DEVELOPMENT OF AN INTEGRATED
HYDRO-CLIMATIC SYSTEMS ANALYSIS FRAMEWORK AND
ITS APPLICATION TO THE ATHABASCA RIVER BASIN, CANADA

A Thesis
Submitted to the Faculty of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy
in Environmental Systems Engineering
at University of Regina

by
Guanhui Cheng
Regina, Saskatchewan
September, 2016

Copyright 2016: G.H. Cheng
Guanhui Cheng, candidate for the degree of Doctor of Philosophy in Environmental Systems Engineering, has presented a thesis titled, *Development of an Integrated Hydro-Climatic Systems Analysis Framework and its Application to the Athabasca River Basin, Canada*, in an oral examination held on August 30, 2016. The following committee members have found the thesis acceptable in form and content, and that the candidate demonstrated satisfactory knowledge of the subject material.

External Examiner: Dr. Chongyu Xu, University of Oslo

Supervisor: Dr. Guohe Huang, Environmental Systems Engineering

Committee Member: Dr. Amornvadee Veawab, Environmental Systems Engineering

Committee Member: Dr. Boting Yang, Department of Computer Science

Committee Member: Dr. Chunjiang An, Environmental Systems Engineering

Chair of Defense: Dr. Bjoern Wissel, Department of Biology
ABSTRACT

Climate change has profound impacts on regional hydrological characteristics in large unregulated continental river basins (LUCRiBs) such as the Athabasca River Basin (ARB), Canada. A systematic analysis of these impacts is confronted with many challenges. For instance, the performances of general circulation models (GCMs) vary with many factors, e.g. climate variables, geographic locations, temporal scales, and evaluation measures. Mesoscale atmospheric features can barely be provided by coarse-resolution GCMs. Filling this gap by statistical downscaling is further challenged by redundant computations, resulting from spatial climatic similarities, and the complexities of data uncertainties, nonlinear correspondences, normality prerequisites, and multivariate dependencies. Climatic projection may lack a solid GCM-evaluation foundation and a high spatial resolution. These complexities in downscaling may also exist and be coupled with massive computations in integer optimization in hydrological simulation. Furthermore, an integration of these challenges would decrease the reliability of long-term streamflow forecastings for guiding socio-economic development and eco-environmental conservation over LUCRiBs such as the ARB under climate change.

To fill the gap of few effective techniques, an integrated hydro-climatic systems analysis framework is developed and applied to the ARB. This framework includes six modules. (a) The multi-dimensional performances of CMIP5 GCMs and their ensemble are evaluated. (b) The climate over the ARB is classified by recursive dissimilarity and similarity inferences. (c) The spatial resolution of GCM is enhanced by recursive multivariate principal-monotonicity inferential downscaling based on (a) and (b). (d) High-resolution climatic projection under four representative concentration pathways (RCPs) are generated by coupling (a) to (c). (e) The
correspondence between climate and streamflow is reproduced by Bayesian principal-
monotonicity inference based on (b). (f) Modules (d) and (e) are integrated for
streamflow forecasting under climate change.

A series of findings are revealed while methodological reliability is verified. For
instance, the multi-model ensemble has a relatively high modeling accuracy. The
climatic conditions over the ARB are classified into 20 classes based on their
dissimilarity and similarity. The overall downscaling accuracies are relatively high for
temperature and acceptable for precipitation although varying with multiple factors.
At the scale of octo-decades, daily minimum temperature would increase by 1.7, 2.3,
2.1 and 3.0 °C, daily maximum temperature by 1.4, 1.8, 1.6 and 2.2 °C, and daily total
precipitation by 0.03, 0.07, 0.08 and 0.16 mm under RCPs 2.6, 4.5, 6.0 and 8.5,
respectively. The approach in module (e) is effective at capturing the temporal
variability and the multi-year averages of streamflow and the uncertainties of climate-
streamflow correspondences. Streamflow tends to increase at the upper and middle
reaches and decline at the lower one. The increments of streamflow would be the
highest in March and the decrements would be dominated by less flow in July or
Summer. Either RCP scenarios or modeling biases are significant for the temporal
variability and trends and are insignificant for the overall magnitudes of streamflow.

The methods and findings in this study would be helpful for gaining insights into
coupled climatic and hydrological systems over the ARB, evaluating the impacts of
climate change, guiding regional socio-economic development and eco-environmental
conservation, and promoting developments of more advanced climatic and hydro-
meteorological systems analysis methods.
ACKNOWLEDGMENTS

First of all, I would like to express the sincerest appreciation to my supervisor, Dr. Gordon Huang, for his guidance, support, encouragement, and patience from beginning to completion of this dissertation. His respectable personality, profound knowledge, excellent abilities, rigorous attitudes, innovative thinking, and sense of responsibility for the research community and the public will be the brightest beacon lightening my career and life. My appreciation must also extend to his family, particularly to Ms. Chunling Ke for her warming help.

Many thanks are due to the Faculty of Graduate Studies and Research and the Faculty of Engineering at the University of Regina for providing research awards, teaching assistantships, and travel awards during my Ph.D study.

Sincere thanks to Dr. Chongyu Xu, the external examiner of my thesis defense, for many constructive suggestions on my future research and career development. My appreciation also goes to Drs. Yongping Li, Ping Guo, Qianguo Lin, Xianghui Nie, Hua Zhang, Yanpeng Cai, Xiaodong Zhang, Qian Tan, Gongchen Li, Wei Sun, Chunjiang An, and Renfei Liao for their advices on my research.

I gratefully acknowledge my dear friends in Regina for their kindly help. They include Yumin Chen, Yao Yao, Shanshan Wu, Shan Zhao, Xiujuan Chen, Jiapei Chen, Seang Cau, Xiupuan Wang, Jinxin Zhu, Zhong Li, Yuanyuan Zhai, Kailong Li, Xiong Zhou, and many others. The friendship between us would be the most precious memory during my Ph.D study at University of Regina.
DEDICATION

This dissertation is dedicated to my entire family, especially my wife (Dr. Cong Dong), my parents (Ms. Guilan Yang and Mr. Baoli Cheng), my uncle (Mr. Cuiliu Yang), my parents-in-law (Ms. Cuiping Zhang and Mr. Shoubin Dong), and my sister (Ms. Fengling Cheng), for their unconditional cares, steadfast support, and endless encouragement during my Ph.D study at Regina. Special thanks must be delivered to my maternal grandmother, my grandmother, and my grandfather who are blessing me the greatest happiness and success.
## TABLE OF CONTENTS

ABSTRACT .......................................................................................................................... i  
ACKNOWLEDGMENTS ..................................................................................................... iii  
DEDICATION ................................................................................................................... iv  
TABLE OF CONTENTS ................................................................................................. v  
LIST OF TABLES ............................................................................................................. x  
LIST OF FIGURES .......................................................................................................... xi  
LIST OF APPENDICES .................................................................................................. xv  
LIST OF ABBREVIATIONS ............................................................................................. xvi  

<table>
<thead>
<tr>
<th>CHAPTER 1 INTRODUCTION</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 IMPORTANCE ..........</td>
<td>1</td>
</tr>
<tr>
<td>1.2 LITERATURE REVIEW</td>
<td>1</td>
</tr>
<tr>
<td>1.2.1 GCM Evaluation</td>
<td>1</td>
</tr>
<tr>
<td>1.2.2 Climate Classification</td>
<td>2</td>
</tr>
<tr>
<td>1.2.3 Statistical Downscaling</td>
<td>3</td>
</tr>
<tr>
<td>1.2.4 Climate Projection</td>
<td>4</td>
</tr>
<tr>
<td>1.2.5 Hydrological Simulation</td>
<td>5</td>
</tr>
<tr>
<td>1.2.6 Streamflow Forecasting under Climate Change</td>
<td>6</td>
</tr>
<tr>
<td>1.3 CHALLENGES AND GAPS</td>
<td>8</td>
</tr>
<tr>
<td>1.3.1 GCM Evaluation</td>
<td>8</td>
</tr>
<tr>
<td>1.3.2 Climate Classification</td>
<td>9</td>
</tr>
<tr>
<td>1.3.3 Statistical Downscaling</td>
<td>9</td>
</tr>
<tr>
<td>1.3.4 Climate Projection</td>
<td>10</td>
</tr>
<tr>
<td>1.3.5 Hydrological Simulation</td>
<td>11</td>
</tr>
<tr>
<td>1.3.6 Streamflow Forecasting under Climate Change</td>
<td>12</td>
</tr>
<tr>
<td>1.4 OBJECTIVES ..........</td>
<td>12</td>
</tr>
<tr>
<td>1.5 ORGANIZATION .........</td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 2 AN EVALUATION OF CMIP5 GCMS OVER ATHABASCA RIVER BASIN, CANADA</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 BACKGROUND ................</td>
<td>19</td>
</tr>
<tr>
<td>2.2 DATA AND METHODOLOGY ....</td>
<td>21</td>
</tr>
<tr>
<td>2.2.1 Overview ................</td>
<td>21</td>
</tr>
<tr>
<td>2.2.2 Data Collection ..........</td>
<td>26</td>
</tr>
</tbody>
</table>
5.3.3 Spatial Averages: Intra-Annual Variability ............................................. 176
5.3.4 Spatial Variability: Bi-Decadal Averages ............................................. 178
5.3.5. Spatial Variability: Inter-Seasonal Differences ...................................... 184
5.4 DISCUSSIONS .............................................................................................. 187
5.4.1 Uncertainty Analysis ............................................................................... 187
5.4.2 Bias Analysis ........................................................................................... 190
5.4.3 Impacts Analysis ...................................................................................... 190
5.4.4 Potential Extensions ................................................................................. 191
5.5 SUMMARY ................................................................................................... 193

CHAPTER 6 BAYESIAN PRINCIPAL-MONOTONICITY INference:
METHODOLOGY DEVELOPMENT AND AN APPLICATION TO A HYDRO-
METEOROLOGICAL SYSTEMS ANALYSIS OVER THE ATHABASCA RIVER
BASIN ........................................................................................................ 196

6.1 BACKGROUND ............................................................................................ 196
6.2 METHODOLOGY DEVELOPMENT .......................................................... 200
6.2.1 Overview.................................................................................................. 200
6.2.2 Normality Analyses ................................................................................. 202
6.2.3 Recursive Dissimilarity Inferences.......................................................... 203
6.2.4 Bayesian Principal Monotonicity Analyses ............................................. 205
6.2.5 Recursive Similarity and Dissimilarity Inferences ......................... 207
6.3 CASE STUDY: ATHABASCA RIVER BASIN ........................................... 209
6.3.1 Study Area ............................................................................................... 209
6.3.2 Data.......................................................................................................... 212
6.3.3 Correlation Analysis ................................................................................ 215
6.3.4 Model Configuration ............................................................................... 218
6.4 RESULTS ANALYSIS .................................................................................. 219
6.4.1 Calibration ............................................................................................... 219
6.4.2 Verification .............................................................................................. 222
6.4.3 Bias Analysis ........................................................................................... 222
6.5 DISCUSSIONS .............................................................................................. 225
6.5.1 Bayesian Gaussian Process Optimization ............................................ 225
6.5.2 Correspondence Analysis ................................................................. 228
6.5.3 Sensitivity Analysis ................................................................................. 230
6.6 SUMMARY ................................................................................................... 233
LIST OF TABLES

Table 2.1. Representative Outputs of CMIP5 GCMs for Continental Impact Studies. ................................................................. 28

Table 2.2. Data Availabilities of CMIP5-GCMs. ................................................................. 29

Table 2.3. Selection of Alternative CMIP5 GCMs for the ARB under Seasonal and 10-km Resolutions. ................................................................. 39

Table 4.1. Selected Predictors.................................................................................. 129

Table 5.1. CMIP5 Models Used in Chapter 5.......................................................... 168

Table 6.1. Selected Gauge Stations around the Outlets of Sub-Catchments. .......... 214

Table 7.1. CMIP5 Models Used in Chapter 7.......................................................... 246

Table 7.2. Octo-Decadal Averaged Discharge Changes. ........................................... 251
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Organization of This Dissertation</td>
<td>18</td>
</tr>
<tr>
<td>2.1</td>
<td>Geographical Conditions of the Athabasca River Basin</td>
<td>24</td>
</tr>
<tr>
<td>2.2</td>
<td>Framework of CMIP5-GCM Evaluation</td>
<td>25</td>
</tr>
<tr>
<td>2.3</td>
<td>Spatial Distributions of Reanalysis and Selected CMIP5-GCM Datasets</td>
<td>31</td>
</tr>
<tr>
<td>2.4</td>
<td>CMIP5 GCMs of The Highest Accuracies in Simulating Seasonal Spatial Distributions of Climate Variables</td>
<td>40</td>
</tr>
<tr>
<td>2.5</td>
<td>CMIP5-GCM Simulation Accuracies: Multi-Year Spatial Distributions</td>
<td>41</td>
</tr>
<tr>
<td>2.6</td>
<td>CMIP5-GCM Simulation Accuracies: Annual Cycle of Spatial Variability</td>
<td>45</td>
</tr>
<tr>
<td>2.7</td>
<td>CMIP5-GCM Simulation Accuracies: Annual Cycle of Spatial Means</td>
<td>46</td>
</tr>
<tr>
<td>2.8</td>
<td>CMIP5-GCM Simulation Accuracies: Monthly and Seasonal Spatial Variability</td>
<td>49</td>
</tr>
<tr>
<td>2.9</td>
<td>CMIP5-GCM Simulation Accuracies: Monthly and Seasonal Spatial Means</td>
<td>51</td>
</tr>
<tr>
<td>2.10</td>
<td>CMIP5-GCM Simulation Accuracies: Multi-Year Spatial Variability and Means</td>
<td>54</td>
</tr>
<tr>
<td>3.1</td>
<td>Geographical Conditions and Climatic Observations of The Athabasca River Basin</td>
<td>85</td>
</tr>
<tr>
<td>3.2</td>
<td>Multi-Year-Averaged Spatial Climatological Variability</td>
<td>87</td>
</tr>
<tr>
<td>3.3</td>
<td>Monthly Spatial Climatological Variability over The ARB in Winter</td>
<td>88</td>
</tr>
<tr>
<td>3.4</td>
<td>Monthly Spatial Climatological Variability over The ARB in Spring</td>
<td>89</td>
</tr>
<tr>
<td>3.5</td>
<td>Monthly Spatial Climatological Variability over The ARB in Summer</td>
<td>90</td>
</tr>
<tr>
<td>3.6</td>
<td>Monthly Spatial Climatological Variability over The ARB in Autumn</td>
<td>91</td>
</tr>
</tbody>
</table>
Figure 3.7. Climate Zones by The *RedSICC* Approach........................................... 94
Figure 3.8. Heterogeneities and Similarities of Climate Zones................................. 97
Figure 3.9. Effectiveness of The *DDT* Approach...................................................... 100
Figure 3.10. Analyses of the *RedSICC* Process.......................................................... 102
Figure 3.11. Sensitivity Analyses for The *RedSICC* Approach................................. 105
Figure 4.1. Framework of This Chapter..................................................................... 114
Figure 4.2. Study Area and Data Collection............................................................... 126
Figure 4.3. Grids-Similarity Analysis Results............................................................. 131
Figure 4.4. Sensitivities of Parameters in *ReMPMID*................................................. 134
Figure 4.5. Accuracies of the Calibrated *ReMPMID* Model......................................... 137
Figure 4.6. Verification Accuracies of *ReMPMID* in Simulating Multi-Year Climate
    Variability and Magnitudes................................................................................... 140
Figure 4.7. Verification Accuracies of *ReMPMID* in Simulating Climate Variability
    and Magnitudes in Winter.................................................................................... 144
Figure 4.8. Verification Accuracies of *ReMPMID* in Simulating Climate Variability
    and Magnitudes in Spring................................................................................... 145
Figure 4.9. Verification Accuracies of *ReMPMID* in Simulating Climate Variability
    and Magnitudes in Summer.................................................................................. 146
Figure 4.10. Verification Accuracies of *ReMPMID* in Simulating Climate Variability
    and Magnitudes in Autumn.................................................................................. 147
Figure 4.11. Optimal Values of *Nmin*...................................................................... 149
Figure 4.12. Uncertainties in the Results of *ReMPMID*.............................................. 151
Figure 5.1. Framework of Chapter 5......................................................................... 161
Figure 5.2. Study Area and Data Collection............................................................... 165
Figure 5.3. Bi-decadal Anomalies.............................................................................. 173
Figure 5.4. Annual Anomalies. (Orange lines: no temperature change; pink lines: no precipitation change). ..................................................................................... 175

Figure 5.5. Intra-Annual Variability of Climate Anomalies........................................ 177

Figure 5.6. Spatial Variability of Climate Anomalies under RCP 2.6......................... 180

Figure 5.7. Spatial Variability of Climate Anomalies under RCP 4.5......................... 181

Figure 5.8. Spatial Variability of Climate Anomalies under RCP 6.0......................... 182

Figure 5.9. Spatial Variability of Climate Anomalies under RCP 8.5......................... 183

Figure 5.10. Inter-Seasonal Differences of Spatial Climate-Anomaly Variability... 186

Figure 5.11. Quartiles of Monthly Climatic Anomalies. (Red: maximum; orange: 75%; black: median; blue: 25%; green: minimum). ............................................. 189

Figure 6.1. Principle, Structure and Significance of BaPMI. ..................................... 201

Figure 6.2. Study Area and Data Collection. ............................................................ 211

Figure 6.3. Correlations between Climatic Conditions and Discharges. (PC: the Pearson product-moment correlation coefficient; SC: the Spearman’s rank correlation coefficient; PCP: the P value of PC; SCP: the P value of SC)... 217

Figure 6.4. Calibration and Verification Results......................................................... 221

Figure 6.5. Modeling Biases. .................................................................................... 224

Figure 6.6. Bayesian Optimization. ........................................................................ 227

Figure 6.7. Criterion Predictors for Every Gauge Station........................................ 229

Figure 6.8. Impacts of Parameters in BaPMI on the Verification Results............. 232

Figure 7.1. Study Area and Data.............................................................................. 242

Figure 7.2. Octo-Decadal Averaged Discharges...................................................... 250

Figure 7.3. Bi-Decadal Variability of Averaged Discharges..................................... 253

Figure 7.4. Annual Flow. (Pink lines: historical averages; OSDA: originally simulated discharge averages; CSDA: corrected simulated discharge averages)........ 255
Figure 7.5. Monthly Variability of Flow. \textit{(CSDR: corrected simulated discharge changing ratios; OSDR: originally simulated discharge changing ratios; Mean: the mean values; Q0.25: the first quartile; Q0.75: the third quartile)}............. 257
LIST OF APPENDICES

APPENDIX A. THE MODIFIED NEL AND VAN DER MERWE TEST
(KRISHNAMOORTHY AND YU, 2004) .................................................... 303

APPENDIX B. THE ZS TEST (ZHOU AND SHAO, 2014) ......................... 305

APPENDIX C. THE SHAPIRO-WILKS TEST (SHAPIRO AND WILK, 1965) .. 308

APPENDIX D. THE ASPIN-WELCH-SATTERTHWAITE T-TEST (JÁCOME ET
AL., 2007) .................................................................................................. 310
LIST OF ABBREVIATIONS

**ARB**  Athabasca River Basin

**BaGPO**  Bayesian Gaussian process optimization

**BaPMI**  Bayesian principal-monotonicity inference

**BO**  Bayesian optimization

**CMIP3**  The third phase of the Coupled Model Intercomparison Project

**CMIP5**  The fifth phase of the Coupled Model Intercomparison Project

**CSDAs**  Corrected simulated discharge averages

**CSDRs**  Corrected simulated discharge changing ratios

**DDT**  Discrete distribution transformation

**DJF**  December, January, and February

**GCMs**  Global climate models

**GP**  Gaussian process

**IHCSA**  Integrated hydro-climatic systems analysis framework

**JJA**  June, July, and August

**KC**  Kendall rank correlation coefficient

**LUCriBs**  Large unregulated continental river basins

**MAE**  Mean absolute error

**MAM**  March, April, and May

**MNV**  Modified Nel and van der Merwe

**Nash**  Nash-Sutcliffe coefficient

**Nmin**  Minimum partition row number

**NRMSE**  Normalized root-mean-square error

**OSDAs**  Originally simulated discharge averages

**OSDRs**  Originally simulated discharge changing ratios
$PC$  Pearson product-moment correlation coefficient

$Prec$  Daily cumulative precipitation

$RAE$  Relative absolute error

$RCP$  Representative concentration pathway

$RDA$  Recursive dissimilarity analysis

$ReDSICC$  Recursive dissimilarity and similarity inferential climate classification

$ReMPMID$  Recursive multivariate principal-monotonicity inferential downscaling

$RMSE$  Root mean square error

$RRSE$  Root relative squared error

$RSA$  Recursive similarity analysis

$SC$  Spearman’s rank correlation coefficient

$SDSM$  Statistical downscaling model

$SI$  Scatter index

$SON$  September, October, and November

$Tmax$  Daily maximum temperature

$Tmin$  Monthly averages of daily minimum temperature
CHAPTER 1
INTRODUCTION

1.1 IMPORTANCE

The planet has experienced climate change throughout geologic time as a result of earth’s natural processes (Fath and Fath, 2014). Global climate change appears to affect most of the world’s water resources by altering the processes in natural ecosystems (Fiseha et al., 2014). One of the most severe impacts will be the changes in regional water availability (Xu and Singh, 2004). In particular, large unregulated continental river basins (LUCRiBs), such as the Athabasca River Basin (ARB) in Canada, tend to be sensitive to climate change due to relatively low intensities of human activities such as hydro-power dam construction. Nevertheless, systematic quantification of the impacts of climate change on hydrological systems in LUCRiBs is challenging because of the complicated coupling of climatic and hydrologic systems. Thus, the task of this research is associated with a variety of modeling processes that are interrelated with each other. They are climate modeling evaluation, climate classification, downscaling, climatic projection, hydrological simulation, and streamflow forecasting. To facilitate examining the impacts of climate change on hydrological systems, a number of studies were carried out on these modeling processes by researchers all over the world (e.g., Lettenmaier et al. 1999; Fowler et al. 2007). A brief review of these studies with an emphasis on the ARB is presented below.

1.2 LITERATURE REVIEW

1.2.1 GCM Evaluation
A widely used tool for climatic projection is general circulation models (GCMs). Since a meeting involving 20 climate modeling groups from around the world in September 2008, a new set of coordinated climate model experiments based on multiple GCMs, comprising the fifth phase of the Coupled Model Intercomparison Project (CMIP5), has been promoted. As a result, a challenging problem for climate-change impact studies in the ARB is to decide which CMIP5 GCMs are desired. Previously, very few studies focused on evaluating the performances of CMIP5 GCMs for the ARB or regions covering the ARB. In (Sheffield et al., 2013), the twentieth-century simulation of intraseasonal to multidecadal variability and teleconnections with North American climate by CMIP5 GCMs were evaluated. As revealed in this publication, the multimodel ensemble was found to be reasonably well on reproducing observed variability in multiple aspects. The performance of 22 GCMs in the Coupled Model Intercomparison Project Phase 3 (CMIP3) over all of North America and its western subregion was analyzed based on multiple evaluation metrics (Radic and Clarke, 2011). For other related studies (e.g., Šeparović et al., 2013; Erler et al., 2015; Jiang et al., 2015; Leong and Donner, 2015), the selection of GCMs in CMIP5 or CMIP3 was mainly based on subjective judgement instead of a systematic comparison of all available models.

1.2.2 Climate Classification

Classifications are an important tool in both general and applied climatology, since a strict assignment of individual objects to disjunctive groups is necessary or at least favorable for achieving clearly structured results from complex data sets (Jacobeit, 2010). Classifications have had a long history in meteorology and climatology, and can be generalized as a task of grouping entities (cases) so that they share similar features within
each group, while being dissimilar between groups (Huth et al., 2008). To achieve this task in many fields of the meteorological or climatological sciences for a large spectrum of purposes, various classification methods were developed. Representative methods consisted of, but were not limited to, the Köppen-Geiger approach (Köppen, 1923), the principal component analysis (Bartzokas and Metaxas, 1996; Huth, 2000; Jacobheit et al., 2003), the Hess-Brezowsky catalog (Baur et al., 1944; Hess and Brezowsky, 1952), the Vangengeim-girs method (Vangengeim, 1935), the k-means cluster analysis (Brinkmann, 1999; Esteban et al., 2006; Stahl et al., 2006), self-organizing maps (Cassano et al., 2006; Reusch et al., 2007), the Lamb (Lamb, 1972), and fuzzy-rules-based classification (Bardossy et al., 1995; Stehlik and Bardossy, 2003). Comprehensive reviews of existing methods were presented in a few of publications (e.g., Huth et al., 2008; Jacobheit, 2010).

