing of the chains; results shown for the shape parameter and the scale parameter for one of the gages in CRB-North. M represents the mean of the samples up to each iteration.*
Figure 2-5: Autocorrelation (i.e. the correlation between every draw and its $i^{th}$ lag) plots for the GPD shape parameter and the scale parameter for one of the gages in CRB- North, after thinning of 30 iterations; low values indicate satisfactory mixing of the MCMC samples.

As mentioned before, in order to determine the true stationary posterior distribution three initially overdispersed parallel chains are generated. (Gelman and Rubin, 1992) defined an MCMC convergence criterion for each parameter of the model by comparing the variation within the chains in relation to the total variation across the chains for the final $n$ iterations. They developed the scale reduction factor ($\hat{R}$) as follows:

$$\sqrt{\hat{R}} = \sqrt{\frac{n-1}{n} + \frac{m+1}{nm} \frac{B}{W}} \frac{df}{df - 2}$$  

(2-15)
When $\hat{R}$ is higher than 1.2 the chains should be run out longer to improve convergence to the stationary distribution. The factor should decline to 1 as n increases indefinitely. m is the number of parallel runs (chains) with different starting points (here m=3) and, n is the number of iterations in each chain after burn in period. B shows the variances between the means of the m parallel chains, W is the mean of the m-within chain variances, df shows the degrees of freedom of a Student-t distribution which approximates the posterior distribution. (Brooks and Gelman, 1998) extended the scale reduction factor to concurrently monitor the convergence of several parameters. The “boa” package (Smith, 2007) was used in this study to evaluate the convergence. The 0.975 quantiles of the scale reduction factors for all parameters are less than 1.20 which indicates the convergence of the model.

2.3.2. Model Structure

In this study the 95th percentile of the streamflow records from December through March at each gage site was selected as the threshold because it meets the threshold selection criteria mentioned previously (e.g. plot of mean excess vs. the threshold). The period of Dec-Mar is chosen for the analysis as the largest streamflows occur during this period. Because of the possible temporal dependencies in the exceeding values, one may consider dependencies in the modeling or provide independent data by discarding some consecutive records.
(declustering). As a result the value of exceedance rates will decrease. After declustering of the extremes over the threshold, the distributions of $\varepsilon$ for the gage sites for the full records of data are shown as histograms in Figure 2-6. Most of the $\varepsilon$ values range from 0.005-0.03 for CRB-N showing the wider range as compared to CRB-S with $\varepsilon = 0.01-0.02$.

![Histograms of Exceedance Rates](image)

*Figure 2-6: The histograms of exceedance rates for a) CRBN and b) CRBS.*

In order to assess the degree of influence of each covariate on the GPD scale parameter, a generalized Pareto distribution (GPD) is fitted to the recorded extremes at each gage site using the maximum likelihood estimation. The correlations between the scale parameters and the covariates are then calculated (Table 2-1). For both regions the linear dependence between the scale parameter and drainage area is higher than the one between the scale parameter and gage
elevation, however all the correlations are significant considering the p-values. The results are in accordance with the previous studies suggesting the close relationship between the parameters of the generalized extreme value distribution and the drainage area (Villarini et al., 2011); (Lima and Lall, 2010).

Table 2-1: Linear correlation between the scale parameter of the fitted generalized Pareto distribution at each gage (ψ) and the corresponding drainage area as well as the gage elevation.

<table>
<thead>
<tr>
<th>Location</th>
<th>Model</th>
<th>r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRB-North</td>
<td>ψ ~ x</td>
<td>0.757</td>
</tr>
<tr>
<td></td>
<td>ψ ~ y</td>
<td>0.166</td>
</tr>
<tr>
<td></td>
<td>ψ ~ x+y</td>
<td>0.759</td>
</tr>
<tr>
<td>CRB-South</td>
<td>ψ ~ x</td>
<td>0.841</td>
</tr>
<tr>
<td></td>
<td>ψ ~ y</td>
<td>0.293</td>
</tr>
<tr>
<td></td>
<td>ψ ~ x+y</td>
<td>0.848</td>
</tr>
</tbody>
</table>

\( x = \log(Area) - \log(\text{Area}) \), \( y = \log(Elev) - \log(\text{Elev}) \)

Several models are generated based on different covariates, including drainage area and elevation) as well as covariograms (exponential, Gaussian, rational quadratic and Matérn at \( \nu = 3/2 \) [Banerjee et al., 2004]) and are compared according to the deviance information criterion (DIC) (Spiegelhalter et al., 2002). DICs are calculated based on the posterior samples of each model. For each of the three model runs (chains) a separate DIC is calculated and the average of the DICs is then obtained. A model with no covariate is considered as the base model which has
merely two parameters of $\psi$ and $\kappa$. The models which contain covariograms and covariates present approximately similar DIC values which are significantly lower compared to the base model. This indicates the importance of the process stage in characterizing the extreme runoffs. Considering the results provided in Table 2-1, the model including the covariates of latitude, longitude, area and elevation is chosen for this study. The exponential family is also selected because of its simplicity and being a valid variogram in all dimensions.

The structure of the hierarchical model is summarized in Figure 2-7 showing the results of USGS gage#14220500 in CRB$_N$ using the full record time series. This figure includes all the prior and posterior distributions for the parameters in different stages of modeling. The posterior distributions are obtained through MCMC sampling as explained in the previous section. The posteriors show normal distributions except for the decay (range) parameter that is positively skewed. The distribution of the sill parameter shows a small positive skewness as well.
Figure 2-7: Spatial Hierarchical Bayesian model; the dark green density functions correspond to priors while the light green ones display the posterior distributions. Uninformative priors (with infinite integral under the distribution) are used for the mean latent parameters ($\lambda$) in (-\infty, \infty) and for the covariance latent parameters ($\chi$) in the ranges shown in this figure.
As shown in Figure 2-8, using the scale parameter $\sigma$ obtained for each gage station combined with the observed exceedance rates and the shape parameter $\kappa$, the $\tau$-year return level flood distribution is then obtained.

![Figure 2-8: Use of the estimated scale and shape parameters along with the exceedance rate values (not shown) to calculate the 100-year return level flood; results shown for USGS gage#14220500, located in the CRBN.](image)

Since the spatial hierarchical model combines extreme data from different locations, one may expect an increase in the precision of the return level estimates. To verify this, the R package “ismev” (Coles, 2001) is used to calculate the maximum
likelihood estimates along with the 95% confidence interval of the GPD parameters for each gage separately. 95% confidence intervals of the 100-year flood obtained from the hierarchical model is compared with the ones from MLEs. As seen in Figure 2-9, considerable reduction of uncertainty is obtained when estimating the return level using the hierarchical modeling since data from different locations are combined in the second stage of the modeling. The results suggest that the hierarchical spatial models can be applied in flood intensity and frequency analyses while the uncertainty can effectively be reduced even at gage sites with lower number of recorded extreme events. The accuracy of the model is further evaluated in the model test section.

Figure 2-9: Analysis for CRBs (a) and CRBs (b) regions: Comparing the 95% confidence interval of 100-year flood by fitting the GPD distribution to each gage separately (left), and the one from the Spatial Hierarchical Bayesian Model (right).
2.3.3. Model Test

In this study the performance of the model is assessed through prediction of flood events at ungaged sites considering the full record of dataset. For this purpose 20% of the gages are randomly selected and left aside from the parameter estimation. Using the rest of the available gage data the distributions of the latent parameters as well as the shape parameter are determined through the MCMC process. The distribution of the scale parameter for each ungaged site is obtained based on the estimated latent parameter distributions. However, generating such a multivariate normal distribution is not straightforward. Considering each latent parameter taken from its posterior distribution and the covariate values for each ungaged station, the mean and covariance of a multivariate normal distribution is obtained. One thousand samples are taken from this distribution and the average of all is considered as one estimate of the scale parameter in the ungaged site. This procedure is repeated for the rest of the latent parameter posterior values in order to obtain the distribution of the scale parameters at each of the ungaged sites. The distribution of the 100-year return levels are then calculated and compared with the return level values obtained from fitting the GPD distribution to each gage separately with its parameters determined through maximum likelihood estimation.

In Figure 2-10 the results for the two study regions demonstrate that the predicted distribution encompasses the MLEs at each gage. Both underestimation and
overestimation of the 100-year floods are seen when comparing the mean of the predictive distribution with the one from MLE.

Figure 2-10: Evaluating the performance of the Spatial Hierarchical Bayesian model in gages not being used during the model setup; comparison is made based on separate GPD fit for each gage station (shown by triangles). The USGS gage numbers are shown on top of each panel followed by N and S representing CRB_N and CRB_S; the vertical lines represent the 95% confidence intervals of the hierarchical model estimations.
2.3.4. Prior Sensitivity

In Bayesian analysis the choice of priors might have significant effect on the final inference. Therefore, it is necessary to perform a sensitivity analysis on priors. The posterior distributions of $\chi_{vi}$ are shown to be sensitive to the priors (Berger et al., 2001). As discussed earlier, bounded uniform priors are considered for the sill and decay latent parameters. Two other uninformative prior distributions are additionally assigned to evaluate the variations in the posterior distributions of the latent parameters. Furthermore, the sensitivity of the scale parameter posterior distributions to the covariance latent parameter prior distributions is assessed. As shown in Figure 2-11 the prior distributions for the two regions are:

$\chi_{v1}$: Inverse-Gamma (shape = 4, scale = 6) and Log-Normal (mean = log3, sd = log3);

$\chi_{v2}$: Inverse-Gamma (shape = 10, scale = 5) and Log-Normal (mean = log1.5, sd = log3);

Application of the above prior distributions shows that the posterior distributions of the latent parameters are sensitive to the choice of the prior distributions. The decay parameter shows higher dependency on priors compared with the sill parameter. The density of the prior distributions in the lower bounds also affects the posterior distributions and cause shifts in the means. No significant change is seen in the scale parameter posterior distribution in both study regions.
Figure 2-11: Sensitivity analysis of the covariance latent parameter ($\chi$) and the scale parameter ($\psi$) to changes in prior distribution in a) CRB-North and b) CRB-South.
2.3.5. Parameter Estimates

The median and the 95% confidence intervals of the latent parameters are shown in Table 2-2.

Table 2-2: The median and 95% confidence intervals of the estimated latent parameters corresponding to the scale parameter; the model has two covariates.

<table>
<thead>
<tr>
<th>Location</th>
<th>Parameter</th>
<th>Variable</th>
<th>Median</th>
<th>95% Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRB-North</td>
<td>Mean Latent ($\lambda$)</td>
<td>Intercept</td>
<td>9.692</td>
<td>(8.921, 10.454)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Area</td>
<td>0.663</td>
<td>(0.488, 0.843)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elev</td>
<td>0.023</td>
<td>(-0.041, 0.083)</td>
</tr>
<tr>
<td></td>
<td>Covariance Latent</td>
<td>sill</td>
<td>0.69</td>
<td>(0.361, 1.184)</td>
</tr>
<tr>
<td></td>
<td>($\chi$)</td>
<td>decay*</td>
<td>0.521</td>
<td>(0.12, 2.732)</td>
</tr>
<tr>
<td>CRB-South</td>
<td>Mean Latent ($\lambda$)</td>
<td>Intercept</td>
<td>7.458</td>
<td>(7.213, 7.682)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Area</td>
<td>0.652</td>
<td>(0.475, 0.835)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elev</td>
<td>-0.157</td>
<td>(-0.599, 0.27)</td>
</tr>
<tr>
<td></td>
<td>Covariance Latent</td>
<td>sill</td>
<td>0.124</td>
<td>(0.047, 0.264)</td>
</tr>
<tr>
<td></td>
<td>($\chi$)</td>
<td>decay*</td>
<td>1.466</td>
<td>(0.15, 2.843)</td>
</tr>
</tbody>
</table>

*range = 3/$\chi_1$: The distance at which correlation is less than 0.05

This corresponds to a spatial hierarchical model with latitude, longitude, area and elevation as covariates with uniform priors for the parameters. The range of the sill parameter is higher in CRB-N as compared to CRB-S due to the larger area and number of gage sites. The latent parameter pertaining to drainage area is positive for both regions while the one related to the gage elevation is positive for CRB-N and
negative for CRB-S. The Median, 2.5% and 97.5% quantiles of the scale parameters for the two regions are shown in Figure 2-12.

![Figure 2-12: 95% confidence intervals and the median of the GPD scale parameter for the CRB- North (top panels) and CRB- South (bottom panels) based on the results of the Spatial Hierarchical Bayesian Model. Triangle marks show the location of the streamflow gages.](image)

Similar results are shown for different return levels for the two regions in Figure 2-13 and Figure 2-14. Each circle represents a gage station with its size relative to the magnitude of the estimated return level.
Figure 2-13: 95% confidence intervals for 25, 50 and 100-Year return level floods for the CRB-North based on the results of the Spatial Bayesian Hierarchical Model, the size of the circles represent the magnitude of the return levels.
Figure 2-14: 95% confidence intervals and the median of the 25, 50 and 100-Year return level floods for the CRB-South based on the results of the Spatial Bayesian Hierarchical Model; the size of the circles represent the magnitude of the return levels.
2.3.6. Analysis of the Temporal Trend

It is widely accepted that hydro-climate events are not stationary due to changes in land use and the impact of climate change (Milly et al., 2008; Najafi et al., 2011b). In this study the changes in extreme events are investigated for the two study regions for time periods of fifteen years. Streamflow records from available gages at each time period are used to perform the MCMC simulations and sample the parameters. The posterior samples of the return levels with different return periods are calculated using the posterior samples of the shape and scale parameters along with the observed exceedance rates. This allows for quantification of the uncertainties associated with the return level assessments. This temporal analysis is conducted for the period of 1905-2011 in CRB-N and 1950-2011 in CRB-S.

The fifteen year time window is selected in this study so that sufficient number of recorded data is incorporated and also the temporal trends in extreme parameters are represented. Other time periods of 20 years (Dias et al., 2012) and 30 years (Najafi et al., 2011b) have also been implemented to compare the future hydro-climate conditions to historical conditions, however still no specific criterion is suggested in literature. The length of the time windows to choose and its influence on the results of climate change studies is subject to further studies. The trends of the scale parameters and 100-year return level floods over the three periods of 1965-80, 1980-95 and 1995-2011 are shown for each gage in CRB-S (Figure 2-15).
Each circle represents the mean of the scale parameter and return level flood distribution at each time period. Different colors correspond to different values and larger circles are linked with higher magnitudes. The ranges of the scale parameters (σ) and the 100-year return level floods (cfs) vary between (150-8000) and (1500cfs-75000cfs) respectively.

![Figure 2-15: Change of the GPD scale parameter (top) and 100-Year return level (bottom) for three periods of 1965-1980, 1980-1995 and 1995-2011 in CRB-South.](image)

The spatial changes of the scale parameter and 100-year return level floods are shown for three periods of 1965-80, 1980-95 and 1995-2011 for CRB-N in Figure 2-16. Both scale...
parameter and 100-year return level floods show increasing trends in the majority of the gages for the latest two periods compared with 1965-1980.

Figure 2-16: Change of the GPD scale parameter (top) and 100-Year return level (bottom) for three periods of 1965-1980, 1980-1995 and 1995-2011 in CRB-North.

The scale parameter reduces in the northern parts and increases in eastern and western gages. The return levels show a general reduction from 1965-80 to 1980-95 but then increases considerably in 1995-2011 in most gages.
In Figure 2-17, each boxplot on top panel represents the medians of the return level distributions for all gages. A cyclic change is detectable for CRB-N where increases in extreme runoff have occurred during 1920-1935; 1950-1965 and 1995-2011. For CRB-S runoff extremes tended to decrease during 1935-1950 to 1965-1980 while they increased from 1965-1980 to 1995-2011. Correspondingly the median and 95% confidence intervals of the shape parameters are shown in Figure 2-17 bottom. The variations in shape parameters show agreements with the ones from the return levels. This is in accordance with the analysis of Naik and Jay (2011) who showed decreased spring peak flow due to flow regulations accompanied by increased peak flows in other seasons since 1970.
2.4. Conclusions

A spatial hierarchical Bayesian method is developed to model the extreme runoffs at USGS gage sites over two regions in the Columbia River basin namely CRB-N and CRB-S. The regions include 31 and 20 streamflow gage stations respectively. Extreme events occurred during the months of December through March are selected for this analysis. Generalized Pareto distribution is the basis of the model with its scale parameter being spatially characterized in a hierarchical Bayesian approach.
Declustering process provides temporally independent data for the hierarchical model. The scale parameter of the extreme distribution is spatially modeled through generalized linear relationships considering covariates of latitude, longitude, drainage area and elevation. The parameters of the spatial hierarchical model are estimated through an MCMC procedure called Metropolis Hastings within Gibbs sampler. (Gelman, 1996) criteria is used to inspect the MCMC convergence. The performance of the model is verified by predicting the 100-year return level floods for several test gages using the hierarchical model and comparing the resultant distributions with the at-site maximum likelihood estimates.

Results show significant increase in the precision of the model compared with a simple maximum likelihood estimator regarding the flood return levels, since information content in the data from different locations are combined using the spatial hierarchical model. Besides, the performance evaluations of the model in predicting ungaged sites show satisfactory results. The spatial dependence among extreme events in different locations is weaker than that of the daily data itself because they do not necessarily occur at the same time (Sang and Gelfand, 2010). The assumption of conditional independence implies that the extreme values in different gages are spatially independent given the GPD parameters. Although the spatial dependence in large scale is considered in the hierarchical parameterization, the dependencies of the data in small scale may not be represented (Sang and
The assumption of conditional independence is widely considered in the spatial hierarchical modeling (Banerjee et al., 2004). Another approach is to explicitly consider the small scale dependencies of the extremes through a copula density function and obtain continuous surface realizations (Renard, 2011; Sang and Gelfand, 2010). This is still an open area of research.

We run the model for historical time periods of 15 years in order to detect possible trends in runoff extremes. The results show cyclic variations in the spatial average of the 100-year return level floods. However for some areas consistent increasing trends are distinguishable. Developing spatio-temporal hierarchical modeling of hydrologic extreme events as indicated in (Sang and Gelfand, 2009; Vanem et al., 2012) is a substitute for time window analysis and is intended for future studies.

3.1. Background

The rising temperature of the earth due to climate change has shown to alter the variations in hydro-climate variables regarding their intensities, frequencies and durations (Halmstad et al., 2012; Hidalgo et al., 2009; Karl and Melillo, 2009). Extreme events, i.e. floods and droughts, are susceptible to any disturbances in climate cycles (Field et al., 2012). As such, it is important to provide the policymakers with sufficient knowledge about the probable impacts of climate change on hydrologic extremes and most importantly floods which cause or threaten damage to properties and lives (Moradkhani et al., 2010). Frei et al. (2006) and Villarini et al. (2011) studied the variability in both observed and climate model simulated extreme event occurrence. The potential for change in the occurrence of extreme events in the future in conjunction with projected climate variability is also prevalent in recent studies (Mote and Salathe, 2010; Tryhorn and DeGaetano, 2011).

Study of the climate change impact on hydrologic variables such as runoff is commonly conducted based on hydrologic simulation because runoff amount is not generally known at the spatial scale of interest (in this study computational grid cells). Therefore, hydrologic model is initially calibrated based on the observed time series, then the model is forced with future climate scenarios (Risley et al., 2011).
The future and historical simulations are compared to assess the relative impact of climate change (Chiew et al., 2009; Jung et al., 2011; Najafi et al., 2011b; Wood et al., 2004a). Global climate models provide historical as well as future assessments of the climatologic variables across the globe. However, due to their coarse resolution, attempts have been made to nest regional climate models within the boundaries of GCMs and provide projections at finer scale resolutions. Several regional programs have been conducted in the last decade focused on addressing the need for appropriate scale level assessment of climate change impacts. Examples include the Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects (PRUDENCE) project (Christensen and Christensen, 2007), followed by the Ensembles-Based Predictions of Climate Changes and Their Impacts (ENSEMBLES) project (Van der Linden and Mitchell, 2009). The North American Regional Climate Change Assessment Program (NARCCAP) provides data from multiple GCM-RCM coupled simulations over the majority of the American continent (Mearns et al., 2009). Simulations of both future (2041-2070) and historic (1971-2004) periods are produced by the NARCCAP modelers at a spatial resolution of 50km and sub-daily temporal resolution.

Extreme value theory is commonly utilized in hydrology and water resources to evaluate the characteristics of extreme events and has seen more applications over the past decade (e.g., (Acero et al., 2010; Aryal et al., 2009; Davison et al., 2012;
Kharin et al., 2010b). Extreme Value Theory (EVT) is a statistical method to model the tails of the probability distribution of a random variable. The parameters of the extreme value distributions fitted to a time series of a hydro-climate variable at a location yield estimates of intensity and frequency of the extreme events. EVT has been widely applied in studies of precipitation (Acero et al., 2010; Katz et al., 2002; Kharin et al., 2010b; Wehner et al., 2010), temperature (Kharin et al., 2010a; Kharin et al., 2010b; Rusticucci and Tencer, 2008), streamflow (Hurkmans et al., 2010; Katz et al., 2002; Lima and Lall, 2010) and wind speed (Brabson and Palutikof, 2010; Caires and Sterl, 2010; Hundecha et al., 2008).