1.2.3 Statistical Downscaling

Within the scope of projecting future climate change, GCMs are commonly used to assess changes resulting from further increases of atmospheric greenhouse gases (Hertig and Jacobheit, 2013). The atmospheric mesoscale features and the land surface heterogeneity are not properly resolved in coarse-resolution GCMs (Xu and Yang, 2015). Statistical downscaling provides a computationally inexpensive technique that can be adapted for a wide range of applications (Hertig and Jacobheit, 2013). Statistical downscaling can be generalized as a process of deducing finer-resolution regional or local climatic variables (i.e., predictands) from related influencing factors (i.e., predictors), which may include large-scale atmospheric variables or physiographic features (e.g., topography and land use), through statistical analyses. In the past decades, a number of statistical downscaling approaches were developed by researchers all over the world. An
overview of these approaches is presented in many publications such as (Maraun et al. 2010), (Wilby and Wigley, 1997), and (Schoof, 2013). Representative methods include but are not limited to the followings. Scaling techniques are perhaps the most intuitive statistical methods for inferring fine scale information from GCMs (Schoof, 2013). For example, spatial interpolation or disaggregation of GCM outputs provides a baseline against which more rigorous downscaling methods can be compared (Wheater et al. 1999). The artificial neural network can approximate nonlinear relations between predictors and predictands and their derivatives without prior knowledge of a specific nonlinear function (Gupta et al., 2014). This approximation is helpful for making accurate forecasting of highly nonlinear climate systems (Schoof, 2013). A growing number of computational learning algorithms, including tree-based methods (Goyal et al. 2012), genetic programming (Pour et al., 2014), support vector machines (Tripathi et al. 2006), and relevance vector machines (Ghosh and Mujumdar 2008), were developed as well. Classification methods based on fuzzy rules (Bárdossy et al. 2005), stepwise cluster analysis (Wang et al., 2013), and self-organizing maps (Huva et al., 2015) were widely applied within a context of downscaling.

1.2.4 Climate Projection

Climate projection is the key driver of long-term hydrological forecasting of which the boundary climate conditions in the future would be provided by climate projection. In particular, high-resolution climate projection based on downscaling would be more helpful because of the capability of characterizing the diversity of regional climate. However, very few studies were conducted to enhance the resolution of climate projection over the ARB. In (Erler et al., 2015), the Weather Research and Forecasting Model
(Skamarock and Klemp, 2008) was applied to support dynamical downscaling of climatic conditions over western Canada (including the ARB). This study was associated with multi-aspect simplifications. Only one GCM, i.e. the Community Earth System Model (Gent et al., 2011), was used to drive the WRF model without a comprehensive evaluation of the multi-dimensional performances of multiple GCMs. The projection period was only from 2045 to 2060; the future status of climate over the ARB in the 21st century could not be provided. Only one emission scenario, RCP 8.5, out of multiple potential scenarios was considered. This scenario corresponded to a relatively extreme case of future emissions and could hardly guarantee the systematization of projection results. These simplifications would not be helpful for increasing the reliability of long-term planning of impact or adaptation practices.

1.2.5 Hydrological Simulation

Hydrological models (Izzard, 1966; Dooge, 1967) can be categorized as metric (also called data-based, empirical or black box) models, parametric (also called conceptual or grey box) models, and mechanistic (also called physically-based or white box) models. Representative ones include, but are not limited to, artificial neural networks (Stelling, 2000; Govindaraju, 2000), recurrent neural networks (Price et al., 1998; Proaño et al., 1998; Van den Boogaard et al., 1998), nonlinear stochastic simulation (Kavvas, 2003), response surface analysis (Box and Wilson, 1951), nonlinear dynamics and chaos (Sivakumar, 2000), supportive vector machines (Cristianini and Shawe-Taylor, 2000), lumped models (Rockwood, 1966; Anderson, 1967), distributed models (Huggins and Monke, 1968; Beasley et. al., 1980; Quick, 1995), semi-distributed models (Hughes and Sami, 1994; Arnold et al., 1998; Saleh et al., 2000; Gassman et al., 2001; Nasr et al., 2005),
the Systeme Hydrologique Européen model (Abbott et al., 1986a, b), and the Data-based Mechanistic model (Young and Beven, 1994).

Reflection of nonlinearities in hydrological systems varies with the hydro-system analysis methods. In metric models, this is commonly derived from analyses of available data series (Wüst, 1995; Minns, 1996; Scardi, 1996; Recknagel et al., 1997; Sanchez et al., 1998; See and Openshaw, 1999; Abrahart et al., 2004; Dong et al., 2015). Especially for classification methods (De Bruin and Stein, 1998; Hong et al., 2004; Minglei et al., 2010; Olden et al., 2012), a particular type of metric models, the analyses are focused on either dissimilarity or similarity of data series based on an artificially designated threshold or criterion. When parametric models are employed, hydrological systems are modeled through analyses of balances among fluxes such as rainfall, infiltration, percolation, evapotranspiration, runoff and drainage. In contrast to metric models, parameter relationships that relate to nonlinearities are defined by the modeller’s understanding. As for mechanistic models, nonlinearities are formulated as equations based on an understanding of the physical processes and spatial discretization of hydrological systems. In general, modeling of nonlinearities is achieved through either one-way empirical evaluation in metric models or predefined functions in parametric or mechanistic models.

1.2.6 Streamflow Forecasting under Climate Change

Previously, many studies were also carried out to achieve streamflow projection under climate change, e.g. (Gutiérrez and Dracup, 2001; Chiew et al., 2003; Chandimala and Zubair, 2007; Ionita et al., 2008; Gámiz-Fortis et al., 2010; Lamb et al., 2010; Oubeidillah et al., 2011). More recently, Sulis et al. (2011) applied a fully integrated surface-subsurface hydrological model for climate change impact analysis. Das et al.
Kalra et al. (2013) used the support vector machine which was driven by annual average oceanic-atmospheric indices to predict annual streamflow volumes for multiple sites in the Gunnison River Basin and the San Juan River Basin. Sittichok et al. (2014) analyzed the ability of various statistical techniques to forecast the July-August-September total rainfall and monthly streamflow in the Sirba watershed (West Africa). Musau et al. (2015) assessed the impact of climate change on the streamflow in Mt. Elgon watersheds using the 10-GCM Special Report on Emissions Scenarios in combination with a hydrological model (SWAT). To examine water resource shortage and salt water intrusion during dry seasons in Pearl River, China, Yan et al. (2015) evaluated the variations in low flow using the Variable Infiltration Capacity model driven by bias-corrected results of five GCMs under scenarios RCP4.5 and 8.5. In (Zhang et al., 2016), climate data was projected from three CMIP5 GCMs under RCP scenarios using a statistical downscaling model, and future streamflow was modeled through coupling climatic projection with SWAT.

Meanwhile, many issues regarding regional socio-economic development and eco-environmental conservation, e.g. the conflict between water consumptions of oil-sands mining and ecosystem health at the downstream, is closely related with the temporal and spatial variability of the streamflow over the ARB. The streamflow is associated with climate change that would significantly alter the hydrological regimes over this river basin. Long-term planning of socio-economic and eco-environmental systems desires reliable projection of the streamflow under climate change. Few studies were carried out to
examine the future changes of streamflow over the ARB. Until recently, two papers presented some related studies. In (Leong and Donner, 2015), the combination of a land surface process model and a hydrological routing model was used to evaluate the influence of water withdrawals and climate change on streamflow in the ARB. In (Sauchyn et al., 2015), the decadal-scale variability in river discharge in the ARB was examined by a generalized least-squares regression analysis of the trend and variability in gauged flow. Sauchyn et al., (2015) claimed that there would be long-term declining flows throughout the ARB.

1.3 CHALLENGES AND GAPS

Nevertheless, hydro-climatic systems analyses are confronting with many challenges in coupled climatic and hydrological systems, leading to decreased reliability of existing related studies. An abstraction of these challenges in various aspects is stated below.

1.3.1 GCM Evaluation

The existing studies can hardly provide a comprehensive evaluation of the multi-dimensional performances of CMIP5 GCMs in modeling regional climatic conditions of the ARB at various spatial and temporal scales. The reliability of the evaluation results is challenged by complexities of regional particularity, multiple temporal and spatial scales, multiple climate variables, and multiple evaluation criteria. The simulation accuracies of CMIP5 GCMs for the ARB may vary with these complexities. Failure to robustly incorporate these complexities into the CMIP5 GCM evaluation processes may lead to a series of severe consequences. For instance, the recommended CMIP5 GCM may not be the most accurate, or may be the least accurate, in simulating a statistical feature of a
climate variable at a particular combination of temporal and spatial scales; accordingly, planning human activities based on a CMIP5 GCM of a low accuracy may pose threats on local socio-economic development and eco-environmental conservation in the ARB.

1.3.2 Climate Classification

Climate classifications are associated with a few of complexities. For instance, the climate-variable observations collected for classification may be of uncertainties due to causes such as instrumental or operational errors. Multiple climate variables selected to represent regional climate may be dependent on each other. A common prerequisite for most statistical climate classification methods is that the climate-variable samples should come from a normally distributed population. The most desired result of climate classification is that climatic conditions are significantly different for any two climate classes and are not of significant differences in the same climate class. Besides, many existing climate classification methods rely on subjective judgement to support screening of climate variables, selection of classification thresholds, and setting of climate class sizes or numbers prior to classification practices. These issues are challenging the effectiveness of existing climate classification methods, the reliability of climate classification results, and the reasonability of climate-change impact studies or the other related research.

1.3.3 Statistical Downscaling

As for statistical downscaling, the historical observations or simulation of climate variables may be of uncertainties. Two atmospheric-variables combinations that are not significantly different may correspond to two different values of a climate variable, in which the significance of the climate-variable difference is unknown. The relationship
between a predictor and a predictand may be highly nonlinear (Weichert and Bürger, 1998), which is further complicated by the multiplicity of predictors and predictands, the uncertainties of observations and simulation, the temporal and spatial heterogeneities of climate systems, as well as other potential complexities. It is possible that such a complicated relationship is hardly quantified by a continuous specific function. Multiple predictands may be dependent on each other. This dependent effect is neglected when a downscaling tool is used for every predictand separately. The effectiveness of most statistical downscaling methods relies on a prerequisite, i.e. the population of predictor or predictand samples coming from a normal distribution, which is merely discussed in downscaling studies or practices. The correspondence between predictors and predictands may be of high similarity among various spatial units, resulting in redundant computations. There may be nonstationarities in the predictors-predictands correspondence between different periods. For a statistical downscaling model, the optimal parameter combination calibrated over a specific period; the modeling accuracy of this model may not be the highest or even possibly be the lowest over another period. A single statistical metrics may not be capable of comprehensively quantifying the multi-aspect accuracies of a statistical downscaling method.

1.3.4 Climate Projection

Reliable high-resolution climate projection is highly desired for impact or adaptation studies over the ARB, Canada. Many issues in the ARB, e.g. water availabilities, wild fires, flooding, droughts and energy, are associated with climate change. GCMs-based climate projection for addressing these issues is challenged by the diverse finer-scale climatic regimes over this region. Previously, very few studies were conducted to
enhance the resolution of climate projection over the ARB. In addition to the challenges in statistical downscaling, existing studies regarding climate projection over the ARB are based on multi-aspect simplifications of the complicated climate system. These simplifications would not be helpful for increasing the reliability of long-term planning of impact or adaptation practices.

1.3.5 Hydrological Simulation

Nonlinearities in hydrological systems may be temporally or spatially heterogeneous, leading to irregularity of the related nonlinear relationships. A particular function generalized over the entire temporal-spatial horizon, which is commonly used in existing parametric or mechanistic methods, may not reasonably simulate the irregularity. In addition, a hydrological system analysis may be further challenged by the coexistence of irregular nonlinearities and other system complexities. For instance, uncertainties may appear in the observations of both independent and dependent factors such as daily precipitation and surface runoff. The same values of independent factors may correspond to significantly different values of a dependent variable. Besides, a hydrological variable of interest may be nonlinear with multiple influencing factors, and these nonlinear relationships are dependent upon each other. The significance of a factor for the hydrological variable may vary with value levels, e.g. low significance of the factor for high values of the hydrological variable and high significance for moderate values. It is possible that a factor is significant for global variation of the hydrological variable and that another factor is only significant for local variation. Furthermore, massive computational loads are a common problem for hydrological modeling, especially for large-scale finer-resolution problems. For existing hydrological systems analysis methods
such as classification methods, incorporating all of these complexities into the modeling process poses a difficult challenge. Failure to address these integrated complexities in hydro-system analyses may lead to a misrepresentation of the hydrological system, distortion of analysis processes and results, unreliability of resultant decision alternatives, and resultant disasters in socio-economic development or eco-environmental qualities.

1.3.6 Streamflow Forecasting under Climate Change

The reliability of existing studies is challenged by multiple complexities in coupled climatic and hydrological systems analyses. In particular, a comprehensive evaluation of the multi-dimensional performances of GCMs, which could build a solid foundation for providing reliable climatic projection to drive streamflow forecasting, was barely carried out in existing studies. These challenges would further propagate into streamflow projection, leading to decreased reliability of projection results and posing threats on socio-economic development and eco-environmental health.

1.4 OBJECTIVES

To fill the gap of insufficient effective techniques, an integrated hydro-climatic systems analysis framework (IHCSA) is developed in this dissertation to quantify the impacts of climate change on hydrological systems in the ARB. This framework includes six modules that are briefly introduced in the following paragraphs.

(1) The accuracies of CMIP5 GCMs and their ensemble in reproducing historical climate conditions in the ARB will be evaluated comprehensively. In section 2.2, thirty representative outputs of CMIP5 GCMs will be identified through a review of existing continental climate-change impact studies; in consideration of data availabilities in these representative GCM outputs and four emission scenarios of CMIP5, six CMIP5 GCMs
and their ensemble will be selected as alternatives for evaluation; based on ten statistical metrics, ten groups of indicators for the simulation accuracies of CMIP5 GCMs in various aspects will be proposed. The evaluation results will be presented in section 2.3; the CMIP5 GCMs of the highest accuracies for diverse impact studies in the ARB will be identified through a series of systematic comparisons. In section 2.4, scaling effects and statistical-metrics interactions will be revealed, a few of suggestions on CMIP5 GCM selection for impact studies in the ARB will be specified, and the potential extensions of this study will be discussed.

(2) A recursive dissimilarity and similarity inferential climate classification (ReDSICC) approach will be developed, providing an additional alternative for enabling effective classification of climate under these challenges. Specifically, the Shapiro-Wilk test (Shapiro and Wilk, 1965; Royston, 1995) will be incorporated into the framework of ReDSICC to evaluate the normality of multi-dimensional samples based on a collection of the historical observations of representative climate variables. For the samples of a climate variable that do not obey a normal distribution, a discrete distribution transformation approach will be developed to enable reversible transformation between the original distribution and a normal distribution. This will build a solid foundation for the subsequent statistical inferences for which the normality of samples is required. A framework of recursive dissimilarity and similarity inferences will be designed to identify the desired classification of regional climate. In this framework, the modified Nel and van der Merwe test (Krishnamoorthy and Yu, 2004) will be employed to quantify the significance of the dissimilarity or similarity of multi-dimensional climate-variables samples in any two climate classes. For the features of obtained climate classes, the gaps
between them due to discreteness of collected observations will be filled by extending their boundaries to achieve systematic classification of regional climate, and various characterization indicators will be proposed to present their rich properties. Through these efforts, the multi-dimensional samples of selected climate variables will be discretized as a series of groups that are internally inseparable, externally divergent, mutually exclusive, and collectively exhaustive. To verify methodological effectiveness and facilitate studies such as downscaling and hydrological simulation, the \textit{ReDSICC} approach will be applied to a case study of climate classification in the Athabasca River Basin, Canada. A few of findings regarding the regional climate in this river basin or the mechanism of \textit{ReDSICC} will be revealed from a series of comparisons.

(3) A recursive multivariate principal-monotonicity inferential downscaling approach (\textit{ReMPMID}) will be developed for supporting climate downscaling under complexities such as data uncertainties, nonlinear predictors-predictands correspondences, predictands’ interactions, non-normal distributions, spatial homogeneities, and temporal nonstationarities. This approach is an advanced framework in which statistical inferential methods are integrated to address these complexities. Specifically, the principle, innovations and technical details of the \textit{ReMPMID} approach will be discussed in section 4.2. In section 4.3, this approach will be applied to the \textit{ARB}, a large river basin on the Canadian prairies that is closely connected with climate change, to verify methodological effectiveness and facilitate local impact or adaptation practices. Specifically, based on predictor and predictand (i.e. \textit{Tmin}, \textit{Tmax} and \textit{Prec}) selection, data collection and processing, a grids-similarity analysis, a sensitivity analysis, and parameter optimization, a \textit{ReMPMID} model will be constructed for every selected grid in the \textit{ARB}. In section 4.4,
the multi-dimensional modeling accuracies of the *ReMPMID* approach will be verified and examined in various aspects. The optimal parameter values and the uncertainties in high-resolution climate simulation will be analyzed through a series of comparisons.

(4) High-resolution projection of 21st century climate over the *ARB, Canada* under the four *RCP* scenarios will be enabled through the *ReMPMID* approach developed in Chapter 4. Based on previous studies such as *GCM* evaluation (Chapter 2), climate classification (Chapter 3) and model assessment (Chapter 4), we will systematically examine the future status of climatic conditions over the *ARB* at a finer spatial resolution. Specifically, section 5.2 will focus on presenting the results related with system analyses, data collection, method application, and climate characterization. In section 5.3, the future climate conditions over the *ARB* will be analyzed from multiple dimensions, e.g. spatial averages, spatial variability, bi-decadal variability, and intra-annual variability. A series of issues regarding uncertainties, modeling biases, climate-change impacts, and potential extensions of this study will be examined in section 5.4.

(5) To mitigate the related challenges, a Bayesian principal-monotonicity inference (*BaPMI*) method will be proposed in Chapter 6. In *BaPMI*, the responsive relationship from influencing factors (named as predictors) to the hydrological variable of interest (named as the predictand) will be discretized as interrelated end nodes (i.e. groups of paired samples of predictors and the predictand) under irregular nonlinearities, data uncertainties, and multivariate dependencies. In detail, a discrete distribution transformation approach will be developed to enable transformation of non-normally distributed predictand samples as a normal distribution and invertible restoration of the simulated predictand values as the original non-normal distribution. To eliminate data
uncertainties, statistical inference will be employed to assess the significance of
differences among groups of predictand samples. The nonlinear responsive relationship
between predictors and the predictand will be interpreted as piecewise monotonicity,
similar to piecewise linearization of a nonlinear function. The piecewise monotonicity will
be further represented as principal monotonicity, representing the most significant
monotonicity between the predictand and one of the predictors, for dealing with existing
predictor correlations. The process of identifying the principal monotonicity, which is
equivalent to an integer optimization problem, will be accelerated by an incorporation of
Gaussian process analyses and Bayesian optimization. Based on a recursive classification
and cluster analysis, all paired samples of predictors and the predictand that represent the
responsive relationship between them will be discretized as a series of end nodes that are
internally inseparable, externally divergent, mutually exclusive and collectively
exhaustive. Given any combination of predictors, the corresponding predictand value will
be estimated through a *BaPMI* prediction scheme. The properties, strengths, extensions
and shortcomings of *BaPMI* will be also examined through an application to streamflow
simulation in the Athabasca River Basin, Canada.

(6) Finally, the *IHCSA* framework will be finally exported in Chapter 7 to mitigate
the related challenges, enhance the reliability of streamflow projection under climate
change, and facilitate climate-change impact or adaptation studies. This framework will
be applied to a case study of streamflow projection in the *ARB*. Specifically, the
characteristics of the *ARB*, the key structure of *IHCSA*, data requirements, and the
streamflow projection analysis method will be presented in section 7.2. In section 7.3, the
multi-dimensional impacts of climate change on streamflow over the *ARB* will be
examined systematically, and the advantages and shortcomings of IHCSA will also be discussed.

1.5 ORGANIZATION

This dissertation is composed of eight interrelated chapters (Figure 1.1). The importance of this dissertation, a review of existing studies, an analysis of challenges and gaps, and the objectives and organization of this dissertation are stated in Chapter 1 (Introduction). Chapters 2 to 7 focus on presenting the six modules in the IHCSA framework. Specifically, the multi-dimensional performances of CMIP5 GCMs and their ensemble are evaluated in Chapter 2. In Chapter 3, the climate over the ARB is classified by recursive dissimilarity and similarity inferences. In Chapter 4, the spatial resolution of GCM is enhanced by recursive multivariate principal-monotonicity inferential downscaling based on Chapters 2 and 3. In Chapter 5, high-resolution climatic projection under four RCP scenarios is generated by coupling the results of Chapters 2 to 4. In Chapter 6, the correspondence between climate and streamflow is reproduced by Bayesian principal-monotonicity inference based on Chapter 3. In Chapter 7, the results of Chapters 5 and 6 are integrated for streamflow forecasting under climate change. Finally, a summary of this dissertation, research achievements, and recommendations for future research will be discussed in Chapter 8.
Figure 1.1. Organization of This Dissertation.
2.1 BACKGROUND

The Athabasca River originates from the Columbia Icefield on the Rocky Mountains in Alberta, flows approximately 1500 kilometres northeast, and enters the Peace-Athabasca Delta, the largest freshwater continental river delta in North America and the home of numerous migratory birds. As the largest undammed catchment on the Canadian prairies, the ARB covers an area of approximately 138,000 km² on which 0.15 to 0.17 million people reside. In the ARB, there are a series of issues, e.g. water scarcities, wild fires, flooding and droughts, which may be associated with climate change. Elimination of these issues through long-term planning of socio-economic activities desires a reliable projection of the future climatic conditions in the ARB. A widely used tool for climatic projection is GCMs. Since a meeting involving 20 climate modeling groups from around the world in September 2008, a new set of coordinated climate model experiments based on multiple GCMs, comprising the CMIP5, have been promoted. As a result, a challenging problem for climate-change impact studies in the ARB is to decide which CMIP5 GCMs are desired.

Previously, very few studies focused on comparisons of CMIP5 GCMs for the ARB or regions covering the ARB. In (Sheffield et al., 2013), the twentieth-century simulation of intraseasonal to multidecadal variability and teleconnections with North American climate by CMIP5 GCMs were evaluated; as revealed in this publication, the multimodel ensemble was found to be reasonably well at reproducing observed variability in multiple
aspects. The performance of 22 GCMs in the CMIP3 over North America and its western subregion was analyzed based on multiple evaluation metrics (Radic and Clarke, 2011). For other related studies (e.g., Šeparović et al., 2013; Erler et al., 2015; Jiang et al., 2015; Leong and Donner, 2015), the selection of GCMs in CMIP5 or CMIP3 was mainly based on subjective judgement instead of a systematic comparison of all available models.

The existing studies can hardly provide a comprehensive evaluation of the multi-dimensional performances of CMIP5 GCMs in modeling regional climatic conditions of the ARB at various spatial and temporal scales. The reliability of the evaluation results is challenged by the complexities of regional particularity, multiple temporal and spatial scales, multiple climate variables, and multiple evaluation criteria. The simulation accuracies of CMIP5 GCMs for the ARB may vary with these complexities. Failure to robustly incorporate these complexities into the CMIP5 GCM evaluation processes may lead to a series of severe consequences. For instance, the recommended CMIP5 GCM may not be the most accurate, or may be the least accurate, in simulating a statistical feature of a climate variable at a particular combination of temporal and spatial scales; accordingly, planning of human activities based on a CMIP5 GCM of low accuracy may pose threats on local socio-economic development and eco-environmental conservation in the ARB.

Therefore, this study aims to comprehensively evaluate the accuracies of CMIP5 GCMs and their ensemble in reproducing historical climate conditions in the ARB. In section 2.2, thirty representative outputs of CMIP5 GCMs will be identified through a review of existing continental climate-change impact studies; in consideration of data availabilities in these representative GCM outputs and four emission scenarios of CMIP5, six CMIP5 GCMs and their ensemble will be selected as alternatives for evaluation; based
on ten statistical metrics, ten groups of indicators for the simulation accuracies of CMIP5 GCMs in various aspects will be proposed. The evaluation results will be presented in section 2.3; the CMIP5 GCMs of the highest accuracies for diverse impact studies in the ARB will be identified through a series of systematic comparisons. In section 2.4, scaling effects and statistical-metrics interactions will be revealed, a few of suggestions on CMIP5 GCM selection for impact studies in the ARB will be specified, and the potential extensions of this study will be discussed.

2.2 DATA AND METHODOLOGY

2.2.1 Overview

As shown in Figure 2.1, the Athabasca River originates from the Columbia Icefield on the eastern slopes of the Rocky Mountains around Jasper in Alberta, Canada. It flows approximately 1500 km northeast before entering the Peace-Athabasca Delta, the largest freshwater continental river delta in North America and the home to numerous migratory birds, and draining into Lake Athabasca. The evaluation drops from approximately 3715 m at Jasper to 211 m at the outlet in the Peace-Athabasca Delta. This river is the longest undammed river in the Canadian prairies. The ARB covers an area of approximately 138,000 km² over 52° ~ 59° N and -119° ~ -107° W. The ARB includes diverse hydro-climatic regimes due to physiographical heterogeneity; snowcapped mountains, coniferous forest, mixed wood and deciduous forest are found in the uplands, whereas willow brush, shrubs, black spruce and sphagnum moss dominate the lowlands (Kerkhoven and Gan, 2006). Around 0.15 to 0.17 million residents distribute over 22 rural or regional municipalities, 1 city, 12 towns, and 14 Aboriginal settlements in this river basin.
A series of issues in the \textit{ARB}, e.g. water availabilities, wild fires, flooding frequencies, drought durations and energy demands, may be associated with climate warming. For instance, the Athabasca Oil Sands located at the downstream of the \textit{ARB} is one of the largest remaining reserves of petroleum on the planet and a central point of friction in Canadian, American and global climate politics and policy (Leong and Donner, 2015). Surface water use by oil sands mining operations accounts for the largest sectoral water allocations (62 \%) and actual water use (57 \%) in the \textit{ARB} (AMEC Earth and Environmental, 2007). As mining activity expands, surface water use demand is projected to rapidly increase, adding pressure to water availability in the \textit{ARB} (Natural Resources Canada, 2009). A relatively large amount of water use for mining in winters is threatening the health of habitats in the lower-stream wetlands. In addition, it appears that temperature is the most important predictor of area burned in Canada with warmer temperatures associated with increased area burned (Flannigan et al., 2005). In 2016, a fire in Fort McMurray, a city at the downstream of the \textit{ARB}, forced more than 80,000 people to flee and is recognized as one of the most devastating disasters in Alberta’s history.

Therefore, provision of scientific support for guiding local socio-economic and eco-environmental activities and eliminating occurrences of related losses under climate change desires a reliable projection of the future climatic conditions in the \textit{ARB}. Correspondingly, this study focuses on evaluating the accuracies of \textit{CMIP5 GCMs} and their ensemble at reproducing the historical climatic conditions over the \textit{ARB}, Canada
under various temporal and spatial scales based on ten statistical metrics. The framework of this evaluation study is presented in Figure 2.2.
Figure 2.1. Geographical Conditions of the Athabasca River Basin.
Figure 2.2. Framework of CMIP5-GCM Evaluation.
2.2.2 Data Collection

The outputs of GCMs involve a number of climate variables. A review of many related publications (e.g., Mandal et al., 2016; Sarhadi et al., 2016; Borges et al., 2016; Mizukami et al., 2016; Singh et al., 2016; Liu et al., 2016; Dong et al., 2015, 2014a, b, c, 2012, 2013a, b, 2011; Cheng et al., 2016a, b) indicates that part of them are of interests for various impact studies regarding large unregulated continental river basins such as the ARB. The selected climate variables are listed in Table 2.1. Among them, the representative ones that are widely used in climate-change adaptation practices are daily minimum near-surface air temperature ($T_{min}$), daily maximum near-surface air temperature ($T_{max}$), and daily precipitation ($Prec$). Meanwhile, four emission scenarios, i.e. RCP2.6, RCP4.5, RCP6.0 and RCP8.5, are considered in the CMIP5. According to an investigation of datasets achieved on the website of World Data Center for Climate (http://cera-www.dkrz.de/WDCC/ui/), the simulation results of all selected climate variables under historical and the four emission scenarios are not completely available for most CMIP5 GCMs. The details are presented in Table 2.2. In consideration of data availabilities under multiple climate variables and emission scenarios, six GCMs are selected for evaluation in this study. They are IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC-ESM-CHEM, MIROC5, GFDL-ESM2G and GFDL-ESM2M.

The climate variables selected for evaluating the performances of the six CMIP5 GCMs are the three representative ones, i.e. $T_{min}$, $T_{max}$ and $Prec$. The corresponding historical observations are derived from a raster-gridded climate dataset. This dataset contains daily estimates of the three climate variables from 1961 to 2003 over 1615 10-km grids in the ARB (Figure 2.3). The simulated historical values of the three climate
variables, in the corresponding GCM grids, with the same temporal resolution and domain, and by the six CMIP5 GCMs are obtained from the World Data Center for Climate. To make our evaluation analysis consistent, we use one run for the selected CMIP5 GCMs that have multiple runs. The spatial distributions of GCM grids are shown in Figure 2.3. In consideration of possibly better performances in reproducing historical climatic conditions according to existing studies (e.g., Sheffield et al., 2013), the ensemble mean of the six GCMs is taken as the seventh GCM to be evaluated.

For the convenience of presentation in the following sections, the observation dataset is expressed as \(OBS_{vjymd}\) where: \(v\) represents the three climate variables (1 for \(T_{min}\), 2 for \(T_{max}\), and 3 for \(Prec\)), \(j\) represents the 1615 10-km grids, \(y\) represents 43 years from 1961 to 2003, \(m\) represents 12 months from January to December, and \(d\) represents all days in month \(m\) (1, 2, …, \(D_m\) and \(D_m\) is the total number of days in this month). Similarly, the GCM dataset is expressed as \(GCM_{gvjymd}\) where: \(g\) represents the selected CMIP5 GCMs and their ensemble (1 for IPSl-CM5A-LR, 2 for IPSL-CM5A-MR, 3 for MIROC-ESM-CHEM, 4 for MIROC5, 5 for GFDL-ESM2G, 6 for GFDL-ESM2M, and 7 for ensemble) and \(v\), \(j\), \(y\), \(m\) and \(d\) are defined as above. For any \(v\), \(j\), \(y\), \(m\) and \(d\), we have

\[
GCM_{7vjymd} = \left( \sum_{g=1,2,\ldots,6} GCM_{gvjymd} \right)/6. \tag{2.1}
\]
Table 2.1. Representative Outputs of CMIP5 GCMs for Continental Impact Studies.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name (Unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>clt</td>
<td>Total cloud fraction (%)</td>
</tr>
<tr>
<td>hfls</td>
<td>Surface upward latent heat flux (W/m²)</td>
</tr>
<tr>
<td>hfss</td>
<td>Surface upward sensible heat flux (W/m²)</td>
</tr>
<tr>
<td>Prec</td>
<td>Precipitation (kg<em>m⁻²</em>s⁻¹)</td>
</tr>
<tr>
<td>psl</td>
<td>Sea level pressure (Pa)</td>
</tr>
<tr>
<td>rds</td>
<td>Surface downwelling shortwave radiation (W/m²)</td>
</tr>
<tr>
<td>rsus</td>
<td>Surface upwelling shortwave radiation (W/m²)</td>
</tr>
<tr>
<td>tas</td>
<td>Near-surface air temperature (K)</td>
</tr>
<tr>
<td>Tmax</td>
<td>Daily maximum near-surface air temperature (K)</td>
</tr>
<tr>
<td>Tmin</td>
<td>Daily minimum near-surface air temperature (K)</td>
</tr>
<tr>
<td>prc</td>
<td>Convective precipitation (kg/m²/s)</td>
</tr>
<tr>
<td>prsn</td>
<td>Snowfall flux (kg/m²/s)</td>
</tr>
<tr>
<td>rlds</td>
<td>Surface downward longwave radiation (W/m²)</td>
</tr>
<tr>
<td>rlus</td>
<td>Surface upwelling longwave radiation (W/m²)</td>
</tr>
<tr>
<td>hus</td>
<td>Specific humidity (1)</td>
</tr>
<tr>
<td>huss</td>
<td>Near-surface specific humidity (1)</td>
</tr>
<tr>
<td>rlut</td>
<td>TOA outgoing longwave radiation (W/m²)</td>
</tr>
<tr>
<td>ta</td>
<td>Air temperature (K)</td>
</tr>
<tr>
<td>ua</td>
<td>Eastward wind (m/s)</td>
</tr>
<tr>
<td>va</td>
<td>Northward wind (m/s)</td>
</tr>
<tr>
<td>wap</td>
<td>omega (=dp/dt) (Pa/s)</td>
</tr>
<tr>
<td>rhs</td>
<td>Near-surface relative humidity (%)</td>
</tr>
<tr>
<td>uas</td>
<td>Eastward near-surface wind speed (m/s)</td>
</tr>
<tr>
<td>vas</td>
<td>Northward near-surface wind speed (m/s)</td>
</tr>
<tr>
<td>hur</td>
<td>Relative humidity (%)</td>
</tr>
<tr>
<td>zg</td>
<td>Geopotential height (m)</td>
</tr>
<tr>
<td>rhsmax</td>
<td>Surface daily maximum relative humidity (%)</td>
</tr>
<tr>
<td>rhsmin</td>
<td>Surface daily minimum relative humidity (%)</td>
</tr>
<tr>
<td>sfcWind</td>
<td>Daily-mean near-surface wind speed (m/s)</td>
</tr>
<tr>
<td>sfcWindmax</td>
<td>Daily maximum near-surface wind speed (m/s)</td>
</tr>
</tbody>
</table>
Table 2.2. Data Availabilities of *CMIP5-GCMs*.

<table>
<thead>
<tr>
<th>CMIP5 GCMs</th>
<th>Historical</th>
<th>RCP 2.6</th>
<th>RCP 4.5</th>
<th>RCP 6.0</th>
<th>RCP 8.5</th>
<th>Unavailable Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCC bcc-csm1-1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>sfcWind, sfcWindmax</td>
</tr>
<tr>
<td>BCC bcc-csm1-1-m</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>sfcWind, sfcWindmax</td>
</tr>
<tr>
<td>BNU BNU-ESM</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>*</td>
</tr>
<tr>
<td>CCCma CanCM4</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>*</td>
</tr>
<tr>
<td>CCCma CanESM2</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>*</td>
</tr>
<tr>
<td>CMCC CMCC-CESM</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>*</td>
</tr>
<tr>
<td>CMCC CMCC-CM</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>*</td>
</tr>
<tr>
<td>CMCC CMCC-CMS</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>*</td>
</tr>
<tr>
<td>CNRM-CERFACS CNRM-CM5</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>*</td>
</tr>
<tr>
<td>CSIRO-BOM ACCESS1-0</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>*</td>
</tr>
<tr>
<td>CSIRO-BOM ACCESS1-3</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>*</td>
</tr>
<tr>
<td>CSIRO-QCCCE CSIRO-Mk3-6-0</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>zg</td>
</tr>
<tr>
<td>ICHEC EC-EARTH</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>*</td>
</tr>
<tr>
<td>INM inmem4</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>*</td>
</tr>
<tr>
<td>INPE HadGEM2-ES</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>*</td>
</tr>
<tr>
<td>IPSL IPSL-CM5A-LR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>■</td>
</tr>
<tr>
<td>IPSL IPSL-CM5A-MR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>■</td>
</tr>
<tr>
<td>IPSL IPSL-CM5B-LR</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>*</td>
</tr>
<tr>
<td>LASG-CESS FGOALS-g2</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>*</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>■</td>
</tr>
<tr>
<td>MIROC-ESM-CHEM</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>■</td>
</tr>
<tr>
<td>MIROC MIROC4h</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>■</td>
</tr>
<tr>
<td>MIROC MIROC5</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>■</td>
</tr>
<tr>
<td>MOHC HadCM3</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>■</td>
</tr>
<tr>
<td>MOHC HadGEM2-CC</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>■</td>
</tr>
<tr>
<td>MOHC HadGEM2-ES</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>■</td>
</tr>
<tr>
<td>MPI-M MPI-ESM-LR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>■</td>
</tr>
<tr>
<td>MPI-M MPI-ESM-MR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>■</td>
</tr>
<tr>
<td>MPI-M MPI-ESM-P</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>■</td>
</tr>
<tr>
<td>MRI MRI-CGCM3</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>■</td>
</tr>
<tr>
<td>MRI MRI-ESM1</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>■</td>
</tr>
<tr>
<td>NASA-GISS GISS-E2-R-CC</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>■</td>
</tr>
<tr>
<td>Model</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>NCAR CCSM4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>rlds, rlus, rlut, uas, vas, rhsmax, rhsmin, sfcWind, sfcWindmax</td>
<td></td>
</tr>
<tr>
<td>NCC NorESM1-M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>rhsmax, rhsmin, sfcWind, sfcWindmax</td>
<td></td>
</tr>
<tr>
<td>NOAA-FGDL GFDL-CM3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>prc, prsn, hur</td>
<td></td>
</tr>
<tr>
<td>NOAA-FGDL GFDL-ESM2G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOAA-FGDL GFDL-ESM2M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSF-DOE-NCAR CESM1-BGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSF-DOE-NCAR-CESM1-FASTCHEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: XAvailable; -Unavailable; *Unselected due to unavailability of outputs under at least one scenario; ■No missing data for all scenarios and selected climate variables.
Figure 2.3. Spatial Distributions of Reanalysis and Selected $CMIP5$-$GCM$ Datasets.
2.2.3 Quantification of Simulation Accuracies

Diverse statistical metrics are available for quantifying the accuracy of GCM simulation. A single one can hardly be more desirable than others in any cases and there are dependencies among alternative metrics (Hayhoe, 2010). Hence, ten metrics are employed in this study: (a) the Pearson product-moment correlation coefficient (PC) (Pearson, 1895), (b) the Kendall rank correlation coefficient (KC) (Kendall, 1938), (c) the Spearman’s rank correlation coefficient (SC) (Spearman, 1904), (d) the root relative squared error (RRSE) (Diaz and Jones, 2004), (e) the root mean square error (RMSE) (Hyndman et al., 2006), (f) the normalized root-mean-square error (NRMSE) (Hyndman et al., 2006), (g) the scatter index (SI) (Hanna and Heinold, 1985), (h) the relative absolute error (RAE) (Armstrong and Collopy, 1992), (i) the mean absolute error (MAE) (Willmott and Matsuura, 2005), and (j) the Nash-Sutcliffe coefficient (Nash) (Nash and Sutcliffe, 1970). For any given paired data of observations ({xₙ}ₙ=1…N) and GCM simulation ({yₙ}ₙ=1…N), metrics (a) to (c) are measures of their linear or nonlinear correlation, while metrics (d) to (j) are measures of their absolute or relative difference. The formulations of these metrics are generalized as

\[ M_c = M_c(\{x_n\}_{n=1\ldots N}, \{y_n\}_{n=1\ldots N}) \]  

in this study where \( c \in \{1, 2, \ldots, 10\} \) represents metrics (a) to (j) and \( M_c(, ) \) represents the corresponding function, respectively. For any particular metrics, its detailed formulation is stated in the corresponding reference.

This study aims to evaluate the multi-aspect performances of seven GCMs, including the ensemble of the selected GCMs, in reproducing historical status of three
climate variables under various temporal-spatial resolutions and windows. To cope with diversity of the performances of GCMs, a suite of scale-varied indices are calculated from the original daily dataset given the ten statistical metrics. The term “simulation accuracy” is incorporated into the names of these GCM-evaluation indices. Let the correspondence between every season \((s = 1, 2, \ldots, 4)\) and three months \((m = 1, 2, 3)\) be denoted as a matrix: \(D(s, m) = \{(12, 1, 2), (3, 4, 5), (6, 7, 8), (9, 10, 11)\}^T\). These GCM-evaluation indices are formulated as follows. These indices are estimated for evaluating the multi-dimensional accuracies of CMIP5 GCMs and their ensemble in reproducing (A1) seasonal averages of \(T_{\text{min}}, T_{\text{max}}\) and \(P_{\text{rec}}\) in every 10-km grids, (A2) multi-year averages of \(T_{\text{min}}, T_{\text{max}}\) and \(P_{\text{rec}}\) in every 10-km grids, (B1) spatial variability of \(T_{\text{min}}, T_{\text{max}}\) and \(P_{\text{rec}}\) in 365 days of the annual cycle, (B2) 1615-grids averages of \(T_{\text{min}}, T_{\text{max}}\) and \(P_{\text{rec}}\) in 365 days of the annual cycle, (C1) spatial variability of \(T_{\text{min}}, T_{\text{max}}\) and \(P_{\text{rec}}\) in 12 months, (C2) spatial variability of \(T_{\text{min}}, T_{\text{max}}\) and \(P_{\text{rec}}\) in 4 seasons, (D1) 1615-grids averages of \(T_{\text{min}}, T_{\text{max}}\) and \(P_{\text{rec}}\) in 12 months, (D2) 1615-grids averages of \(T_{\text{min}}, T_{\text{max}}\) and \(P_{\text{rec}}\) in 4 seasons, (E1) spatial variability of \(T_{\text{min}}, T_{\text{max}}\) and \(P_{\text{rec}}\) in 43 years, and (E2) 1615-grids averages of \(T_{\text{min}}, T_{\text{max}}\) and \(P_{\text{rec}}\) in 43 years, respectively.

(A1) Seasonal distributed simulation accuracies:

\[
DDS_{gvjsc} = Mc\{GCM_{gvjymd} | y = 1961, 1962, \ldots, 2003; m = D(s, 1), D(s, 2), D(s, 3); d = 1, 2, \ldots, Dm\}, \{OBS_{vjymd} | y = 1961, 1962, \ldots, 2003; m = D(s, 1), D(s, 2), D(s, 3); d = 1, 2, \ldots, Dm\} \tag{2.3}
\]

for any \(g \in \{1, 2, \ldots, 7\}, v \in \{1, 2, 3\}, j \in \{1, 2, \ldots, 1615\}, s \in \{1, 2, \ldots, 4\}, \text{ and } c \in \{1, 2, \ldots, 10\};\)
(A2) Multi-year distributed simulation accuracies:

\[ DDO_{gvjc} = M_c(\{ GCM_{gvjymd} \mid y = 1961, 1962, \ldots, 2003; m = 1, 2, \ldots, 12; d = 1, 2, \ldots, D_m \}, \{ OBS_{vjymd} \mid y = 1961, 1962, \ldots, 2003; m = 1, 2, \ldots, 12; d = 1, 2, \ldots, D_m \}) \] (2.4)

for any \( g \in \{1, 2, \ldots, 7\}, v \in \{1, 2, 3\}, j \in \{1, 2, \ldots, 1615\}, \) and \( c \in \{1, 2, \ldots, 10\}; \)

(B1) Annual cycle of spatial-variability simulation accuracies:

\[ DCD_{gymdc} = M_c(\{ \sum_{y=1961, 1962, \ldots, 2003} GCM_{gvjymd}/43 \mid j = 1, 2, \ldots, 1615 \}, \{ \sum_{y=1961, 1962, \ldots, 2003} OBS_{vjymd}/43 \mid j = 1, 2, \ldots, 1615 \}) \] (2.5)

for any \( g \in \{1, 2, \ldots, 7\}, v \in \{1, 2, 3\}, m \in \{1, 2, \ldots, 12\}, d \in \{1, 2, \ldots, D_m \}, \) and \( c \in \{1, 2, \ldots, 10\}; \)

(B2) Annual cycle of spatial-means simulation accuracies:

\[ DMD_{gymdc} = M_c(\{ \sum_{j=1, 2, \ldots, 1615} GCM_{gvjymd}/1615 \mid y = 1961, 1962, \ldots, 2003 \}, \{ \sum_{j=1, 2, \ldots, 1615} OBS_{vjymd}/1615 \mid y = 1961, 1962, \ldots, 2003 \}) \] (2.6)

for any \( g \in \{1, 2, \ldots, 7\}, v \in \{1, 2, 3\}, m \in \{1, 2, \ldots, 12\}, d \in \{1, 2, \ldots, D_m \}, \) and \( c \in \{1, 2, \ldots, 10\}; \)

(C1) Monthly spatial-variability simulation accuracies:

\[ DCM_{gvmc} = M_c(\{ \sum_{y=1961, 1962, \ldots, 2003} \sum_{d=1, 2, \ldots, D_m} GCM_{gvjymd} / (43 \cdot D_m) \mid j = 1, 2, \ldots, 1615 \}, \{ \sum_{y=1961, 1962, \ldots, 2003} \sum_{d=1, 2, \ldots, D_m} OBS_{vjymd} / (43 \cdot D_m) \mid j = 1, 2, \ldots, 1615 \}) \] (2.7)

for any \( g \in \{1, 2, \ldots, 7\}, v \in \{1, 2, 3\}, m \in \{1, 2, \ldots, 12\}, \) and \( c \in \{1, 2, \ldots, 10\}; \)

(C2) Seasonal spatial-variability simulation accuracies:

\[ DCS_{gvs} = M_c(\{ \sum_{y=1961, 1962, \ldots, 2003} \sum_{m=D(s, 1), D(s, 2), D(s, 3)} GCM_{gvjymd} / (43 \cdot \sum_{p=1, 2, 3} D_{D(s, p)}) \mid j = 1, 2, \ldots, 1615 \}, \{ \sum_{y=1961, 1962, \ldots, 2003} \sum_{m=D(s, 1), D(s, 2), D(s, 3)} OBS_{vjymd} / (43 \cdot \sum_{p=1, 2, 3} D_{D(s, p)}) \mid j = 1, 2, \ldots, 1615 \}) \]
\[
\sum_{d=1,2,\ldots,Dm} \frac{OBS_{jy/md}}{(43 \cdot \sum_{p=1,2,3} D_{D(s,p)})} \bigg| j = 1, 2, \ldots, 1615 \biggr) \tag{2.8}
\]

for any \( g \in \{1, 2, \ldots, 7\}, v \in \{1, 2, 3\}, s \in \{1, 2, \ldots, 4\}, \) and \( c \in \{1, 2, \ldots, 10\}; \)

(D1) Monthly spatial-means simulation accuracies:

\[
DMM_{gvmc} = Mc\big\{\sum_{j=1,2,\ldots,1615} GCM_{gvyjymd}/1615 \bigg| y = 1961, 1962, \ldots, 2003; \nonumber
d = 1, 2, \ldots, Dm\big\}, \big\{\sum_{j=1,2,\ldots,1615} OBS_{jy/md}/1615 \bigg| y = 1961, 1962, \ldots, 2003; \nonumber
d = 1, 2, \ldots, Dm\big\} \biggr) \tag{2.9}
\]

for any \( g \in \{1, 2, \ldots, 7\}, v \in \{1, 2, 3\}, m \in \{1, 2, \ldots, 12\}, \) and \( c \in \{1, 2, \ldots, 10\}; \)

(D2) Seasonal spatial-means simulation accuracies:

\[
DMS_{gsvsc} = Mc\big\{\sum_{j=1,2,\ldots,1615} GCM_{gvyjymd}/1615 \bigg| y = 1961, 1962, \ldots, 2003; \nonumber\nonumber
m = D(s, 1), D(s, 2), D(s, 3); d = 1, 2, \ldots, Dm\big\}, \big\{\sum_{j=1,2,\ldots,1615} OBS_{jy/md}/1615 \bigg| y = 1961, 1962, \ldots, 2003; \nonumber\nonumber
m = D(s, 1), D(s, 2), D(s, 3); d = 1, 2, \ldots, Dm\big\} \biggr) \tag{2.10}
\]

for any \( g \in \{1, 2, \ldots, 7\}, v \in \{1, 2, 3\}, s \in \{1, 2, \ldots, 4\}, \) and \( c \in \{1, 2, \ldots, 10\}; \)

(E1) Multi-year spatial-variability simulation accuracies:

\[
DCS_{gvc} = Mc\big\{\sum_{j=1,2,\ldots,1615} \sum_{m=1,2,\ldots,12} \sum_{d=1,2,\ldots,1615} GCM_{gvyjymd}/(43 \cdot 365) \bigg| y = 1961, 1962, \ldots, 2003; \nonumber\nonumber
j = 1, 2, \ldots, 1615\big\}, \big\{\sum_{j=1,2,\ldots,1615} \sum_{m=1,2,\ldots,12} \sum_{d=1,2,\ldots,1615} OBS_{jy/md}/(43 \cdot 365) \bigg| y = 1961, 1962, \ldots, 2003; \nonumber\nonumber
j = 1, 2, \ldots, 1615\big\} \biggr) \tag{2.11}
\]

for any \( g \in \{1, 2, \ldots, 7\}, v \in \{1, 2, 3\}, \) and \( c \in \{1, 2, \ldots, 10\}; \) and

(E2) Multi-year spatial-means simulation accuracies:

\[
DMS_{gvc} = Mc\big\{\sum_{j=1,2,\ldots,1615} \sum_{m=1,2,\ldots,12} \sum_{d=1,2,\ldots,1615} GCM_{gvyjymd}/(1615 \cdot 365) \bigg| y = 1961, 1962, \ldots, 2003; \nonumber\nonumber
j = 1, 2, \ldots, 1615\big\}, \big\{\sum_{j=1,2,\ldots,1615} \sum_{m=1,2,\ldots,12} \sum_{d=1,2,\ldots,1615} OBS_{jy/md}/(1615 \cdot 365) \bigg| y = 1961, 1962, \ldots, 2003; \nonumber\nonumber
j = 1, 2, \ldots, 1615\big\} \biggr) \tag{2.12}
\]

for any \( g \in \{1, 2, \ldots, 7\}, v \in \{1, 2, 3\}, \) and \( c \in \{1, 2, \ldots, 10\}. \)
2.3 EVALUATION RESULTS

2.3.1 Spatial Heterogeneities

The results of distributed seasonal simulation accuracies ($DDS_{gyse}$) are presented in Figure 2.4. For any combinations of climate variables ($T_{min}$, $T_{max}$ and $Prec$), seasons ($DJF$, $MAM$, $JJA$ and $SON$), statistical metrics ($PC$, $KC$, …, and $RRSE$) and 10-km grids (1, 2, …, 1615) in the ARB, the GCM of which the simulation is of the highest agreement with observations is identified. For a climate variable in a 10-km grid and a season, the skill of a GCM at capturing its intraseasonal variability is quantified by the metrics of $PC$, $KC$ and $SC$, while other metrics focus on the absolute or relative magnitude of the climate variable. The significantly different patterns between (, a, ) to (, c, ) and (, d, ) to (, j, ) in Figure 2.4 demonstrates the difference between the roles of the two groups of metrics.