Fitting of extreme distributions at individual locations is restricted by the limited records of hydro-climatic extremes in space and time (Fuentes et al., 2010). Therefore, methods have been developed to combine extreme data from different locations to provide a more robust statistical model. Index flood method (Dalrymple and Survey, 1960) is one approach to improve the accuracy of the estimates and to predict flood at ungagged sites. A homogeneous region is first defined and the extreme data at each gage are divided by the index flood followed by fitting a distribution to the combined data from all gages. RFA however, does not consider the spatial components of the point data and cannot incorporate additional variables (i.e. covariates) into the analysis. Besides it is not possible to explicitly estimate the uncertainties using this approach.
Spatial hierarchical Bayesian method is an alternative approach which has been recently introduced in the literature. It considers a spatial model on the parameters of a univariate extreme value distribution (such as GEV) at each point (or grid cell). This approach has gained much attention in climate extreme analyses as it effectively combines spatially distributed data. The studies include the analysis of wind (Fawcett and Walshaw, 2006; Li and Shi, 2012), precipitation (Aryal et al., 2009; Atyeo and Walshaw, 2012; Cooley et al., 2007; Ghosh and Mallick, 2011; Renard, 2011; Sang and Gelfand, 2010; Schliep et al., 2010) and temperature (Brynjarsdóttir and Berliner, 2011; Lemos et al., 2009). (Cooley and Sain, 2010) studied the changes in precipitation extremes using regional climate model data for historical and future periods. (Sang and Gelfand, 2009; Schliep et al., 2010) utilized a spatial autoregressive model of the annual maximum rainfall. (Cooley and Sain, 2012; Davison et al., 2012) reviewed and compared several modeling approaches for spatial extremes with the applications to rainfall events. Fewer attempts, however, have been made to enhance the hydrologic extreme analysis. (Lima and Lall, 2010) employed a hierarchical Bayesian model for streamflow analysis using the drainage area as the only explanatory variable with no spatial model. (Najafi and Moradkhani, under review) analyzed the historical extreme runoff data recorded at gage sites based on a spatial hierarchical Bayesian approach incorporating the elevation and drainage area as additional covariates. They considered geostatistical
modeling for spatially characterizing the extreme parameters. Nevertheless, analyzing grid-based hydrologic variables (e.g. when obtained from distributed hydrologic modeling) is a challenging task by using the geostatistical models (i.e. variograms) especially for large regions (Schliep et al., 2010).

Hydropower is the major energy source in the Pacific Northwest (PNW) region of USA; hence the possible impacts of climate change on its hydro-climate characteristics is of concern (Hamlet et al., 2010; Miles et al., 2000; Snover et al., 2003). Winter (October-March) precipitation dominates the hydrologic characteristics of PNW (Miles et al., 2000). Precipitation in the mountains is higher especially on the western slopes of the Olympics and Coast mountain range and the Cascades (Figure 3-1).
The region is climatically divided by the Cascade Mountains. In low level elevations such as the low-lying valleys west of the Cascades precipitation falls as rain in winter and little snow pack is accumulated. In intermediate elevations precipitation falls as rain during fall and early winter and transits to snowfall during winter followed by spring/summer snowmelt. East of Cascades is snow dominated with spring and summer snowmelt resulting in high runoff rates. Consequently the temperature fluctuations due to climate change have significant impacts on the intensity, frequency and seasonality of the streamflow (Hamlet and Lettenmaier,
Columbia River is a predominant hydrologic and water resource feature in the PNW. The Columbia River is the fourth largest river in the United States in terms of runoff which is strongly influenced by climate patterns such as ENSO (El Nino Southern Oscillation) and PDO (Pacific Decadal Oscillation). The wet western sub-basins cover 8% of the CRB area while contributing almost 25% of the total streamflow [Naik and Jay, 2011]. It originates from the Columbia Lake in Canada and discharges in the North Pacific Ocean at its mouth in Astoria, Oregon. The Columbia River Basin (CRB) covers parts of seven states along with the province of British Columbia in Canada. It has a drainage area of 238,000 mi² which is surrounded by the Rocky Mountains on the east and north and the Cascade Mountains on the west. Climate change impact studies on the hydrologic components of PNW (and CRB) region have indicated increased winter streamflow along with earlier snowmelt which results in decreased spring snow pack and reduction of summer flows (Lee et al., 2009). The flow management of the Columbia River can be divided into three periods of: pre-regulation (<1900), historic (1900-1969) showing increased human intervention and modern (1970-date) which reflects strong flow regulation (Naik and Jay, 2005).

Jay and Naik (2011) provided a detailed study of the Columbia River spring and winter freshet styles and the related effects of climate change. Naik and Jay (2011) separated the effects of human and climate change on Columbia River flow.
variations for a period of 150 years. They considered three distinct flow measurements: observed, adjusted and virgin flow and concluded that the reduction of mean annual river virgin flow (at The Dalles gage station) is the result of climate change (8-9%) and of human impacts such as irrigation usage (7-8%). They detected three principal factors affecting the hydrology of the Columbia River: direct flow manipulation such as dam constructions; climate change and climate cycles such as ENSO (El Nino Southern Oscillation) and PDO (Pacific Decadal Oscillation). Their results show decreased mean annual flow from 1879-1899 to 1945-2004, and significant decrease in observed maximum spring freshet flow.

This study addresses the climate change impact on runoff extremes over the PNW. Individual analyses are performed for each season using the complete regional climate model simulations over the US available at the time when this research was performed. In section 3.2 of this study the major steps of the hydrologic extreme analysis are illustrated including the distributed hydrologic modeling and spatial hierarchical Bayesian analysis of extremes. Section 3.3 describes the application of the method over the PNW study region starting with a description of the grid based observed data, climate scenarios provided by the North American Regional Climate Change Assessment Program (NARCCAP) simulations followed by the downscaling procedure, hydrologic modeling and parameterization, extreme analysis based on hierarchical Bayesian, and the
discussion of results (section 3.4). Summary and conclusions are provided in section 3.5.

**3.2. Methodology**

**3.2.1. Hydrologic Modeling**

The variable infiltration capacity (VIC) hydrologic model is used in this study for hydrologic simulation using observational and RCM gridded datasets. VIC (Liang et al., 1994; Liang et al., 1996) is a semi-distributed land surface model which has been successfully applied for climate change studies over the PNW region (Hamlet and Lettenmaier, 1999; Hidalgo et al., 2009; Lettenmaier et al., 1992; Matheussen et al., 2000; Payne et al., 2004). The forcing data including precipitation, maximum and minimum temperature and wind speed are required to perform the hydrologic modeling along with the watershed characteristics such as land cover, soil and elevation. Simulations are independently implemented for each grid cell at a daily or sub-daily time step without considering the horizontal flow. For routing purposes a separate model is needed to transport the surface runoff to the grid outlet and to the flow channel (Lohmann et al., 1996; Lohmann et al., 1998). The VIC model can be run in water balance and energy balance modes of operations. The water balance does not consider the surface energy balance and assumes that the soil surface temperature and the surrounding air temperature are equal. Although more computationally intensive, the energy balance mode solves the complete water
balance while simulating the surface energy fluxes such as sensible heat, latent heat, and ground heat to account for the total incoming radiation fluxes. VIC model includes a snow algorithm which considers a two layer formulation of surface and pack layers. It models the surface energy balance at the snow/air interface along with the ground heat flux from the snowpack into the ground.

The VIC model was initially developed as a soil-vegetation-atmosphere transfer scheme to be incorporated in GCMs. The current VIC model considers sub-grid variability in land cover, soil-moisture storage, precipitation and elevation. Each grid cell is partitioned into snow bands (elevation bands) containing a number of land cover tiles. Each land cover type consists of a vegetation layer, and several soil layers (commonly three soil layers are used (Cherkauer et al., 2003)). The land cover tiles are described based on leaf area index (LAI), albedo, canopy resistance, roughness length etc. Spatial variability of infiltration and runoff generation is simulated using the variable infiltration curve within each vegetation class. Penman-Monteith equation governs the potential and actual evapotranspiration which is dependent on net radiation and vapor pressure deficit. The dynamic behavior of soil response to the rainfall event is signified by the upper soil layers. Gravity controls the flow from upper layers to lower ones. Empirically based Arno curve is used to model baseflow.
3.2.2. Spatial Hierarchical Bayesian Model

Extreme value theory is commonly adopted in order to describe the tail of a distribution (Coles, 2001). Based on asymptotic outcomes it provides different classes of distributions for characterizing extremes. The extreme value theory (EVT) considers a sequence of independent and identically distributed (iid) random variables of \( Y_1, Y_2, \ldots \) with \( M_n = \max\{Y_1, Y_2, \ldots, Y_n\} \). It states that for large \( n \) there exists normalizing constants of \( a_n > 0 \) and \( b_n \), such that \( \lim_{n \to \infty} P\left( \frac{M_n - b_n}{a_n} \leq y \right) = F(y) \). \( F(y) \) is a non-degenerate distribution function (i.e. probability distribution of a random variable 'y' which takes a continuum range of possible states for it) which forms a family of distribution called the generalized extreme value (GEV) distribution:

\[
F(y|\mu, \sigma, \kappa) = \exp\left\{-\left[1 + \kappa\left(\frac{y-\mu}{\sigma}\right)\right]^{-\frac{1}{\kappa}}\right\},
\]

(3-1)

where \( \mu, \sigma \) and \( \kappa \) are the location, scale and shape parameters respectively and \( 1 + \kappa(y - \mu)/\sigma > 0 \). Depending on different values of \( \kappa \) the GEV distribution will be Gumbel, Frechet or Weibull distributions. For \( \kappa > 0 \) the upper tail of the distribution decreases slowly with a power function and never reaches zero, for \( \kappa < 0 \) the GEV distribution has a bounded upper tail and for \( \kappa = 0 \) it has an exponentially decreasing tail. GEV distribution models the maximum of sufficiently long series of data. In hydrologic applications it is commonly used when data are considered in
annual, seasonal or monthly blocks where their block maxima are taken as iid random variables. Based on this approach, only the block maximum events are considered and other extremes with lower magnitudes are disregarded. Generalized Pareto distribution (GPD) which uses data over a specified threshold overcomes this issue, however setting a proper threshold and coping with the temporal dependencies between sequential extremes are of concern.

Based on equation below the \( \tau \)-year return level corresponding to an extreme event which is expected to be exceeded once every \( \tau \) years (i.e. expected value of an event with an exceedance probability of \( p=1/\tau \)) can be obtained by:

\[
 z_{\tau} = \mu - \frac{\sigma}{k} \left[ 1 - (-\log(1-1/\tau))^{1/k} \right], \tag{3-2}
\]

In order to deal with the limited records of hydro-climatic extremes in space and time spatial hierarchical Bayesian model is developed to collect information from different locations to increase the reliability of the estimates. In addition this model aims to describe the spatial variations of marginal distributions. It consists of a latent process in which a spatial model characterizes the parameters of the marginal distribution. The spatial model is commonly based on variograms in applications where hydro-climate data are point-referenced (i.e. geostatistical) and based on the conditional, intrinsic or simultaneous autoregressive models (CAR, IAR, SAR) for
data recorded at areal units of grids (i.e. lattice) (Banerjee et al., 2004; Cressie, 1992).

The spatial hierarchical model consists of the so called data, process and prior stages which are linked through conditional distributions. In this study the hierarchical modeling process is performed separately for each RCM (runoff outputs from VIC driven by RCM) and season as well as the historical and future time periods.

The data stage of the hierarchical model defines the likelihood function. Let’s denote $Y_{i,t}$ as the maximum annual runoff (mm) at the $i^{th}$ cell in time $t$. $Y_{i,t}$ follows the GEV distribution with parameters $(\mu_i, \sigma_i, \kappa_i)$ that makes it conditionally independent of $Y_{j,t}$ ($i \neq j$). This implies that the maximum series in each grid cell would follow their own GEV parameters. Therefore the likelihood function is the product of GEV equation (above) over each grid cell:

$$p_t(y_i|\mu_i,\sigma_i,\kappa_i) = \prod_{i=1}^{n} \prod_{t=1}^{T} \left[ \frac{1}{\sigma_i} \right]^{\frac{1}{\kappa_i}} \left[ 1 + \frac{1}{\kappa_i} \left( \frac{y_{it} - \mu_i}{\sigma_i} \right) \right]^{-\frac{1}{\kappa_i}} \times \exp \left\{ - \left[ 1 + \kappa_i \left( \frac{y_{it} - \mu_i}{\sigma_i} \right) \right]^{1/\kappa_i} \right\}$$ (3-3)

where $T$ is the total number of years for each of the historical and future periods and $n$ the total number of cells. In order to improve the maximum likelihood estimate of extreme distributions for a short record data (Martins and Stedinger, 2000) recommended a Bayesian prior which restricts the values of $\kappa$ to a reasonable
range. Similarly (Cooley and Sain, 2010; Schliep et al., 2010) considered this prior in the data stage of their hierarchical model for precipitation extreme analysis. This prior function is a beta distribution, supported in the [-0.5,0.5] interval:

\[
\pi(\kappa_i) = \frac{\Gamma(15)}{\Gamma(9)\Gamma(6)} (0.5 + \kappa_i)^5 (0.5 - \kappa_i)^5
\]

(3-4)

In which \( \Gamma \) is the gamma function. The product of this prior and the likelihood function defined in equation above constitutes the data stage of the hierarchical model (i.e. \( p_i \times \pi(\kappa_i) \)).

The process stage of the model tries to combine extreme data from different cells. The cell-wise extreme runoff data obtained from VIC suggests using the conditionally autoregressive models.

Here we assume that each GEV parameter follows a normal distribution:

\[
p_2(\mu | \lambda_\mu, U_{i,\mu}, \tau^2) = N\left(X_{i,q}^{\prime} \lambda_{i,q}^{\prime} + U_{i,\mu}, \frac{1}{\tau^2_{\mu}}\right)
\]

\[
p_2(\psi | \lambda_\psi, U_{i,\psi}, \tau^2) = N\left(X_{i,q}^{\prime} \lambda_{i,q}^{\prime} + U_{i,\psi}, \frac{1}{\tau^2_{\psi}}\right)
\]

\[
p_2(\kappa | \lambda_\kappa, U_{i,\kappa}, \tau^2) = N\left(X_{i,q}^{\prime} \lambda_{i,q}^{\prime} + U_{i,\kappa}, \frac{1}{\tau^2_{\kappa}}\right)
\]

with \( \psi_i = \log (\sigma_i) \). \( i \) denotes grid location, \( q \) the number of covariates and \( q' \) the number of their associated factors (\( q' = q+1 \)). \( X_{i,q} \) and \( \lambda_{i,q}^{\prime} \) with \( \theta \sim \mu, \psi, \kappa \)
represent the covariates (i.e. explanatory variables associated with each cell including the geographic coordinates, physical and climatological characteristics etc.) and their corresponding factors respectively; $\tau_\theta^2$ is the precision parameter. $U_{i,\theta}$ are spatial random effect parameters which account for the dependencies between GEV parameters.

Consider a random vector of $Y = (Y_1, Y_2, ..., Y_N)^T$ in which $Y_i$ pertains to the $i^{th}$ cell. One can generate the joint distribution $p(y)$ using ‘n’ full conditional distributions of $p(y_i|y_j)$ where $j \neq i$. This Markov random field approach calls the $j^{th}$ cell a neighbor of cell ‘i’ if $p(y_i|y_j)$ depends on $Y_j$.

The standard conditional autoregressive model (CAR) is a Gaussian MRF which defines the conditional distributions for each of the random variables (RV) given the RVs in neighboring cells (Besag, 1974; Rue and Held, 2005). Gaussian MRFs can be described by the precision matrix (Q) which is the inverse of the covariance matrix. The most popular CAR employment is the pairwise difference formulation known as the intrinsic autoregressive (IAR) model (Besag and Kooperberg, 1995). In this study the spatial effect parameter is modeled using the multivariate intrinsic autoregressive spatial model which is based on the proximity matrix of $W$ with components of $w_{ij}$ (Banerjee et al., 2004; Besag, 1974; Schliep et al., 2010). The
proximity matrix is used to weight the similarities between values at different cells. It is calculated by the neighborhood structure of the data in a way that if two cells have a common border line or notch, the corresponding values in the proximity matrix become one otherwise they are set to zero. Row standardization can be performed to increase the influence of cells with fewer neighbors (If \( p(\theta_i | \theta_j) \) where \( j \neq i \) depends on \( \theta_j \) then \( j \) is called a neighbor of cell \( i \)). The precision is defined by 

\[
Q = T \otimes Q_i
\]

where \( T \) is a \( 3 \times 3 \) positive definite matrix reflecting the dependence information between spatial effect parameters corresponding to \( \mu, \sigma \) and \( \kappa \).

\[
Q_i = D_w - W
\]  \hspace{1cm} (3-6)

\( D_w \) is a diagonal matrix with elements of \( d_{wij} \):

\[
d_{wij} = \sum_{j=1}^{n} w_{ij}
\]  \hspace{1cm} (3-7)

In the third stage of the hierarchical Bayesian model the priors are designated for the latent parameters which are defined in the process stage including \( \lambda_{q,q} \) and \( T \).

### 3.3. Application to the Runoff Extremes over the Pacific Northwest

The process of hydrologic extreme analysis over PNW is summarized in Figure 3-2. Each step is explained in the following sections.
3.3.1. Data

The observational gridded meteorological data covering the period of January 1949 through July 2000 was considered in this study (Maurer et al., 2002). It includes daily total precipitation (mm/day), maximum and minimum temperature (°C), and the average wind speed (m/s) at an average height of two meters above the surface. 6392 cells with 1/8th degree resolutions cover the PNW region including the CRB.

North American Regional Climate Change Assessment Program (NARCCAP) provides a suite of regional climate model simulations on a 50km by 50km spatial resolution and daily/sub-daily time scale based on the SRES-A2 emission scenario.

Figure 3-2: Flowchart of the Modeling Procedure.
(Mearns et al., 2012). The A2 scenario predicts large population increases, high carbon dioxide emissions, and weak environmental concerns. The RCMs include CRCM, ECP2, MM5I, RCM3, WRFG, and HRM3. These are nested inside general circulation models of GFDL, CGCM3, HADCM3, and CCSM. Data is available for the historical and future periods of 1971-2000 and 2041-2070 respectively. For hydrologic simulations we were interested in the 3-hourly precipitation (kg/m²s), daily maximum and minimum temperature (K), zonal and meridional wind speeds (m/s) from NARCCAP. At the time when this research was conducted, data from eight RCM_gcms were available covering the historical and future time periods and one RCM_gcm for the historical period. These include CRCM_cgcm3, CRCM_ccsm, ECP2_gfdl (for the historical period), HRM3_gfdl, HRM3_hadcm3, RCM3_gfdl, RCM3_cgcm3, WRFG_cgcm3 and WRFG_ccsm.

3.3.2. Downscaling the NARCCAP Climate Variables

The process of downscaling has been established to address the inadequacies of large scale resolution models. There are two main classes of downscaling procedures: statistical and dynamical. Several studies over the last several years have provided detailed comparisons of both downscaling types (Caldwell, 2010; Di Luca et al., 2011; Giorgi, 1990; McGregor, 1997; Murphy, 1999; Najafi et al., 2011c). Statistical approaches involve determining reliable statistical relationships between large-scale climate variables, and local scale variables, such as temperature or
precipitation. Dynamic downscaling approaches, on the other hand, are based on the same numerical integration of differential equations, as in GCMs, but over a smaller spatial and temporal domain. Furthermore, dynamical downscaling approaches can include modified physical schemes that have been demonstrated to better address topographical variations. RCMs provided by NARCCAP have 50km resolutions, though hydrologic models commonly require meteorological variables at a finer resolution. Further downscaling is therefore required to develop the high resolution data (Najafi et al., 2011c; Samadi et al., 2013). In this study the quantile-mapping (QM) approach (Wood et al., 2004a) was used for downscaling and bias-correcting the individual RCM variables.

In QM method the observed and simulated datasets are each characterized in terms of their full distributions of daily values, a so-called non-parametric approach since it does not rely on adjusting simply the mean, standard deviation, or other standard statistical parameters. QM corrects the model outputs distribution to exclude systematic biases. The correction is made for each cell at the resolution of the observed data. It considers each percentile of the variable’s distribution and compares the 1/8° observed grid cell with the closest 50km resolution RCM cell (Quintana Seguí et al., 2010). For both the observed and simulated datasets, the cumulative distribution functions (CDFs) are computed on a monthly basis. After computing the CDFs, the scaling factor determined based on the respective quantile
values during the observed period are applied for the future period. This is based on the assumption that the correction function remains constant in time. In this study the downscaling of NARCAAP product was performed for each model following the above procedure.

3.3.3. Hydrologic Modeling

The hydrologic simulation was carried out using the VIC model driven by the 1/8th degree resolution observational and downscaled NARCCAP meteorological data. Modeling was performed in a daily time step for the historical time period of 1971-2000 and future period of 2041-2070 by considering three layers of soil. Vegetation, soil and snow band data were obtained from the Land Surface Hydrology Research Group, University of Washington. The geographic information of each grid cell, soil parameters and the initial soil moisture conditions are stored in the soil parameter file. The land cover types, number of vegetation tiles and their coverage in each grid cell as well as other vegetation parameters such as root depth etc. are contained in the vegetation parameter file. The VIC model is capable of disaggregating each grid cell into elevation bands in order to lapse the related average temperature, pressure and precipitation. The snow band file contains information on each elevation band used by the snow model to account for the topographical influences on snow pack accumulation and ablation. Five elevation (snow) bands were considered to better characterize the snow processes at each
grid cell. Simulations were made based on the water balance mode indicating that the surface temperature was set to the surrounding air temperature. Several variables were estimated for this study region including the evapotranspiration, canopy interception of liquid water, moisture content of each soil layer, snow water equivalent, snow pack depth, fractional area of snow cover. In this study we analyzed the resulting runoff (mm) at each cell for the hydrologic extreme analysis.

3.4. Results and Discussion

The grid based annual maximum runoffs obtained from each RCM data were separately extracted for the historical and future periods and all seasons including winter (DJF), spring (MAM), summer (JJA), and fall (SON). The spatial hierarchical Bayesian modeling of extremes was then performed for each dataset. Both the historical (1971-2000) and future (2041-2070) time periods consist of 29 years of data (i.e. $T = 29$) while there are in total 6392 cells with $1/8^\circ$ resolution in the study region (i.e. $n = 6392$). Maximum likelihood estimates of the GEV parameters for each cell were considered as starting points for the inference process in the spatial hierarchical Bayesian model. As discussed in the previous sections the location $\mu$, scale $\sigma$ and shape $\kappa$ parameters of the generalized extreme value distribution were spatially characterized by a multivariate intrinsic autoregressive (IAR) model. Furthermore the latitude, longitude, 30-yr average daily precipitation and temperature at each cell were considered as potential covariates. These variables
were all transformed to have a mean of zero and standard deviation of one. Values of precision parameters were set constant \( \tau^2_\theta = (4,100,1600) \) which allowed the spatial effect parameters \((U_{i,\theta})\) to define most of the variability in the spatial model.