It is revealed that any GCM cannot be more accurate than any others in any case at the resolutions of seasons and 10 kms. Whether a GCM is of the highest simulation accuracy varies with seasons, climate variables, statistical metrics, and grid locations. For instance, for the season of $DJF$, ESM2G is the most capable at modeling the magnitudes of $T_{min}$ and $T_{max}$ in the upper and middle stream and of $Prec$ in the middle stream; CM5A-LR shows the highest accuracy of capturing the intraseasonal variability of $T_{min}$ and $T_{max}$ in the entire river basin, while in comparison with other GCMs or Ensemble, ESM-CHEM is more reliable in approaching the intraseasonal variability of daily precipitation; and ESM2M is the most accurate in simulating both the intraseasonal variability and the magnitudes of daily precipitation in the downstream. In comparison, for the season of $JJA$, the intraseasonal variability of $T_{min}$ is well simulated by Ensemble in the upstream and by ESM-CHEM in the downstream; the modeling of magnitudes of
Tmax in the upper, middle and downstreams should rely on MIROC5, Ensemble and CM5A-LR, respectively; and Ensemble is more reliable than others in reflecting the magnitude of Prec.

Based on an analysis of the GCMs of the most satisfactory performances under all combinations, a preliminary scheme of selecting alternative GCMs for the related impact studies in the ARB is designed for the convenience of practices (see Table 2.2). The ARB is divided as three regions, i.e. the upper, middle and downstreams. The intraseasonal variability and magnitudes of climate variables in four seasons and three regions are considered in this scheme. It should be noted that, in some cases such as precipitation magnitudes in the upstream, the suggested GCM is not more accurate than any others for all grids in this region. The GCM corresponding to the most grids on which it shows the highest accuracy is selected in these cases. In consideration of the overall performances in all cases, the sequence of the seven GCMs would be Ensemble, ESM2G, CM5A-LR, ESM-CHEM, ESM2M, MIROC5 and CM5A-MR corresponding to high to low simulation accuracies. If a single GCM is preferred, we would suggest the ensemble of the six selected CMIP5 GCMs; however, it should be noted that this ensemble is not competitive in many cases, e.g. magnitudes of the three climate variables in DJF, intraseasonal variability of Prec in SON, and both the intraseasonal variability and the magnitudes of Tmax in the downstream in JJA.

Shown in Figure 2.5 is the multi-year distributed simulation accuracies (DDO_{gye}), i.e. multi-year simulation accuracies of the selected six CMIP5 GCMs and their ensemble in reproducing Tmin, Tmax and Prec over 1615 10-km grids in the ARB. In consideration of the similarity among the ten statistical metrics as displayed in Figure 2.4, three
representative metrics, i.e. $SC$, $RMSE$ and $Nash$, are used. It is revealed that, at the spatial resolution of ten kilometres, ESM2G shows the highest accuracy in modeling the multi-year magnitudes of $Tmin$ compared with other $GCM$s. Due to the significant superiority of ESM2G in this aspect, the $Nash$ metrics of multi-year $Tmin$ simulation, equivalent with coupling of variability and magnitudes, is maximized by ESM2G over all grids. This reveals that ESM2G is the most accurate $GCM$ in reproducing the multi-year overall conditions of $Tmin$ at the spatial resolution of ten kilometers. Meanwhile, it is illustrated in Figure 2.5 that Ensemble is more accurate than other $GCM$s in reflecting the multi-year variability and the magnitudes of $Tmax$ and $Prec$ as well as the multi-year variability of $Tmin$ for every 10-km grids. Furthermore, the superiority of either Ensemble or ESM2G over other $GCM$s is significant for all grids. Accordingly, our suggestion on selection of $CMIP5$ $GCM$s for the multi-year spatial distributions of climate variables would be Ensemble for the variability and the magnitudes of $Tmax$ and $Prec$ and the variability of $Tmin$ and ESM2G for the magnitudes of $Tmin$. 
Table 2.3. Selection of Alternative CMIP5 GCMs for the ARB under Seasonal and 10-km Resolutions.

<table>
<thead>
<tr>
<th>Seasons</th>
<th>Climate Variables</th>
<th>Upstream</th>
<th>Middle Stream</th>
<th>Downstream</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Intraseasonal Variability</td>
<td>Magnitude</td>
<td>Intraseasonal Variability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DMF</td>
<td>DJF</td>
<td>MAM</td>
</tr>
<tr>
<td></td>
<td>Tmin</td>
<td>CM5A-LR</td>
<td>CM5A-LR</td>
<td>ESM2G</td>
</tr>
<tr>
<td></td>
<td>Tmax</td>
<td>CM5A-LR</td>
<td>CM5A-LR</td>
<td>ESM2G</td>
</tr>
<tr>
<td></td>
<td>Prec</td>
<td>ESM-CHEM*</td>
<td>ESM-CHEM*</td>
<td>ESM2G**</td>
</tr>
<tr>
<td></td>
<td>Tmin</td>
<td>Ensemble</td>
<td>ESM2G**</td>
<td>Ensemble*</td>
</tr>
<tr>
<td></td>
<td>Tmax</td>
<td>Ensemble</td>
<td>ESM2G**</td>
<td>Ensemble*</td>
</tr>
<tr>
<td></td>
<td>Prec</td>
<td>ESM2M*</td>
<td>ESM2M*</td>
<td>ESM2G*</td>
</tr>
<tr>
<td></td>
<td>Tmin</td>
<td>ESM-CHEM*</td>
<td>ESM-CHEM*</td>
<td>ESM2G*</td>
</tr>
<tr>
<td></td>
<td>Tmax</td>
<td>Ensemble</td>
<td>ESM2M*</td>
<td>Ensemble*</td>
</tr>
<tr>
<td></td>
<td>Prec</td>
<td>MIROC5**</td>
<td>MIROC5**</td>
<td>ESM2M**</td>
</tr>
</tbody>
</table>

Note: *the other GCMs may be more accurate than the selected GCM for a small part of this region; **the other GCMs may be more accurate than the selected GCM for a large part of this region.
Figure 2.4. CMIP5 GCMs of the highest accuracies in simulating seasonal spatial distributions of climate variables.
Figure 2.5. CMIP5-GCM Simulation Accuracies: Multi-Year Spatial Distributions.
2.3.2 Annual Cycles

The annual cycle of spatial-variability simulation accuracies \( (DCD_{gymdc}) \) is illustrated in Figure 2.6. These indices are proposed to quantify the multi-dimensional performances of a \( GCM \) in reproducing the multi-year-averaged magnitude and pattern of a spatially various climate variable in any one of 365 days. For any \( v \in \{1, 2, 3\} \), \( m \in \{1, 2, \ldots, 12\} \), \( d \in \{1, 2, \ldots, D_m\} \), and \( c \in \{1, 2, \ldots, 10\} \), the \( GCM \) of the highest accuracy is identified from seven alternatives and displayed in a particular color in the figure. In consideration of all statistical metrics and climate variables, any single \( GCM \) fails to be of the highest accuracy in modeling the multi-year-averaged spatial variability for every day in the annual cycle. Corresponding to the twelve months, the \( GCMs \) showing the highest accuracies for the most combinations of statistical metrics, climate variables, and days are ESM2G, ESM2G, MIROC5, ESM-CHEM, CM5A-LR, CM5A-LR, CM5A-LR, CM5A-LR, CM5A-LR, CM5A-MR, CM5A-LR, and ESM2G, respectively. For any particular combination of statistical metrics and climate variables, the intramonth differences of the skills of the selected \( CMIP5 \) \( GCMs \) at reflecting multi-year averaged spatial variability are relatively insignificant for January, February, May, June, September and December and significant for other months.

The superiority of \( GCMs \) varies with statistical metrics and periods for \( Tmin \) and \( Tmax \). For the overall magnitude of spatially various \( Tmin \), the annual cycle can be divided as six periods: late November to early March, middle March to middle April, late April, early May to middle July, late July to late August, and early September to middle November. The corresponding \( GCMs \) of the highest accuracies are ESM2G, MIROC5, ESM-CHEM, Ensemble, CM5A-MR, and CM5A-LR, respectively. For the pattern of
spatially various $T_{min}$, the division is changed to: late October to early April, middle April to late May, early June to early July, late July to middle August, late August to middle September, and late September to middle October. The corresponding GCMs are MIROC5, ESM-CHEM, ESM2M, ESM2G, CM5A-MR, and CM5A-LR. The selection of GCMs for $T_{max}$ is identical with $T_{min}$ in most cases except that: ESM-CHEM, CM5A-LR and CM5A-MR are desired for modeling magnitudes of spatially various $T_{max}$ in the periods of late March to late April, late July to middle August, and middle October to late October, respectively, while ESM2G and CM5A-LR should be relied on to reproduce the spatial pattern of $T_{max}$ in the periods of early June to early July and late August to middle September, respectively. In addition, any GCM is not significantly more accurate than any others in modeling either the pattern or the magnitude of spatially various $Prec$ within a continuous period.

It is also illustrated in Figure 2.6 that the ensemble of multiple GCMs is not more accurate than any GCMs in the task of reproducing spatial pattern and magnitudes of climate variables at the daily scale. Taking the above analyses into account, our suggestions on selecting CMIP5 GCMs for this task are: not neglecting the variation of the desired GCM with climate variables, periods and statistical metrics, combining CM5A-LR, ESM2G, MIROC5 and CM5A-MR according to their performances in different periods, and using CM5A-LR if only one GCM is allowed.

The results of the annual cycle of spatial-means simulation accuracies are displayed in Figure 2.7. These indices are proposed to quantify the multi-dimensional performances of a GCM in reproducing both magnitude and variability of grids-averaged yearly-various climate variables in any one of 365 days. For any $v \in \{1, 2, 3\}$, $m \in \{1, 2, \ldots, 12\}$, $d \in \{1,$
2, ..., \(D_m\), and \(c \in \{1, 2, ..., 10\}\), the GCM of the highest accuracy is identified from seven alternatives and displayed in a particular color in the figure. It is illustrated that the multi-model ensemble shows the highest accuracy in modeling the magnitude of grids-averaged yearly-various climate variables in most days. In approximately ninety days concentrating in November, December, January, March and May, ESM2G is more accurate than Ensemble, especially for \(T_{min}\). In addition, any GCM is not significantly more accurate than any others in modeling the yearly variability of grids-averaged climate variables at the daily scale within a continuous period. Thus, our suggestions on selecting GCMs for estimating the annual cycle of grids-averaged yearly-various climate variables at the daily scale are: prioritizing the multi-model ensemble and, if a higher accuracy is required, comparing the simulation results of Ensemble and ESM2G in November, December, January, March and May.
Figure 2.6. *CMIP5-GCM* Simulation Accuracies: Annual Cycle of Spatial Variability.
Figure 2.7. CMIP5-GCM Simulation Accuracies: Annual Cycle of Spatial Means.
2.3.3 Monthly and Seasonal Spatial Variability

To evaluate the multi-aspect performances of the selected CMIP5 GCMs in modeling the multi-year-averaged pattern and magnitude of spatially various climate variables at the monthly and seasonal scales, the monthly spatial-variability simulation accuracies (DCM$_{gvmc}$) and the seasonal spatial-variability simulation accuracies (DCS$_{gsyc}$) are calculated. The related results are shown in Figure 2.8. It is revealed that, in comparison with Prec, the selected CMIP5 GCMs are more accurate at modelling the multi-year-averaged patterns of Tmin and Tmax. Due to the relatively low magnitude of Prec, it seems that the selected CMIP5 GCMs are of higher accuracies at simulating the multi-year-averaged absolute magnitudes of Prec than that of Tmax and Tmin. After removal of the difference of the original magnitudes of Prec, Tmin and Tmax, however, the overall accuracies of GCMs at reproducing multi-year-averaged relative magnitudes are maximized for Tmax and minimized for Prec.

The multi-year-averaged patterns, absolute magnitudes, and relative magnitudes are translated as three groups of metrics, i.e. KC, PC and SC, MAE and RMSE, and the others, respectively. As a result, there are significant differences among the estimated monthly or seasonal spatial-variability simulation accuracies by the three groups of metrics. As shown in Figure 2.8(a), (f) and (j), the accuracy of GCMs in reproducing the multi-year-averaged spatial patterns varies with climate variables, months and seasons, in which MIROC5, CM5A-LR, Ensemble and CM5A-MR show significant advantages. For Tmax and Tmin, the simulation accuracies of GCMs are decreased from DJF, SON, MAM to JJA, and the corresponding GCMs of the best performances are MIROC5, CM5A-LR, CM5A-LR, and ESM2G, respectively; for Prec, the sequence of the four seasons is changed to MAM, JJA,
SON and DJF, and the corresponding GCMs are Ensemble, CM5A-LR, CM5A-LR, and CM5A-MR, respectively.

As for simulation of the multi-year-averaged magnitudes of climate variables, CM5A-LR, ESM2G, Ensemble and CM5A-MR show relatively high accuracies. In terms of the absolute magnitudes, the overall accuracies of GCMs are decreased from SON, MAM, JJA to DJF for Tmax, from JJA, SON, MAM to DJF for Tmin, and from SON, DJF, MAM to JJA for Prec; correspondingly, the GCMs of the highest accuracies in the sequenced seasons are CM5A-LR, CM5A-LR, CM5A-LR and ESM2G for Tmax, CM5A-LR, CM5A-LR, ESM2G and ESM2G for Tmin, and CM5A-LR, ESM2G, ESM2G and CM5A-MR for Prec. By contrast, the overall accuracies of GCMs in simulating the relative magnitudes are decreased from MAM, SON, JJA to DJF for Tmax, from MAM, SON, JJA to DJF for Tmin, and from SON, JJA, MAM to DJF for Prec; correspondingly, the GCMs of the highest accuracies are Ensemble, CM5A-LR, Ensemble and ESM2G for Tmax, ESM2G, CM5A-LR, CM5A-LR and ESM2G for Tmin, and CM5A-LR, CM5A-LR, ESM2G and ESM2G for Prec.

As a summary of the analyses in the two paragraphs above, CM5A-LR, Ensemble and CM5A-MR are advantageous at reproducing the multi-year-averaged patterns and magnitudes of spatially various climate variables at the monthly and seasonal scales, and MIROC5 and ESM2G are only suitable for modeling the patterns and the magnitudes, respectively.
Figure 2.8. CMIP5-GCM Simulation Accuracies: Monthly and Seasonal Spatial Variability.
2.3.4 Monthly and Seasonal Spatial Means

Figure 2.9 illustrates the monthly spatial-means simulation accuracies \((DMM_{gvmc})\) and the seasonal spatial-means simulation accuracies \((DMS_{gvsc})\) in the ARB. The performances of GCMs in modeling variability, absolute magnitudes, and relative magnitudes of grids-averaged climate variables at the monthly or seasonal scale are quantified by the metrics of \(KC\), \(PC\) and \(SC\), \(MAE\) and \(RMSE\), and the others, respectively. It is revealed that the selected CMIP5 GCMs are more accurate at modelling the grids-averaged monthly or seasonal variability of \(T_{min}\) and \(T_{max}\) than that of \(Prec\). In comparison with \(T_{max}\) and \(T_{min}\), the relative magnitude of grids-averaged \(Prec\) is more difficult for the selected CMIP5 GCMs to model, while the absolute magnitude is reproduced at higher accuracies.

Taking the variation of the performances of the selected CMIP5 GCMs with statistical metrics, months, seasons and climate variables into account, ESM2G and Ensemble are more competitive than others in modeling the monthly or seasonal variability and magnitudes of grids-averaged \(T_{min}\), \(T_{max}\) and \(Prec\). For instance, in terms of reproducing seasonal variability from \(DJF\), \(MAM\), \(JJA\) to \(SON\), the GCMs of the highest accuracies are ESM-CHEM, ESM2G, ESM2M and MIROC5 for \(Prec\), CM5A-LR, ESM2G, Ensemble and ESM2G for \(T_{min}\), and CM5A-LR, ESM2G, ESM2M and ESM2G for \(T_{max}\), respectively. As for simulation of the seasonal absolute or relative magnitudes, ESM2G is the most accurate for \(Prec\) in \(DJF\) and \(MAM\), \(T_{min}\) in \(DJF\), \(MAM\) and \(SON\), and \(T_{max}\) in \(DJF\), while Ensemble shows the highest accuracy for \(Prec\) in \(JJA\) and \(SON\) and for \(T_{max}\) in \(MAM\), \(JJA\) and \(SON\).
Figure 2.9. *CMIP5-GCM* Simulation Accuracies: Monthly and Seasonal Spatial Means.
2.3.5 Multi-Year Spatial Variability and Means

The multi-year spatial-variability simulation accuracies ($DCS_{gvc}$) and the multi-year spatial-means simulation accuracies ($DMS_{gvc}$) are illustrated in Figure 2.10. The indices of $DCS_{gvc}$ are used to quantify the skills of the selected CMIP5 GCMs at reproducing the spatial variability and magnitude of multi-year-averaged $T_{max}$, $T_{min}$ and $Prec$, while $DMS_{gvc}$ focus on the inter-annual variability and magnitude of grids-averaged annual means of $T_{max}$, $T_{min}$ and $Prec$. The performances of GCMs vary with climate variables and statistical metrics.

In terms of the spatial variability of multi-year-averaged climate variables, the overall simulation accuracies of GCMs are decreased from $T_{min}$, $T_{max}$ to $Prec$; for the three climate variables, the GCMs of the highest accuracies are MIROC5, MIROC5, and Ensemble, respectively. The abilities of GCMs at modeling the absolute magnitudes of multi-year-averaged spatially-distributed climate variables increase from $T_{min}$, $T_{max}$ to $Prec$; the corresponding GCMs of relative high accuracies are ESM2G, CM5A-LR, and CM5A-MR, respectively. As for the relative magnitudes of multi-year-averaged spatially-distributed climate variables, the accuracies of GCMs are the highest for $T_{max}$ and the lowest for $Prec$; the GCMs of the highest accuracies for $T_{max}$, $T_{min}$ and $Prec$ are CM5A-LR, ESM2G, and CM5A-MR. Therefore, for the spatially various multi-year averages of $T_{min}$, $T_{max}$ and $Prec$, the desired GCMs are MIROC5, MIROC5 and Ensemble for their spatial variability and ESM2G, CM5A-LR and CM5A-MR for their magnitudes, respectively.

For the simulation of the inter-annual variability of grids-averaged annual means of $T_{max}$, $T_{min}$ and $Prec$, the overall accuracies of GCMs are decreased from $T_{max}$, $T_{min}$ to
Prec for which the most accurate GCMs are ESM2G, Ensemble, and Ensemble, respectively. For the absolute magnitudes of grids-averaged annual means, Prec is modeled at relatively high accuracies and the simulation deviations are the highest for Tmin; the multi-model ensemble is more accurate than any others for the three climate variables. Corresponding to Tmax, Tmin and Prec, the overall accuracies of the selected CMIP5 GCMs at simulating the relative magnitudes of their grids-averaged annual means are gradually decreased; the GCM of the highest accuracy for any one of the three climate variables is the multi-model ensemble. It is revealed that, excluding the inter-annual variability of grids-averaged annual means of Tmax which is well reproduced by ESM2G, the multi-model ensemble shows significant advantages of modeling the inter-annual variability and magnitudes of grids-averaged annual means of the three climate variables.
Figure 2.10. *CMIP5-GCM* Simulation Accuracies: Multi-Year Spatial Variability and Means.
2.4 DISCUSSIONS

2.4.1 Scaling Effects

In this study, the simulation accuracies of the selected CMIP5 GCMs for three climate variables in the ARB are evaluated at multiple spatial and temporal scales. As revealed by a comparison of the evaluation results, there are scaling effects for the performances of CMIP5 GCMs in modeling climate conditions of the ARB in some cases. A GCM of the highest accuracies at a higher spatial or temporal scale may not be the GCM which is more accurate than any other GCMs at a coarser scale, and vice versa. This mainly occurs when the most accurate GCMs are significantly heterogeneous at the finer scale. For instance, according to the KCs of multi-year-averaged spatially-distributed simulation and observations of Tmin at the monthly or seasonal scale, the GCMs of the highest accuracies in DJF, December, January and February are CM5A-MR, MIROC5, ESM2G and MIROC5, respectively; although MIROC5 is more reliable than others including CM5A-LR for two months in DJF, the desired GCM for DJF is CM5A-LR rather than MIROC5. The scaling effects also exist in the process of averaging all 10-km grids in the ARB. For example, as shown in Figure 2.4(a, d, c), the number of grids that are mainly in the downstream of the ARB and in which ESM2G is of the highest MAEs is the most. CM5A-LR is the most accurate GCM for some grids in the upper and middle stream, of which the number is less than that of ESM2G. It can be concluded that the overall accuracy of ESM2G in modeling Tmin in JJA over all grids is higher than other GCMs including CM5A-LR. On the contrary, it is displayed in Figure 2.9(b) that the GCM of the highest accuracy, i.e. the lowest value of MAE, in simulating grids-averaged Tmin in JJA is CM5A-LR rather than ESM2G. Meanwhile, as shown in Figure 2.4(a, d, a), ESM2G is
the most accurate for almost all grids except a few ones in the downstream in modeling the absolute magnitude of distributed $T_{min}$ in DJF; it is illustrated in Figure 9(b) that the GCM of the highest accuracy in modeling the grids-averaged $T_{min}$ in DJF is also ESM2G; a comparison with the example stated above reveals that the occurrence of the scaling effects is associated with the heterogeneity of the performances of the selected CMIP5 GCMs at the finer scale.

### 2.4.2 Similarity and Dissimilarity of Metrics

In this study, ten statistical metrics are employed to quantify the accuracies of the selected CMIP5 GCMs in reproducing climate conditions in the ARB at various temporal and spatial scales. Based on these metrics, ten groups of simulation-accuracy indices are proposed. A comparison of these metrics in every group of indices reveals that there are significant similarity and dissimilarity among these metrics. The metrics of $KC$, $PC$ and $SC$ focus on reflecting the spatial or temporal correlation of climate-variable observations and simulation; a representative difference among them is that $SC$ is able to quantify the nonlinear correlation which is hardly achievable for $KC$ or $PC$. For any climate variable at a particular combination of spatial and temporal scales, the accuracy of a GCM at modeling absolute magnitudes can be quantified by the metrics of $MAE$ and $RMSE$, while quantification of the relative-magnitude modeling accuracy can be relied on the metrics of $Nash$, $NRMSE$, $SI$, $RAE$ or $RRSE$. Furthermore, there are significant differences among the results of the three groups of metrics. A relatively high accuracy of a CMIP5 GCM in reproducing one aspect of variability, absolute magnitude or relative magnitude of a climate variable cannot guarantee the same accuracy in another aspect. For example, as shown in Figure 2.4, ESM2G shows the highest accuracy in modeling either the relative
or the absolute magnitudes of $T_{min}$ in $DJF$ over almost all grids; however, its accuracy in modeling the seasonal variability of $T_{min}$ in $DJF$ is less than CM5A-LR. This implies that selection of multiple metrics in quantifying the multi-aspect performances of $CMIP5$ $GCM$s is highly desired. In consideration of the similarity and dissimilarity of the selected metrics, the statistical metrics that we suggest to be used for $GCM$ evaluations are $SC$, $RMSE$ and $Nash$.

2.4.3 Concluding Remarks

Through this study, a series of findings related with the multi-dimensional performances of six $CMIP5$ $GCM$s and their ensemble in reproducing $T_{min}$, $T_{max}$ and $Prec$ in the $ARB$ at various temporal and spatial scales are revealed. These findings are helpful for guiding researchers or practisers to select the desired $CMIP5$ $GCM$ for climate-change impact studies in the $ARB$ or neighbouring regions. For the convenience of readers, these findings are specified as follows. To sum up, the most reasonable scheme of selecting $CMIP5$ $GCM$s for impact studies in the $ARB$ is to integrate them into a general framework through taking the variation of their accuracies with climate variables, geographical locations, temporal windows, statistical metrics, and temporal and spatial scales into account. The overall simulation accuracy of $CMIP5$ $GCM$s for climatic conditions in the $ARB$ can be optimized by this scheme. If only one set of climate simulation is demanded due to reasons such as limitation of computational capacities, the multi-model ensemble is recommended. If only one $CMIP5$ $GCM$ is preferred in some cases, ESM2G would be the desired one.

(a) For the simulation of intraseasonal variability or magnitudes of $T_{min}$, $T_{max}$ or $Prec$ in the $ARB$ at finer spatial resolutions such as ten kilometres, any $GCM$ can barely
be more accurate than the others in all cases; taking all cases into account, the overall simulation accuracies of Ensemble, ESM2G, CM5A-LR, ESM-CHEM, ESM2M, MIROC5 and CM5A-MR are gradually decreased; if a single GCM is preferred, the multi-model ensemble is suggested although its accuracy is not the highest in many cases such as the intraseasonal variability of Prec over all grids in SON.

(b) At the spatial resolution of ten kilometres, ESM2G shows the highest accuracy in modeling the multi-year-averaged magnitudes of Tmin, while Ensemble is more accurate than others in reflecting the inter-yearly variability and the multi-year-averaged magnitudes of Tmax and Prec as well as the inter-yearly variability of Tmin for every 10-km grids.

(c) For the simulation of the multi-year-averaged spatial variability of Tmin, Tmax and Prec in the annual cycle, any single GCM including the multi-model ensemble fails to be of the highest accuracy in every day; a combination of CM5A-LR, ESM2G, MIROC5 and CM5A-MR according to their performances under different periods, statistical metrics and climate variables is recommended; and CM5A-LR is the most desired if only one GCM is allowed.

(d) The multi-model ensemble shows the highest accuracy in modeling the magnitude of grids-averaged yearly-various climate variables in most days; ESM2G is more accurate than Ensemble, especially for Tmin, in approximately ninety days; any GCM is not significantly more accurate than any others in modeling the yearly variability of grids-averaged climate variables at the daily scale within a continuous period.

(e) In the aspect of modelling the multi-year-averaged spatial patterns of Tmin and Tmax, the accuracies of GCMs are higher than that of Prec; in comparison with Tmax and
$T_{min}$, the accuracy of GCMs at simulating the multi-year-averaged absolute magnitudes of $Prec$ is increased; the overall accuracies of GCMs for the multi-year-averaged relative magnitudes are maximized for $T_{max}$ and minimized for $Prec$; CM5A-LR, Ensemble and CM5A-MR are advantageous at reproducing the multi-year-averaged spatial patterns and magnitudes of spatially various climate variables at the monthly and seasonal scales, and MIROC5 and ESM2G are only suitable for modeling the patterns and the magnitudes, respectively.

(f) The selected CMIP5 GCMs are more accurate at modeling the grids-averaged monthly or seasonal variability of $T_{min}$ and $T_{max}$ than that of $Prec$; in comparison with $T_{max}$ and $T_{min}$, the relative magnitude of grids-averaged $Prec$ is more difficult for the selected CMIP5 GCMs to model, while the absolute magnitude is reproduced at higher accuracies; ESM2G and Ensemble are more competitive than others in modeling the monthly or seasonal variability and magnitudes of grids-averaged $T_{min}$, $T_{max}$ and $Prec$.

(g) For the spatially various multi-year averages of $T_{min}$, $T_{max}$ and $Prec$, the desired GCMs are MIROC5, MIROC5 and Ensemble for their spatial variability, respectively, and ESM2G, CM5A-LR and CM5A-MR for their magnitudes, respectively; excluding the inter-annual variability of grids-averaged annual means of $T_{max}$ which is well reproduced by ESM2G, the multi-model ensemble shows significant advantages of modeling the inter-annual variability and magnitudes of grids-averaged annual means of the three climate variables.

(h) A CMIP5 GCM of the highest accuracies in modeling climate conditions of the ARB at a higher spatial or temporal scale may not be the GCM which is more accurate
than any other GCMs at a coarser scale, and vice versa, which mainly occur when the most accurate GCMs are significantly heterogeneous at the finer scale.

(i) In consideration of the similarity and dissimilarity of the selected metrics, the statistical metrics that we suggest to be used for CMIP5 GCM evaluations are SC, RMSE and Nash.

2.4.4 Potential Extensions

This study is the first attempt to comprehensively evaluate the multi-dimensional performances of CMIP5 GCMs at reproducing climatic conditions in the ARB at various temporal and spatial scales. In addition to a suite of findings as discussed above, the framework of CMIP5 GCM evaluation can be improved in some aspects and provides scientific support for climate-change impact studies in other regions. A few of representative examples are discussed as follows. The identification of relatively accurate CMIP5 GCMs for the ARB can be directly used for a variety of impact or adaptation studies, e.g. streamflow simulation, energy availability analysis, and air quality control, to facilitate socio-economic development and eco-environmental conservation. A total of thirty outputs of CMIP5 GCMs are selected as representative ones due to their extensive usage in various existing continental impact studies; accordingly, six CMIP5 GCMs are selected for evaluation, while the others are challenged by the existence of missing data for the representative outputs under at least one emission scenarios. If only a part of the selected CMIP5 GCM outputs or other outputs are preferred, the results of the screening of CMIP5 GCMs based on analyses of data availabilities may be different, which may lead to the change of CMIP5 GCM evaluation results. The evaluation of CMIP5 GCMs in this study focuses on the averaged performances of these GCMs at modeling climate
conditions in the ARB under various scales. An extension of this evaluation to extreme climatic conditions which are desired for some particular impact studies such as flooding control and drought elimination is helpful for enhancing the applicability of this study.

The principal objective of this study is to evaluate the performances of CMIP5 GCMs in reproducing historical climatic conditions in the ARB; analyses of climatic conditions projected by the selected CMIP5 GCMs under different emission scenarios as well as of the impacts of various CMIP5 GCM selection schemes on future climatic projection results, which are not involved in this study, deserve further efforts. The developed framework can be applied to the evaluation of CMIP5 GCMs for diverse impact studies in other river basins or regions worldwide, providing specified suggestions on selection of CMIP5 GCMs according to particular requirements of climatic conditions.

2.5 SUMMARY

In this study, the accuracies of six selected CMIP5 GCMs and their ensemble at reproducing the historical (from 1961 to 2003) climatic conditions (i.e. Tmin, Tmax and Prec) over the ARB, Canada under various temporal and spatial scales were evaluated based on ten statistical metrics. A series of analyses and comparisons facilitated revealing a few of findings as specified in section 2.4 for readers’ convenience. As a short summary of these findings, the most reasonable scheme of selecting CMIP5 GCMs for impact studies in the ARB was to integrate them into a general framework because their accuracies varied with climate variables, geographical locations, temporal windows, statistical metrics, and temporal and spatial scales. If only one set of climate simulation was demanded, the multi-model ensemble was recommended. If only one CMIP5 GCM was
preferred in some cases, ESM2G would be the desired one. In addition, a \textit{CMIP5 GCM} of the highest accuracies in modeling climate conditions of the \textit{ARB} at a higher spatial or temporal scale might not be the \textit{CMIP5 GCM} which was the most accurate at a coarser scale, and vice versa. This scaling effect mainly occurred when the most accurate \textit{CMIP5 GCMs} were significantly heterogeneous at the finer scale. In consideration of the similarity and dissimilarity of the selected statistical metrics, \textit{SC}, \textit{RMSE} and \textit{Nash} were recommended for evaluation of \textit{CMIP5 GCMs} in multiple aspects such as correlations and magnitudes. These findings would be much helpful for guiding researchers or practisers to select the desired \textit{CMIP5 GCM} for climate-change impact or adaptation studies in the \textit{ARB} or neighbouring regions.

In the future, the framework of \textit{CMIP5 GCM} evaluation employed in this study can be improved in some aspects and provides scientific support for climate-change impact or adaptation studies in other regions. For example, analyses of climatic conditions projected by the selected \textit{CMIP5 GCMs} under different emission scenarios as well as of the impacts of various \textit{CMIP5 GCM} selection schemes on future climatic projection results deserve further efforts. The identification of relatively accurate \textit{CMIP5 GCMs} for the \textit{ARB} can be used for facilitating socio-economic development and eco-environmental conservation. An extension of this evaluation to extreme climatic conditions which are desired for some particular impact studies such as flooding control and drought elimination is helpful for enhancing the applicability of this study. The developed framework can be applied to the evaluation of \textit{CMIP5 GCMs} for diverse impact studies in other river basins or regions.
worldwide, providing specified suggestions on selection of CMIP5 GCMs according to particular requirements of climatic conditions.
CHAPTER 3
DEVELOPMENT OF A RECURSIVE DISSIMILARITY AND SIMILARITY INFERENTIAL CLIMATE CLASSIFICATION APPROACH

3.1 BACKGROUND

Classifications are an important tool in both general and applied climatology, since often a strict assignment of individual objects to disjunctive groups is necessary or at least favorable for achieving clearly structured results from complex data sets (Jacobeit, 2010). Classifications have had a long history in meteorology and climatology and can be generalized as a task of grouping entities (cases) so that they share similar features within each group, while being dissimilar between groups (Huth et al., 2008). To achieve this task in many fields of the meteorological or climatological sciences for a large spectrum of purposes, various classification methods were developed. Representative ones consisted of but were not limited to the Köppen-Geiger approach (Köppen, 1923), the principal component analysis (Bartzokas and Metaxas, 1996; Huth, 2000; Jacobeit et al., 2003), the Hess-Brezowsky catalog (Baur et al., 1944; Hess and Brezowsky, 1952), the Vangengeim-Girs method (Vangengeim, 1935), the k-means cluster analysis (Brinkmann, 1999; Esteban et al., 2006; Stahl et al., 2006), self-organizing maps (Cassano et al., 2006; Reusch et al., 2007), the Lamb (Lamb, 1972), and fuzzy-rules-based classification (Bardossy et al., 1995; Stehlik and Bardossy, 2003). Comprehensive reviews of existing methods were presented in a few of publications (e.g., Huth et al., 2008; Jacobeit, 2010).

However, climate classifications are associated with a few of issues. For instance, the climate-variable observations collected for classification may be of uncertainties due to causes such as instrumental or operational errors. Multiple climate variables selected to
represent regional climate may be dependent with each other. A common prerequisite for most statistical climate classification methods is that the climate-variable samples should come from a normally distributed population. The most desired result of climate classification is that climatic conditions are significantly different for any two climate classes and are not of significant differences in the same climate class. Besides, many existing climate classification methods rely on subjective judgement to support screening of climate variables, selection of classification thresholds, and setting of climate class sizes or numbers prior to classification practices. These issues are challenging the effectiveness of existing climate classification methods, the reliability of climate classification results, and the reasonability of climate-change impact studies or the other related ones (Cheng et al., 2015a,b,c, 2016a,b; Dong et al., 2011, 2012, 2013, 2014a,b,c, 2015).

In this study, we aim to develop a recursive dissimilarity and similarity inferential climate classification (ReDSICC) approach, providing an additional alternative for enabling effective classification of climate under these challenges. Specifically, the Shapiro-Wilk test (Shapiro and Wilk, 1965; Royston, 1995) will be incorporated into the framework of ReDSICC to evaluate the normality of multi-dimensional samples based on a collection of the historical observations of representative climate variables. For the samples of a climate variable that do not obey a normal distribution, a discrete distribution transformation approach will be developed to enable reversible transformation between the original distribution and a normal distribution. This will build a solid foundation for the subsequent statistical inferences for which the normality of samples is required. A framework of recursive dissimilarity and similarity inferences will be designed to identify
the desired classification of regional climate. In this framework, the modified Nel and van der Merwe test (Krishnamoorthy and Yu, 2004) will be employed to quantify the significance of the dissimilarity or similarity of multi-dimensional climate-variables samples in any two climate classes. For the features of obtained climate classes, the gaps between them due to discreteness of collected observations will be filled by extending their boundaries to achieve systematic classification of regional climate, and various characterization indicators will be proposed to present their rich properties. Through these efforts, the multi-dimensional samples of selected climate variables will be discretized as a series of groups that are internally inseparable, externally divergent, mutually exclusive, and collectively exhaustive. To verify methodological effectiveness and facilitate studies such as downscaling and hydrological simulation, the ReDSICC approach will be applied to a case study of climate classification in the Athabasca River Basin, Canada. A few of findings regarding the regional climate in this river basin or the mechanism of ReDSICC will be revealed from a series of comparisons.

3.2 METHODOLOGY DEVELOPMENT

3.2.1 Normality Analyses

Let the climate variables on which classification is based, e.g. daily minimum temperature, daily maximum temperature and daily cumulative precipitation in this study, be denoted as \( x_j \) \((j = 1, 2, \ldots, n)\) where \( n \) is the number of climate variables. They compromise a vector \( X = (x_1, x_2, \ldots, x_n) \). The data series of these variables can be expressed as \( \{X_t\}_{t=1}^T = \{X_1, X_2, \ldots, X_T\} \) where \( t \) indicates a temporal unit and \( T \) is the temporal duration. For any \( t \in \{1, \ldots, T\} \), we have \( X_t \) is an \( n \)-dimensional vector \((x_{t1}, x_{t2}, \ldots, x_{tn})\).
Due to reliance on statistical inferences, the ReDSICC method requires one prerequisite that any climate variable should obey a normal distribution. However, this prerequisite may not hold in reality, e.g. daily precipitation tending to obey a Gamma distribution (McMahon and Srikanthan, 1981). To avoid incorrectness of the classification result resulting from dissatisfaction of the normal-distribution prerequisite, any climate variable is tested in terms of normality. Approximately forty methods are available for normality tests (e.g., Kolmogorov, 1933; Massey Jr, 1951; Lilliefors, 1967; Dufour et al., 1998; Oztuna et al., 2006; Peat and Barton, 2005). In this study, the Shapiro-Wilk (abbreviated as SW) test (Shapiro and Wilk, 1965; Royston, 1995) is employed because it has a competitive power performance and can be adopted for both symmetric non-normal and asymmetric distributions (Royston, 1995; Razali and Wah, 2011).

If the normality condition does not hold, a transformation from an abnormal distribution to a normal distribution is required to guarantee the reliability of the ReDSICC results. The process must be reversible so that the original values of the climate variables can be restored, avoiding distortion of the similarity and dissimilarity among samples of climate variables. To achieve this, a discrete distribution transformation (DDT) approach is proposed based on convertibility of any distribution to the 0-1 uniform distribution. The details of the DDT approach are presented below.

Let the samples of the climate variable of which the distribution has to be transformed as a normal distribution be denoted as \( \{x_t\}_{t=1}^{T} \) where: \( t \) indicates a temporal unit and \( T \) is the temporal duration. For any sample \( x_t \) (\( t \in \{1, 2, \ldots, T\} \)), \( F(x_t) \) is the proportion of samples for which values are less than or equal to \( x_t \). For instance,
\[ F(x_{\min}) \approx 0 \quad (3.1) \]

and

\[ F(x_{\max}) = 1 \quad (3.2) \]

where

\[ x_{\min} = \min(\{x_t\}_{t=1}^T) \quad (3.3) \]

and

\[ x_{\max} = \max(\{x_t\}_{t=1}^T). \quad (3.4) \]

We have:

\[ F(x_t) \in [0, 1] \text{ for any } t \in \{1, 2, \ldots, T\} \quad (3.5) \]

and

\[ F(x_{t_1}) < F(x_{t_2}) \text{ if } x_{t_1} < x_{t_2}. \quad (3.6) \]

Let \( F^{-1}(\cdot) \) be the inverse function of \( F(\cdot) \) and \( \alpha \) be any cumulative frequency. We have \( F^{-1}(\alpha) \) is the threshold under which samples account for \( \alpha \) of sample size \( T \). If \( \alpha_1 < \alpha_2 \), then

\[ F^{-1}(\alpha_1) \leq F^{-1}(\alpha_2). \quad (3.7) \]

We denote \( F(x_t) \) as \( \alpha_t \) for any \( t \in \{1, 2, \ldots, T\} \). It can be proved that \( \{\alpha_t\}_{t=1}^T \) obeys the 0-1 uniform distribution.
The target is normally distributed samples through transformation of the original samples \( \{x_t\}_{t=1}^T \), which can be denoted as \( \{x_t^*\}_{t=1}^T \). Suppose that \( \{x_t\}_{t=1}^T \) and \( \{x_t^*\}_{t=1}^T \) are consistent in their mean and their standard deviation. Namely,

\[
\text{mean}(\{x_t^*\}_{t=1}^T) = \text{mean}(\{x_t\}_{t=1}^T) \tag{3.8}
\]

and

\[
\text{sdv}(\{x_t^*\}_{t=1}^T) = \text{sdv}(\{x_t\}_{t=1}^T) \tag{3.9}
\]

where \( \text{mean()} \) and \( \text{sdv()} \) are the operators of mean values and standard deviations, respectively. Let the normal distribution of which the mean is \( \mu \) and the standard deviation is \( \sigma \) be expressed as \( N(\mu, \sigma^2) \) where

\[
\mu = \text{mean}(\{x_t\}_{t=1}^T) \tag{3.10}
\]

and

\[
\sigma = \text{sdv}(\{x_t\}_{t=1}^T). \tag{3.11}
\]

The cumulative distribution function of \( N(\mu, \sigma^2) \), denoted as \( \Psi(x) \), is

\[
\Psi(x) = \left( \int_{-\infty}^{x} \exp\left(-((y - \mu)/\sigma)^2/2\right) dy \right)/2\pi \tag{3.12}
\]

which can be approximated as:

\[
\Psi(x) \approx 0.5 + 0.5 \cdot \text{sgn}(x - \mu)/(2^{0.5} \cdot \sigma)(1 - \exp(-(x - \mu)/(2^{0.5} \cdot \sigma)))(4/\pi + a((x - \mu)/(2^{0.5} \cdot \sigma))^2/(1 + a((x - \mu)/(2^{0.5} \cdot \sigma))^2))^{0.5} \tag{3.13}
\]
(Vazquez-Leal et al., 2012) where $a \approx 0.140012$ and $\text{sgn}(\cdot)$ and $\exp(\cdot)$ are the sign function and the exponential function, respectively. The inverse function of $\psi(x)$, i.e. the normal quantile function, denoted as $\psi^{-1}(\alpha)$, can be approximated as:

$$
\psi^{-1}(\alpha) \approx \mu + 2^{0.5} \sigma \text{sgn}(2\alpha - 1)(((2/(a \cdot \pi)) + (\ln(1 - (2\alpha - 1)^2))/2)^2 - (\ln(1 - (2\alpha - 1)^2))/a)^{0.5} - (2/(a \cdot \pi)) + (\ln(1 - (2\alpha - 1)^2))/2)^{0.5} \quad (3.14)
$$

(Vazquez-Leal et al., 2012) where $\ln(\cdot)$ is the natural logarithm function. It is verified in (Vazquez-Leal et al., 2012) that the functions $\psi(x)$ and $\psi^{-1}(\alpha)$ are reversible. We denote $\psi(x_t^*)$ as $\alpha_t^*$ for any $t \in \{1, 2, \ldots, T\}$. We can have $\{\alpha_t^*\}_{t=1}^T$ also obey the 0-1 uniform distribution in the same manner as $\{\alpha_t\}_{t=1}^T$ does.

The reversible transformation between non-normal samples ($\{x_t\}_{t=1}^T$) and normal samples ($\{x_t^*\}_{t=1}^T$) is enabled through the equivalence of 0-1 uniform distributions. In detail, for any original sample ($x_t$), its cumulative frequency in $\{x_t\}_{t=1}^T$ is equal to $F(x_t)$. Assuming the cumulative frequency of the corresponding normal sample ($x_t^*$) is $F(x_t)$ as well, i.e.:

$$
\psi(x_t^*) = F(x_t). \quad (3.15)
$$

We can have:

$$
x_t^* = \psi^{-1}(F(x_t)) \text{ for any } t \in \{1, 2, \ldots, T\}. \quad (3.16)
$$

Samples $\{\psi^{-1}(F(x_t))\}_{t=1}^T$ are normally distributed, which can be tested through the $SW$ test.

In the process of ReDSICC, the original samples ($\{x_t\}_{t=1}^T$) are replaced with normally-distributed samples ($\psi^{-1}(F(x_t))$), providing inputs for statistical inference. At the
end of classification, these inputs are restored as the original samples. Take any transformed sample \((x_t^*)\) as an example. Its cumulative frequency equals to \(\Psi(x_t^*)\). Similarly, suppose the cumulative frequency of the original sample \((x_t)\) is also equal to \(\Psi(x_t^*)\), i.e.:

\[
F(x_t) = \Psi(x_t^*). \quad (3.17)
\]

We can have:

\[
x_t = F^{-1}(\Psi(x_t^*)) \text{ for any } t \in \{1, 2, \ldots, T\}. \quad (3.18)
\]

Due to reversibility of the functions \(F()\) and \(F^{-1}()\) and of functions \(\Psi()\) and \(\Psi^{-1}()\) and the equivalence of the intermediate 0-1 uniform distribution, the transformation between the original distributions \((\{x_t\}_{t=1}^T)\) and the normal distributions \((\{x_t^*\}_{t=1}^T)\) is approximately reversible. This avoids distortion of information in the normalization and restoration processes, and thus enhances the robustness of the ReDSICC method.

Through this approach, the original samples are replaced with normally-distributed samples. At the end of ReDSICC, the values of the climate variables in the classification result are restored to the original values. For the sake of clarity, the two-way transformation processes are not explicitly repeated in the following sections.

3.2.2 Dissimilarity Inferences

The multi-dimensional samples of the selected climate variables constitute an initial node of which the sample size is \(T\). Many nodes will also be gradually generated in the process of ReDSICC. The principle of the ReDSICC approach is evaluating the dissimilarity and similarity between the samples of climate variables in any two separated
nodes. To illustrate the technical details in the dissimilarity inferences process, the initial node is taken as an example.

Let the initial node be denoted as

\[ N_1 = \{X_t\}_{t=1}^T \]

(3.19)

where \( T \) is the amount of samples in the node. For the \( j \)th climate variable \( (x_j) \) where \( j \in \{1, \ldots, n\} \), the corresponding samples in \( N_1 \), i.e. \( \{x_{jt}\}_{t=1}^T \), are sequenced from the lowest value to the highest. The samples of \( x_j \) after sequencing are denoted as \( \{sx_{jt}\}_{t=1}^T \), the sequence numbers of \( \{x_{jt}\}_{t=1}^T \) corresponding to \( \{sx_{jt}\}_{t=1}^T \) as \( \{s(j, 1), s(j, 2), \ldots, s(j, T)\} \), and the sequence numbers of \( \{sx_{jt}\}_{t=1}^T \) corresponding to \( \{sx_{jt}\}_{t=1}^T \) as \( \{\hat{s}(j, 1), \hat{s}(j, 2), \ldots, \hat{s}(j, T)\} \). Corresponding to the sequence of \( \{x_{jt}\}_{t=1}^T \) from \( \hat{s}(j, 1) \) to \( \hat{s}(j, T) \), the multi-dimensional samples in the node are reordered accordingly. Let the multi-dimensional samples of the selected climate variables after sequencing be denoted as \( \{SX_t\}_{t=1}^T \) where

\[ SX_t = X_{(j, t)} \] for any \( t \in \{1, \ldots, T\} \).

(3.20)

To support identifying the most significant dissimilarity in the node, the multi-dimensional samples are cut into two groups, row by row. The node is thus divided as two child nodes. The partition row number is denoted as \( u \ (u \in \{1, \ldots, T\}) \) and the two child nodes as \( N_{1u_1} \) and \( N_{1u_2} \), respectively. That is,

\[ N_1 = N_{1u_1} \cup N_{1u_2} \]

(3.21)

and

\[ N_{1u_1} \cap N_{1u_2} = \emptyset \]

(3.22)
where \( N_{1j} = \{ S_{X_{ij}} \}_{t=1}^{u} \), \( N_{1j} = \{ S_{X_{ij}} \}_{t=u+1}^{T} \), \( j \in \{ 1, \ldots, n \} \), and \( u \in \{ 1, \ldots, T \} \). An analysis of the dissimilarity between climate-variables samples in the two child nodes is challenged by data uncertainties and multivariate dependencies. To mitigate this challenge, statistical inference is incorporated into the framework of ReDSICC.

The known features of sample sets \( \{ S_{X_{ij}} \}_{t=1}^{u} \) and \( \{ S_{X_{ij}} \}_{t=u+1}^{T} \) consist of: (1) two unpaired multivariate samples from normal populations and assumed as independent; (2) both variances are unknown and may not be equal, and (3) sample sizes are uncertain. The task of dissimilarity inferences is equivalent to the multivariate Behrens-Fisher problem (Christensen and Rencher, 1997), i.e. testing the significance of the inequality or equality of the population mean vectors of the two sample sets under these conditions. There are various hypothesis tests that are potentially effective in this task. In this study, the modified Nel and van der Merwe (MNV) test (Krishnamoorthy and Yu, 2004) is employed because, as revealed in (Krishnamoorthy and Yu, 2004), it outperforms most of the other related tests. The technical details of the MNV test are presented in the appendix. In brief, given an acceptable level of statistical significance (\( \alpha \)), the climate-variables samples in the two child nodes are of a significant difference if the \( p \) value of the MNV statistic, which is a function of the two groups of samples, is less than \( \alpha \).