In the third stage of the hierarchical Bayesian model priors were defined over the latent parameters of \( \lambda_{\theta,q} \) and \( T \). Uninformative Gaussian distributions were considered for regression coefficients \( \lambda_{\theta,q} \) with zero mean and variance of 5. The intercept terms \( \lambda_{\theta,0} \) were assigned Gaussian distributions with means equal to the maximum likelihood estimates of the corresponding GEV parameters at each cell with variance 50. Vague Wishart prior with three degrees of freedom was considered for the precision matrix \( T \) with diagonal of \((0.04, 8, 80)^T\) (Cooley and Sain, 2010).

Metropolis-Hastings within Gibbs sampler was used for parameter estimation (Banerjee et al., 2004; Casella and Robert, 1999; Rue and Held, 2005). To ensure converging to stationary posterior distributions, simulations were performed for 65,000 iterations with a burn in period of 50,000 iterations. In order to break the dependencies between draws and improve the mixing of the posterior samples in the Markov Chain, thinning was considered where every 15\(^{th}\) sample draw was kept. Two parallel chains were generated each having different initial values. MCMC convergence was assessed based on trace plots (i.e. plot of the iteration number
versus the value of the sampled parameter) and the scale reduction factor which should decline to one as the number of iterations increases indefinitely (Gelman and Rubin, 1992). The chains were merged to produce the posterior distributions of the GEV parameters as well as the latent parameters. The probability distributions of the 100-yr return level runoffs were calculated based on the estimated posterior distributions of the GEV parameters.

Several models were constructed using different covariates in the process stage including temperature, precipitation, latitude, and longitude. Model selection was performed using the deviance information criteria (DIC) (Spiegelhalter et al., 2002). The posterior distribution of the deviance statistic is obtained using the likelihood function of:

\[ D(Y_{it} | \theta_i) = -2 \sum_{i} \sum_{t} \log p(Y_{it} | \mu_i, \psi_i, \kappa_i) \]

where \( \theta_i = (\mu_i, \psi_i, \kappa_i) \)

An arbitrary scaling function can be added to this equation which is not considered here as it has no influence on the relative differences between DICs across the models (Banerjee et al., 2004). The posterior expectation of the deviance \( \bar{D} \) and the effective number of parameters \( p_D \) are calculated by:

\[ \bar{D} = \frac{1}{n} \sum_{i} D(Y_{it} | \theta_i) \]

(3-9)
\[ p_D = \overline{D} - D(Y_n|\theta) \]  

(3-10)

The deviance information criterion (DIC) is then determined by:

\[ DIC = \overline{D} + p_D \]  

(3-11)

DICs were calculated based on the posterior samples of each model. After comparing models with different covariates the model with latitude, longitude and 30yr average precipitation presented the lowest DIC, hence selected for extreme runoff analysis driven by different RCMs (Table 3-1).

Table 3-1: Model selection based on the deviance information criterion.

<table>
<thead>
<tr>
<th>Models</th>
<th>DIC</th>
<th>pD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model1</td>
<td>290071.9</td>
<td>10746.96</td>
</tr>
<tr>
<td>Model2</td>
<td>287712.3</td>
<td>10650.56</td>
</tr>
<tr>
<td>Model3</td>
<td>288995</td>
<td>10700.37</td>
</tr>
<tr>
<td>Model4</td>
<td>288478.2</td>
<td>10529.89</td>
</tr>
</tbody>
</table>

Model1: includes latitude and longitude as the covariates  
Model2: includes latitude, longitude and 30-yr average precipitation as the covariates (having the lowest DIC value therefore chosen for the rest of the analyses)  
Model3: includes latitude, longitude and 30-yr average temperature as the covariates  
Model4: includes latitude, longitude and 30-yr average precipitation and temperature as the covariates

3.4.1. Analysis of GEV Parameters

Figure 3-3 shows the posterior mean of the GEV location parameter of runoff extremes for each 1/8th degree grid cell. Legend is truncated for better visualization. On the left, results show the historical period of 1971-2000 and on the right the future period of 2041-2070. The estimated location parameter values are higher on the west side of the Cascade Range and on the Coast Range (geographical locations...
are shown in Figure 3-1). They show the decrease from winter to summer and then increase during fall. In the Rocky Mountains from Canada down to Idaho and Montana the values are lowest in winter and as the snowmelt starts they increase during spring and reach the highest in summer followed by a decrease in fall. For other parts of the PNW the models show different results but overall the values of the location parameters are higher on the east of Cascade Range during winter and on the east of the PNW during spring and summer. The results for future time period (figures on the right column) indicate an overall increase for location parameter for all seasons except for summer. In particular, the parameter values increase on the west of Cascade Range during winter and on the Rocky Mountains in northern parts during spring. The results show a decrease in parameters during summer especially on high elevations including the Cascade Range, Coast Range and Rocky Mountains. The Olympics Mountains on the northwest of the state of Washington shows the highest values during the winter compared to the other seasons except for the HRM3_hadcm3 model data. Increasing trend is seen for the future time period during winter in most of the models except the one mentioned previously along with RCM3_cgcm3, RCM3_gfdl and WRFG_ccsm which indicate a decrease in the winter season and an increase in fall for the latter.
Figure 3-3: Posterior estimate mean of the location parameter for each grid shown for the current period of 1971-2000 on left and future period of 2041-2070 on right for (a) CRCM_ccsm (b) CRCM_cgcm3 (c) HRM3_gfdl (d) HRM3_hadcm3 (e) RCM3_cgcm3 (f) RCM3_gfdl (g) WRFG_ccsm and (h) WRFG_cgcm3
The posterior distribution mean of the GEV scale parameter follows a similar pattern to the location parameter (Figure 3-4). Accordingly the HRM3_hadcm3, RCM3_cgcm3 and RCM3_gfdl show contradictory results on the Coast Range for the future time period when compared with other model RCM data inputs.
Figure 3-4: Posterior estimate mean of the scale parameter for each grid shown for the current period of 1971-2000 on left and future period of 2041-2070 on right for (a) CRCM_ccsm (b) CRCM_cgcm3 (c) HRM3_gfdl (d) HRM3Hadcm3 (e) RCM3_cgc3m (f) RCM3_gfdl (g) WRFG_ccsm and (h) WRFG_cgcm3

Figure 3-5 shows the posterior mean of the GEV shape parameter for each grid cell. Most of the models agree that on the west side of Cascade Range the shape parameters are lowest during winter, except for the Olympics Mountains, and highest during fall. However, results show highest values on the Rocky Mountains on north and east of the region during winter which reaches the lowest in summer. East of the Cascade Range, presents highest values during fall and lowest in spring. The model results using input data from HRM3_hadcm3, however, shows lower values in the west of Cascades and the Rocky Mountains during spring and higher values on the east of Cascade during summer while they reach the lowest in fall. The
future results of the estimated shape parameter for 2041-2070, when compared to the historical period, vary in different areas of the basins using different RCM input data. Overall the models indicate an increase in the GEV shape parameter on the west and east side of Cascade Range in all seasons. Almost all the models show decreased values in winter season over the Rocky Mountains which then increases in spring through fall.

![Maps of estimated shape parameter for different seasons](image-url)
Figure 3-5: Posterior estimate mean of the shape parameter for each grid shown for the current period of 1971-2000 on left and future period of 2041-2070 on right for (a) CRCM_ccsm (b) CRCM_cgcm3 (c) HRM3_gfdl (d) HRM3_hadcm3 (e)RCM3_cgcm3 (f) RCM3_gfdl (g) WRFG_ccsm and (h) WRFG_cgcm3

3.4.2. Spatial and Temporal Variations of Runoff Extremes

As explained previously, the RCM results were downscaled and bias-corrected using the UW observed data. These were used as forcing data to the VIC model and, the annual maxima of the estimated runoff (mm) over each grid cell were modeled using a GEV distribution based on a spatial hierarchical Bayesian model. The GEV parameters including the location, scale and shape parameters were then used to calculate the 100-yr return level runoff for each grid cell. Figure 3-6 shows the posterior mean of the 100-yr return level runoff for different RCMs compared with the observational data inputs. Overall models seem to agree well for all seasons except HRM3_hadcm3 which differs from others in the Rockies during winter and
west of Cascades during spring and summer. Also CRCM_cgcm3 and RCM3_cgcm3 overestimate the 100-yr return level runoffs in the Rockies during spring, however, marginal differences are still seen.
Figure 3-6: Posterior mean of the 100-year return level runoffs (mm) for each grid cell estimated based on observation versus downscaled RCM data which were used as VIC model forcing (a) Winter (b) Spring (c) Summer (d) Fall

The differences between the posterior means of the 100-yr return level runoffs using observation data and RCMs are better seen in Figure 3-7 (shown as Obs-RCM). During winter season the RCMs slightly underestimate the runoff values over the west of the Cascades except for WRFG_ccsm which indicates overestimation. On the north of the Rockies the RCMs slightly overestimate the runoffs except HRM3_hadcm3 which underestimates them and on the east they mostly underestimate the values. Over the east of the Cascades they show underestimation. During spring season the RCMs underestimate the runoff values over the west of the Cascades except for HRM3_hadcm3 which indicates overestimation. The situation is reversed on the Olympics Mountains. On the north of the Rockies the RCMs
overestimate the runoffs except CRCM_ccsm which underestimates them and on the east they mostly underestimate the values except for the HRM3_hadcm3. Over the east of the Cascades they show underestimation except for the HRM3_hadcm3. During summer season the RCM3 underestimate the runoff values over the west of the Cascades except for CRCM_cgcm3, HRM3_gfdl and HRM3_hadcm3 which indicate overestimation. On the north of the Rockies the RCMs underestimate the runoffs except HRM3_hadcm3 which overestimates them and on the east they mostly underestimate the values. Over the east of the Cascades they show underestimation. During fall season the RCM3 underestimate the runoff values over the west of the Cascades except for the south of the state of Oregon where they overestimate the return levels. On the north of the Rockies the RCMs underestimate the runoffs except HRM3_hadcm3 which overestimates them. Over the east of the Cascades they show underestimation. The overall biases of the RCMs are low which are given in Table 3-2.

Table 3-2: Overall bias (obs-sim) of the regional climate models compared with the observation.

<table>
<thead>
<tr>
<th></th>
<th>CRCM_ccsm</th>
<th>CRCM_cgcm3</th>
<th>HRM3_gfdl</th>
<th>HRM3_hadcm3</th>
<th>ECP2_gfdl</th>
<th>RCM3_gfdl</th>
<th>RCM3_cgcm3</th>
<th>WRFG_ccsm</th>
<th>WRFG_cgcm3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>0.123</td>
<td>0.31</td>
<td>0.109</td>
<td>2.561</td>
<td>0.326</td>
<td>0.096</td>
<td>0.655</td>
<td>-0.041</td>
<td>0.376</td>
</tr>
<tr>
<td>Spring</td>
<td>0.692</td>
<td>0.518</td>
<td>0.782</td>
<td>-2.697</td>
<td>0.617</td>
<td>1.115</td>
<td>0.822</td>
<td>1.093</td>
<td>0.941</td>
</tr>
<tr>
<td>Summer</td>
<td>0.085</td>
<td>-0.041</td>
<td>-0.138</td>
<td>-0.535</td>
<td>0.142</td>
<td>0.128</td>
<td>0.071</td>
<td>0.206</td>
<td>0.051</td>
</tr>
<tr>
<td>Fall</td>
<td>-0.088</td>
<td>0.021</td>
<td>0.082</td>
<td>-0.002</td>
<td>0.079</td>
<td>0.097</td>
<td>0.149</td>
<td>0.035</td>
<td>0.058</td>
</tr>
</tbody>
</table>

Each value reflects the median of biases for all cells. In winter, spring and summer HRM3_hadcm3 has the largest bias and in fall the RCM3_cgcm3 has the
largest one. Considering relative seasonal changes, highest values are seen during spring for all models, while lowest ones occur primarily in fall and occasionally in summer (for CRCM_ccsm, RCM3_cgcm3, and WRFG_cgcm3).
Figure 3-7: Difference between the posterior mean of the 100-year return level runoffs (mm) for each grid cell estimated based on observation versus downscaled RCM data which were used as VIC model forcing (a) Winter (b) Spring (c) Summer (d) Fall

Figure 3-8 shows the posterior mean of the change in 100-yr return level runoff (future-historic) for all models. Although the results vary between models for different regions especially the ones that are close to the coast, they show overall increase during winter, spring and fall and decrease during summer. On winter the models predict increase over the Cascades, the Rockies in particular on south of Canada and north of US. Models show contradictory results for the Olympics Mountains and Coast Range, for example CRCM_ccsm, CRCM_cgcm3 and WRFG_cgcm3 present increasing runoff over Olympics Mountains while others present reversed scenario. The models agree on the reduction of runoff extremes over the Cascades and the Rockies in summer. They also project an increase over the west of the Cascades and the Rockies during fall. The projections show increases in runoff extremes over the Rockies and west of Cascades especially in Canada during spring. In fact the highest predicted increase is seen for this region for this period.
Figure 3-8: Change in the posterior mean of 100-year return level runoff (mm) estimated for each grid cell for the future period of 2041-2070 versus the current period of 1971-2000 (a) CRCM_ccsm (b) CRCM_cgcm3 (c) HRM3_gfdl (d) HRM3_hadcm3 (e) RCM3_cgcm3 (f) RCM3_gfdl (g) WRFG_ccsm and (h) WRFG_cgcm3
The posterior standard deviations of the 100-yr return level runoffs are shown in Figure 3-9. Overall the values are higher over the west of Cascades during winter and over the Rockies on the north during summer. The results are lowest on the center and south of the region. In the future time period the standard deviations of the posterior distributions increase. In particular the results indicate increases over the west of Cascades and the Rockies during spring and fall.
(b)
Figure 3-9: Posterior standard deviation of the estimated 100-year return level runoff (mm) for each grid shown for the current period of 1971-2000 on left and future period of 2041-2070 on right for (a) CRCM_ccsm (b) CRCM_cgcm3 (c) HRM3_gfdl (d) HRM3_hadcm3 (e) RCM3_cgcm3 (f) RCM3_gfdl (g) WRFG_ccsm and (h) WRFG_cgcm3.

The posterior distributions of the precipitation coefficient (λ) for the observational data inputs are shown in Figure 3-10. The values are highest for the GEV location parameter and lowest for the shape parameter. They also vary between seasons. For the location parameter highest values occur during summer and lowest during fall. For the scale parameter highest values occur during summer however lowest occur during winter. For the shape parameter overall the precipitation coefficients are close to zero and during the summer and fall, in part, they are negative.
Figure 3-10: Posterior distribution of the precipitation coefficient for the mean (top), scale (middle) and shape (bottom) parameters for each season.

Figure 3-11 (left) presents the posterior means of the precipitation coefficients for all models, seasons and time periods. Comparison between the results for the historical versus future time periods indicates overall increase during the winter, spring and fall and decrease during summer for GEV location and scale parameters. Nevertheless HRM3_gfdl, HRM3_hadcm3, RCM3_cgcm3 and RCM3_gfdl show different trends depending on the season and parameter. This could be one of the reasons why these models presented contrary results regarding the estimates of the 100-yr return level runoffs. Considering the shape parameter, the results are more diverse than the previous two GEV parameters. Precipitation coefficient decreases on winter except for HRM3_hadcm3, RCM3_cgcm3, RCM3_gfdl and WRFG_cgcm3. On
spring it increases except for CRCM_ccsm, CRCM_cgcm3 and WRFG_cgcm3. On summer it decreases except for HRM3_hadcm3, RCM3_gfdl and WRFG_cgcm3, and on fall it increases except for HRM3_gfdl. Figure 3-11 (right) presents the posterior standard deviations of the precipitation coefficients. With regard to the location and shape parameters the overall seasonal trends correspond to the ones for the posterior means, although a model which shows increase in mean does not necessarily do so for the standard deviation. Furthermore for the shape parameter the overall standard deviation values decrease in the future during spring. Considering the scale parameter the standard deviations of the precipitation coefficients are highest for CRCM_cgcm3 during winter in the historical period and RCM3_gfdl during fall for the future period.
(a)

122
(b)
Figure 3.11: Posterior distribution mean (Top) and standard deviation (Bottom) of the precipitation coefficient for all RCMs confronting current and future time periods; results shown for (a) mean, (b) scale and (c) shape parameters.
3.4.3. **Hydro-Climate Analysis of the Pacific Northwest (PNW):**

In this section an analysis of the precipitation mean and extreme precipitation and runoff in PNW is performed. Aridity index is also assessed for current and future periods to reflect the impact of temperature (and the resulting evaporation) changes on the PNW. Aridity index is defined as the ratio of the annual potential evaporation to precipitation (Budyko, 1974) which describes the ratio of annual evapotranspiration to precipitation. Aridity index higher than unity ($\alpha > 1$) implies that the evaporative demand of the region under study is not met by precipitation. The aridity index classifies the regions into climate regimes of arid ($12 > \alpha \geq 5$), semi-arid ($5 > \alpha \geq 2$), sub-humid ($2 > \alpha \geq 0.75$) and humid ($0.75 > \alpha \geq 0.375$) (Arora, 2002).

The aridity index is calculated for each $1/8^\circ$ cell of the PNW for two periods of 1971-2000 and 2041-2070. Data from the CRCM-ccsm climate model is used for this analysis. As shown in Figure 3-12 (top) the areas in the west close to the Pacific Ocean and the Rockies on the north and east have humid climates. The areas on the east of the Cascades and west of the Rockies are dry. The precipitation mean shown in Figure 3-12 (bottom) agrees with the aridity indices showing higher precipitation rates for regions with $\alpha<1$. However while the mean precipitation increases in the future period compared to the historical one for most areas, the aridity index also increases especially for humid regions. This indicates that the rate of evaporation
increase due to the temperature rise is higher than the precipitation rate for these areas.

The seasonal spatial variation of mean precipitation also agrees well with the 100-yr return level precipitation (Figure 3-13). With regard to the extreme runoff (100-yr return level runoff) the timing of the snowmelt is an important factor in causing differences between spatial distribution of precipitation and runoff.

Figure 3-14 shows the change of each variable from the historical period to the future period. The spatial variation of climate variables indicates higher chance of having extreme precipitation in most areas with either increase or decrease of the mean precipitation. For example on the west of Cascades in Oregon during winter the mean precipitation is expected to decrease while the extreme precipitation increases.
Figure 3-12  Aridity index (top) and mean precipitation (bottom) for historical period of 1971-2000 and future period of 2041-2070 as well as the changes over the PNW.

Figure 3-13: Mean precipitation, 100-yr return level precipitation and runoff for winter and summer over the historical period of 1971-2000.
3.5. Conclusions

The impact of climate change on runoff extremes over the Pacific Northwest was studied in this study. Nine RCMs including CRCM_cgcm3, CRCM_ccsm, ECP2_gfdl (for the historical period), HRM3_gfdl, HRM3_hadcm3, RCM3_gfdl, RCM3_cgcm3, WRFG_cgcm3 and WRFG_ccsm covering the historical and future time periods of 1971-2000 and 2041-2070 were considered for this analysis. Hydrologic modeling was performed after downscaling of the precipitation, maximum and minimum temperature and wind speed using the quantile-mapping approach. Variable Infiltration Capacity (VIC) which is a distributed hydrologic model was used to provide daily runoff estimates (mm) for each cell of 1/8th degree resolution. Spatial
hierarchical Bayesian model was then applied on the cell-wise extreme runoff (mm) for both time periods and for all seasons.

The comparison between different RCMs regarding the changes in extreme runoff showed varying outcomes especially for the HRM3-hadcm3. This highlights the importance of multi-modeling techniques by evaluating the performance of model ensembles during a training period, assigning spatially distributed weights to each model and provide a multi-model average result for future projections. The hierarchical Bayesian model identified the spatial variations in the marginal distributions of the GEV parameters and the corresponding 100-yr return level runoffs. The posterior distributions of the latent variables provided information about the significance of each covariate on the extreme analysis in each season. Overall outcomes showed increases in the estimated 100-yr return level runoffs for most seasons particularly over the high elevation areas during winter. The Canadian portions of the study region reflected higher increases during spring. Summer indicated reduction of extreme events in most areas.

In this study stationarity assumption was assumed for downscaling by applying similar delta changes for each quantile obtained in historical period to the future RCMs. The spatial hierarchical Bayesian model was based on the conditional independence of runoff extremes at each cell.
This study was conducted based on the CMIP3 climate data under A2 emission scenario. In future studies, other climate scenarios from CMIP5 (Taylor et al., 2012) could be analyzed after dynamically downscaling the coarse resolution climate scenarios. The analyses similar to current work, may be used for risk assessment and storm water system designs and eventually adaptation strategies.
4. Chapter 4. Multi-Model Ensemble Analysis of the runoff Extremes for Climate Change Assessments

4.1. Background

Extensive global climate model (GCM) simulations are conducted until recently including the WCRP Coupled Model Intercomparison Project – Phase 3 (CMIP3) (Meehl et al., 2007a) with a coordinated set of global coupled climate model experiments which are also assessed in the IPCC fourth assessment report. Additionally CMIP5 experiments, incorporating about 20 modeling groups located around the world, is another continued effort for multi-model intercomparison and analysis (Taylor et al., 2012). These are based on pre-defined scenarios of future greenhouse gas emissions in order to assess changes in the oceanic-atmospheric quantities such as the special report on emission scenarios (SRES) and representative concentration pathways (RCP) corresponding to CMIP3 and CMIP5 respectively. Aside from a plethora of statistical downscaling approaches currently available (Najafi et al., 2011c; Wood et al., 2004a), several regional programs have been established in order to meet the increasing demand for finer spatio-temporal resolution datasets, such as the Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects (i.e. PRUDENCE) project (Christensen and Christensen, 2007), followed by the Ensembles-Based Predictions of Climate Changes and Their Impacts (i.e. ENSEMBLES) project (Van der Linden and Mitchell, 2009) which provides a suite of regional datasets for future climate
variations; the STAtistical and Regional dynamical Downscaling of EXtremes for European regions (i.e. STARDEX) project; and the North American Regional Climate Change Assessment Program (NARCCAP).