For any \( j \in \{ 1, \ldots, n \} \) and \( u \in \{ 1, \ldots, T \} \), the climate-variables samples in \( N_{1} \) are sequenced according to the \( j \)th climate variable, the sequenced samples are partitioned as two groups of samples, i.e. rows 1 to \( u \) and rows \( u + 1 \) to \( T \), and the \( p \) value of the corresponding MNV statistic, denoted as \( p_{ju} \), is estimated. A \( p_{ju} \) value less than the given significance level (\( \alpha \)) implies that their dissimilarity is statistically significant. For any two groups of samples after sequencing and cutting, the \( p \) value enables quantification of
their dissimilarity and reflects the monotonic variation of climate-variables samples between them. As the decrease of the value of $p_{ju}$, the significance of their dissimilarity is increased. The minimal value of $p_{ju}$ for all $j \in \{1, \ldots, n\}$ and $u \in \{1, \ldots, T\}$ represents the most significant monotonic variation of climate-variables samples between two groups after classification. Let the minimal $p_{ju}$ value be denoted as $p_{\text{min}}(j_0, u_0)$ where $j_0$ is the corresponding climate variable (named as the criterion climate variable) and $u_0$ is the corresponding row number (named as the partition row number). The average of $x_{j_0, u_0}$ and $x_{j_0, u_0+1}$ is named as the partition threshold. The minimal $p_{ju}$ value also implies that the $j_0$th climate variable dominates the variation of the multi-dimensional climate-variables samples in $N_1$. As a representative feature of ReDSICC, the classification of multi-dimensional samples in any node is based on the minimal value of $p_{ju}$. The multi-dimensional samples ($\{X_t\}_{t=1}^{T}$) in the node of $N_1$ are classified as two groups in two child nodes, denoted as

$$N_2 = \{X_{\tilde{j}(j, u)}\}_{t=1}^{u_0}$$

and

$$N_2 = \{X_{\tilde{j}(j, u)}\}_{t=u_0+1}^{T}, \text{ if } p_{\text{min}}(j_0, u_0) < \alpha;$$

otherwise, these samples are not classified due to insignificant dissimilarity among them for any $j \in \{1, \ldots, n\}$ and $u \in \{1, \ldots, T\}$.

For any node, the involved multi-dimensional climate-variables samples of which the size is assumed as $T$ are tested through a series of dissimilarity inferences: (1) sequencing of a climate variable ($j \in \{1, \ldots, n\}$), (2) partition of sequenced samples at a
row \((u \in \{1, \ldots, T\})\) as two groups, (3) calculate the \(p_{ju}\) value corresponding to the significance of the difference between the two groups of samples, (4) cyclization of (2) and (3) for any row, (5) cyclization of (1) to (4) for any climate variable, (6) identification of the minimal \(p_{ju}\) value \((p_{\text{min}}(j_0, u_0))\), and (7) classification of the samples in the parent node into two groups in two child nodes if

\[
p_{\text{min}}(j_0, u_0) < \alpha. \tag{3.25}
\]

These dissimilarity inferences are conducted on every newly generated node. As a result, all nodes are internally inseparable; in other words, any node cannot be further classified because it does not involve two significantly different groups of climate-variables samples.

The number of nodes after dissimilarity inferences can be assumed as integer \(D\). Let the nodes be denoted as \(N^d\) where \(d = 1, 2, \ldots, D\), the number of samples in the \(d\)th node as \(T(d)\), the index for samples in the \(d\)th node as \(p\) (i.e., \(p = 1, 2, \ldots, T(d)\)), and the original sequencing numbers of samples in the initial node \((N_1)\) as \(U(d, p)\) for the \(d\)th node. We can have

\[
N^d = \{X_{U(d, p)}\}_{p=1}^{T(d)} \text{ for any } d \in \{1, 2, \ldots, D\}. \tag{3.26}
\]

Since all nodes are originated from the initial node \((N_1)\), we can have

\[
N^d \subseteq N_1 \text{ for any } d \in \{1, 2, \ldots, D\}. \tag{3.27}
\]

For any sample in \(N_1, X_t\) \((t \in \{1, 2, \ldots, T\})\), there exists a \(d \in \{1, 2, \ldots, D\}\) such that \(X_t \in N^d\). Therefore:

\[
N_1 \supseteq \bigcup_{d=1}^{D} N^d, \tag{3.28}
\]
\[ N_1 \subseteq \cup_{d=1}^{D} \mathcal{N}_d, \]  

and equivalently:

\[ N_1 = \cup_{d=1}^{D} \mathcal{N}_d. \]

Due to the independence between any pair of child nodes in any classification, we can have

\[ \mathcal{N}_{d_1} \cap \mathcal{N}_{d_2} = \emptyset \text{ for any } d_1 \in \{1, 2, \ldots, D\} \text{ and any } d_2 \in \{1, 2, \ldots, D\} \]

where \( d_1 \neq d_2 \). Namely, the obtained nodes are mutually exclusive and collectively exhaustive.

### 3.2.3 Similarity Inferences

Through the dissimilarity inferences, the samples in any node cannot be further classified. However, it is possible that the climate-variables samples in two nodes are of significant similarity, which cannot be avoided in the dissimilarity inferences focusing on the dissimilarity of samples within every node. Any two nodes which are of significant similarity should be clustered into a parent node through the following similarity inferences.

Let \( \mathcal{N}_{d_1} \) and \( \mathcal{N}_{d_2} \) be any two nodes of \( \{\mathcal{N}_d\}_{d=1}^{D} \). The involved climate-variables samples are \( \{X_{t(d_1, p)}\}_{p=1}^{T(d_1)} \) and \( \{X_{t(d_2, p)}\}_{p=1}^{T(d_2)} \), respectively. An analysis of their similarity is also enabled through the statistical inference method: the MNV test (Krishnamoorthy and Yu, 2004). The significance of the difference between population mean vectors of \( \{X_{t(d_1, p)}\}_{p=1}^{T(d_1)} \) and \( \{X_{t(d_2, p)}\}_{p=1}^{T(d_2)} \) is quantified by the \( p_{12} \) value. Given an acceptable level of statistical significance \( (\alpha) \), the difference is of statistical significance.
if \( p_{12} < \alpha \). In this case, the two nodes should be clustered. Let the new node be denoted as \( N^{d_1 \cup d_2} \). That is,

\[
N^{d_1 \cup d_2} = N^{d_1} \cup N^{d_2}.
\]  

(3.32)
The samples in \( N^{d_1 \cup d_2} \) are \( \{\{X_{U(d_1, \rho)}\}_{p=1}^{\tilde{T}(d_1)}\}, \{X_{U(d_2, \rho)}\}_{p=1}^{\tilde{T}(d_2)}\} \). The node list is updated correspondingly. The number of nodes is decreased from \( D \) to \( D - 1 \). On the other hand, the difference of the samples in nodes \( N^{d_1} \) and \( N^{d_2} \) is statistically insignificant if \( p_{12} \geq \alpha \). Updating is not required for the node list in this case. A similarity inference is conducted for any two nodes in the node list. The list is updated once any two nodes should be clustered. As the result of similarity inferences, any two nodes in the node list cannot be clustered, i.e. they are externally divergent.

The number of nodes after similarity inferences can be assumed as an integer (\( \check{D} \)). Since the number of nodes is decreased by one in a cluster analysis, we can have

\[
\check{D} \leq D.
\]  

(3.33)

Let these nodes be denoted as \( \check{N}^{d} \) where \( d = 1, 2, \ldots, \check{D} \), the number of climate-variables samples in the \( d \)th node as \( \check{T}(d) \), the index for climate-variables samples in the \( d \)th node as \( p \) (i.e., \( p = 1, 2, \ldots, \check{T}(d) \)), and the original sequencing numbers of climate-variables samples in the initial node (\( N_1 \)) as \( \check{U}(d, p) \) for the \( d \)th node. We can have

\[
\check{N}^{d} = \{X_{U(d, \rho)}\}_{p=1}^{\check{T}(d)} \text{ for any } d \in \{1, 2, \ldots, \check{D}\}.
\]  

(3.34)

All nodes originate from the initial node (\( N_1 \)) and any climate-variables sample in \( N_1 \) belongs to a node of \( \{\check{N}^{d}\}_{d=1}^{\check{D}} \), thus
\[ N_1 \supseteq \bigcup_{d=1}^{D} \tilde{N}^d, \quad (3.35) \]

\[ N_1 \subseteq \bigcup_{d=1}^{D} \tilde{N}^d \quad (3.36) \]

and

\[ N_1 = \bigcup_{d=1}^{D} \tilde{N}^d. \quad (3.37) \]

For any \( d_1 \in \{1, 2, \ldots, D\} \) and any \( d_2 \in \{1, 2, \ldots, D\} \) where \( d_1 \neq d_2 \), we can have

\[ \tilde{N}^{d_1} \cap \tilde{N}^{d_2} = \varnothing, \quad (3.38) \]

resulting from the independence of any two nodes after dissimilarity inferences. Namely, all nodes obtained from similarity inferences are mutually exclusive and collectively exhaustive.

As the result of dissimilarity inferences, all nodes are internally inseparable. This may be disabled by the potential generation of new nodes after similarity inferences. Meanwhile, the nodes that are externally divergent after similarity inferences may need to be further classified because one of them may not be internally inseparable. Therefore, a recursion of dissimilarity inferences and similarity inferences is desired for all nodes until they cannot be further classified or clustered. The number of nodes after recursive dissimilarity and similarity inferences can be assumed as \( D \). Let the nodes be denoted as \( N^d \) and named as end nodes \( (d = 1, 2, \ldots, D) \), the number of climate-variables samples in the \( d \)th end node as \( T(d) \), the index for climate-variables samples in the \( d \)th end node as \( p \) (i.e., \( p = 1, 2, \ldots, T(d) \)), and the original sequencing numbers of climate-variables samples in the initial node \( (N_1) \) as \( U(d, p) \) for the \( d \)th end node. We can have
\[ N^d = \{X_{U(d, p)}\}_{p=1}^{T(d)} \text{ for any } d \in \{1, 2, \ldots, D\} \]. \hspace{1cm} (3.39) 

All end nodes are originated from the initial node \(N_1\) and any climate-variables sample in \(N_1\) belongs to an end node of \(\{N^d\}_{d=1}^{D}\), thus

\[ N_1 \supseteq \bigcup_{d=1}^{D} N^d \] \hspace{1cm} (3.40)  
\[ N_1 \subseteq \bigcup_{d=1}^{D} N^d \] \hspace{1cm} (3.41)

and

\[ N_1 = \bigcup_{d=1}^{D} N^d \] \hspace{1cm} (3.42)

For any \(d_1 \in \{1, 2, \ldots, D\}\) and any \(d_2 \in \{1, 2, \ldots, D\}\) where \(d_1 \neq d_2\), we can have

\[ N^{d_1} \cap N^{d_2} = \emptyset \] \hspace{1cm} (3.43)

resulting from the independence of any two nodes in both dissimilarity inferences and similarity inferences. Namely, all end nodes that are obtained from recursive dissimilarity and similarity inferences are internally inseparable, externally divergent, mutually exclusive, and collectively exhaustive.

### 3.2.4 Climate Classification Results

Let the lower bound, upper bound and ranges of any climate variable be denoted as \(X_L, X_U, \) and \(X_{LU}\), respectively. Namely,

\[ X_L = (x_{1L}, x_{2L}, \ldots, x_{nL}) \] \hspace{1cm} (3.44)  
\[ X_U = (x_{1U}, x_{2U}, \ldots, x_{nU}) \] \hspace{1cm} (3.45)

and
\[ X_{LU} = [X_L, X_U]. \] (3.46)

We have

\[ x_{jL} \leq x_{jU}, \] (3.47)
\[ x_{jL} \leq x_{jt}, \] (3.48)
\[ x_{jU} \geq x_{jt}, \] (3.49)

and

\[ x_{jt} \in [x_{jL}, x_{jU}] \] (3.50)

for any \( j \in \{1, 2, \ldots, n\} \) and any \( t \in \{1, 2, \ldots, T\} \). Let the lower bound, upper bound, and ranges of samples of every selected climate variables in any end node be denoted as \( X_{dL}, X_{dU}, \) and \( X_{dLU} \), respectively. Namely,

\[ X_{dL} = (x_{1dL}, x_{2dL}, \ldots, x_{ndL}), \] (3.51)
\[ X_{dU} = (x_{1dU}, x_{2dU}, \ldots, x_{ndU}), \] (3.52)

and

\[ X_{dLU} = [X_{dL}, X_{dU}] \] (3.53)

for any \( d \in \{1, 2, \ldots, D\} \). We have

\[ x_{jdL} \leq x_{jdU}, \] (3.54)
\[ x_{jdL} \leq x_{jU(d, p)}, \] (3.55)

and

\[ 80 \]
\[ x_{jU} \geq x_{jU(d,p)} \quad (3.56) \]

for any \( j \in \{1, 2, \ldots, n\} \), any \( d \in \{1, 2, \ldots, D\} \) and any \( p \in \{1, 2, \ldots, T(d)\} \). End nodes \( \{N^d\}_{d=1}^D \) are mutually exclusive, thus we have

\[ X_{d_1LU} \cap X_{d_2LU} = \emptyset \quad (3.57) \]

for any \( d_1 \in \{1, 2, \ldots, D\} \) and any \( d_2 \in \{1, 2, \ldots, D\} \) where \( d_1 \neq d_2 \). Since end nodes \( \{N^d\}_{d=1}^D \) are collectively exhaustive, we should have

\[ X_{LU} = \bigcup_{d=1}^{D} X_{dLU}. \quad (3.58) \]

However, due to discreteness of climate-variable samples, there is a gap between \( X_{LU} \) and \( \bigcup_{d=1}^{D} X_{dLU} \). For example, the initial node \( (N_1) \) is classified as nodes \( N_2 \) and \( N_3 \) where

\[ N_2 = \{X_{j^0U}\}_{u=1}^{u^0} \quad (3.59) \]

and

\[ N_3 = \{X_{j^0U}\}_{u^0+1}^{u+1}. \quad (3.60) \]

The range of the \( j^0 \)th climate variable is \([x_{j^0(1)}, x_{j^0(T)}]\), while the ranges of the \( j^0 \)th climate variable in nodes \( N_2 \) and \( N_3 \) are \([x_{j^0(1)}, x_{j^0(u^0)}]\) and \([x_{j^0(u^0+1)}, x_{j^0(T)}]\), respectively. Thus, we have the gap is:

\[
[x_{j^0(1)}, x_{j^0(T)}] - (\bigcup [x_{j^0(1)}, x_{j^0(u^0)}] \cup [x_{j^0(u^0+1)}, x_{j^0(T)}]) = (x_{j^0(u^0)}, x_{j^0(u^0+1)}).
\]

\[ (3.61) \]
To fill the gap (e.g. \((x_{j_i^0}, u^y_{j_i^0}), x_{j_i^0}, u^{y+1}_{j_i^0})\)), the bounds of the criterion climate variable in two child nodes should be extended. In detail, the upper bound of the criterion climate variable in the lower node (i.e. \(N_2\)) is equal to the mid point of the gap (i.e. \((x_{j_i^0}, u^{y+1}_{j_i^0}) + x_{j_i^0}, u^y_{j_i^0})/2\), and so is the lower bound of the criterion climate variable for the upper node (i.e. \(N_3\)). For the given example of the initial node \(N_1\), the related revisions on bounds of samples in the child nodes are:

\[
X_{2U} = (x_{j_i^0}, u^{y+1}_{j_i^0}) + x_{j_i^0}, u^y_{j_i^0})/2 \tag{3.62}
\]

and

\[
X_{3L} = (x_{j_i^0}, u^{y+1}_{j_i^0}) + x_{j_i^0}, u^y_{j_i^0})/2. \tag{3.63}
\]

Through this correction, the end nodes obtained from recursive dissimilarity and similarity inferences are continuously inclusive, resulting in a systematic classification of regional climate reflected by the collected climate-variable samples.

Any end node after classification, i.e. \(N^d (d \in \{1, 2, \ldots, D\})\), represents a climate class in which the climatic conditions are significantly dissimilar to those in other classes and are significantly similar to each other. Characterization of the features of these climate classes is challenged by multiplicity of climate-variables samples in end nodes. For the \(d\)th climate class \((d \in \{1, 2, \ldots, D\})\), the corresponding climate-variables samples are \(\{X_{U(d, p)}\}_{p=1}^{T(d)}\), and their ranges after correction are \([X_{dL}, X_{dU}]\). These samples are not significantly different, but are not exactly identical for most cases. There are multiple alternative schemes of representing climate-class features. The mean values of samples \(\{X_{U(d, p)}\}_{p=1}^{T(d)}\) represent the average climatic conditions in this class, and the extremeness
or spatial heterogeneity of climatic conditions can be quantified by the minimum or maximum values, the variations or standard variations, or the coefficients of variation of these samples. In addition, the distribution of any climate variable in this class can be presented as an interval, discrete quartiles, a probabilistic density function or other means, which is helpful for providing rich information for related studies or practices. In the following case study, the mean, minimum and maximum values are employed.

3.3 CASE STUDY: ATHABASCA RIVER BASIN

3.3.1 Study Area and Data Collection

As shown in Figure 3.1, the Athabasca River originates from the Columbia Icefield on the eastern slopes of the Rocky Mountains around Jasper in Alberta, Canada. It flows approximately 1500 km northeast before entering the Peace-Athabasca Delta, the largest freshwater continental river delta in North America and the home to numerous migratory birds, and draining into Lake Athabasca. The evaluation drops from approximately 3715 m at Jasper to 211 m at the outlet in the Peace-Athabasca Delta. This river is the longest undammed river in the Canadian prairies. The ARB covers an area of approximately 138,000 km² over 52° ~ 59° N and -119° ~ -107° W. The ARB includes diverse hydro-climatic regimes due to physiographical heterogeneity; snowcapped mountains, coniferous forest, mixed wood and deciduous forest are found in the uplands, whereas willow brush, shrubs, black spruce and sphagnum moss dominate the lowlands (Kerkhoven and Gan, 2006). Around 0.15 to 0.17 million residents distribute over 22 rural or regional municipalities, 1 city, 12 towns, and 14 Aboriginal settlements in this river basin. Due to diversity of climate over the ARB and limitation of computational capacities in some cases, a classification of the regional climate is required for studies such as
downscaling, hydrological simulation, and energy systems management. Hence, the developed ReDSICC approach is applied to achieve this objective.

To support climate classification, the observations of historical climatic conditions should be collected. Multiple climate variables are available for characterizing regional climate. In this case study, three climate variables, i.e. daily minimum near-surface air temperature (\(T_{\text{min}}\)), daily maximum near-surface air temperature (\(T_{\text{max}}\)), and daily precipitation (\(P_{\text{rc}}\)), are selected as representative ones. Correspondingly, the historical observations at finer resolutions are derived from a raster-gridded climate dataset. The dataset contains daily estimates of the three climate variables from 1961 to 2003 over 1615 10-km grids in the ARB (Figure 3.1). Furthermore, the multi-year-averaged monthly, seasonal and yearly means of the three climate variables over all grids are calculated from these daily-resolution observations and used as the inputs of climate classification in this study. Prior to the classification study, the spatial variability of the regional climate over the ARB is analyzed in the followings.
Figure 3.1. Geographical Conditions and Climatic Observations of The Athabasca River Basin.
3.3.2 Spatial Climatological Variability

The multi-year-averaged spatial, monthly, seasonal and overall variability of climate over the ARB is illustrated in Figure 3.2. Figure 3.2(a) and (b) reveal that, on average, the day or night is warmest, around 5.5 to 6.5 °C, in the southwest and the north central, is coldest, around 0.69 to 2 °C, in the northeast, and gradually becomes cold from southwest to northeast across the ARB. It is different with the spatial distribution of temperature that a southwest-northeast-direction strip-shaped region in the central is the wettest, the northeast is moderate wet, the north end, the south end and the north central is the driest, and the other regions are moderate dry. The monthly or seasonal variability of either temperature or precipitation is very significant. Both the night and the day are coldest in January, reaching -28 and -17, respectively, and are hottest in July, reaching 11.5 and 23, respectively. As for precipitation, it is the lowest in February, around 0.7 mm, and the highest in July, around 2.4 mm; approximately 43.20% is in JJA, 26.19% in SON, 16.33% in MAM, while 14.29% in DJF. The spatial variability of temperature and precipitation at the monthly scale is presented in Figure 3.3 to Figure 3.6. These figures show that the spatial variability of both temperature and precipitation vary with months. For instance, the day temperature gradually decreases from south to north in August, and the precipitation center moves to the south end in April. To sum up, temperature and precipitation are of significant variability temporally and spatially.
Figure 3.2. Multi-Year-Averaged Spatial Climatological Variability.
Figure 3.3. Monthly Spatial Climatological Variability over The ARB in Winter.
Figure 3.4. Monthly Spatial Climatological Variability over The ARB in Spring.
Figure 3.5. Monthly Spatial Climatological Variability over The ARB in Summer.
Figure 3.6. Monthly Spatial Climatological Variability over The *ARB* in Autumn.
3.3.3 Model Configuration

The objective of climate classification over the ARB is to identify multiple climate zones of which the climatic conditions are significantly dissimilar among different zones and similar for grids in the same zone. The inputs of the ReDSICC approach are multi-year-averaged $T_{min}$, $T_{max}$ and $Prec$ in twelve months over 1615 10-km grids in the ARB.

Parameters in this approach consist of: a level of statistical significance ($\alpha$) and the minimum partition row number ($N_{min}$). Normally, the alternative values of $\alpha$ are 0.01, 0.05 and 0.10 from the viewpoint of statistical analyses. The parameter of $N_{min}$ is introduced to control the values of the partition row number in dissimilarity analyses due to requirement of statistical inferences on sample sizes. Its values range from 3 to 30 and are higher for low- and large-sample-size climate classifications, respectively.

Theoretically, the number of climate classes after ReDSICC is increasingly monotonic with $\alpha$ and decreasingly monotonic with $N_{min}$. The higher the number of climate classes is, the more local characteristics of regional climate are captured. Meanwhile, the classification result is relatively less reliable due to the decrease of average sample sizes for every climate class from the viewpoint of statistical analysis. Therefore, there is an interactive tradeoff among $\alpha$, $N_{min}$, the number of climate classes, and the statistical reliability of classification results.

In this study, the parameters of $\alpha$ and $N_{min}$ are set as 0.05 and 20, respectively. A program is developed using the open-source R language (http://www.r-project.org/) in consideration of its advantage at statistical analyses. The climate classification result is analyzed in the following section. In addition, climate classification under other parameter
combinations are also obtained, which is for analyzing parameter sensitivities and facilitating parameter calibration in application of the developed ReDSICC approach. The related results are presented in the section of discussions.

3.4 RESULTS ANALYSIS

3.4.1 Climate Zones

The climate classification results through the ReDSICC approach are presented in Figure 3.7. A total of 20 climate zones are identified under the settings of the parameters of ReDSICC. The corresponding grids of most zones are geographically adjacent, while the locations of grids are discontinuous for some zones such as the one corresponding to the 20th end node. This is resulted from coexistence of overall spatial continuity and local spatial discontinuity of climate over the ARB as illustrated in Figure 3.2. In addition, the number of grids in the zones approaching the boundary of the ARB tends to be lower than that in the central zones, which may be related with the increase of spatial gradients from the latter types of zones to the former types. A comparison of Figure 3.7(b) to (d) with Figure 3.2(a) to (c) reveals that the overall spatial patterns of Tmin, Tmax and Prec are well captured by the obtained climate classification results, especially for Tmin and Tmax. The spatial resolution is decreased in the process of climate classification; as a result, the ranges of climate variables are shrunk; for example, the original range of Tmin is from -9.33 to -5.17 °C, and it is changed to from -9.21 to -5.30 °C after classification. The quartiles and size of samples for every selected climate variable in every climate zones will be deliberated in the section of discussions. In general, the spatial dissimilarity of climatic conditions among all grids over the ARB is reflected in the classification result.
Figure 3.7. Climate Zones by The ReDSICC Approach.
3.4.2 Zonal Heterogeneities and Similarities

For any climate variable in every climate zone, the range of multi-year averages over the corresponding grids is calculated and presented in Figure 3.8(a) to (c). This range quantifies the spatial heterogeneity of the climate variable within the climate zone. It is revealed that, for every selected climate variable, there are slight heterogeneities among multiple grids in the same zone, although the multi-grid climate-variable ranges are relatively low in comparison with the ranges of monthly climate-variable observations. Furthermore, these heterogeneities vary with climate zones, of which the variation pattern is different for the selected three climate variables. Specifically, these heterogeneities gradually decrease from the lower and upstreams to the middle stream for $T_{max}$ and $Prec$, while presenting complicated discontinuous characteristics for $T_{min}$.

In addition, the similarities among multiple grids in any climate zone are analyzed. For any grid in any climate zone, its similarities with the zonal average for every selected climate variable are quantified by the Nash-Sutcliffe coefficient (Nash and Sutcliffe, 1970). The related results are displayed in Figure 3.8(d) to (f). A comparison of the selected climate variables reveals that the overall zonal similarities are relatively high for and vary with climate variables. The ranges of the Nash-Sutcliffe coefficient are from 0.925 to 1.000 for $T_{min}$, from 0.933 to 1.000 for $T_{max}$, and from 0.431 to 0.993 for $Prec$. This implies that the similarities of temperature among all grids are well captured by the developed $ReDSICC$ approach and, on the contrary, those of precipitation are captured at a relatively lower efficiency due to a lower magnitude of precipitation. Besides, it is shown in these figures that the grids of relatively low similarities with the corresponding zonal averages concentrate in five regions, i.e. regions A to E as marked. This may be related
with the higher spatial heterogeneities of climatic conditions in these regions compared with those in other regions. Attention should be paid on these regions when the climate classification result of the ReDSICC approach is used for the related studies.

To sum up, the analyses in this section support us reach a conclusion that the complicated dissimilarities and similarities of climatic conditions among all grids over the ARB are effectively reflected in the climate classification results obtained from the developed ReDSICC approach. In brief, this approach is a relatively reliable tool of supporting various related studies such as downscaling, hydrological simulation, urban planning, and ecological systems management in the ARB.
Figure 3.8. Heterogeneities and Similarities of Climate Zones.
3.5 DISCUSSIONS

3.5.1 Effectiveness of DDT

As one of innovations of this study, the DDT approach is developed to enable reversible transformation between an abnormal distribution and a normal distribution. The Wilks statistics and the corresponding $p$ value are introduced to evaluate the normality of a probabilistic distribution. Their values for the collected observation samples of every selected climate variables before and after distribution transformation through 50 times of resampling are presented in Figure 3.9. According to the theory of normality test, given a level of statistical significance, e.g., $\alpha = 0.01$, the null hypothesis that a dataset is from a normally distributed population should be rejected if the calculated $p$ value is less than $\alpha$; on the other hand, the null hypothesis cannot be rejected if the $p$ value is greater than $\alpha$. Namely, a group of samples tend to come from a normal distribution if the $p$ value is relatively high in comparison with the given significance level. As shown in Figure 3.9, the Wilks statistics are significantly increased for every selected climate variables after distribution transformation. Prior to transformation, the $p$ values of the Wilks statistics for the samples of every selected climate variable are nearly zero which is absolutely less than $\alpha$. After transformation, the $p$ values are increased to 0.183 for $T_{min}$, 0.148 for $T_{max}$, and 0.013 for Prec, respectively. If the significance level is set as 0.01, the samples of every selected climate variable after distribution transformation through the DDT approach obey a normal distribution. The difference of $p$ values among the three climate variables after transformation also verifies a fact that, compared with precipitation, temperature observations tend to follow a normal distribution with a higher likeness. As a conclusion of this subsection, the objective of enabling reversible transformation between an
abnormal distribution and a normal distribution is effectively achieved by the proposed

$DDT$ approach.
Figure 3.9. Effectiveness of The DDT Approach.
3.5.2 ReDSICC Process Analyses

The key of the ReDSICC approach is a recursive process of analyzing dissimilarities and similarities of multi-dimensional climate-variables samples under data uncertainties and multivariate dependencies. An analysis of the ReDSICC process and intermediate results can help reveal a few of laws in the regional climate of the ARB. Hence, for every climate zone and every selected climate variable, the quartiles and size of samples are calculated and presented in Figure 3.10(a) to (c). It is revealed that, as an expected result of ReDSICC, there is a significant difference between the sample distributions of at least one selected climate variable for any two climate zones. For $T_{min}$ or $T_{max}$ of which the samples tend to obey a bell shaped distribution, the sample sizes of the climate zones where the climate-variable values are extremely high or low are less than those of other climate zones. A comparison of Figure 3.10(a) to (c) discloses a monotonically increasing relationship between $T_{min}$ and $T_{max}$ and, in consideration of the variation of Prec sample distributions with $T_{min}$ and $T_{max}$ in the last 11 climate zones, a globally monotonically decreasing relationship between temperature and precipitation. In addition, for every classification operation in the dissimilarity analyses module of ReDSICC, the sample size of the parent node, the criterion climate variable, and the partition threshold value are generated (Figure 3.10(d)). It is shown that the criterion climate variables are mainly $T_{min}$ and $T_{max}$, implying that the spatial heterogeneity of temperature is higher than that of precipitation. An analysis of the correspondence between sample sizes and criterion climate variables reveals that the spatial heterogeneity of $T_{max}$ is the highest over the ARB.
Figure 3.10. Analyses of the ReDSICC Process.
3.5.3 Sensitivity Analyses

When the proposed ReDSICC approach is applied to classification of climate in various cases, the classification result and effectiveness is associated with multiple factors, e.g. distribution transformation, climate variable selection, and parameter settings. To disclose the impacts of these factors, a series of sensitivity analyses are conducted accordingly. The related results are presented in Figure 3.11.

As shown in Figure 3.11(a), a total of 17 climate zones are generated if the original sample distribution is not transformed as a normal distribution for every selected climate variable. In comparison with the classification result from ReDSICC with distribution transformation, the spatial distribution of climate zones is significantly changed. To reveal whether this change is positive or negative, the root mean square error (RMSE) (Hyndman et al., 2006) is introduced to quantify the overall distance between the original multi-year-averaged distribution and the regionalized distribution of every selected climate variables. If the module of distribution transformation is activated, , the RMSE equals to 0.125, 2.003 and 0.139 for Tmin, Tmax and Prec, respectively; on the contrary, the RMSE is increased to 0.155, 10.676 and 3.014, respectively, if climate-variable samples are not transformed by the DDT approach. This implies that, although the number of climate zones is not changed significantly, the effectiveness of the climate classification result at reflecting the spatial dissimilarity and similarity of regional climate over the ARB is decreased if the DDT approach is not employed.

In this study, the selected climate variables are Tmin, Tmax and Prec of which Tmin and Tmax are of relatively significant correlations. Another experiment of climate classification is attempted through replacing Tmin and Tmax with their average, i.e.
Tmean = (Tmin + Tmax)/2. As shown in Figure 3.11(b), the corresponding result is different with the original one, especially for local features of some climate zones, although the global distribution of climate zones does not change significantly. It is implied that the classification result of ReDSICC varies with changes of representative climate variables. When the ReDSICC is applied to the climate-change impact studies or other related ones, attentions should be paid on climate-variable selection.

The parameters of the minimum partition row number (Nmin) and the level of statistical significance (α) in ReDSICC are set as 20 and 0.05, respectively in the above case study. To analyze their impacts on the classification result, four experiments in which Nmin is set as 10 and 30 and α is set as 0.01 and 0.10, respectively are conducted. The corresponding results are shown in Figure 3.11(c) to (f). As revealed by a comparison of Figure 3.11(c) and (d), an increase of Nmin may lead to a decrease of the obtained climate zones, which mainly occurs for zones of which the frequency of the corresponding climate-variable samples is relatively low. On the other hand, a comparison of Figure 3.11(e) and (f) shows that, as α decreases, an increased number of climate zones are merged. Theoretically, these revelations are of generality for various climate classification practices from the viewpoint of statistical analyses. In climate classification practices, a representative problem is that the number of climate zones is too large or small to satisfy expectations. When the developed ReDSICC approach is applied to these practices, adjusting the values of the two parameters, i.e. Nmin and α, within reasonable ranges is an effective solution for resolving this problem.
Figure 3.11. Sensitivity Analyses for The ReDSICC Approach.
3.5.4 Potential Extensions

The developed ReDSICC approach is advantageous at multiple aspects. For instance, the module of DDT eliminates the restriction of samples being normally distributed which is a common prerequisite for most statistical climate classification methods. Incorporation of advanced statistical inferential methods into the framework of ReDSICC enables classification of regional climate under complexities of data uncertainties and multivariate dependencies. The recursive process of dissimilarity and similarity inferences facilitates identifying the most desired climate classification result in which climatic conditions are significantly different for any two climate zones and are not of significant differences for grids in the same climate zone. The whole process of ReDSICC is independent with subjective judgement such as selections of climate variables, classification thresholds, and class sizes prior to classification practices; instead, these work are automatically accomplished by statistical inferences, e.g. identifying the climate variables of the highest spatial heterogeneity from multiple alternatives. These features of ReDSICC are much helpful for enhancing the reliability of climate classification results, improving the effectiveness of existing climate classification methods, and providing scientific support for climate-change impact studies or the other related ones.

On the other hand, the developed ReDSICC approach is challenged by multiple issues that deserve further efforts to mitigate in the subsequent research. For example, due to reliance on discrete historical observations over all finer-resolution grids, the feasibility of the DDT approach is restricted to cases in which these observations are many enough to avoid loss of representativeness; for this issue, continuous transformation functions may be more effective than discrete proximity searching. Tests on inequality or equality of the
covariance matrices of two multi-dimensional climate-variables samples may be helpful for enhancing the reliability of dissimilarity or similarity inference results. In the module of dissimilarity inferences, there may be another scheme of identifying the two groups of climate-variables samples that are of the most significant dissimilarity to replace the current one of stepwise sequencing and classification. Necessity of normalization of climate-variable samples should be verified in case that the difference between magnitudes of these samples influences selection of criterion climate variables. A dynamic version of ReDSICC may be more desired for the related practices because of, as shown in the third section, the seemingly significant variation of spatial distributions of climatic conditions with months or seasons. An optimal combination of diverse values of the parameters ($\alpha$ and $N_{\text{min}}$) instead of two identical values for the whole process of ReDSICC may lead to enhanced effectiveness of the climate classification result. Applications to other various climate-classification case studies may help disclose other opportunities of improving the developed ReDSICC approach.

3.6 SUMMARY

In this study, the ReDSICC approach was developed to provide an alternative for enabling effective classification of climate under data uncertainties and multivariate dependencies. Based on development of a discrete distribution transformation method and integration of advanced statistical inferential methods, a recursive framework of dissimilarity and similarity inferences was proposed for gradually grouping multi-dimensional climate-variables observations. In comparison with existing climate classification methods, the developed ReDSICC approach was advantageous at multiple aspects. For instance, the module of DDT eliminated the restriction of samples being
normally distributed. Statistical inferential methods enabled classification of regional climate under data uncertainties and multivariate dependencies. The recursive process of dissimilarity and similarity inferences facilitated identifying the most desired climate classification result in which climatic conditions were significantly different for any two climate zones and were not of significant differences for grids in the same climate zone. The whole process of ReDSICC was independent with subjective judgement which was replaced with statistical inferences.

To verify methodological effectiveness and facilitate local studies such as downscaling and hydrological simulation, the ReDSICC approach was applied to a case study of climate classification in the ARB, Canada. It was revealed that the complicated dissimilarities and similarities of climatic conditions among all grids over the ARB were effectively reflected in the climate classification results obtained from ReDSICC. The objective of enabling reversible transformation between an abnormal distribution and a normal distribution was effectively achieved by the DDT approach. The effectiveness of the climate classification result at reflecting the spatial dissimilarity and similarity of regional climate over the ARB was decreased if the DDT approach was not employed. In addition, a series of comparisons helped gain insights into the regional climate in the ARB and the mechanism of ReDSICC. For instance, in comparison with Tmin, the spatial heterogeneity of Tmax was higher while that of Prec was lower over the ARB. The classification result of ReDSICC varied with changes of representative climate variables. An increase of Nmin might lead to a decrease of the obtained climate zones, which mainly occurred for zones of which the frequency of the corresponding climate-variable samples
was relatively low. As $\alpha$ decreased, an increased number of climate zones would be merged.

These advantages of ReDSICC and revelations from this case study were much helpful for enhancing the reliability of climate classification results, improving the effectiveness of existing climate classification methods, and providing scientific support for climate-change impact studies or the other related ones. Meanwhile, the developed ReDSICC approach was challenged by multiple issues that deserved further efforts to mitigate in the subsequent research. Representative ones included but were not limited to the followings: developing continuous distribution transformation functions, adding a module of covariance-matrices inequality or equality tests, improving the scheme of identifying the most significant dissimilarity, evaluating the necessity of removing the different magnitudes of climate-variable samples, developing a dynamic version of ReDSICC, optimizing the values of the parameters ($\alpha$ and $N_{\text{min}}$), and disclosing other opportunities of improving ReDSICC through extensive applications.
CHAPTER 4

RECURSIVE MULTIVARIATE PRINCIPAL-MONOTONICITY

INFERENTIAL DOWNSCALING: METHODOLOGY DEVELOPMENT

AND APPLICATION TO ATHABASCA RIVER BASIN, CANADA

4.1 BACKGROUND

Within the scope of projecting future climate change, general circulation models (GCMs) are commonly used to assess changes resulting from further increases of atmospheric greenhouse gases (Hertig and Jacobeit, 2013). The atmospheric mesoscale features and the land surface heterogeneity are not properly resolved in coarse-resolution GCMs (Xu and Yang, 2015). Statistical downscaling provides a computationally inexpensive technique that can be adapted for a wide range of applications (Hertig and Jacobeit, 2013). Statistical downscaling can be generalized as a process of deducing finer-resolution regional or local climatic variables (i.e., predictands) from related influencing factors (i.e., predictors), which may include large-scale atmospheric variables or physiographic features (e.g., topography and land use), through statistical analyses. This process is challenged by the existence of various complexities in the correspondence between predictors and predictands.

For instance, the historical observations or simulation of climate variables may be of uncertainties. Two atmospheric-variables combinations that are not significantly different may correspond to two different values of a climate variable, in which the significance of the climate-variable difference is unknown. The relationship between a predictor and a predictand may be highly nonlinear (Weichert and Bürger, 1998), which is further complicated by the multiplicity of predictors and predictands, the uncertainties
of observations and simulation, the temporal and spatial heterogeneities of climate systems, as well as other potential complexities. It is possible that such a complicated relationship is hardly quantified by a continuous specific function. Multiple predictands may be dependent with each other. This dependent effect is neglected when a downscaling tool is used for every predictand separately. The effectiveness of most statistical downscaling methods relies on a prerequisite, i.e. the population of predictor or predictand samples coming from a normal distribution, which is merely discussed in downscaling studies or practices. The correspondence between predictors and predictands may be of high similarity among various spatial units, resulting in redundant computations. There may be nonstationarities in the predictors-predictands correspondence between different periods. For a statistical downscaling model, the optimal parameter combination calibrated over a period may not be of the highest accuracy or even possibly be the least accurate over another period. A single statistical metrics may not be capable of comprehensively quantifying the multi-aspect accuracies of a statistical downscaling method.

In the past decades, a number of statistical downscaling approaches were developed by researchers all over the world. An overview of these approaches is presented in many publications such as (Maraun et al. 2010), (Wilby and Wigley, 1997), and (Schoof, 2013). Representative ones include but are not limited to the followings. Scaling techniques are perhaps the most intuitive statistical methods for inferring fine scale information from 

GCMs (Schoof, 2013). Spatial interpolation or disaggregation of GCM outputs, for example, provides a baseline against which more rigorous downscaling methods can be compared (Wheater et al. 1999). The artificial neural network can approximate nonlinear
relations between predictors and predictands and their derivatives without prior knowledge of a specific nonlinear function (Gupta et al., 2014), which is helpful for making accurate forecasting of highly nonlinear climate systems (Schoof, 2013). A growing number of computational learning algorithms including tree-based methods (Goyal et al. 2012), genetic programming (Pour et al., 2014), support vector machines (Tripathi et al. 2006), and relevance vector machines (Ghosh and Mujumdar 2008) were developed as well. Classification methods based on fuzzy rules (Bárdossy et al. 2005), stepwise cluster analysis (Wang et al., 2013), and self-organizing maps (Huva et al., 2015) were widely applied within a context of downscaling. However, few studies were dedicated to incorporating all of these complexities into the downscaling process without unreasonable simplifications based on proposition of an effective downscaling approach. Correspondingly, the downscaled results are questioned in terms of reliability.

Therefore, this study aims to develop a recursive multivariate principal-monotonicity inferential downscaling approach (ReMPMID) for supporting climate downscaling under complexities such as data uncertainties, nonlinear predictors-predictands correspondences, predictands’ interactions, non-normal distributions, spatial homogeneities, and temporal nonstationarities. This approach is an advanced framework in which statistical inferential methods are integrated to address these complexities. The framework of this study is presented in Figure 4.1. Specifically, the principle, innovations and technical details of the ReMPMID approach will be discussed in section 4.2. In section 4.3, this approach will be applied to the ARB, a large river basin on the Canadian prairies that is closely connected with climate change, to verify methodological effectiveness and facilitate local impact or adaptation practices. Specifically, based on predictor and
predictand (i.e. $T_{min}$, $T_{max}$ and $Prec$) selection, data collection and processing, a grids-similarity analysis, a sensitivity analysis, and parameter optimization, a ReMPMID model will be constructed for every selected grid in the ARB. In section 4.4, the multi-dimensional modeling accuracies of the ReMPMID approach will be verified and examined in various aspects. The optimal parameter values and the uncertainties in high-resolution climate simulation will be analyzed through a series of comparisons.
Figure 4.1. Framework of This Chapter.
4.2 METHODOLOGY DEVELOPMENT

4.2.1 Normality Analyses

As an advanced statistical downscaling tool, the proposed ReMPMID approach focuses on characterizing the correspondence between finer-resolution regional/local climatic conditions (i.e., predictands) and the related influencing factors (i.e. predictors), e.g. coarser-resolution climatic simulation as well as physiographic features, through a framework of recursive multivariate principal-monotonicity inferences. Let the predictands and the predictors be denoted as \( \{y_i\}_{i=1}^m \) or \( Y = (y_1, y_2, \ldots, y_m) \) and \( \{x_j\}_{j=1}^n \) or \( X = (x_1, x_2, \ldots, x_n) \), respectively where \( m \) (or \( n \)) is the number of predictands (or predictors). The corresponding datasets of the predictands and the predictors are \( \{Y_t\}_{t=1}^T \) and \( \{X_t\}_{t=1}^T \), respectively, which constitute the multi-dimensional paired samples of predictors and predictands:

\[
\{XY_t\}_{t=1}^T = \{(X_t, Y_t)\}_{t=1}^T. \tag{4.1}
\]

It is identical with other statistical downscaling approaches that, in ReMPMID, the population of predictand samples should be normally distributed. Invalidity of this prerequisite may introduce significant errors into the downscaling process and results and lead to misrepresentation of the correspondence between predictors and predictands. A normality analysis is desired prior to applying any statistical downscaling approaches including ReMPMID. Hence, a high-power multivariate normality test technique, i.e. the ZS test (Zhou and Shao, 2014), is employed to examine the normality of predictand samples in this study. Taking advantages of the Shapiro-Wilks test (Shapiro and Wilk, 1965) and the skewness and kurtosis test (Joanes and Gill, 1998), this test shows relatively
high reliability in comparison with other alternatives. A brief summary the ZS test is stated in the appendix.

If the normality prerequisite does not hold, a transformation from an abnormal distribution to a normal distribution is required to guarantee the reliability of the ReMPMID results. The process must be reversible so that the original values of the climate variables can be restored, avoiding distortion of the similarity and dissimilarity among samples of climate variables. To achieve this, a discrete distribution transformation (DDT) approach (Cheng et al., 2016b) which was developed based on the reversible convertibility of any distribution to the 0-1 uniform distribution is incorporated into the framework of ReMPMID. The related details are presented in the appendix. Through this approach, the original predictand samples are replaced with normally-distributed ones. At the end of ReMPMID, the predictand values in the constructed model are restored to the original values. For the sake of clarity, the two-way transformation processes are not explicitly repeated in the following sections.

4.2.2 Recursive Dissimilarity Inferences

All paired samples of predictors and predictands constitute an initial node of which the sample size is \( T \). Let this node be denoted as \( N_1 \). In the process of ReMPMID, many intermediate nodes would be generated. For any of them, a recursive dissimilarity analysis (RDA) is conducted to reveal the dissimilarity of the predictors-predictands correspondence. The initial node (\( N_1 \)) is taken as an example to present the procedures of ReMPMID.
Let $x_j$ be the $j$th predictor. The samples of $x_j$ in $N_1$, i.e. \{x_{jt}\}_{t=1}^T$, are sequenced from the lowest value to the highest. We denote the samples of $x_j$ after sequencing as \{s_jx_{jt}\}_{t=1}^T, the sequence numbers of \{x_{jt}\}_{t=1}^T corresponding to \{s_jx_{jt}\}_{t=1}^T as \{s(j, 1), s(j, 2), \ldots, s(j, T)\}, and the sequence numbers of \{s_jx_{jt}\}_{t=1}^T corresponding to \{x_{jt}\}_{t=1}^T as \{s(j, 1), s(j, 2), \ldots, s(j, T)\}. Namely,

\[
x_{jt} = s_jx_{s(j, t)}
\]  \hspace{1cm} (4.2)

and

\[
s_jx_{jt} = x_{s(j, t)}
\]  \hspace{1cm} (4.3)

for any $t \in \{1, \ldots, T\}$. Accordingly, the paired samples of predictors and predictands, the predictors samples, the predictands samples, and the local paired samples of predictor $x_j$ and predictands $(\{(x_{jt}, Y_t)\}_{t=1}^T)$ are sequenced as $\{XY_{s(j, t)}\}_{t=1}^u$, $\{X_{s(j, t)}\}_{t=1}^u$, $\{Y_{s(j, t)}\}_{t=1}^u$, and $\{(x_{s(j, t)}, Y_{s(j, t)})\}_{t=1}^u$, respectively.

The sequenced paired samples are cut as two groups in two child nodes from a row. We denote the partition row numbers as $u$ and the two child nodes as $N_{1/u_1}$ and $N_{1/u_2}$, respectively. A parameter ($N_{\text{min}}$) which is named as the minimum partition row number and which represents the minimum value of $u$ in the process of ReMPMID is introduced to reflect the requirement of degrees of freedom in multivariate hypothesis testing. The partition row numbers ranges from $N_{\text{min}}$ to $T - N_{\text{min}}$. According to the partition operation, we have

\[
N_{1/u_1} = \{XY_{s(j, t)}\}_{t=1}^u,
\]  \hspace{1cm} (4.4)

\[
N_{1/u_2} = \{XY_{s(j, t)}\}_{t=u+1}^T,
\]  \hspace{1cm} (4.5)
\[ N_1 = N_{1 \mu_1} \cup N_{1 \mu_2}, \]  \hspace{1cm} (4.6) \\
and \\
\[ N_{1 \mu_1} \cap N_{1 \mu_2} = \emptyset. \] \hspace{1cm} (4.7) \\

Meanwhile, the sequenced predictands in nodes \( N_{1 \mu_1} \) and \( N_{1 \mu_2} \) are \( \{ Y_{\tilde{j}(j, t)} \}_{t=1}^{u} \) and \( \{ Y_{\tilde{j}(j, t)} \}_{t=u+1}^{T} \), respectively. \\

Subsequently, the modified Nel and van der Merwe (MNV) test (Krishnamoorthy and Yu, 2004) is employed to analyze the significance of the difference between \( \{ Y_{\tilde{j}(j, t)} \}_{t=1}^{u} \) and \( \{ Y_{\tilde{j}(j, t)} \}_{t=u+1}^{T} \). The technical details of the MNV test are presented in the appendix. \\

For any \( j \in \{1, 2, \ldots, n\} \) and \( u \in \{1, 2, \ldots, T\} \), the \( p \) value of the MNV statistic which is a function of \( \{ Y_{\tilde{j}(j, t)} \}_{t=1}^{u} \) and \( \{ Y_{\tilde{j}(j, t)} \}_{t=u+1}^{T} \) is calculated. Let the result be expressed as \( p_{ju} \). \\

An indicator, \( w_{ju}(\alpha) \), defined as \( \alpha/p_{ju} \) and named as the marginal monotonicity significance, is introduced to quantify the significance of the local marginal monotonicity of predictands \( \{ y_i \}_{i=1}^{m} \) with predictor \( x_j \) between nodes \( N_{1 \mu_1} \) and \( N_{1 \mu_2} \) where \( \alpha \) is an acceptable level of statistical significance. The multi-dimensional samples of predictands in \( N_{1 \mu_1} \) and \( N_{1 \mu_2} \) are significantly different if \( p_{ju} < \alpha \), i.e. \( w_{ju}(\alpha) > 1 \). \\

For the initial node \( (N_i) \), there can be many alternative schemes to partition it as a pair of child nodes. In the framework of ReMPMID, the principal local marginal monotonicity which corresponds to the most significant difference of predictands between samples in two child nodes after predictor-oriented sequencing and partitioning operations is used to support the discretization of all parent nodes including \( N_i \). Let the criterion
predictor and the partition row numer corresponding to the highest \( w_{j_0}(\alpha) \) be denoted as \( j_0 \) and \( u_0 \), respectively. The predictor \((x_{j_0})\) is defined as the criterion predictor. This predictor dominates the variation of the predictands within the initial node. Accordingly, all paired samples in the initial node are classified as two child nodes: \( N_2 \) and \( N_3 \) where

\[
N_2 = \{XY_{j_0,n}\}_{t=1}^{u_0}
\]

and

\[
N_3 = \{XY_{j_0,n}\}_{t=u_0+1}^{T}.
\]

In the process of ReMPMID, parent nodes are recursively replaced by child nodes. As a result, all nodes are internally inseparable, i.e. cannot be further classified in accordance with the principal local marginal monotonicity between any predictor and the predictands. The number of nodes after RDA can be assumed as an integer \( D \). These nodes constitute a node list of which the length is \( D \). Let the nodes be denoted as \( N^d \) where \( d = 1, 2, \ldots, D \). We have

\[
N^d = \{XY_{U(d,p)}\}_{p=1}^{T(d)}
\]

for any \( d \in \{1, 2, \ldots, D\} \) where \( T(d) \) is the number of included paired samples, \( p \) is the index for paired, and \( U(d, p) \) is the original sequencing numbers of paired samples in the initial node.