These experiments set the ground for climate change projection assessments as well as the uncertainty estimations of the individual and combined regional climate model simulations (Mearns et al., 2009). (Fowler et al., 2007b) Combined the results of RCMs from PRUDENCE ensemble assuming the models have equal skill. Nonparametric bootstrapping was used for probabilistic estimates of extreme precipitation. (Fowler and Ekström, 2009) further proposed a weighting scheme based on the spatial characteristics and extreme precipitation statistics of the RCMs and observations. Their results indicated that when all RCMs are driven by a single GCM the weighting model does not show significant impact. Studies on multi-modeling of RCMs suggest that a weighting procedure based on the model performance during a training period would enhance the simulation results (Casanova and Ahrens, 2009; Kug et al., 2008). Currently multi-model ensemble studies using RCM simulations are more focused on climate variables such as temperature (Kang et al., 2012; Smith et al., 2009; Tebaldi et al., 2005) and precipitation (Fowler and Ekström, 2009; Halmstad et al., 2012; Mailhot et al., 2011). (Kang et al., 2012) proposed a Bayesian hierarchical model to combine winter temperature from six RCMs in Phase-I of NARCCAP experiments. (Zhu et al.,
performed a Bayesian model averaging of the extreme rainfall intensities using the RCM simulations provided by NARCCAP.

BMA is a probabilistic weighting scheme for multi-model ensemble prediction. It considers a PDF for each model and provides a mixture PDF based on the optimized weights and variances. BMA has been successfully applied in various applications including the regional weather forecast of surface temperature and sea level pressure (Raftery et al., 2005), hydrologic model combination (Duan et al., 2007; Najafi et al., 2011b), precipitation and wind forecasting (Sloughter et al., 2010; Sloughter et al., 2007; Zhu et al., 2012) as well as the humidity field (Yang and Wang, 2012). These have motivated us to consider this method for probabilistic multi-modeling of runoff extremes from different RCMs.

The analyses are performed over the Pacific Northwest region. PNW covers the Columbia River Basin having the Columbia River as its predominant hydrologic and water resource feature. Winter (October-March) precipitation dominates the hydrologic characteristics of PNW (Miles et al., 2000). Precipitation in the mountains is higher especially on the western slopes of the Olympics and Coast mountain range and the Cascades. The region is climatically divided by the Cascade Mountains.

We performed analysis on multi-modeling of eight downscaled RCMs for the runoff extreme ensembles. RCM outputs are used to force the VIC hydrologic model
followed by hierarchical Bayesian modeling of extreme runoffs. Bayesian model averaging of precipitation, temperature, wind, and runoff time series are separately conducted. Three combination scenarios are defined and compared based on the simulated multi-model extreme runoff performances. BMA weights associated with each RCM are used to assess the RCMs performances as well. The uncertainties related to different models are also studied for the historical and projected runs.

In section 4.2 of this study a discussion of the NARCCAP datasets and the corresponding regional climate models is provided along with an introduction to the observational data. In section 4.3 the steps for multi-modeling of runoff extremes are explained including a brief introduction to the downscaling approach and VIC hydrologic model, spatial hierarchical Bayesian modeling of extremes, Bayesian model averaging along with the defined scenarios for multi-modeling. Section 4.4 provides the results of multi-modeling scenarios, spatially distributed weights corresponding to each model, average future projections of runoff extremes, uncertainties and between model standard deviations. A summary of the study along with the concluding remarks are provided in section 4.5.

4.2. Regional Climate Models

The North American Regional Climate Change Assessment Program (NARCCAP) (Mearns et al., 2009; Mearns et al., 2012) comprises several regional climate models (RCMs). It aims at producing higher resolution datasets for climate change impact
assessment and its corresponding uncertainties. RCMs generally have varying
dynamical formulations and physical parameterizations including different sub-grid
scale turbulence, radiative transport, boundary layer effects and moist processes.
NARCCAP experiments are performed over a similar spatio-temporal domain. The
spatial domain contains the United States, Canada and Northern Mexico. Phase-I of
the experiments were performed over the time period of 1979-2004 using the
National Center for Environmental Prediction (NCEP-2) reanalysis boundary
conditions taken as a substitute for observations. The lateral boundary conditions
include energy, moisture and momentum fluxes. In the second phase of the
experiments fully coupled global climate models replaced the NCEP reanalysis
boundary conditions. Simulations are performed in 50-km horizontal resolution,
over the periods of 1971-2000 and 2041-2070 based on the SRES A2 emission
scenario. In the A2 scenario the CO$_2$ concentration is expected to reach 575 and
870ppm (parts per million) by the middle and the end of the 21st century
respectively.

NARCCAP includes six RCMs (CRCM, ECP2, HRM3_MM5I, RCM3, and WRFG) that
are driven by four GCMs including the Canadian Global Climate Model version 3
(CGCM3, (Flato, 2005)), the NCAR Community Climate Model version 3
(CCSM3, (Collins et al., 2006)), the Geophysical Fluid Dynamics Laboratory (GFDL)
Climate Model version 2.1 (CM2.1, (Anderson et al., 2004)), and the United Kingdom
(UK) Hadley Centre Climate Model version 3 (HadCM3, (Gordon et al., 2000)). Each RCM is driven by only a portion of the GCMs. In this study data from eight RCM_AOGCMs combinations were considered containing CRCM_cgcm3, CRCM_ccsm, HRM3_gfdl, HRM3_hadcm3, RCM3_gfdl, RCM3_cgcm3, WRFG_cgcm3 and WRFG_ccsm.

Each RCM simulation provides 3-hourly as well as daily variable time series at each cell. In this study the 3-hourly precipitation (kg/m2s), daily maximum and minimum temperature (K), zonal and meridional wind speeds (m/s) were considered for the analysis.

It should be noted that commonly RCMs have systematic biases because of several reasons including the lateral boundary forcing and the model structures. In this study bias-correction was performed at the resolution of the observational data based on the quantile mapping approach. The CDFs of each 1/8° cell and the closest RCM grid were matched for the historical period of 1971-2000. The delta changes at each quantile were applied on the future CDFs of the corresponding RCM grid cell.

The observed climatological fields were taken from the retrospective analysis of (Maurer et al., 2002). These data are mostly calculated values based on other observations rather than direct measurements. The data consists of 1/8° gridded
precipitation, temperature and wind time series covering the period of January 1949 until July 2000.

4.3. Methodology

4.3.1. Runoff Extreme Assessment

Using the 1/8-degree observed climatological data as well as downscaled NARCCAP simulations the Variable Infiltration Capacity model was driven by precipitation, temperature and wind time series. The VIC hydrologic model estimates snow accumulation and melt, soil moisture, evapotranspiration, surface runoff as well as baseflow. Modeling is performed in a daily time step for the historical time period of 1971-2000 and future period of 2041-2070 by considering three layers of soil. Five elevation (snow) bands are considered to better characterize the snow processes at each grid cell. Simulations are made based on the water balance mode indicating that the surface temperature is set to the surrounding air temperature. The resulting runoff (mm) at each cell is considered for the hydrologic extreme analysis. Further information is provided in (Najafi and Moradkhani, under review).

Extreme value analysis has been developed to characterize the tail of a distribution (Coles, 2001). Generalize extreme value distribution is one of the distributions derived based on asymptotic results which is commonly used in hydrologic applications:
\[ F(y|\mu, \sigma, \kappa) = \exp \left\{ - \left[ 1 + \kappa \left( \frac{y - \mu}{\sigma} \right) \right]^{-\frac{1}{\kappa}} \right\}, \]

(4-1)

in which \( \mu, \sigma \) and \( \kappa \) are the location, scale and shape parameters respectively and \( 1 + \kappa(y - \mu)/\sigma > 0 \). GEV is applied when data are considered as annual, seasonal or monthly blocks in which the block maxima are taken as \( iid \) random variables.

The \( \tau \)-year return level extreme event can be obtained from:

\[ z_\tau = \mu - \frac{\sigma}{\kappa} \left[ 1 - \left( -\log \left( 1 - \frac{1}{\tau} \right) \right)^{\frac{1}{\kappa}} \right]. \]

(4-2)

Spatial hierarchical Bayesian models are recently proposed to combine rare extreme data from different locations. Three model stages constituting the hierarchy include the data, process and prior stage which are connected through conditional distributions.

The data stage of the hierarchical model defines the likelihood function over time and space:

\[ p_i(y_{i,t}|\mu_i, \sigma_i, \kappa_i) = \prod_{t=1}^{T} \prod_{i=1}^{S} \left[ 1 + \kappa_i \left( \frac{y_{i,t} - \mu_i}{\sigma_i} \right) \right]^{-\frac{1}{\kappa_i}} \times \exp \left\{ - \left[ 1 + \kappa_i \left( \frac{y_{i,t} - \mu_i}{\sigma_i} \right) \right]^{-\frac{1}{\kappa_i}} \right\} \]

(4-3)

\( Y_{i,t} \) is the annual maximum runoff at cell ‘i’ in time ‘t’. T is the total number of years for each of the historical and future periods and ‘S’ the total number of cells. (Martins and Stedinger, 2000) recommended a Bayesian prior to restrict the values
of the shape parameters ‘κ’ to a reasonable range. The function is a beta distribution with a mean of 0.1 which is supported in the [-0.5,0.5] interval:

\[
\pi(\kappa_i) = \frac{\Gamma(15)}{\Gamma(9)\Gamma(6)} (0.5 + \kappa_i)^5 (0.5 - \kappa_i)^5
\]  

(4-4)

Here the product of this prior and the likelihood function defined in equation above constitutes the data stage of the hierarchical model (i.e. \( p_i \times \pi(\kappa_i) \)).

The process stage of the hierarchy is defined based on the conditionally autoregressive (CAR) spatial model. The most popular CAR employment is the pairwise difference formulation known as the intrinsic autoregressive (IAR) model which is defined by the proximity matrix (Banerjee et al., 2004; Cooley and Sain, 2010).

It is assumed that each GEV parameter follows a normal distribution:

\[
p_x(\mu_i, \lambda_{\theta,i}, U_{i,\theta}, \tau_{\theta}^2) = \mathcal{N}\left(\lambda_{\theta,i} \cdot U_{i,\theta} + X_{i,q}, \frac{1}{\tau_{\theta}^2}\right)
\]  

(4-5)

with \( \theta = \{\mu, \psi, \kappa\} \), \( i \) denotes each grid cell, \( q \) the number of covariates and \( q' \) the number of their associated factors (\( q' = q+1 \)). \( X_{i,q} \) and \( \lambda_{\theta,q} \) with \( \theta \sim \mu, \psi, \kappa \) represent the covariates (i.e. explanatory variables such as the geographic coordinates, physical and climatologic characteristics that are associated with each cell) and their corresponding factors respectively; \( \tau_{\theta}^2 \) is the precision parameter.
$U_{i, \theta}$ represent the spatial random effect parameters which account for the dependencies between GEV parameters.

In the third stage of the hierarchical Bayesian model the priors are designated for the latent parameters including $\lambda_{\theta, \varphi}$ and $\Gamma$ which are defined in the process stage (Najafi and Moradkhani, under review).

4.3.2. Bayesian Model Average

Bayesian model averaging is a probabilistic mixture model based on post-processing of deterministic ensemble models over a training period (Duan et al., 2007; Najafi et al., 2011a; Raftery et al., 2005). In BMA each ensemble member $m_n$ is linked to an individual probability distribution function $p(y|m_n)$, which is the PDF of the hydro-climate variable $y$ conditional on model $m_n$ being the best ensemble member. The posterior probability of each model $m_n$ given the measured data $O$ (i.e. $p(m_n|O)$) is considered as the associated weight with that model reflecting its relative performance in the training period. Hence the BMA weights are probabilities and are summed up to unity. The resulting PDF associated with $y$ is a weighted average of each constituent PDF. The probability density function of the forecast variable $y$ can be assessed based on the total probability:

$$p(y) = \sum_{n=1}^{N} p(y|m_n, O) p(m_n|O) \tag{4-6}$$
where $O$ is the estimated runoff using observed data. By assuming a Gaussian distribution $p(y|m_n,O)$ can be considered as $g(y|m_n^2,\sigma^2_n)$ with $m_n$ as the mean and $\sigma_n$ as the standard deviation. In order to satisfy the Gaussianity assumption the Box-Cox transformation was applied over the estimated runoffs prior to the BMA implementation. The log-likelihood function is defined by:

$$l(w_1,\ldots,w_n,\sigma_n^2) = \log \left( \sum_{n=1}^{N} w_n \sum_{t=1}^{T} g(y_{n,t}^{obs}|m_{n,t},\sigma_{n}^{iter}) \right)$$  \hspace{1cm} (4-7)

The ML estimates of the unknown values of $p(m_n|O)$ (or $w_n$) and $\sigma_n$ are those that maximize the likelihood function. Here the likelihood is maximized using the iterative procedure of EM algorithm which swaps between Expectation and Maximization steps (Duan et al., 2007; McLachlan and Krishnan, 2007; Raftery et al., 2005). In the E step a latent variable is defined based on the values of the parameters $\theta = w_n, \sigma_n \{n = 1,2,3,\ldots,N\}$:

$$z_{n,t}^{iter} = \frac{g(y_{n,t}^{obs}|m_{n,t},\sigma_{n}^{(iter-1)})}{\sum_{n=1}^{N} g(y_{n,t}^{obs}|m_{n,t},\sigma_{n}^{(iter-1)})}$$  \hspace{1cm} (4-8)

where $g(y_{n,t}^{obs}|m_{n,t},\sigma_{n}^{(iter-1)})$ is a normal PDF with the mean of $m_{n,t}$ and standard deviation of $\sigma_{n}^{(iter-1)}$. In the M step the parameters ‘$\theta$’ are estimated given the values of $z_{n,t}$:

141
The convergence criterion is defined as \( l(\theta^{\text{iter}}) - l(\theta^{\text{iter-1}}) \leq \varepsilon \) where \( \varepsilon \) is the specified threshold.

The mean and variance of the BMA are obtained from equations below:

\[
E(y|\theta) = \sum_{n=1}^{N} w_n m_n \quad (4-11)
\]

\[
Var(y|\theta) = \sum_{n=1}^{N} w_n [m_n - E(y|\theta)]^2 + \sum_{n=1}^{N} w_n \sigma_n^2 \quad (4-12)
\]

The variance consists of two constituents, the between model variance and the within model variance.

In this study BMA is considered to provide a multi-model average result on climate variables as well as the runoff obtained from VIC hydrologic model runs.

4.3.3. Multi-Modeling of Extreme Runoff

It has been shown that multi-modeling of ensemble predictions can enhance the accuracy and interpretation of the analyses (Duan et al., 2007; Kharin and Zwiers, 2002; Najafi et al., 2011a; Najafi et al., 2012; Raftery et al., 2005). The ensemble merging method of Bayesian model averaging (BMA) has gained much attention in
the ensemble hydro-climate predictions. BMA is a statistical technique to provide a weighted average of the models based on their performances in the observational period.

For multi-modeling of extreme runoffs three scenarios are defined for Bayesian model averaging. First BMA is performed separately on the RCM output time series including the daily precipitation, maximum and minimum temperature and wind speed from eight regional climate models. The multi-model averaged climate variables then drive the variable infiltration capacity hydrologic model for the historical period of 1971-2000. Extreme analysis is performed on the runoff estimates on each 1/8° grid cell for all seasons. Second, climate inputs from each RCM drive the VIC model to assess the daily runoff time series at each cell. BMA of runoff estimates from the eight RCMs is followed by the extreme analysis based on the hierarchical Bayesian model over the averaged runoff. And finally, hydrologic modeling of each RCM input is followed by extreme analysis of the individual models. The 100-yr return level runoffs are then obtained for each model. The estimated weights corresponding to the daily runoff data are applied on the extreme runoffs from each model. The estimated weights for each cell were then considered on the 100-yr return level runoffs for the corresponding RCMs in order to obtain the multi-model average results for the future time periods.
BMA was separately performed for the individual 1/8° grid cells over the PNW. This allows for spatial assessment of the model performances using the estimated weights at each cell. The probabilistic approach of BMA also allows for the estimation of the between model variance (and corresponding standard deviation) as discussed previously. A comparative analysis of the between model standard deviation in the historical as well as future period is conducted in the following section.

4.4. Results

Bayesian model averaging was conducted over the downscaled daily precipitation data from eight RCMs on the historical period of 1971-2000. Individual analyses were performed on each 1/8° grid cell. The spatial distributions of the estimated weights are shown in Figure 4-1. In multi-modeling of precipitation data the RCM3_gfdl and HRM3_gfdl are assigned higher weights overall compared to the other RCMs. Lowest weights are given to RCM3_cgcm3. Regarding the Rockies on the north, west of the Cascades and the regions on the west of the Cascades higher weights are allotted to HRM3_hadcm3, HRM3_gfdl and RCM3_gfdl respectively.
Similar simulations were conducted over the maximum and minimum temperatures and wind speed that are the forcing inputs to the VIC hydrologic model. Analysis of the maximum temperature (Figure 4-2) shows that the RCMs are weighted closely which are higher than 0.1 close to the average weight of $1/8 = 0.125$. Overall the HRM3_hadcm3 and the WRFG_ccsm on the north of Rockies and west of the Cascades are assigned higher weights compared to the others.

*Figure 4-1: Spatially distributed weights obtained for each RCM based on daily precipitation.*
BMA weights obtained based on the minimum temperature values tend to be very close among the eight RCMs as well (Figure 4-3). Only exception is HRM3_hadcm3 which is assigned a small level of higher weights than the other RCMs. Similar results are observed for the wind speed (Figure 4-4).

The BMA weights corresponding to the temperatures and wind speed suggest that in a Bayesian view point all the RCMs are performing similarly well in estimating these variables. Therefore the multi-model average results would be similar to a scenario were all RCMs are equally weighted. This is in contrary to the precipitation estimates (Figure 4-1) where differences between model weights on various RCMs and spatial changes within each RCM are detectable.
Figure 4-3: Spatially distributed weights obtained for each RCM based on daily minimum temperature.

Figure 4-4: Spatially distributed weights obtained for each RCM based on daily wind speed.
After forcing the VIC hydrologic model with climate variables BMA was conducted on the resultant simulated runoffs. The assigned weights based on daily runoff estimates are much more spatially dispersed between RCMs than the ones related to the precipitation (Figure 4-5). Overall, the results indicate that CRCM_ccsm and WRFG_ccsm receive the highest weights for most of the grid cells, while HRM3_hadc3m, RCM3_cgcm3 and RCM3_gfdl have the lowest values.

*Figure 4-5: Spatially distributed weights obtained for each RCM based on daily runoff from the VIC hydrologic model.*

The spatial mean weights for each RCM can be regarded as an average performance of that RCM. The boxplots in Figure 4-6 show the ranges of the spatial mean weights based on the eight RCMs. Each boxplot corresponds to a hydro-climate variable. The spatial average performances between the models are close
regarding the maximum and minimum temperatures along with the wind speed. The mediums are close to 0.125 meaning that the models be equally weighted. The dispersions between models become more distinguished while considering precipitation and runoff. This emphasizes on the dissimilarities between regional climate models in simulating the precipitation which is followed by variations in hydrologic estimates.

As previously discussed three scenarios are considered for multi-modeling of 100-yr return level runoffs. In the first scenario individual Bayesian model averaging is performed over the daily precipitation, maximum and minimum temperature and wind speed from eight regional climate models. The multi-model
averaged climate variables then force the variable infiltration capacity hydrologic model for the historical period of 1971-2000. Extreme analysis based on the spatial hierarchical Bayesian model is performed on the runoff estimates on each 1/8° grid cell for all seasons. The estimated 100-yr return level runoffs are shown in Figure 4-7, fourth row. In the second scenario climate inputs from each RCM force the VIC model to assess the daily runoff timeseries at each cell. Bayesian model averaging of runoff estimates from the eight RCMs is followed by the extreme analysis based on the hierarchical Bayesian model over the averaged runoff. The estimated 100-yr return level runoffs are shown in Figure 4-7, third row. In the third scenario hydrologic modeling of each RCM input is followed by extreme analysis of the individual models. The 100-yr return level runoffs are then estimated for each model. The weights obtained from applying BMA on daily runoff data found in the second scenario are applied on the extreme runoffs from each model (Figure 4-7, second row). Results from all scenarios are compared with the estimated 100-yr return level runoffs using observed gridded data shown in Figure 4-7, first row. Overall the results suggest that the third scenario is the closest to observation while the other two scenarios tend to under estimate the extreme runoffs in most of the spatial domain in all seasons. However all the scenarios tend to capture the seasonal trend of the extreme events showing highest extremes on the west of the Cascades along with the Olympics during winter which tends to decrease by spring and
summer and again increases during fall. On the Rockies in the north because of the
snowmelt, the extreme runoffs tend to increase in spring and summer while
decrease in fall and winter when most of the precipitation falls and stays as snow.

Figure 4-7: Comparing the 100-yr return level runoffs obtained from the observed forcing data (top)
with scenarios of (from top to bottom respectively): assigning the BMA weights on estimated extremes,
BMA of daily runoff estimates followed by extreme assessment, and BMA of RCM data (precip, Tmax,
Tmin and wind) followed by hydrologic modeling and extreme analysis.
The biases of each multi-modeling scenario are shown in Figure 4-8 (scenarios three to one shown from top to bottom). The two scenarios in which Bayesian model averaging is applied before the extreme analyses provide negatively biased extreme estimations (under estimating the 100-yr return level runoffs).

Figure 4-8: Biases of different multi-modeling approaches (sim-obs) including (top to bottom): assigning the BMA weights on estimated extremes, BMA of daily runoff estimates followed by extreme assessment, and BMA of RCM data (precip, Tmax, Tmin and wind) followed by extreme assessment, and BMA of RCM data (precip, Tmax, Tmin and wind) followed by hydrologic modeling and extreme analysis.
The biases are strong on the west of Cascades and north in the Rockies during spring through fall. Biases of the scenario where the obtained weights are assigned to extremes are considerably lower than the ones from the other scenarios. The results suggest that in multi-modeling of extreme events application of the model averaging before extreme analysis would result in under estimated outcomes. The mean absolute errors of the estimated 100-yr return level runoffs (mm) from the Bayesian model averaging and its constituting models are compared in Table 4-1. The results indicate that the BMA outperforms the best model during the spring and summer. The results of BMA are also satisfactory for the rest of seasons.

**Table 4-1:** Comparison between the mean absolute errors of the estimated 100-yr return level runoffs (mm) from each model and from the Bayesian model averaging.

<table>
<thead>
<tr>
<th>Seasons</th>
<th>BMA</th>
<th>WRFG_cgcm3</th>
<th>WRFG_ccsm</th>
<th>RCM3_cgcm3</th>
<th>RCM3_gfdl</th>
<th>HRM3_hadcm3</th>
<th>HRM3_gfdl</th>
<th>CRCM_cgcm3</th>
<th>CRCM_ccsm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>2.84</td>
<td>3.26</td>
<td>3.17</td>
<td>3.37</td>
<td>3.27</td>
<td>7.98</td>
<td>3.03</td>
<td>3.8</td>
<td>2.78</td>
</tr>
<tr>
<td>Spring</td>
<td>2.48</td>
<td>3.02</td>
<td>3.05</td>
<td>3.64</td>
<td>3.25</td>
<td>5.65</td>
<td>3.27</td>
<td>3.24</td>
<td>2.92</td>
</tr>
<tr>
<td>Summer</td>
<td>1.99</td>
<td>2.16</td>
<td>2.03</td>
<td>2.68</td>
<td>2.5</td>
<td>3.48</td>
<td>2.65</td>
<td>2.64</td>
<td>2.42</td>
</tr>
<tr>
<td>Fall</td>
<td>2.11</td>
<td>2.49</td>
<td>1.76</td>
<td>2.61</td>
<td>2.9</td>
<td>2.18</td>
<td>2.54</td>
<td>2.62</td>
<td>2.66</td>
</tr>
</tbody>
</table>

Considering the third scenario of multi-modeling of extremes the between model variance is calculated for each cell over the historical period of 1971-2000 and future period of 2041-2070 according to:

\[ Var_y = \sum_{m=1}^{N} w_m \left[ m_n - E(y|O) \right]^2 \]

In which \( w_m \) is the weight assigned to each model based on the daily runoff timeseries, \( m_n \) is the estimated 100-yr return level runoff from each model and
$E(y|O)$ is the expected value of the extreme runoff simulations obtained through BMA. The squared root of the between model variances (i.e. standard deviations) are shown in Figure 4-9 for the two historical (red) and future (blue) time periods. The two peaks on the left side of each figure correspond to the Coast Range and the Cascades on the west, showing high standard deviations between models. From Figure 4-9 the between model deviations considerably increase over the future period especially on the west of Cascades and close to the Pacific coast. This indicates an increased model uncertainty in hydrologic extreme analysis for the future period compared to the historical one.

*Figure 4-9: Between model standard deviation over 1971-2000 (red) and 2041-2070 (blue); results obtained from the estimated BMA weights on the daily runoff data and 100-yr return level runoffs; each circle (star) corresponds to a 1/8 ° cell.*
The ensemble model average of the estimated 100-yr return level runoff for all seasons is shown in Figure 4-10. The results show that in winter time extreme runoff increases, in particular over the Cascades and Olympics Mountains while it decreases for the eastern Oregon and west of Idaho. In spring the results show increases over the Rockies especially for the portion in Canada as well as the west of Cascades and the Olympics Mountains. The 100-yr return level runoffs decrease in summer, while showing slight increases on the west coast. For the fall there are increases in extreme runoff on the west of Cascades and Rockies. There are slight increases in almost all other locations as well. Over the east of the Cascades and the west of the Rockies the values decrease from winter to summer followed by increase in fall.
Figure 4-10: Multi-model average results showing the estimated 100-year return level runoff (mm) obtained for the current and future periods (top) and their corresponding differences (bottom).

Figure 4-11 corresponds to Figure 4-10 where it compares the 100-yr return level runoffs for the historical and future time periods using the Bayesian model average results. Each circle represents a grid cell in the region with the ones above the splitting line showing increases in runoff extremes. The highest increases occur during spring and fall and the lowest in summer where most grid cells indicate decreases in runoff extremes.
Figure 4-11: Bayesian model average results of the estimated 100-year return level runoff (mm) obtained from the hierarchical Bayesian model for the current versus future time periods; each circle shows one of the 6392 grid cells with 1/8 degree resolution in the Columbia River basin.

4.5. Conclusions

Eight RCMs including CRCM_cgcm3, CRCM_ccsm, ECP2_gfdl (for the historical period), HRM3_gfdl, HRM3_hadcm3, RCM3_gfdl, RCM3_cgcm3, WRFG_cgcm3 and WRFG_ccsm covering the historical and future time periods of 1971-2000 and 2041-2070 were considered for this analysis. Downscaling of the precipitation, temperature and wind time series was performed using the quantile mapping approach. Observational as well as downscaled RCM data were used to drive the VIC
hydrologic model to provide daily runoff estimates. Spatial hierarchical Bayesian model was then applied on the cell-wise extreme runoff for both time periods and for all seasons. Multi-model average results provided by Bayesian model averaging were then compared to each constituting model.

Three scenarios were considered for multi-modeling of 100-yr return level runoffs. In the selected scenario Bayesian model averaging was applied on the daily runoff results obtained from the VIC simulations using different RCMs. The estimated weights for each cell were then considered on the 100-yr return level runoffs for the corresponding RCMs in order to obtain the multi-model average results for both the historical and future time periods. BMA produced satisfactory performance while it overtook the best model during spring and summer seasons. Overall outcomes showed increases in the estimated 100-yr return level runoffs for most seasons particularly over the high elevation areas during winter. The Canadian portions of the study region reflected higher increases during spring. Summer indicated reduction of extreme events in most areas.
5. Chapter 5. Assessing the Uncertainties of Hydrologic Model Selection in Climate Change Impact Studies

5.1. Background

For hydrologic climate change impact assessment, the hydrologic models are essential tools to assess runoff changes in the watershed of interest. To date, several models have been used to predict the potential consequences of climate change. These might range from simple conceptual lumped models to the comprehensive, physically-based, distributed ones – e.g. PDM (Kay, 2009), NWSRFS (Nash and Gleick, 1991), WatBal (Yates, 1996), macro-scale hydrological model (Arnell, 1999), VIC (Lettenmaier, 1999), HBV (Gardelin et al., 2002), MODFLOW (Kirshen, 2002), CATCHMOD (Wilby, 2005), SWAT (Jha et al., 2006), monthly water balance model (Jiang et al., 2007), ARC-SWAT (Moradkhani et al., 2010), PRMS (Jung et al., 2010a).

Although the hydrologic models may produce consistent results under historical climate conditions, they have often projected uncertain results regarding the probable changes in spatial and temporal distribution of runoff. This would be attributed to the structural differences in the hydrologic models and uncertainties in parameter estimation (Bae et al., in review; Jiang et al., 2007; Kay, 2009; Wilby and Harris, 2006). Previous studies have investigated the uncertainties in the climate change impact assessment, nonetheless none has quantified all sources of uncertainties. In addition, it is not clear what kind of hydrologic model with what
level of complexity is more suitable for climate change impact assessment and how much uncertainty is associated with the model. (Jiang et al., 2007) showed that the runoff change simulated by six water balance models differed up to 20% under an increasing temperature of (+4°C) and decreasing precipitation of (-20%). (Bae et al., in review) employed three semi-distributed models to investigate the uncertainty associated with hydrologic model structure using 13 GCM simulations with 3 GHG emission scenarios. They showed that monthly and seasonal runoff change, simulated by a single hydrologic model, is within ±10% of a multi-model ensemble result, except in the low flow season. However, the uncertainty attributed from the hydrologic model structure was not represented since the employed models had almost the same level of complexity.

This study investigates the uncertainty stemming from hydrologic model structure and parameters. The sections are organized as follows: in section 5.2, the characteristics of the basin and in-situ data used in the study are described; section 5.3 provides a detailed description of the hydrologic models, calibration process, and GCM downscaled data; the results presented in section 5.4 include the assessment of uncertainty in climate data, generation of plausible hydrologic models based on different objective functions and a procedure to statistically combine the results of all hydrologic models using Bayesian Model Averaging (BMA) method which is used. Section 5.5 provides the conclusion.
5.2. **Study Area and Data**

The Tualatin River Basin (TRB), located in the northwestern part of Oregon, USA (Figure 5-1), is one of the major tributaries to the Willamette River. The TRB, with an area of 1847 km², is surrounded by the Tualatin Mountains on the east and north east, the coast range on the west and north-west and the Chehalem Mountains on the south. Around 50 percent of the drainage basin is covered with forest, 35 percent is used for agricultural purposes and 15 percent is urbanized (Risley and Doyle, 1997). The cities are situated primarily in the middle and eastern edge of the TRB. The basin has a modified-maritime climate characterized by cool/wet winters (December, January, and February) and warm/dry summers (June, July, and August). Annual precipitation is about 1,140 mm, but about 75 percent of the precipitation falls in the form of rain during November through April. Mean temperature is 4°C in the winter and 18°C in the summer. River flow in the TRB reflects this seasonal variation in the precipitation. River flow at the mouth of the basin typically decreases from more than 5.7 m³/s in early May to less than 0.3 m³/s during July through October (Rounds and Wood, 2001).

The measured climate and hydrology data were collected to calibrate the parameters of each hydrologic model. Maximum and minimum temperature, and precipitation data were obtained from the National Oceanic and Atmospheric Administration Cooperative Observer Program (NOAA-coop, 2009) for 1973-2006.
The streamflow gauging station data at the mouth of TRB (USGS #14207500) was used to estimate the performance of the hydrologic models (USGS-NWIS, 2009) (see Table 5-1). The Digital Elevation Model (DEM), soil map, land cover, and geology were obtained from other sources (Fegeas et al., 1983; McFarland, 1983; NRCS, 1986; USGS, 1990).

Figure 5-1: Tualatin River Basin, river network, climate stations, and the USGS streamflow gauging station.

Table 5-1: Climate and streamflow stations used in this study

<table>
<thead>
<tr>
<th>Data type</th>
<th>ID</th>
<th>Name</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Elevation</th>
<th>Data period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climate</td>
<td>COOP 350595</td>
<td>Beaverton 2 SSW</td>
<td>45.455</td>
<td>-122.820</td>
<td>82.3</td>
<td>1973-2006</td>
</tr>
<tr>
<td>Climate</td>
<td>COOP 352325</td>
<td>Dilley 1 S</td>
<td>45.483</td>
<td>-123.124</td>
<td>50.3</td>
<td>1973-2006</td>
</tr>
<tr>
<td>Climate</td>
<td>COOP 352997</td>
<td>Forest Grove</td>
<td>45.524</td>
<td>-123.103</td>
<td>54.9</td>
<td>1973-2006</td>
</tr>
<tr>
<td>Climate</td>
<td>COOP</td>
<td>Glenwood 2</td>
<td>45.656</td>
<td>-123.311</td>
<td>196.3</td>
<td>1973-2006</td>
</tr>
</tbody>
</table>
Climate COOP

<table>
<thead>
<tr>
<th>Climate</th>
<th>353318</th>
<th>WNW</th>
<th>Hillsboro</th>
<th>45.514</th>
<th>-122.990</th>
<th>48.8</th>
<th>1973-2006</th>
</tr>
</thead>
</table>
| Streamflow USGS

| Streamflow | USGS 14207500 | Tualatin River at West Linn | 45.351 | -122.510 | 26.4 | 1928-2006 |

### 5.3. Methodology

Four hydrologic models of varying complexities are employed, from a simple conceptual model to a physically-based, semi-distributed model. The models are applied to the historical streamflow data of the TRB. The model parameters are calibrated using the Shuffle Complex Evolution algorithm – University of Arizona (SCE-UA) (Duan et al., 1993) method, as a global optimization procedure (Duan et al., 1993). To consider the uncertainty of hydrologic model parameter estimation, three objective functions are employed, including the Root Mean Square Error (RMSE), Heteroscedastic Maximum Likelihood Estimator (HMLE) and the absolute BIAS. The combination of eight GCMs with two emission scenarios is used to account for the uncertainties associated with the GCM structures and emission scenarios (Table 5-2). Three future time slices, 2010-2039, 2040-2069, and 2070-2099 were considered and the analyses were compared to 1960-1989 as the reference period.
Table 5-2: Description of the Global Climate Models used in this study ([Randall et al., 2007])

<table>
<thead>
<tr>
<th>Model ID</th>
<th>Country</th>
<th>Resolution</th>
<th>Emission scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Atmosphere</td>
<td>Ocean</td>
</tr>
<tr>
<td>CCSM3</td>
<td>USA</td>
<td>1.4°×1.4°</td>
<td>1.0°×1.0°</td>
</tr>
<tr>
<td>CNRM-CM3</td>
<td>France</td>
<td>1.9°×1.9°</td>
<td>2.0°×2.0°</td>
</tr>
<tr>
<td>ECHAM5/MPI-OM</td>
<td>Germany</td>
<td>1.9°×1.9°</td>
<td>1.5°×1.5°</td>
</tr>
<tr>
<td>ECHO-G</td>
<td>Germany/Korea</td>
<td>3.9°×3.9°</td>
<td>2.8°×2.8°</td>
</tr>
<tr>
<td>IPSL-CM4</td>
<td>France</td>
<td>2.5°×3.75°</td>
<td>2.0°×2.0°</td>
</tr>
<tr>
<td>MIROC3.2(hires)</td>
<td>Japan</td>
<td>1.1°×1.1°</td>
<td>0.2°×0.3°</td>
</tr>
<tr>
<td>PCM</td>
<td>USA</td>
<td>2.8°×2.8°</td>
<td>0.7°×1.1°</td>
</tr>
<tr>
<td>UKMO-HadCM3</td>
<td>UK</td>
<td>2.5°×3.75°</td>
<td>1.25°×1.25°</td>
</tr>
</tbody>
</table>

5.3.1. Hydrologic models

Three conceptual lumped hydrologic models with varying complexities and one semi-distributed model were selected for this study, including: Thornthwaite-Mather model (TM) ([Thornthwaite and Mather, 1957]), HYdrologic MODel (HYMOD) ([Boyle et al., 2000]), Sacramento Soil Moisture Accounting model (SAC-SMA) ([Burnash et al., 1973]) and, Precipitation and Runoff Modeling System (PRMS) ([Leavesley et al., 1983]). The parameters of each model were calibrated using SCE-UA method based on the three distinct objective functions. Therefore three different
parameter sets were used in this study. The models are briefly described in the order of increasing complexity (see Table 5-3).

*Table 5-3: a) Lumped hydrologic models parameter description*

<table>
<thead>
<tr>
<th>SAC-SMA</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>UZTWM</td>
<td>Upper zone tension water capacity (mm)</td>
</tr>
<tr>
<td>UZFWM</td>
<td>Upper zone free water capacity (mm)</td>
</tr>
<tr>
<td>UZK</td>
<td>Fractional daily upper zone free water withdrawal rate</td>
</tr>
<tr>
<td>PCTIM</td>
<td>Minimum impervious area (decimal fraction)</td>
</tr>
<tr>
<td>ADIMP</td>
<td>Additional impervious area (decimal fraction)</td>
</tr>
<tr>
<td>ZPERC</td>
<td>Maximum percolation rate coefficient</td>
</tr>
<tr>
<td>REXP</td>
<td>Percolation equation exponent</td>
</tr>
<tr>
<td>LZTWM</td>
<td>Lower zone tension water capacity (mm)</td>
</tr>
<tr>
<td>LZFSM</td>
<td>Lower zone supplemental free water capacity (mm)</td>
</tr>
<tr>
<td>LZFPM</td>
<td>Lower zone primary free water capacity (mm)</td>
</tr>
<tr>
<td>LZSK</td>
<td>Fractional daily supplemental withdrawal rate</td>
</tr>
<tr>
<td>LZPK</td>
<td>Fractional daily primary withdrawal rate</td>
</tr>
<tr>
<td>PFREE</td>
<td>Fraction of percolated water going directly to lower zone free water storage</td>
</tr>
<tr>
<td>KQ</td>
<td>Storage constant</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HyMod</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cmax</td>
<td>Maximum storage capacity within the watershed</td>
</tr>
<tr>
<td>bexp</td>
<td>Degree of spatial variability of the soil moisture capacity within the watershed</td>
</tr>
<tr>
<td>alpha</td>
<td>A factor for partitioning the flow between two series of tanks</td>
</tr>
<tr>
<td>Rq</td>
<td>Residence time parameter of quick-flow</td>
</tr>
<tr>
<td>Rs</td>
<td>Residence time parameter of slow-flow</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Thornthwaite-Mather</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWC</td>
<td>Available water capacity</td>
</tr>
<tr>
<td>f</td>
<td>Reservoir coefficient</td>
</tr>
</tbody>
</table>
a) The Thornthwaite-Mather (TM) model (Thornthwaite and Mather, 1957) is a simple hydrologic model and is composed of only one soil reservoir and one watershed storage reservoir. It has two parameters: available water capacity (AWC) and routing parameter (f), which transfers water from watershed storage to the river. The precipitation increases water in the soil reservoir and the potential evapotranspiration (PET) decreases it. When the soil water exceeds AWC, it is stored in the watershed storage. The stored water flows to the river based on a linear equation.

b) The HYdrologic MODel (HYMOD) is a rainfall excess model which consists of a nonlinear tank connected with three identical quick-flow tanks, which are in parallel to a slow-flow tank representing the groundwater flow (Boyle et al., 2000; Moradkhani et al., 2005b). The three quick-flow tank storages represent the temporary detentions. The soil moisture content is simulated as the relation of the maximum storage capacity and the degree of spatial variability of the soil moisture capacity within the watershed. The excess water from the soil zone flows into the quick-flow tanks and the groundwater, according to a partitioning factor (α). The quick-flow and the slow-flow from the two series of tanks route by the residence time parameters (Rq, Rs).

c) The Sacramento Soil Moisture Accounting (SAC-SMA) model is a spatially-lumped, continuous hydrologic model, which divides the basin into two soil zones,
an upper zone and a lower zone (Burnash et al., 1973). The upper zone simulates the short term storage of the basin and the lower zone represents the underground soil in the long-term storage. Each zone includes tension water, which produces no runoff and is influenced by the evapotranspiration, and free water, which represents the water that flows, or percolates, downward and the water that evaporates. Runoff generation in the SAC-SMA consists of direct runoff from the impervious part of the land, surface runoff, interflow from the upper zone, and both primary and supplemental baseflow from the lower zone.

d) Precipitation Runoff Modeling System (PRMS) is a deterministic physically-based, distributed-parameter modeling system (Leavesley et al., 1983). The model is designed to analyze the effects of changes in precipitation, climate, and land use based on hydrologic response units (HRU), which are assumed to be homogeneous with respect to each unit’s hydrologic response. The model requires daily precipitation and minimum and maximum daily air temperature as inputs. It generates the water and energy balances for each HRU, at each time step.

5.3.2. GCM Derived Climate Data

This study uses 16 statistically downscaled climate model simulations (8 GCMs with 2 GHG emission scenarios) for the period of 1960-2099 generated by the Climate Impacts Group (CIG) at the University of Washington for integrated assessment of climate change impacts in the Pacific Northwest (see Table 5-1). The
GHG emission scenarios as explained in the Intergovernmental Panel on Climate Change 4th Assessment Report (IPCC-AR4) include: A1FI, A2, A1B, B2, A1T and B1 from the warmest to the coolest scenarios (Solomon, 2007). The CIG chose A1B and B1 scenarios because those are used widely at global modeling groups and implemented in the IPCC-AR4 (Mote and Salathé Jr, 2009). The CIG disaggregates monthly mean GCM simulations of precipitation, maximum and minimum temperature and wind speed to daily variables at 1/16 degree grid spacing based on gridded measured data with the same resolution. To remove the bias of the GCM simulations, the quantile mapping method (Wood et al., 2002) is employed. Then, these data are spatially downscaled by using scale factors obtained from the gridded measured data (Widmann, 2003). The spatially downscaled GCM simulations are finally extended to daily time series based on historical data, by considering the changes in the climate variables (Salathe, 2005).

5.3.3. Hydrologic Modeling and Calibration

The four hydrologic models described above were calibrated using the measured streamflow data at the watershed outlet. Areal-averaged climate data were used to reduce the impact of uncertain spatial forcing data in hydrologic modeling, although PRMS model can use the distributed climate forcing for each HRU. Each model is calibrated separately using the Shuffle Complex Evolution algorithm – University of Arizona (SCE-UA) (Duan et al., 1993) based on three distinct objective functions.
which accordingly results in three different sets of optimum parameters for each model. The models were calibrated and validate for the periods of 1973-1989 and 1990-2006 respectively. SCE-UA has been used extensively and verified to be an effective global optimization method for the calibration of hydrologic models (Muttil and Jayawardena, 2008). However, emerging approaches based on Markov Chain Monte Carlo (Vrugt et al., 2003), sequential Monte Carlo methods (Leisenring and Moradkhani, 2012; Moradkhani et al., 2005a; Moradkhani et al., 2005b) and block bootstrap resampling to enhance the robustness of the SCE-UA method (Ebtehaj et al., 2010), have been introduced which can be considered in future studies. To address the uncertainty in parameter estimation, three distinct objective functions were used, including the Root Mean Square Error (RMSE), Heteroscedastic Maximum Likelihood Estimator (HMLE) and the absolute BIAS. The RMSE is an appropriate measure when the measurement errors are known to be uncorrelated and homoscedastic, or when the properties of the measurement errors are unknown (Gupta et al., 1998). On the other hand, the HMLE is a goodness of fit estimate when the measurement errors are believed to be heteroscedastic (Sorooshian and Dracup, 1980). These objective functions force the hydrologic models to favor different phases of the hydrograph. The RMSE and Bias force the models to fit the high flows and low flows respectively, while the HMLE places an equal emphasis on all parts of the streamflow hydrograph, which compromises
between RMSE and Bias (Duan et al., 2007). The calibrated models are then evaluated by comparing the simulation with the measured streamflow using the Nash Sutcliffe Efficiency (NSE) and RMSE goodness of fit measures (see Table 5-4).

Table 5-4: Hydrologic model performance in the calibration and evaluation periods.

<table>
<thead>
<tr>
<th>Hydrologic Model</th>
<th>Objective Function</th>
<th>Calibration (C.) Evaluation (E.) RMSE</th>
<th>NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAC-SMA</td>
<td>RMSE</td>
<td>C. 14.97</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 17.61</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>BIAS</td>
<td>C. 14.00</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 15.81</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>HMLE</td>
<td>C. 19.26</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 20.37</td>
<td>0.89</td>
</tr>
<tr>
<td>HyMod</td>
<td>RMSE</td>
<td>C. 16.69</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 19.66</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>BIAS</td>
<td>C. 17.34</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 20.65</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>HMLE</td>
<td>C. 19.08</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 23.11</td>
<td>0.86</td>
</tr>
<tr>
<td>TM</td>
<td>RMSE</td>
<td>C. 21.56</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 23.69</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>BIAS</td>
<td>C. 22.00</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 24.24</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>HMLE</td>
<td>C. 25.44</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 28.42</td>
<td>0.79</td>
</tr>
<tr>
<td>PRMS</td>
<td>RMSE</td>
<td>C. 23.01</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 25.10</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>BIAS</td>
<td>C. 21.29</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 24.01</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>HMLE</td>
<td>C. 26.01</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E. 29.01</td>
<td>0.78</td>
</tr>
</tbody>
</table>
To analyze the uncertainty of the projected streamflow, all models simulated the long-term daily streamflow for the period of 1960-2099 using 16 downscaled climate data. The annual mean and the uncertainty in seasonal runoff projection are assessed for three future time slices. The total variance of the combined hydrologic models is then obtained by the Bayesian Model Averaging (BMA) method. The range of the obtained total variance for each GCM, from the reference period to the future periods, is then investigated.

5.4. Results and Discussion

5.4.1. Precipitation and Temperature

To analyze the projection trends of the climate and hydrologic variables, the reference period (1960-1989) and the projected time series were partitioned into time slices of thirty years and the average variations of these variables were examined. The two emission scenarios (A1B and B1) were analyzed separately and the uncertainties associated with the 8 GCMs were then investigated. Figure 5-2 shows the precipitation and temperature changes in the three future time slices relative to the reference period for the 8 GCMs. The relative changes are calculated as the percentage change for the precipitation and as the simple subtraction for the temperature. The uncertainties and the trends of the climate variables expressed for each season were used to explain the associated uncertainties in the streamflow.
To assess the sensitivity of the runoff in the TRB to precipitation and temperature, the aridity index, first presented by (Budyko, 1974), was used. The aridity index \((a)\) is a function of the ratio of the actual average evapotranspiration to the average precipitation in a long term period.

\[
\frac{\bar{E}}{\bar{P}} = a \left( \tanh \left( \frac{1}{a} \right) \right) \left( 1 - \frac{\cosh a + \sinh a}{\cosh a} \right)^{\frac{1}{2}}
\]  

where \(\bar{E}\) and \(\bar{P}\) represent the actual evaporation and the average precipitation over the basin, respectively. \(a > 1\) defines a water-limited region indicating that the basin is more sensitive to the changes in radiation and temperature than the precipitation, while \(a < 1\) is an indicator for an energy-limited basin. In other words, for basins that have aridity index of less than 1.0, annual evaporation is constrained by the annual supply of energy. Conversely, for basins that have the aridity index equal to 1.0 or greater, annual evaporation is constrained by the annual supply of water (Milly and Dunne, 2002; Risley et al., in review). In the TRB, using the measured data, the water balance approach and equation above, we estimated the actual evaporation and an aridity index value of \(a = 0.4\), indicating that the runoff in the basin is more sensitive to the precipitation changes (Milly and Dunne, 2002).

In general the uncertainties associated with GCMs increase by moving further in the future time periods, which is more significant in the A1B emission scenario than in the B1 scenario, especially for the temperature (Figure 5-2).
As expected, the ensemble temperature projections increase in all seasons with the A1B emission scenario showing more increase than the B1 scenario in 2050s and 2080s. The highest temperature increase occurs in summer and fall respectively. The precipitation projections, however, show irregularities considering the trends and uncertainties. Some GCMs show contradictory results in that both an increase and a decrease in the projections are seen. The highest uncertainties in the precipitation projections occur in winter and summer.

The mean of the relative changes of precipitation from the 8 GCMs shows an increase in all seasons except in the summer (Table 5-5). As mentioned before, the precipitation rate is low in the summer for the TRB, hence the overall impact of the
temperature rise would be an increase in the annual precipitation. This is shown in Table 5-6, which represents the relative changes of the average annual temperature and precipitation of all GCMs. In order to calculate the relative precipitation change, the mean precipitation of each GCM is summed over all seasons and is then averaged over the whole GCMs for each time period. As expected, the results of the B1 emission scenario are close to the A1B scenario in 2020s and become more apart in 2050s and 2080s. Also higher temperature and precipitation changes are seen for the A1B compared to the B1.

Table 5-5: Seasonal precipitation change relative to the reference period

<table>
<thead>
<tr>
<th>Season</th>
<th>A1B (bias %)</th>
<th>B1 (bias %)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Winter</td>
<td>Spring</td>
</tr>
<tr>
<td>2020s</td>
<td>1.71</td>
<td>1.79</td>
</tr>
<tr>
<td>2050s</td>
<td>3.29</td>
<td>6.67</td>
</tr>
<tr>
<td>2080s</td>
<td>5.08</td>
<td>5.85</td>
</tr>
</tbody>
</table>

These results do not consider the uncertainties associated with the GCMs and the downscaling method. Although all GCMs agree on the temperature increase for the projected scenarios, they sometimes show contradictory results for the
precipitation, and the uncertainties appear to be high where further consideration is necessary. The focus of this study is on the uncertainties associated with the hydrologic models and a comparison with those arising from the GCMs. For this purpose, the downscaled climate data from all GCMs are used as forcing inputs to each hydrologic model. The Bayesian Model Averaging (BMA) method is then implemented to obtain a combined hydrologic prediction, where the degree of contribution of each hydrologic model (i.e., model weight) in combined model result is obtained. In addition to model weights, the associated uncertainty reflected in total variance of combined model is also calculated.

5.4.2. Hydrologic Impact Assessment

Three lumped hydrologic models along with one semi-distributed model, each calibrated with three distinct objective functions are employed to allow investigating the uncertainties related to the hydrologic model parameters. The downscaled data from each of the 8 GCMs were used as the forcing input to the calibrated hydrologic models, hence for every emission scenario (A1B and B1) a total number of 12(models) × 8(GCMs) = 96 runoff simulations were computed. Considering that the basin is energy limited and the precipitation uncertainty is high in both summer and winter seasons, the hydrologic impact assessment is focused on these two seasons while the model uncertainties in the fall season is also presented.
The analysis of each hydrologic model simulation, based on different GCMs, provided the uncertainties associated with the GCMs assuming the model performed best. The results suggest that the uncertainties in the GCMs vary between different hydrologic models in the summer; however the variations are not significant during the winter. This is shown in Figure 5-3 where the results of each model forced by 8 GCM data are expressed as percent bias. The bias is obtained from the 30 years average runoff in the future time periods and the ones in the reference time period. The graphs show the runoff results for the two A1B and B1 emission scenarios. The TM model, having the simplest structure of the models used, results in the highest uncertainty in the summer. The lowest GCM uncertainties in 2020s and 2050s time periods are seen in the HYMOD and PRMS model results, and for the 2080s in the SAC-SMA and PRMS results.

The discrepancies between the models in the dry season can be explained in terms of soil moisture change, which is higher in the summer because of the elevated radiation, which directly influences the hydrologic model states, and therefore the streamflow simulations. This indicates that the model state responses are more sensitive to the dry extremes than to the wet ones. The model results also vary according to the different objective functions. For example in the dry season, the lumped models based on RMSE show more runoff reduction, and the ones based
on the HMLE function have the lowest reductions. For the semi-distributed model the highest reduction is related to the BIAS objective function.

The relative changes in the streamflow for three future time slices (2020s, 2050s, 2080s), relative to the reference time slice (1970s), are shown as probability density function (PDF) plots for summer and fall seasons in Figure 5-4 for the A1B and B1 emission scenarios. In this figure, 8 downscaled GCM data are incorporated as forcing to each hydrologic model calibrated based on the RMSE objective function. To develop these figures kernel density function is used as explained in detail by (Silverman, 1998).
Figure 5-3: Runoff changes relative to the reference period showing the uncertainties in the GCMs for each hydrologic model for winter and summer seasons; RMSE, BIAS and HMLE reflect the hydrologic model parameter set.

Figure 5-4: Probability density functions of runoff change in fall and summer seasons for 8 GCMs under A1B and B1 emission scenarios. The models were calibrated using the RMSE objective function.
The resulting performance indicates that differences between the hydrologic models increase in the summer. The largest difference is seen in the TM model, which performs poorly in the low flow conditions. Also larger differences of PDFs for different models are seen in the summer season as compared to fall season. Overall the HYMOD and PRMS models show less uncertainties in the summer compared to the other models. Although in fall the uncertainties as reflected in the spread (standard deviation) of different models seem to be somewhat similar, the mode of the distributions show considerable differences as pronounced more clearly in the 2050s period.

The uncertainties in the extreme flow rates for the highest 80% of the measured flow in 1977-2006 are also analyzed (Figure 5-5). Results show increases in high flow rates for the future time periods compared with the observed data. The model performances are similar in the high flow.
Figure 5-5: Hydrologic model uncertainty in high flow using the csm downscaled forcing data

5.4.3. Bayesian Model Averaging

The current practice in hydrologic modeling is to rely on a single model to perform the simulation and prediction of land surface properties. Despite the tremendous progress, efforts and investment put into developing more hydrologic models, there is no plausible reason that any particular model in existence is superior to other models for various applications and under all circumstances.
(Ajami et al., 2007; Hsu et al., 2009; Raftery et al., 2005). This results in reducing the size of the plausible model space and often leads to predictions that may well-represent some phenomena or events at the expense of others. In addition, assessment of predictive uncertainty, based on a single model, is subject to statistical bias and the structural error inherent in any single model. To assess the uncertainties in the hydrologic model selection and to separate them from the GCM uncertainties, the Bayesian Model Averaging (BMA) (Hoeting et al., 1999), which has been applied successfully in various studies, e.g. (Ajami et al., 2007; Duan et al., 2007; Raftery et al., 2005) among others, was implemented. BMA is a statistical scheme to infer a combined probabilistic prediction, which possesses more reliability and skill than several competing models used in the model combination. The BMA predictive probability distribution function (pdf) of a quantity of interest is a weighted average of the individual model pdfs, providing that the individual forecasts are bias-corrected. The weights presented by $p(m_n|O)$ are the posterior probability of each model given the observed data $O$. The sum of the weights for all models then equals to 1. One can assess the probability density function of the forecast variable $y$ based on the law of total probability:

$$p(y) = \sum_{n=1}^{N} p(y|m_n, O) p(m_n|O)$$

(5-2)
where $O$ is the observed streamflow. According to (Raftery et al., 2005), $p(y|m_n, O)$ can be simplified by a Gaussian distribution like $g(y|m_n, \sigma^2)$, where $m_n$ is the mean and $\sigma_n$ is the standard deviation, which can be obtained using an optimization procedure such as the Expectation Maximization (EM) algorithm. Considering the non-Gaussian distribution of streamflow error, (Duan et al., 2007) pre-processed both modeled and measured streamflow data using the Box-Cox transformation prior to the BMA implementation to make the transformed variables close to the Gaussian distribution. In the analysis presented here, the same procedure is followed. The simulation results from each model are first bias-corrected using the linear regression. The measurement and the bias-corrected model simulations are then transformed using the Box-Cox equation. In order to find the unknown values of $p(m_n|O)$ (or the model weights) and $\sigma_n$ (in $g(y|m_n, \sigma^2)$) the log-likelihood function is then defined as:

$$
I(w_1, \ldots, w_n, \sigma^2) = \sum_{i=1}^{X} w_i \sum_{t=1}^{T} g(y_{it}^{ob} \mid m_{it}, \sigma^{2_{it}}) 
$$

(5-3)

The unknown values are calculated using the iterative procedure of EM (Raftery et al., 2005), which swaps between Expectation and Maximization steps (For a detailed explanation of this procedure, please see (Duan et al., 2007; Raftery et al., 2005)).

The mean and variance of the BMA can then be obtained from:
\[ E(y|O) = \sum_{n=1}^{N} w_n m_n \]  

\[ \text{Var}(y|O) = \sum_{n=1}^{N} w_n [m_n - E(y|O)]^2 + \sum_{n=1}^{N} w_n \sigma_n^2 \]

in that \( m_n \) is replaced by \( a_n + b_n m_n \) \( a_n + b_n m_n \) if the model is bias-corrected. The variance consists of two constituents, the between model variance and the within model variance. In our study the total variance is obtained as 127.4 (standard deviation of 11.3), which represents the total uncertainty associated with the combined hydrologic model during the reference period.

The EM algorithm is a very efficient method for finding the optimal value. The model converged after 13 iterations. The final weights show that the Sacramento model calibrated based on HMLE function has the highest weight (Figure 5-6). The HYMOD model although simple in structure, has gained the second highest weight after Sacramento model. However, the TM model is assigned the lowest weight, which is due to its very simple and incomplete structure leading to poor performance in low flow periods.
After determining the weight of each model the expected value of the BMA probability distribution can be obtained. The result is compared to the Max-Min range of models ensemble and the measurement in a two year period of 1983-1985 in Figure 5-7. It is expected that the BMA be closer to the measurement compared to each of the model results. This figure shows that the performance of the BMA has been acceptable.
As seen the result of the BMA model is very close to the best model result (e.g. BMA-NSE= 0.938 vs Best-model-NSE= 0.948). (Weigel et al., 2008) provided a detailed discussion on the Multi Model Result (MMR), which is applicable to any multi model procedure. He showed that the MMR outperforms the Single Model Result (SMR) if the single model ensemble is too sharp and the ensemble member are scattered, i.e. they have centers of wrong values. In this case the MMR moves the mean of the resulting MMR distribution toward truth and increases the accuracy accordingly. Also, MMR improves the prediction if the SMRs do not completely cover the forecast uncertainty. This is true for any multi model approach. On the other hand the SMRs may depend on the region, season and other factors. Considering that the most variations of the models were seen during the dry season, we took the data for that period and performed the Bayesian Model Averaging in comparison to the single model runs for various scenarios (Table 5-7). It was found that by removing the models which performed not as good as the best model the BMA result outperformed the best model result. This shows that one should apply the BMA approach cautiously considering the fact that models with poor performances may reduce the effectiveness of multi-modeling. This confirms that although BMA result can get close to the best model result, removing the poor models result in even
better performance of the BMA. In addition, BMA provides the uncertainty range in terms of within model and between model variances (Raftery et al., 2005).

The relative changes in streamflow for the future time periods for each hydrologic model and each GCM were obtained. The relative changes were also computed with the Bayesian Model Average for the average of the 8 GCMs. Figure 5-8 shows the BMA results and the Max-Min relative changes amongst the 12 (models) × 8 (GCMs) results. The extremes show the high uncertainties associated with the GCMs and hydrologic models.
Figure 5-8: Relative changes of streamflow for the winter (wet season, upper panel) and the summer (dry season, lower panel) obtained from the BMA and the average GCMs; the bars show the extreme changes observed in the 12 hydrologic model results for all GCMs.

The variation of the total standard deviation (square root of the total variance) in an annual basis for each GCM is shown as the 95% interval in Figure 5-9. This indicates the uncertainties in the total standard deviation of the models due to the GCMs. The bound change is not considerable within the GCMs, however they show discrepancies between the two emission scenarios after the reference period. The two emission scenarios follow the same bound in the reference period.

Table 5-7: Scenario analysis for BMA Nash Sutcliffe performance for summer and all seasons. Note that for each model included, three objective functions were considered to create plausible models.

<table>
<thead>
<tr>
<th>Models included</th>
<th>BMA NSE - summer</th>
<th>BMA NSE - all seasons</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM, HyMOD, SAC-SMA, PRMS</td>
<td>0.734</td>
<td>0.938</td>
</tr>
<tr>
<td>HyMOD, SAC-SMA, PRMS</td>
<td>0.75</td>
<td>0.941</td>
</tr>
<tr>
<td>TM, HyMOD, SAC-SMA</td>
<td>0.779</td>
<td>0.937</td>
</tr>
<tr>
<td>HyMOD, SAC-SMA</td>
<td>0.795</td>
<td>0.939</td>
</tr>
</tbody>
</table>
Figure 5-9: The range of total standard deviation of the 12 hydrologic models for each GCM in the reference and future time periods in A1B (upper panel) and B1 (lower panel) emission scenarios

5.5. Conclusions

The combined uncertainties in projecting runoff using GCM forcings and 4 hydrologic models were assessed. The statistically downscaled climate data from 8 global climate models and 2 emission scenarios drove the hydrologic models where each was calibrated using 3 objective function, in total resulting in 192 model runs. The average changes of temperature, precipitation and streamflow were estimated. Three lumped and one semi-distributed hydrologic models with different
complexities were employed. Each model was calibrated using three objective functions to create plausible models that could target different portions of hydrograph. The uncertainties associated with the climate data, as well as streamflow, were investigated at seasonal time scales for each emission scenario namely A1B and B1. The BMA method was then applied to statistically combine the hydrologic model simulations based on their performances in the observed period.

The results indicate that although temperature will consistently increase in the future periods the variability in both the projected precipitation and runoff changes are considerable. The analysis indicates that difference in projected runoff is directly dependent on the choice of model structure. For example, the differences between models' behaviors in dry season can be explained in terms of sensitivity of model states (i.e., soil moisture) to small changes of precipitation, if any. It is noted that the uncertainties are presented in percentage, therefore, a small difference in streamflow magnitude resulted from different models will show large relative percent difference. Therefore, analysis of the projected streamflow changes shows that uncertainties in the dry season are higher in both GCMs and the hydrologic models than in the wet season.

The TM model shows the highest model uncertainty compared to the other models in the summer. However, the uncertainties associated with the GCMs are more significant than those of the hydrologic models. The hydrologic model
uncertainties become important when analyzing the dry season. The hydrologic models perform differently according to each distinct objective function. Therefore, each hydrologic model calibrated by different objective functions is treated as a separate hydrologic model.

There are varieties of hydrologic models with different levels of complexities, in order to study the hydrologic climate change impacts but no model is perfect in characterizing the real physical interactions. The models are supposed to be chosen according to the study region and basin characteristics, available data and study purposes but many times the model selection is subject to the taste of the modeler, and rarely an objective model selection is conducted. It is shown that except for the summer, the hydrologic models perform closely in projected runoff change prediction. However there are still uncertainties associated with the choice of the model. Considering the broad range of models and scenarios, the climate change impact assessment is a matter of probabilistic approach, which necessitates the provision of several scenarios, models and quantification of the uncertainties based on the defined criteria. BMA allows for quantifying the model structural uncertainties and performing ensemble estimation. This method is successfully applied in this study for the hydrologic models variance estimation and weight assignment. Quantifying the uncertainties related to different steps of the climate change impact assessment and analyzing the probability of the outcomes seem very
appealing to water managers and policy makers to make suitable strategies in adapting to changing climate and this study demonstrated a step forward toward this goal.
6. Chapter 6. Conclusions and Future Research:

For the study of climate change impact on hydrologic extremes (such as floods) we propose a spatial hierarchical Bayesian method which combines extreme data from different locations. In the first study we use data recorded at gage sites (point-referenced data) in two regions in the Columbia River basin. Extreme events occurred during the months of December through March are selected for this analysis. Generalized Pareto distribution is the basis of the model with its scale parameter being spatially characterized in a hierarchical Bayesian approach. Declustering process provides temporally independent data for the hierarchical model. The scale parameter of the extreme distribution is spatially modeled through generalized linear relationships including covariates of latitude, longitude, drainage area and elevation. The parameters of the spatial hierarchical model are estimated through Metropolis Hastings within Gibbs sampler. The model is capable of predicting the 100-year return level floods for un-gaged basins. Results show significant increase in the precision of the model compared with a simple maximum likelihood estimator regarding the flood return levels, since information content in the data from different locations are combined using the spatial hierarchical model. Historical changes in extreme events are performed by considering time periods of 15 years. The results show cyclic variations in the spatial average of the 100-year
return level floods. However for some areas consistent increasing trends are distinguishable.

The second study introduces the application of the Bayesian hierarchical model in the analysis of climate change impact on runoff extremes. The study is performed over the whole Pacific Northwest using 1/8° grid based data. Climate change data is provided by NARCCAP through several regional climate models nested within four GCMs. Hydrologic modeling is performed after downscaling of the precipitation, maximum and minimum temperature and wind speed using the quantile-mapping approach. Variable Infiltration Capacity (VIC) which is a distributed hydrologic model is used to provide daily runoff estimates (mm) for each cell. Spatial hierarchical Bayesian model is then applied on the cell-wise extreme runoff (mm) for both time periods and for all seasons. Generalized extreme value distribution comprises the likelihood stage of the model. The results indicate that different RCMs have varying behaviors in projecting the extreme runoff. This highlights the importance of multi-modeling by considering an ensemble of climate model simulations. The hierarchical Bayesian model identifies the spatial variations in the marginal distributions of the GEV parameters and the corresponding 100-yr return level runoffs. The posterior distributions of the latent variables provide information about the significance of each covariate on the extreme analysis in each season. Overall outcomes show increases in the estimated 100-yr return level runoffs for
most seasons particularly over the high elevation zones during winter. The Canadian portions of the study region reflect higher increases during spring. Summer indicates reduction of extreme events in most areas.

The third study involves multi-modeling analysis of the runoff extremes for 6392 grid cells in the Pacific Northwest region. Three scenarios are considered for multi-modeling of 100-yr return level runoffs. In the selected scenario Bayesian model averaging is applied on the daily runoff results obtained from the VIC simulations using different RCMs. The estimated weights for each cell are then considered on the 100-yr return level runoffs for the corresponding RCMs in order to obtain the multi-model average results for both the historical and future time periods. BMA performed satisfactorily while it overtook the best model during spring and summer seasons. The multi-modeling outcomes show increases in the estimated 100-yr return level runoffs for most seasons particularly over the high elevation areas during winter. The Canadian portions of the study region reflected higher increases during spring. Summer indicated reduction of extreme events in most areas.

In the fourth study the combined uncertainties in projecting runoff using GCM forcings and 4 hydrologic models are assessed. The statistically downscaled climate data from 8 global climate models and 2 emission scenarios drove the hydrologic models where each was calibrated using 3 objective function, in total resulting in 192 model runs. The average changes of temperature, precipitation and streamflow
are estimated. The uncertainties associated with the climate data, as well as streamflow, are investigated at seasonal time scales for each emission scenario namely A1B and B1. The BMA method is then applied to statistically combine the hydrologic model simulations based on their performances in the observed period. The results indicate that although temperature will consistently increase in the future periods the variability in both the projected precipitation and runoff changes are considerable. The analysis indicates that difference in projected runoff is directly dependent on the choice of model structure. Therefore, analysis of the projected streamflow changes shows that uncertainties in the dry season are higher in both GCMs and the hydrologic models than in the wet season. The uncertainties associated with the GCMs are more significant than those of the hydrologic models. The hydrologic model uncertainties become important when analyzing the dry season. The hydrologic models perform differently according to each distinct objective function.

6.1. Future Research:

Similar analyses proposed in this thesis can be conducted using newly available datasets. A plethora of climate modeling is currently performed in order to provide sufficient datasets for climate change impact studies. These models are improved in spatiotemporal resolutions, as well as the dynamical formulations and physical parameterizations. CMIP5 experiments, incorporating about 20 modeling groups
located around the world, is an example of a continued effort for multi-model intercomparision and analysis (Taylor et al., 2012). Related to this, with the growing computational facilities it is now possible to perform higher resolution regional climate modeling which has several advantages over statistical downscaling approaches (such as the one presented in Appendix A). These are valuable data sources which can be used for hydrologic climate change impact studies. Additionally comparative studies using different datasets can be similarly conducted.

While conditional independence assumption is widely considered in the spatial hierarchical modeling (Banerjee et al., 2004), copula density function can be incorporated into the likelihood function to explicitly consider the small scale dependencies of the extremes. Also developing spatio-temporal hierarchical modeling of hydrologic extreme is a substitute for time window analysis performed in the first study and is recommended for future analyses. Other suggestions for the spatial hierarchical Bayesian modeling are: to spatially model the shape parameter of the GPD distribution (which was assumed homogenous in the first study); incorporate additional covariates such as slope, soil moisture, aspect, elevation, etc.; compare different spatial models within the hierarchy using various covariograms as well as autoregressive models, further study on adding more informative prior distributions. Furthermore a comprehensive study on the multi-modeling
techniques would be valuable by comparing the precisions and accuracies between these models and the individual simulations. In addition studies are needed to speculate on using sequential data assimilation in connection with spatial hierarchical Bayesian modeling in seeing the time evolution of parameters, in light of the recent publications (Leisenring and Moradkhani, 2012; Liu et al., 2012; Moradkhani et al., 2012).

Interdisciplinary studies on the climate change impact are further needed to conduct better risk analyses and take reliable adaptation strategies. These studies will incorporate hydrologic as well as environmental, agricultural, economic, social and ecological impact analyses.
References


Caires, S., Sterl, A., 2010. 100-year return value estimates for ocean wind speed and significant wave height from the ERA-40 data.


208


Kundzewicz, Z.W. et al., 2008. The implications of projected climate change for freshwater resources and their management.


213


Van der Linden, P., Mitchell, J., 2009. ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project. Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK, 160.


Appendices

Appendix A: Statistical Downscaling of Precipitation using Machine Learning with Optimal Predictor Selection

Background

Statistical downscaling methods are dependent on the choice of predictors that would provide accurate information of the regional precipitation. (Huth, 2004) showed that the downscaled temperature is highly sensitive to the predictor combinations. Although there are several studies on various downscaling procedures (Fowler et al., 2007c), the question of determining which GCM predictors can best explain the regional variations of precipitation still remains.

Several studies on precipitation downscaling suggest the application of different methods rather than a single one in order to generate more robust results (Hashmi et al., 2009; Wetterhall et al., 2007). (Salathé Jr, 2003) used three simple methods of local scaling, dynamical scaling and analog method, to downscale the NCEP/NCAR reanalysis data. Other studies include the comparison of statistical versus dynamical downscaling (Haylock et al., 2006; Schmidli et al., 2007; Spak et al., 2007), Support Vector Machine versus Relevance Vector Machine (RVM) (Ghosh and Mujumdar, 2008), and multiple regression analysis versus Canonical Correlation Analysis (CCA) (Hertig and Jacobeit, 2008; Huth, 2004).
Investigating different downscaling procedures allows for characterization of the underlying uncertainties that influence climate change impact and adaptation strategies. (Khan et al., 2006) compared SDSM, LARS-WG and ANN in downscaling NCEP data and showed the associated uncertainties. (Raje and Mujumdar, 2009) proposed the conditional random field, as a probabilistic approach in downscaling GCMs. They utilized PCA to summarize the spatial information from 25 (and 30) NCEP grid points.

In this study an objective procedure for selecting GCM predictors is first presented. We then compare downscaling methods with different structures and complexities and investigate the influence of predictor selection on the performance of each method. The procedure for GCM predictor selection in the ‘linear’ downscaling process is based on the method proposed by (Moradkhani and Meier, 2010). In their study the Independent Components (IC) of the climate signals (ENSO, PDO, etc.) were determined to perform the predictor selection for long-lead statistical water supply forecast. Similarly in the current study the independent components of each GCM predictor were obtained, followed by selection of ICs that were most correlated with regional precipitation. The selected ICs were then analyzed by a linear cross validation procedure to determine which predictor combinations best described the monthly and seasonal precipitation. We found that each model performed best based on specific predictor combination which could be
different from the predictors selected for the other models. The method is described in detail in section 0. Once the predictor selection is completed, three statistical downscaling methods are implemented: linear regression of the independent components of GCM data, Support Vector Machine (SVM) learning model and Adaptive-Network-Based Fuzzy Inference System (ANFIS). The linear regression model finds the linear factor attributed to each predictor to best estimate the precipitation. ANFIS is a nonlinear regression model which is a sugeno type fuzzy system with its parameters tuned based on a hybrid training algorithm. This model has been widely used in flood forecasting and rainfall-runoff modeling (Chen et al., 2006; Firat and Gungor, 2008; Nayak et al., 2007; Vernieuwe et al., 2006). In this study the initial fuzzy system was constructed using three algorithms, namely grid partitioning, subtractive clustering and fuzzy c-means clustering. The three different model structures were then calibrated using ANFIS training process. The third method, (SVM) learning, is widely used for finding non-linear relationships between predictors and predictands. SVM achieves nonlinear regression by first mapping the input vector into a higher-dimensional feature space using a nonlinear transformation function ($\Phi$), followed by determination of the optimal hyperplane that separates the features of the input between “positive” and “negative” examples (Haykin, 1999; Tripathi et al., 2006). For the purpose of precipitation downscaling, the hyperplane was constructed to separate predictor input features and observed
precipitation values. In addition to their abilities to determine non-linear relationships, ANFIS and SVM models, unlike the artificial neural networks (ANN), do not get trapped in local minima during the model parameter optimization process (Ghosh and Mujumdar, 2008; Jang et al., 1997).

**Study area and Data**

The study area chosen for analysis is the Upper Willamette sub-basin in Central Coastal Oregon located within 43.91°N to 45.15°N latitude and 123.75°W to 122.25°W longitude (Figure 01). The sub-basin is 4,848 km² in area and contains the urban cities of Corvallis, Eugene and Albany. The Willamette River, one of Oregon's primary tributaries to the Columbia River, flows for 61.5 miles through this sub-basin converging with Long Tom River, Mary's River and the Calapooia River along the way. The precipitation gage in the basin has a long complete record from 1950-2008. Additionally, since the sub-basin includes Corvallis, Eugene and Albany, understanding climate change impacts on water resources and storm event frequency is of the utmost importance.
The GCM potential predictor data was obtained from the IPCC data distribution centre’s 4th assessment report. The Canadian Center for Climate Modeling and Analysis (CGCM3) model with T47 resolution (~3.71° lat × 3.75° long) was selected as the GCM model. A2 emission scenario which represents the highest greenhouse gas emission, due to the technological advances and population growth, was selected in order to determine the regional precipitation changes resulting from extreme climatological conditions. The CGCM3 model variables considered are shown in Table 1. Some of the variables (wind speed, temperature, geopotential height) were considered for different pressure elevations of 200, 500 and 850 hPa.
Table 1: Potential predictors with their respective elevations. Elevations listed as pressure correspond to the height at which the air pressure occurs.

<table>
<thead>
<tr>
<th>Potential Predictor</th>
<th>Acronym</th>
<th>Elevation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific Humidity</td>
<td>SH</td>
<td>Surface</td>
</tr>
<tr>
<td>Total Soil Moisture</td>
<td>TSM</td>
<td>Surface</td>
</tr>
<tr>
<td>Total Precipitation</td>
<td>TP</td>
<td>Surface</td>
</tr>
<tr>
<td>Convective Precipitation</td>
<td>CP</td>
<td>Surface</td>
</tr>
<tr>
<td>Sea Level Pressure</td>
<td>SLP</td>
<td>Surface</td>
</tr>
<tr>
<td>Downwelling Shortwave Flux</td>
<td>RSDS</td>
<td>Surface</td>
</tr>
<tr>
<td>Snow Area Fraction</td>
<td>SAF</td>
<td>Surface</td>
</tr>
<tr>
<td>Snow Depth</td>
<td>SD</td>
<td>Surface</td>
</tr>
<tr>
<td>Snow Melt</td>
<td>SM</td>
<td>Surface</td>
</tr>
<tr>
<td>Air Temperature</td>
<td>T</td>
<td>Surface, Near Surface, 200 hPa, 500 hPa, 850 hPa</td>
</tr>
<tr>
<td>Zonal Wind</td>
<td>ZW</td>
<td>Surface, 200 hPa, 500 hPa, 850 hPa</td>
</tr>
<tr>
<td>Meridional Wind</td>
<td>MW</td>
<td>Surface, 200 hPa, 500 hPa, 850 hPa</td>
</tr>
<tr>
<td>Geopotential Height</td>
<td>GH</td>
<td>200 hPa, 500 hPa, 850 hPa</td>
</tr>
</tbody>
</table>

The strength of the GCM predictors is time and space dependent. (Wilby and Wigley, 2000) found that the predictor domain is a crucial factor that determines how reliable and robust the downscaled precipitation would be. The application of ICA allows for consideration of a wide range of grid cells over the region. Therefore twenty-eight grid points of GCM data over the region of 38.97°N to 50.1°N latitude and 135°W to 112.5°W longitude were used as potential predictors, for the period of 1950-2100 (Figure 01). The regional observed precipitation data was acquired from the Western Regional Climate Center (WRCC) at Corvallis gage station. Total
monthly precipitation over the period of 1950-2008 from this gage was used for model construction and calibration.

In the following sections the nonlinear models are first described and the predictor selection procedure is illustrated. The results of the models are then followed by the discussion on the proposed methodology.

**Nonlinear Models**

**Adaptive Network Based Fuzzy Inference System (ANFIS)**

In the fuzzy logic concept, unlike the classical algebra, an object is assigned a degree of membership to one or more fuzzy sets or classes. This concept has evolved from its first introduction by (Zadeh, 1965) enhancing the data learning capabilities along with other mathematical integrations (Jang et al., 1997; Kosko and Burgess, 1998; Sugeno and Kang, 1988; Zadeh, 1988; Zadeh, 1994; Zadeh, 1999). The capability of the fuzzy logic to account for the underlying uncertainties in many real world problems has brought it to the context of various hydrologic applications (Alvisi et al., 2006; Bárdossy, 1996; Liong et al., 2000; Samanta and Mackay, 2003; Tayfur and Singh, 2006). Adaptive Network Based Fuzzy Inference System (ANFIS) proposed by (Jang, 1993) is a Sugeno type fuzzy inference system (Sugeno and Kang, 1988) put into the framework of adaptive networks that can compute gradient vectors systematically (Jang, 1996). This system needs to be calibrated
using a training data set as seen in various hydrologic applications (Chang and Chang, 2006; Keskin et al., 2006; Mousavi et al., 2007; Terzi et al., 2006).

Figure 2 (a) shows a sugeno type fuzzy inference system with two fuzzy If-Then rules:

Rule 1: If \text{huss} is High and \text{psl} is High Then \text{prec}_1 = p_1 huss + q_1 psl + r_1

Rule 2: If \text{huss} is Low and \text{psl} is Low Then \text{prec}_2 = p_2 huss + q_2 psl + r_2

where \text{huss} represents the surface specific humidity and \text{psl} represents the pressure at sea level. The adaptive network for this system is shown in Figure 2(b). The layer functions are as follow: **Layer 1**: GCM predictors are the inputs to the nodes in this layer. Each node represents one fuzzy set of one predictor which contains a linguistic label (e.g. Low). The output is the degree to which the given GCM predictor belongs to the fuzzy set. The maximum and minimum membership degrees (e.g. \mu_{Low}(huss)) are 1 and 0 respectively:

\[ O_i^1 = \mu_{Low}(huss) \]  

(1)

where \( O_i^1 \) is the membership function (MF) (\( i \) is the number of the node). Any continuous and piecewise differentiable functions are qualified for node MF. We used the Gaussian and bell-shaped MFs for our study. The generalized bell function is defined as:

\[ O_i^2 = \mu_{Low}(huss) \]  

(2)
\[ \mu_{\text{Low}}(\text{huss}) = \frac{1}{1 + \left( \frac{\text{huss} - c_i}{a_i} \right)^{2b_i}} \]  

(2)

where \{a_i, b_i, c_i\} are the premise parameters.

Layer 2: The node function in this layer is a T-norm fuzzy operator such as a minimum or a multiplication which performs generalized AND in each rule e.g. \( W_1 = \mu_{\text{Low}}(\text{huss}) \times \mu_{\text{Low}}(\text{psl}) \). Each node’s output represents the firing strength of a rule.

Layer 3: The normalized firing strength of each rule is calculated in this layer:

\[ \bar{W}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \]  

(3)

Layer 4: The normalized firing strength from layer 3 is multiplied by the linear functions of the input GCM predictors:

\[ O_i^4 = \bar{W}_i \cdot \text{prec}_i = \bar{W}_i (p_i \cdot \text{huss} + q_i \cdot \text{psl} + r_i) \]  

(4)

The parameters \( p_i, q_i \) and \( r_i \) are the consequent parameters.

Layer 5: All the results from layer 4 are summed up in the single node in this layer and the downscaled precipitation is determined:

\[ O_i^5 = \text{prec} = \sum_i \bar{W}_i \cdot \text{prec}_i \]  

(5)
Figure 2: a) "Sugeno" fuzzy model with two inputs (surface_specific_humidity(huss) and pressure at sea level (psl)): b) The ANFIS model architecture.
Before training the fuzzy system and optimizing the parameters, the initial Sugeno fuzzy model structure was constructed; that is to find an appropriate number of rules and a proper partitioning of the feature space.

**ANFIS Initialization**

Three approaches were adopted to initialize the ANFIS structure called grid partitioning, subtractive clustering and fuzzy c-means clustering. Grid partitioning is the simplest input partition style. In this method the number and types of the membership functions are chosen manually. The feature space is divided into equal spaces, according to the defined range and the number of partitions; the number of rules is the number of membership functions per input to the power of the total number of inputs of the system. The type and the number of membership functions for each model are determined by trial and error. In order to find the optimized ANFIS structure we used the technique presented by (Jang, 1996) in which the model was trained for just one epoch utilizing the least square method. This was based on the assumption that the ANFIS model with the lowest RMSE in the first training epoch would have the lowest RMSE after the converged training epochs. Cross validation was also implemented where ANFIS was trained 20 times, each time with a different train and check dataset. The test dataset remained unchanged. The Root Mean Square Error (RMSE) and Nash Sutcliffe Efficiency (NSE) criterion (Nash and Sutcliffe, 1970) were taken as the performance measure of the model.
structure. The two bell-shape MFs per input results produced lower testing error compared to the others implying that the model performance does not necessarily improve by increasing the number of membership functions and rules.

Although the grid partitioning method worked well for monthly downscaling with at most four predictors, for the seasonal downscaling there was the potential of curse of dimensionality using this method; that is the number of fuzzy rules increases exponentially with the number of inputs of the fuzzy system. In this case the large number of fuzzy rules not only increases the computational complexity of the fuzzy system, but also lessens the interpretability (Jin, 2003). Thus two clustering methods were used to effectively partition the seasonal model input space and decrease the number of rules.

One of the methods to provide some dimension reduction is Subtractive Clustering, which is the modified form of the Mountain Method for cluster estimation proposed by (Yager and Filev, 1994). In the subtractive algorithm each data point, rather than grid points used in the mountain method, is considered as a potential cluster center making the computation dependent on the number of data points and independent of the dimension of the problem. In this study subtractive clustering was examined with different radii, the values between 0 and 1 for each data dimension. The variable radius close to 1 provides fewer rules and consequently fewer parameters. In order to optimize the radius values for each
model, the same procedure performed in grid partitioning was implemented. ANFIS was trained 20 times, each time with one epoch and a different train and check dataset. The results for the winter analysis are shown in Table 2. The Gaussian membership function, with parameters $c$ and $\sigma$ (the center and the width of the membership function), was utilized.

*Table 2: Optimization process of the cluster radius in ANFIS initialized by subtractive clustering (e.g. winter predictors).*

<table>
<thead>
<tr>
<th>Cluster radius</th>
<th>RMSE</th>
<th>NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>61.1904</td>
<td>0.523</td>
</tr>
<tr>
<td>0.8</td>
<td>54.237</td>
<td>0.6278</td>
</tr>
<tr>
<td>0.7</td>
<td>50.4135</td>
<td>0.6788</td>
</tr>
<tr>
<td>0.6</td>
<td>50.2799</td>
<td>0.6806</td>
</tr>
<tr>
<td>0.5</td>
<td>53.4194</td>
<td>0.6354</td>
</tr>
</tbody>
</table>

Fuzzy c-means clustering was the other method used for partitioning the input space. This clustering algorithm (Bezdek, 1981) is a multivariate data analysis technique that partitions a dataset into overlapping clusters. Each cluster is identified by its cluster center. The number of clusters is specified manually and the clustering of a dataset is performed by minimizing an objective function. In order to optimize the number of clusters for each model the same procedure performed in grid partitioning and subtractive clustering algorithm was implemented. The results for the seasonal modeling of the final predictor combination are shown in Table 3. The Gaussian membership function was also used in this algorithm.
Table 3: Number of clusters used for the final predictor combination (ANFIS initialized by fuzzy c-means clustering)

<table>
<thead>
<tr>
<th>Season</th>
<th>No. of Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring</td>
<td>2</td>
</tr>
<tr>
<td>Summer</td>
<td>2</td>
</tr>
<tr>
<td>Fall</td>
<td>3</td>
</tr>
<tr>
<td>Winter</td>
<td>4</td>
</tr>
</tbody>
</table>

Support Vector Machine

Support Vector Machine (SVM) which was developed by (Vapnik and Chervonenkis, 1971) shares the artificial neural networks ability to determine nonlinear relationships. SVM is a universal feed-forward network which is used for pattern classification and nonlinear regression (Haykin, 1999). Nonlinear regression is achieved by the model's ability to find linear classification planes in high-dimensional feature space. Additionally, SVM has the advantage of not being trapped in local minima during calibration.

Independent Component Analysis for Predictor Selection

Since there are twenty-five variables in the CGCM3 A2 dataset, it is necessary to choose the predictors from the dataset that best represent the regional precipitation. However, using all of the potential predictors results in redundancy in the input data for the models. GCM predictors are commonly selected based on their physical relevance, and the measure of their strengths in some form of correlation with the predictand. (Wilby and Wigley, 2000) analyzed the correlation between the predictors and the predictand and selected the potential GCM
predictors that explained the highest variance in precipitation occurrence and amount. The predictors were then used to downscale the HadCM2 GCM variables to daily precipitation. (Anandhi et al., 2009) utilized scatter plots and cross-correlation to evaluate the performance of the GCM predictors in downscaling CGCM3 to river scale surface temperature. Other similar approaches can be found in (Crawford et al., 2007; Haylock et al., 2006; Hessami et al., 2008; Huth, 2004; Schmidli et al., 2007; Spak et al., 2007). Although in this method the predictors with the highest correlation with the predictand are selected, the correlation between the predictors is not accounted for which increases the risk of data redundancy. In a few studies, other approaches have been analyzed, such as utilizing F-test to examine the level of statistical significance of the predictors (Hessami et al., 2008), and accuracy assessment of the downscaling method in the verification period utilizing different predictor combinations (Hertig and Jacob, 2008). However, an objective procedure for selecting GCM predictors for any downscaling method is still lacking. In this study an algorithm for selecting predictors which most suitably downscale to gauge station precipitation is discussed. The application of Independent Component Analysis (ICA) along with the cross-validation based on prediction residual sum of squares (PRESS) is presented to spatially summarize the GCM variables, and choose the optimum predictor combination.
In our method each potential GCM predictor was considered independently for each month and for all 28 grid cells over the entire period (1950-2100). The independent component (IC) GCM variables that could best describe the potential predictors were then found by the process presented in Figure 3.

Figure 3: The procedure for predictor selection.
The first step of this process was to determine the principal components of the dataset that describe the amount of variance over the spatial and temporal scale. Principal component (PC) analysis determines the weight matrix, \( w \), that when applied to the dataset, \( x \), produces a new matrix, \( y \), of uncorrelated variable vectors:

\[
y = wx
\]  

(6)

The weight matrix is determined by finding the variance-covariance matrix of the dataset which identifies how much the input vectors (GCM variable grid cells) vary with respect to each other. The unique eigenvectors and eigenvalues of the variance-covariance matrix are then determined. Sorting the eigenvalues from largest to smallest and also their corresponding eigenvectors determines which components describe the most to the least variance. One can multiply the original dataset by loading eigenvectors matrix to determine the PCs in an order that goes from the PC describing the most variance to the PC describing the least variance. The PCs determined are uncorrelated and redundancy has been removed. Then, taking the first PC that describes the greatest variance of the dataset, independent component analysis is performed to determine the independent GCM variables present. Using the fastICA algorithm developed by (Hyvärinen and Oja, 2000), the independent GCM variables, \( s \), can be found once the unknown mixing matrix, \( A \), is found:

\[
y = As
\]  

(7)
The mixing matrix is found using an iterative procedure that maximizes the non-gaussianity of the estimated ICs. Following the predictor selection approach developed by (Moradkhani and Meier, 2010), the Pearson’s correlation coefficient is calculated between each IC (over the period 1950-2008) and monthly observed gauged precipitation data. The IC that has the highest correlation coefficient with the observed precipitation is then retained.

Once the ICs for each of the twenty-five GCM variables and for each month are found over the period of 1950-2100, it is necessary to determine which combinations of ICs would fully describe the regional precipitation. We performed this procedure by considering all combinations of one, two, three and four IC predictors to reduce the computational burden of the algorithm. Also the inclusion of more than four predictor variables showed marginal or no improvement in the accuracy of the procedures. For each combination we performed the linear regression described in equation below. The column vectors represent the predictor ICs; where $\alpha$ represents the appropriate weight, $x$ represents the predictor value at time interval $i$ (out of $m$ total time intervals) and GCM variable $j$ (out of $n$ variables, ranging from 1 to 4 for our purposes) and $y$ represents observed precipitation at time interval $i$:
Once the $\alpha$ values were determined for a particular combination, we performed a leave-one-out (LOO) cross-validation by leaving one year of predictor IC and observed precipitation out and using the model to simulate a precipitation vector based on the remaining years of ICs and observed data. We performed the LOO over the entire period of the particular combination so that every year of IC and observed data would be left out just once. The prediction residual sum of squares (PRESS) statistic is used to assess each combination by comparing the entire set of observed and simulated values generated through the LOO method:

$$PRESS = \sum (p_{obs} - p_{sim})^2$$  \hspace{1cm} (9)

For each month, the combination of predictor ICs with the smallest PRESS statistic value was selected as the best combination for describing the observed precipitation. The determined predictor ICs for the corresponding months are presented in Table 4. The seasonal combinations of predictor ICs are obtained by lumping together the monthly predictors pertaining to a particular season. These results are also presented in Table 4.
Predictors in precipitation downscaling should encompass the variables that affect the precipitation. This includes the circulation indices like sea level pressure, wind, geopotential height; thermodynamics; and water vapor content of the atmosphere (Cavazos and Hewitson, 2005). In this study, the wind component is one of the main predictors in almost all months and seasons. Besides the temperature and geopotential height for various pressure levels are found to be
effective in the precipitation downscaling process. The other selected component of the GCM variables is the total precipitation. Specific humidity has been chosen as the spring and summer predictor, while components of snow (e.g. snow depth and snow area fraction) are selected as fall and winter predictors, which could implicitly represent the radiation and temperature effects in these seasons.

The IC predictor combinations presented in Table 4 were then used to train and test the SVM and ANFIS models for downscaling monthly and seasonal IC predictors to precipitation. This method allows for considering several GCM grids and several predictor combinations, so that the best combination is selected for each month.

In this method, the predictors selected for each month and season were based on a multi-linear cross validation (MLR-CV) technique, which may not produce satisfactory results when used in a nonlinear regression model such as ANFIS or SVM. Therefore the nonlinear models were trained with each predictor set (Table 4) and their performances were evaluated on the test dataset for each month and season. For each of the selected predictors a unique monthly and seasonal model was constructed and evaluated. The models’ performances were compared to the test data set and the resulting error (RMSE) was calculated (Figure 4). Based on this technique the predictor combinations are first determined in a linear regression model for each month and season. The predictor combinations are then replaced for different time periods according to their performances in the nonlinear models.
Figure 4: Root Mean Square Error (RMSE) of monthly and seasonal precipitation with respect to their corresponding MLR-IC predictors: a) ANFIS (grid partitioning) simulation; b) ANFIS (subtractive clustering) simulation; c) ANFIS (fuzzy c-means clustering) simulation; d) SVM simulation.

Results and Discussion

In order to calibrate the ANFIS, data was divided into train (50%), check (30%) and test datasets (20%). Each model was trained based on the train dataset for
which the error tends to decrease by increasing the number of epochs. This increases the risk of overtraining, meaning that the model is trained to fit to the noise in the train dataset. Therefore, the check dataset was also evaluated during the training process. Consequently the check data error decreases in the initial epochs, however after several iterations it starts to increase. The increase in the error after the minimum error point in the check dataset is due to the model overfitting. For this reason the optimized model parameters were chosen based on the minimum check data error.

Having found the predictors for the two monthly and seasonal time periods and after the models were trained based on the independent components for the observed period (1950-2008) the performance of each model was compared to the observed data for the test dataset. Each model (linear and nonlinear) was then used to downscale the future GCM predictors, providing various scenarios.

Table 5: Comparison of NSE values for training, checking and testing datasets for sigma=4 and C=50, 100, 1000.

<table>
<thead>
<tr>
<th>Season</th>
<th>C Value</th>
<th>NSE_training</th>
<th>NSE_checking</th>
<th>NSE_testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.942</td>
<td>0.767</td>
<td>0.384</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.918</td>
<td>0.732</td>
<td>0.514</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>0.970</td>
<td>0.643</td>
<td>0.382</td>
<td></td>
</tr>
<tr>
<td>Spring</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.650</td>
<td>0.763</td>
<td>0.497</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.873</td>
<td>0.830</td>
<td>0.589</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>0.956</td>
<td>0.778</td>
<td>0.531</td>
<td></td>
</tr>
<tr>
<td>Summer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.848</td>
<td>0.640</td>
<td>0.386</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.894</td>
<td>0.785</td>
<td>0.270</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>0.937</td>
<td>0.551</td>
<td>-0.584</td>
<td></td>
</tr>
<tr>
<td>Fall</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.799</td>
<td>0.632</td>
<td>0.717</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.848</td>
<td>0.640</td>
<td>0.806</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>1.00</td>
<td>0.436</td>
<td>0.583</td>
<td></td>
</tr>
</tbody>
</table>
The determined monthly predictors were used to train the nonlinear models by cross validation. Each model (based on Jan-Dec predictors) was trained 50 times, with different train and check datasets; however, the test dataset remained unchanged. Therefore 50 separate SVM and fuzzy models (grid partitioning, subtractive clustering and fuzzy c-means) with the same structure but different parameters were constructed. The average of all the results over the 50 runs was then calculated in order to compare to the observed precipitation.

The models incorporated in the monthly downscaling include ANFIS initialized by grid partitioning and multi-linear regression model. Comparison was made based on the monthly mean observed and simulated precipitation in the test dataset (Figure 5). The results show the good performance of the two models.

![Figure 5: Downscaled monthly precipitation using ANFIS (grid partitioned) and MLR-IC for the test data set.](image)

243
Figure 6 compares the three nonlinear models with the observed data for the seasonal downscaling. The models include SVM, ANFIS based on subtractive clustering and ANFIS based on fuzzy c-means. The figure indicates that the performances of the models in simulating the observed precipitation data were acceptable.

The performances of all the models were also compared for each season by means of correlation coefficient, RMSE and BIAS (Table 6). The results indicate that the linear model although simple in structure, could reasonably perform the downscaling and was the best overall model for spring months. Amongst the nonlinear models the results of SVM were best for winter and fall season while ANFIS based on subtractive clustering was best for summer season. The results also
indicate that the structure of the fuzzy model is an important factor in model performance, even with the same training process.

Table 6: Models’ performances for each season

<table>
<thead>
<tr>
<th>Model</th>
<th>Winter</th>
<th>Correlation</th>
<th>RMSE</th>
<th>BIAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seasonal SVM</td>
<td></td>
<td>0.78</td>
<td>64.2</td>
<td>26.9</td>
</tr>
<tr>
<td>Seasonal ANFIS (S.C)</td>
<td></td>
<td>0.77</td>
<td>60.25</td>
<td>27.26</td>
</tr>
<tr>
<td>Monthly (MLR)</td>
<td></td>
<td>0.76</td>
<td>62.95</td>
<td>27.62</td>
</tr>
<tr>
<td>Monthly ANFIS (G.P)</td>
<td></td>
<td>0.66</td>
<td>71.57</td>
<td>31.32</td>
</tr>
<tr>
<td>Seasonal ANFIS (FCM)</td>
<td></td>
<td>0.65</td>
<td>76.26</td>
<td>34.51</td>
</tr>
<tr>
<td><strong>Spring</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly (MLR)</td>
<td></td>
<td>0.73</td>
<td>32.41</td>
<td>28.76</td>
</tr>
<tr>
<td>Seasonal SVM</td>
<td></td>
<td>0.72</td>
<td>33.86</td>
<td>28.14</td>
</tr>
<tr>
<td>Seasonal ANFIS (S.C)</td>
<td></td>
<td>0.61</td>
<td>39.33</td>
<td>36.72</td>
</tr>
<tr>
<td>Seasonal ANFIS (FCM)</td>
<td></td>
<td>0.7</td>
<td>12.4</td>
<td>52.24</td>
</tr>
<tr>
<td>Monthly (MLR)</td>
<td></td>
<td>0.61</td>
<td>14</td>
<td>56.4</td>
</tr>
<tr>
<td>Seasonal ANFIS (S.C)</td>
<td></td>
<td>0.61</td>
<td>17.3</td>
<td>73</td>
</tr>
<tr>
<td>Seasonal ANFIS (FCM)</td>
<td></td>
<td>0.59</td>
<td>19.6</td>
<td>74.71</td>
</tr>
<tr>
<td><strong>Summer</strong></td>
<td></td>
<td>0.35</td>
<td>35.5</td>
<td>66.92</td>
</tr>
<tr>
<td>Seasonal SVM</td>
<td></td>
<td>0.61</td>
<td>33.1</td>
<td>27.2</td>
</tr>
<tr>
<td>Seasonal ANFIS (S.C)</td>
<td></td>
<td>0.86</td>
<td>36.65</td>
<td>31.44</td>
</tr>
<tr>
<td>Monthly (MLR)</td>
<td></td>
<td>0.87</td>
<td>36.92</td>
<td>30</td>
</tr>
<tr>
<td>Monthly ANFIS (G.P)</td>
<td></td>
<td>0.85</td>
<td>39.34</td>
<td>29.21</td>
</tr>
<tr>
<td>Seasonal ANFIS (FCM)</td>
<td></td>
<td>0.83</td>
<td>43.13</td>
<td>33.68</td>
</tr>
<tr>
<td><strong>Fall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seasonal SVM</td>
<td></td>
<td>0.9</td>
<td>33.1</td>
<td>27.2</td>
</tr>
<tr>
<td>Seasonal ANFIS (S.C)</td>
<td></td>
<td>0.86</td>
<td>36.65</td>
<td>31.44</td>
</tr>
<tr>
<td>Monthly (MLR)</td>
<td></td>
<td>0.87</td>
<td>36.92</td>
<td>30</td>
</tr>
<tr>
<td>Monthly ANFIS (G.P)</td>
<td></td>
<td>0.85</td>
<td>39.34</td>
<td>29.21</td>
</tr>
<tr>
<td>Seasonal ANFIS (FCM)</td>
<td></td>
<td>0.83</td>
<td>43.13</td>
<td>33.68</td>
</tr>
</tbody>
</table>

The calibrated models were then utilized to downscale all the GCM data from 1950-2100 using the independent components of the GCM predictors of 28 grids. Employing the selected predictors along with the model structure and associated calibrated parameters, the future coarse scale GCM precipitation data was downscaled to the gauge station monthly and seasonal data. The monthly mean
precipitation over the periods of 2010-2029, 2030-2049, 2050-2069 and 2070-2089 are shown as box plots in Figure 7 based on all the monthly and seasonal models. The length of the box plots show the uncertainties related to different downscaling techniques.

Figure 7: Uncertainty range obtained from the ensemble of downscaled data using various techniques at the monthly scale compared with mean observed value shown by straight solid line.
Similarly Figure 8 displays the ensemble seasonal simulations for the future time periods based on all the downscaling techniques.

Figure 8: Uncertainty range in downscaled precipitation data at seasonal time scales compared with the mean observed shown by straight solid line.
The straight line shows the observed mean value for the period of 1989-2008. The analysis of the standard deviations of the model results as a measure of spread show that the highest monthly uncertainty occurs in January for all periods while the lowest is in September. In the seasonal time scale the winter period analysis has the highest uncertainty while spring, fall and summer analyses have lower uncertainties respectively as shown in the figures. This implies that during precipitation downscaling in wet season, models tend to diverge and increase the uncertainty. Generally the spread of the downscaling results is not significantly different at monthly and seasonal time scales. Also no specific trend in the uncertainty change is seen for future time periods. Therefore, application of several methods for downscaling precipitation seems necessary for all months and seasons and different future time periods. Attention should be made particularly for wet seasonal/monthly precipitation downscaling.

**Summary and Conclusion**

The spotlight of this study has been the GCM predictor selection, and uncertainty analysis of the downscaling techniques. The proposed method based on GCM independent components and the MLR leave-one-out-cross-validation provided satisfactory predictors for the linear and nonlinear models without any subjectivity in the selection process. Hence, we considered different predictor sets for each month/season. The predictor selection depends on the downscaling methodology.
Selection of predictors for the nonlinear models was performed based on their performances in the test dataset. This approach was also implemented by (Khan et al., 2006; Schmidli et al., 2007), however comparing the performance of all possible predictor combinations in the models for the test dataset is not efficient. Therefore we applied ICA and cross validation in a linear mode to select the potential monthly/seasonal predictors for the nonlinear models. The use of independent component analysis determined climate variables that were independent and uncorrelated with each other in order to find relationships between the true GCM and observed, regional variables. Based on the results from ANFIS training (i.e. while the three models were based on sugeno fuzzy system and were trained based on the same algorithm) we determined that the model structure and input predictors were important factors in the downscaling process and were the main sources of uncertainties. This study demonstrates the accuracy gained from the predictor selection technique in three downscaling models, which provides useful model inputs for the purpose of downscaling. Additionally, we can use the results from all the models to determine confidence bounds for future predictions by using the model outputs together to describe the potential variability of precipitation in the future. The multimodel ensemble downscaling process also allows for the analysis and quantification of the underlying uncertainty needed by the decision makers. Individual models may indicate different future changes in regional precipitation,
which signifies the importance of the incorporation of several methods. The combination of the ensemble members will most probably produce better results than the individual realizations (Coelho et al., 2006; Duan et al., 2007; Luo et al., 2007). Also the proposed technique allowed for extraction of information from a sufficiently large grid domain.

The analyses performed in this study were focused on the downscaling of the independent components of a large GCM dataset covering a wide area (the entire Pacific Northwest) to a gauge station. Since the independent components were determined for the entire GCM dataset, those values could be used to downscale to any other gauge or gridded area within the extent of the GCM dataset. In this way, the independent component analysis and predictor selection only needs to be performed once.
Appendix B: Metropolis Hastings within Gibbs Sampler

MCMC implementation to obtain the GPD and latent parameters posterior distributions

1- Calculate the mean and the covariance of the scale parameter \( (\varphi) \):

\[
\mu_\varphi = \lambda_{\varphi_1} + \lambda_{\varphi_2} \times (\text{Area})_{\text{est}} + \lambda_{\varphi_3} \times (\text{Elevation})_{\text{est}} \\
\Sigma_\varphi = \chi_{\varphi} \times \exp(-\chi_{\varphi} d)
\]

(10) (11)

2- Find the scale parameter for the new iteration \( (\varphi^{\text{new}}) \):

\[
\mu_{\varphi^{\text{new}}} = \mu_\varphi + \Sigma_\varphi \left( \chi^{-1} + \Sigma_\varphi \right)^{-1} \left( \varphi - \mu_\varphi \right)
\]

(12)

\[
\Sigma_{\varphi^{\text{new}}} = \Sigma_\varphi - \Sigma_\varphi \left( \chi^{-1} + \Sigma_\varphi \right)^{-1} \Sigma_\varphi \text{ with } \chi^{-1} \text{ as the Fisher information matrix.}
\]

(13)

\[\blacktriangleright\text{ Draw } \varphi' \text{ from the proposal distribution } N(\mu_{\varphi^{\text{new}}}, \Sigma_{\varphi^{\text{new}}}).\]

\[
r = \left[ \sum_r \sum_i \log p(Y_{\text{exc}}|\varphi^*, \kappa) + \log N(\varphi^* | \mu_\varphi, \Sigma_\varphi) + \log N(\varphi_{\varphi^{\text{new}}} | \mu_{\varphi^{\text{new}}}, \Sigma_{\varphi^{\text{new}}}) \right]
\]

\[
- \left[ \sum_r \sum_i \log p(Y_{\text{exc}}|\varphi, \kappa) + \log N(\varphi | \mu_\varphi, \Sigma_\varphi) + \log N(\varphi_{\varphi^{\text{new}}} | \mu_{\varphi^{\text{new}}}, \Sigma_{\varphi^{\text{new}}}) \right]
\]

(14)

where \( Y_{\text{exc}} = Y - \nu \) and \( p \) is the GPD function.

\[\blacktriangleright\text{ Draw } (0 < u < 1) \text{ from the uniform distribution.}\]

If \( r > \log(u) \), \( \varphi^{\text{new}} = \varphi^* \) otherwise \( \varphi^{\text{new}} = \varphi \)

3- Find the ”scale” mean latent parameter \( (\lambda_{\varphi}) \) corresponding to each covariate:

\[
\lambda_{\varphi_{\text{new}}} = \lambda_{\varphi} + e_{\lambda} \text{ with } e_{\lambda} \text{ as the jump rate.}
\]

\[
\mu_{\varphi^{\text{prop}}} = f(\text{covariates}) \text{ with } \lambda_{\varphi} = \lambda_{\varphi_{\text{new}}}
\]

(15)
\[ r = \log N(\psi^{\text{new}}|\mu_{\psi}, \Sigma_{\psi}) - \log N(\psi^{\text{rev}}|\mu_{\psi}, \Sigma_{\psi}) \]  

(16)

- **Draw** \(0 < u < 1\) from the uniform distribution.

  \[ \text{If } r > \log(u), \quad \lambda_{\psi}^{\text{new}} = \lambda_{\psi}^{\text{rev}}, \quad \text{otherwise } \lambda_{\psi}^{\text{new}} = \lambda_{\psi}. \]

4. **Calculate the “scale” covariance latent parameter** \(\chi_{\psi}^{\text{new}}\):

\[
\chi_{\psi}^{\text{prop}} = \chi_{\psi} + \varepsilon_{\chi} \quad \text{with } \varepsilon_{\chi} \quad \text{as the vector of jump rates for} \quad (\chi_{\psi}).
\]

(17)

\[
\mu_{\psi} = f(\text{covariates}) \quad \text{with } \lambda_{\psi} = \lambda_{\psi}^{\text{new}}
\]

(18)

\[
\Sigma_{\psi_{\text{prop}}} = \chi_{\psi_{\text{prop}}} \times \exp(-\chi_{\psi_{\text{prop}}} d)
\]

(19)

\[
r = \left\{ \log N(\psi^{\text{new}}|\mu_{\psi}, \Sigma_{\psi_{\text{prop}}}) + \log f(\chi_{\psi_{\text{prop}}}) \right\} - \left\{ \log N(\psi^{\text{rev}}|\mu_{\psi}, \Sigma_{\psi}) + \log f(\chi_{\psi}) \right\}
\]

with \(f(\chi_{\psi}) = K(\chi_{\psi})F(\varepsilon_{\chi})\) as the prior probability density function of \(\chi_{\psi}\).

- **Draw** \(0 < u < 1\) from the uniform distribution.

  \[ \text{If } r > \log(u), \quad \chi_{\psi}^{\text{new}} = \chi_{\psi_{\text{prop}}}, \Sigma_{\psi} = \Sigma_{\psi_{\text{prop}}}, \quad \text{otherwise } \chi_{\psi}^{\text{new}} = \chi_{\psi}. \]

5. **Find the value of shape parameter for the new iteration**:

\[
\kappa_{\text{prop}} = \kappa + \varepsilon_{\kappa} \quad \text{with } \varepsilon_{\kappa} \quad \text{as the jump rate for} \quad \kappa.
\]

(20)

\[
r = \sum_{i} \sum_{t} \log p(Y_{i,t}|\psi^{\text{new}}, \kappa_{\text{prop}}) - \sum_{i} \sum_{t} \log p(Y_{i,t}|\psi^{\text{rev}}, \kappa)
\]

(21)

- **Draw** \(0 < u < 1\) from the uniform distribution.

  \[ \text{If } r > \log(u), \quad \kappa^{\text{new}} = \kappa_{\text{prop}}, \quad \text{otherwise } \kappa^{\text{new}} = \kappa. \]

Repeat until convergence.