4.2.3 Recursive Similarity and Dissimilarity Inferences

Due to the nonlinear correspondence between predictors and predictands, it is possible that a significant difference of predictors between samples in different nodes
corresponds to an insignificant difference of predictands. To reveal the nonlinear correspondence and avoid redundant computation, a recursive similarity analysis (RSA) is conducted on every pair of nodes obtained from RDA.

Let \( N^{d_1} \) and \( N^{d_2} \) be any pair of nodes in the node list \( \{N^d\}_{d=1}^D \). The involved paired samples are \( \{XY^{U(d_1, p)}\}_{p=1}^{T(d_1)} \) and \( \{XY^{U(d_2, p)}\}_{p=1}^{T(d_2)} \), respectively. Following the related definition in RDA, the marginal monotonicity significance is used to quantify the statistical significance of the difference between predictands samples in \( N^{d_1} \) and \( N^{d_2} \). If the difference is statistically insignificant, the nodes \( N^{d_1} \) and \( N^{d_2} \) are merged as and replaced by a new node:

\[
N^{d_1 \cup d_2} = N^{d_1} \cup N^{d_2}.
\]  

(4.11)

The difference of predictands in every pair of nodes in the node list is evaluated until any pair of nodes cannot be merged (i.e. they are externally divergent). Suppose the number of nodes in the node list after RSA is the integer \( \hat{D} \). Let these nodes be denoted as \( \tilde{N}^d \) where \( d = 1, 2, \ldots, \hat{D} \). We have

\[
\tilde{N}^d = \{XY^{U(d, p)}\}_{p=1}^{\hat{T}(d)}
\]

(4.12)

for any \( d \in \{1, 2, \ldots, \hat{D}\} \) where \( \hat{T}(d) \) is the number of paired samples, \( p \) is the index for paired samples, and \( U(d, p) \) is the original sequencing numbers of paired samples in the initial node.

The result of RSA is \( \hat{D} \) nodes for which the samples of predictands in any two nodes are not significantly different. However, it is possible that there is significant local marginal monotonicity between a predictor and the predictands within a node \( d \in \{1, 2, \ldots, \hat{D}\} \).
To completely reveal the dissimilarity and similarity of the correspondence between the predictors and the predictands, the modules of RSA and RDA are alternately conducted on all nodes in the node list until the list does not change any more. Assume the number of nodes in the finalized node list be \( D \). Let these nodes be denoted as \( N^d \) and named as *end nodes* \((d = 1, 2, \ldots, D)\. We have

\[
N^d = \{XY_{U(d,p)}\}_{p=1}^{T(d)}
\]  

(4.13)

for any \( d \in \{1, 2, \ldots, D\} \) where \( T(d) \) is the number of paired samples in the \( d \)th end node, \( p \) is the index for paired samples, and \( U(d,p) \) is the original sequencing numbers of paired samples in the initial node.

**4.2.4 System Characterization and Projection**

In the process of RSA, a discontinuity between the sequenced values of the criterion predictor may be ignored. For example, the initial node \((N_1)\) is partitioned as child nodes \( N_2 \) and \( N_3 \) where

\[
N_2 = \{XY_{[j^0, 0]}\}_{T=1}^{u^0}
\]  

(4.14)  

and

\[
N_3 = \{XY_{[j^0, 0]}\}_{T=u^0+1}^{T}.
\]  

(4.15)

The range of the \( j^0 \)th predictor is \([x_{j^0, 1}], x_{j^0, T}\] for \( N_1 \), while it is \([x_{j^0, 1}], x_{j^0, u^0}\] and \([x_{j^0, u^0+1}], x_{j^0, T}\] for \( N_2 \) and \( N_3 \), respectively. There is a gap, i.e. \((x_{j^0, 1}], x_{j^0, u^0}\) and \([x_{j^0, u^0+1}], x_{j^0, T}\) \(\cup\) \([x_{j^0, 1}], x_{j^0, u^0}\). To fill this gap, the bounds of the criterion predictor in two child nodes should be extended. For the criterion predictor, both the upper bound in \( N_2 \) and the lower bound in \( N_3 \) should be
adjusted as the middle point of the gap, i.e. \((x(j, u^+1) + x(j, u)) / 2\). Through this correction, the end nodes obtained from \textit{ReMPMID} are continuously inclusive, the predictands corresponding to any combination of predictors can be projected, and an infinite loop of the \textit{ReMPMID} program can be avoided.

As the expected result of \textit{ReMPMID}, all paired samples (\(\{XY_i\}_{i=1}^T\)) representing the response from predictors (i.e., \(\{x_j\}_{j=1}^n\)) to the predictand (i.e., \(\{y_i\}_{i=1}^m\)) are discretized as a series of end nodes (\(\{N^d\}_{d=1}^D\)) that are internally inseparable, externally divergent, mutually exclusive and collectively exhaustive. Specifically, all end nodes originate from the initial node (\(N_1\)) and any paired sample in \(N_1\) belongs to a node of \(\{\tilde{N}^d\}_{d=1}^D\), thus

\[
N_1 \supseteq \cup_{d=1}^D \tilde{N}^d, \quad (4.16)
\]

\[
N_1 \subseteq \cup_{d=1}^D \tilde{N}^d \quad (4.17)
\]

and

\[
N_1 = \cup_{d=1}^D \tilde{N}^d. \quad (4.18)
\]

For any \(d_1\) and \(d_2\) in \(\{1, 2, \ldots, D\}\) where \(d_1 \neq d_2\), we have

\[
\tilde{N}^{d_1} \cap \tilde{N}^{d_2} = \emptyset, \quad (4.19)
\]

resulting from the independence of any two nodes after \textit{RDA}. Since all nodes originate from the initial node (\(N_1\)), we have

\[
N^d \subseteq N_1 \quad (4.20)
\]
for any $d \in \{1, 2, \ldots, D\}$. For any paired sample in $N_1$, e.g. $X_Y(t \in \{1, 2, \ldots, T\}$, there
must exist a $d \in \{1, 2, \ldots, D\}$ such that $XY \in N^d$. Consequently,

$$N_1 \supseteq \bigcup_{d=1}^D N^d,$$  \hspace{1cm} (4.21)

$$N_1 \subseteq \bigcup_{d=1}^D N^d,$$  \hspace{1cm} (4.22)

and equivalently

$$N_1 = \bigcup_{d=1}^D N^d.$$  \hspace{1cm} (4.23)

These characteristics of end nodes indicate that a systematic discretization of the
complicated correspondence between predictors and predictands, which was hardly
achieved by existing statistical downscaling methods, can be effectively enabled by the
ReMPMID approach.

For any given coarser-resolution climate information, the corresponding finer-
resolution climatic conditions are obtained from a clustering analysis based on the
constructed discrete statistical model, which is challenged by sample multiplicity. Let $X^*$
be any $n$-dimensional sample of predictors $(x_1, x_2, \ldots, x_n)$ within $[X_L, X_U]$ where $X^* = (x_1^*, \ x_2^*, \ \ldots, \ x_n^*)$, $x_j^* \in [x_{L_j}, x_{U_j}]$, and $j \in \{1, 2, \ldots, n\}$. Since the end nodes are collectively
exhaustive, there must be an end node in $\{N^d\}_{d=1}^D$, e.g. $N^d_{d_0} (d_0 \in \{1, 2, \ldots, D\})$, such that
$X^*$ does not exceed the ranges of predictors, i.e. $X^* \in [X_{Ld}, X_{Ud}]$. Assume the
corresponding values of the predictands in end node $N^d_{d_0}$ be samples $\{Y_{U(d_0, \rho)}\}_{p=1}^{T(d_0)}$. These samples are not significantly different, but are not exactly identical as well. There
are multiple alternative schemes to represent them. For instance, a statistic such as the
mean vector, the covariation matrix, the median vector, the minimum or maximum vector
can reflect a statistical feature of \( \{ Y_{d,p} \}_{p=1}^{T(d)} \). The distribution of \( \{ Y_{d,p} \}_{p=1}^{T(d)} \) can be translated as marginal quartiles, a joint frequency distribution function, a marginal frequency distribution function, or any other distribution fitness mean. The predictand samples in every end node can also be represented as a rough set (Orlowska, 2013), an interval set (Cheng et al., 2015b), or any other uncertainty analysis index (Sheather and Jones, 1991; Russell et al., 1995; Hair et al., 2006).

4.3 APPLICATION: ATHABASCA RIVER BASIN, CANADA

4.3.1 Study Area

The Athabasca River originates from the Columbia Icefield on the eastern slopes of the Rocky Mountains around Jasper in Alberta, Canada. It flows approximately 1500 km northeast before entering the Peace-Athabasca Delta, the largest freshwater continental river delta in North America and the home to numerous migratory birds, and draining into Lake Athabasca. This river is the longest undammed river in the Canadian prairies. The ARB covers an area of approximately 138,000 km\(^2\) over 52\(^\circ\) ~ 59\(^\circ\) N and -119\(^\circ\) ~ -107\(^\circ\) W (Figure 4.2a). Around 0.15 to 0.17 million residents distribute over 22 rural or regional municipalities, 1 city, 12 towns, and 14 Aboriginal settlements in this region. Many issues in the ARB, e.g. water availabilities, wild fires, flooding frequencies, drought durations and energy demands, are associated with climate change. Provision of scientific support for guiding local socio-economic and eco-environmental activities and eliminating occurrences of related losses under climate change desires a reliable projection of the future climatic conditions over the ARB.

A widely used tool for climatic projection is GCMs of coarser spatial resolutions, especially ones in the CMIP5. On the contrary, the ARB includes diverse hydro-climatic
regimes due to physiographical heterogeneity; snowcapped mountains, coniferous forest, mixed wood and deciduous forest are found in the uplands, whereas willow brush, shrubs, black spruce and sphagnum moss dominate the lowlands (Kerkhoven and Gan, 2006). The region can be divided into nine sub-catchments (Figure 4.2b) and twenty climate zones (Figure 4.2d) (Chapter 3). There are significant differences among the local climatic conditions over the whole region (Chapter 2), of which the spatial resolution is finer than that of CMIP5 GCMs. Failure to capture the finer-resolution regional climate may decrease the reliability and applicability of CMIP5 GCMs for the related impact studies over the ARB. Hence, the developed ReMPMID approach is applied to the ARB to enable reproducing regional climate at finer spatial resolutions, building a solid foundation for long-term reliable projection of climate change, and enhancing the reliability of decision support for human activities over this river basin.
Figure 4.2. Study Area and Data Collection.
4.3.2 Data Collection

Three climate variables, i.e. daily minimum near-surface air temperature \((T_{\text{min}})\), daily maximum near-surface air temperature \((T_{\text{max}})\), and daily precipitation \((\text{Prec})\), which are of relatively high interests for impact studies are selected as predictands to represent the regional climate over the ARB. The corresponding predictors are outputs of CMIP5 GCMs. Due to the diversity of CMIP5 GCMs, we comprehensively evaluated the multi-dimensional accuracies of them and their ensemble in reproducing historical \(T_{\text{min}}, T_{\text{max}}\) and \(\text{Prec}\) in the ARB. The related results were presented in another paper (Chapter 2). As revealed by a series of comparisons in Chapter 2, the multi-model ensemble shown relatively high accuracies, although the accuracies of any CMIP5 GCM or their ensemble varied with climate variables, geographical locations, statistical metrics, and temporal or spatial scales. Hence, the multi-model ensemble is used in this study to provide predictors.

Correspondingly, there are multiple potential GCM outputs that can be taken as the predictors in downscaling, e.g. geopotential heights, specific humidity, \(v/\text{zonal wind speeds}\), air temperature, relative humidity, \(u/\text{meridional wind speeds}\), sea level pressure, vorticity, divergence, wind speeds, wind directions, airflow strength, precipitation, clouds, thickness, latent heat flux, sensible heat flux, and snow. For most existing statistical downscaling approaches, screening of predictors is required based on techniques such as correlation analyses (Whan et al., 2014). However, this is unnecessary for the developed ReMPMID approach in which, as one advantage of this approach, predictors of significant impacts on predictands can be automatically identified from all potential ones. This advantage is helpful for avoiding biases being introduced into the downscaling process and results due to predictor screening. Based on a comprehensive review of existing
downscaling studies, a total of 31 predictors, as listed in the appendix, at different heights such as surface, 200 hpa, 300 hpa, 500 hpa, 700 hpa, 850 hpa, or 1000 hpa are used in this study.

The monthly averages of the selected predictands from 1961 to 2003 over 1615 10-km grids in the ARB (Figure 4.2c) are derived from a raster-gridded climate dataset. In consideration of data availabilities (Chapter 2), six CMIP5 GCMs are selected to generate the multi-model ensemble in this study: IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC-ESM-CHEM, MIROC5, GFDL-ESM2G and GFDL-ESM2M. Corresponding to the predictand observations, the simulation of the selected predictors by the six CMIP5 GCMs are downloaded from the website of World Data Center for Climate (http://cera-www.dkrz.de/WDCC/ui/). Accordingly, the historical ensemble means are generated based on regridding.
Table 4.1. Selected Predictors

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>clt</td>
<td>Total cloud fraction</td>
<td>%</td>
</tr>
<tr>
<td>hfls</td>
<td>Surface upward latent heat flux</td>
<td>W/m²</td>
</tr>
<tr>
<td>hfss</td>
<td>Surface upward sensible heat flux</td>
<td>W/m²</td>
</tr>
<tr>
<td>pr</td>
<td>Precipitation</td>
<td>kg<em>m⁻²</em>s⁻¹</td>
</tr>
<tr>
<td>psl</td>
<td>Sea level pressure</td>
<td>Pa</td>
</tr>
<tr>
<td>rlds</td>
<td>Surface downwelling shortwave radiation</td>
<td>W/m²</td>
</tr>
<tr>
<td>rsus</td>
<td>Surface upwelling shortwave radiation</td>
<td>W/m²</td>
</tr>
<tr>
<td>tas</td>
<td>Near-surface air temperature</td>
<td>K</td>
</tr>
<tr>
<td>tasmx</td>
<td>Daily maximum near-surface air temperature</td>
<td>K</td>
</tr>
<tr>
<td>tasmin</td>
<td>Daily minimum near-surface air temperature</td>
<td>K</td>
</tr>
<tr>
<td>prc</td>
<td>Convective precipitation</td>
<td>kg/m²/s</td>
</tr>
<tr>
<td>prsn</td>
<td>Snowfall flux</td>
<td>kg/m²/s</td>
</tr>
<tr>
<td>rlds</td>
<td>Surface downward longwave radiation</td>
<td>W/m²</td>
</tr>
<tr>
<td>rlus</td>
<td>Surface upwelling longwave radiation</td>
<td>W/m²</td>
</tr>
<tr>
<td>hus</td>
<td>Specific humidity</td>
<td>1</td>
</tr>
<tr>
<td>huss</td>
<td>Near-surface specific humidity</td>
<td>1</td>
</tr>
<tr>
<td>rlut</td>
<td>TOA outgoing longwave radiation</td>
<td>W/m²</td>
</tr>
<tr>
<td>ta</td>
<td>Air temperature</td>
<td>K</td>
</tr>
<tr>
<td>ua</td>
<td>Eastward wind</td>
<td>m/s</td>
</tr>
<tr>
<td>va</td>
<td>Northward wind</td>
<td>m/s</td>
</tr>
<tr>
<td>wap</td>
<td>omega</td>
<td>Pa/s</td>
</tr>
<tr>
<td>rhs</td>
<td>Near-surface relative humidity</td>
<td>%</td>
</tr>
<tr>
<td>uas</td>
<td>Eastward near-surface wind speed</td>
<td>m/s</td>
</tr>
<tr>
<td>vas</td>
<td>Northward near-surface wind speed</td>
<td>m/s</td>
</tr>
<tr>
<td>hur</td>
<td>Relative humidity</td>
<td>%</td>
</tr>
<tr>
<td>zg</td>
<td>Geopotential height</td>
<td>m</td>
</tr>
<tr>
<td>rhsmax</td>
<td>Surface daily maximum relative humidity</td>
<td>%</td>
</tr>
<tr>
<td>rhsmin</td>
<td>Surface daily minimum relative humidity</td>
<td>%</td>
</tr>
<tr>
<td>sfcWind</td>
<td>Daily-mean near-surface wind speed</td>
<td>m/s</td>
</tr>
<tr>
<td>sfcWindmax</td>
<td>Daily maximum near-surface wind speed</td>
<td>m/s</td>
</tr>
<tr>
<td>snw</td>
<td>Surface snow amount</td>
<td>kg/m²</td>
</tr>
</tbody>
</table>
4.3.3 Similarity Analysis

As shown in Figure 4.2b and d, the ARB includes 9 sub-catchments and 20 climate zones (Chapter 3). There are a total of 71 combinations of sub-catchments and climate zones. Every combination, named as a downscaling unit in this study, involves a few of grids. The grids-averaged observations are calculated for every predictand in every downscaling unit. For every downscaling unit, its similarity with every involved grid in every predictand is quantified by the Nash-Sutcliffe coefficient (Nash and Sutcliffe, 1970).

For every predictand that is taken as the grid-selection criterion, the grid of the highest similarity with the grids average in any one of 71 downscaling unit is identified. Correspondingly, the similarity between every downscaling unit and the selected grid is estimated for every predictand. The related results are shown in Figure 4.3. It is indicated that the Nash coefficients are higher than 0.99 for most downscaling units. This implies that the selected grids are of very high similarity with the corresponding downscaling units which constitute the entire river basin. A comparison of the selection results for different predictands reveals that the spatial heterogeneity of precipitation is higher than that of temperature and that there are differences among selected grids among predictands. In consideration of all predictands, there are a total of 137 different grids. In the following studies, climate downscaling is conducted on these grids representing the regional climate over the ARB and is extended to the entire river basin according to the results of similarity analysis.
Figure 4.3. Grids-Similarity Analysis Results.
4.3.4 Sensitivity Analysis

The ReMPMID approach includes two parameters: the statistical significance level ($\alpha$) and the minimum partition row number ($N_{\text{min}}$). From the viewpoint of statistical analysis, the conventional values of $\alpha$ are 0.01, 0.05 and 0.10. The theoretical value range of this parameter is from 0 to 1. As for the second parameter ($N_{\text{min}}$), it ranges from 3 to a larger positive integer that depends on the total sample size in various cases. Calibration of both parameters are required for identifying the highest accuracy of the ReMPMID approach in climate downscaling. In order to eliminate redundant computations, a parameter screening study is conducted. The alternative values of the two parameters are selected as \{0.01, 0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90\} and \{3, 4, ..., 30\}, respectively, constituting a set of 308 parameter combinations. In consideration of the potential variation of the optimal parameter combinations with spatial locations and predictands, three representative grids distributed across the ARB are selected for every predictand. The dataset is divided as two parts: [1961, 1990] for constructing a downscaling model and [1991, 2003] for verifying the model’s accuracy. The modeling accuracy corresponding to every period is quantified as the Nash-Sutcliffe coefficient ($Nash$ and Sutcliffe, 1970). The related results are displayed in Figure 4.4.

It is revealed that the modeling accuracies are not sensitive to the parameter of $\alpha$ for any predictands and significantly change with $N_{\text{min}}$. There is a significantly decreasing monotonic relationship between period 1 modeling accuracies and $N_{\text{min}}$. This is related with the law that a lower value of $N_{\text{min}}$ can help the ReMPMID approach capture more local characteristics of the correspondence between predictands and predictors in the first period. In addition, it seems that the modeling accuracies vary with $N_{\text{min}}$ nonlinearly in
the second period. The optimal parameter combinations are not identical for all predictands and grids. There are significant differences among the downscaling efficiencies of every predictands at different grids. Therefore, the parameter of $\alpha$ is set as 0.10 due to its insignificant impacts, and the integer parameter of $N_{min}$ ranges from 3 to 30 from which the optimal value is detected.
Figure 4.4. Sensitivities of Parameters in \textit{ReMPMID}.
4.3.5 Model Configuration

It is also illustrated in Figure 4.4 that the parameter combination maximizing the modeling accuracy in the first period can hardly guarantee maximization of the modeling accuracy in the second period. There is a tradeoff between modeling accuracies in different periods. Correspondingly, the conventional parameter calibration strategy that the parameter combination of the highest accuracy in one period is selected as the most desired one is unreasonable in this case. Instead, a bi-period calibration strategy is used in this study. An ReMPMID model is built for the first calibration period ([1961, 1990]) under every parameter combinations, its effectiveness is estimated as the simulation accuracy over the second calibration period ([1991, 1995]), and the optimal parameter combination is the one showing the highest accuracy in the latter period. Subsequently, the accuracy of the constructed ReMPMID model under the optimal parameter combination is verified in another period ([1996, 2003]).

Multiple statistical metrics are available for quantifying the multi-dimensional performances of a downscaling model. Few of them can outperform any others in various aspects such as reproducing the temporal trends, the absolute magnitudes, and the relative magnitudes of a climate variable. Hence, three statistical metrics are selected as the indicators for the modeling accuracies of the ReMPMID approach in the three aspects. They are the Pearson product-moment correlation coefficient (PC) (Pearson, 1895), the Nash-Sutcliffe coefficient (Nash) (Nash and Sutcliffe, 1970), and the mean absolute error (MAE) (Willmott and Matsuura, 2005). Let \( \{s_t\}_{t=1}^T \) and \( \{o_t\}_{t=1}^T \) represent the simulation and observations of a predictand, respectively where \( T \) is the number of temporal units. The selected metrics can be formulated as \( PC = \frac{\sum_{t=1}^T(s_t - \sum_{s=1}^T s_t/T)(o_t - \sum_{o=1}^T o_t/T)}{\sum_{t=1}^T (s_t - \sum_{s=1}^T s_t/T)^2} \).
\[
- \frac{\sum_{t=1}^{T} t \left( s_t - \sum_{t=1}^{T} T \cdot s_t / T \right)}{\left( \sum_{t=1}^{T} t \left( o_t - \sum_{t=1}^{T} T \cdot o_t / T \right)^2 \right)^{0.5}}, \quad \text{Nash} = 1 - \frac{\sum_{t=1}^{T} t \left( o_t - s_t \right)^2}{\sum_{t=1}^{T} t \left( o_t - \sum_{t=1}^{T} T \cdot o_t / T \right)^2}, \quad \text{and} \quad \text{MAE} = \frac{\sum_{t=1}^{T} t \left( o_t - s_t \right)}{T}, \quad \text{respectively}.
\]

4.4 RESULTS ANALYSIS AND DISCUSSIONS

4.4.1 Calibration

The accuracies of the developed ReMPMID approach in reproducing the selected predictands over all selected grids at a higher spatial resolution in the calibration period are relatively high as shown in Figure 4.5. Corresponding to **Tmin**, **Tmax** and **Prec**, the lowest **PC** reaches 0.9928, 0.9931 and 0.9466, while the lowest **Nash** reaches 0.9857, 0.9863 and 0.8960, respectively. It is illustrated that the two modeling-accuracy indicators are of an increasing monotonic relationship in this case study. In the calibration process, an ReMPMID model of high accuracies in simulating the temporal trends of predictands may also be capable of well reproducing the relative magnitudes of predictands. In addition, the accuracies of ReMPMID in the calibration period vary with grids for every predictand, and these spatial variations are the most significant for **Prec** while being the least significant for **Tmax**. Furthermore, it is also revealed that the modeling accuracies of the ReMPMID approach decrease from **Tmax**, **Tmin** to **Prec** for most grids in the ARB. For some other grids such as ones at the south of Fort McMurray, the ReMPMID approach shows relatively higher accuracies for **Prec** than those for either **Tmax** or **Tmin**.
Figure 4.5. Accuracies of the Calibrated ReMPMID Model.
4.4.2 Verification: Multi-Year Climate Variability and Magnitudes

The accuracies of the developed ReMPMID model for downscaling climate over the ARB are verified. The results related with simulation of the multi-year variability and magnitudes of the three selected predictands at a spatial resolution of 10 kilometres are displayed in Figure 4.6.

It is revealed that, in general, the modeling accuracies of ReMPMID are of significant similarity among the selected modeling-accuracy indicators for most grids, while showing dissimilarity for some local zones. For instance, the accuracies of ReMPMID in modeling either the multi-year variability, the multi-year relative magnitude, or the multi-year absolute magnitude of $T_{min}$ is relatively low for grids around High Prairie, moderate for grids at the upstream of Fort McMurray, and high for grids around Whitecourt. On the other hand, the multi-year variability of $T_{min}$ is reproduced with relatively high accuracies for grids at the northeast of Fort McMurray; for these grids, the accuracies of ReMPMID in simulating the relative magnitudes of $T_{min}$ are moderate, while those for the absolute magnitudes of $T_{min}$ are relatively low. This implies that, although a single statistical metrics is capable of evaluating the multi-dimensional performances of the ReMPMID approach over most geographic locations, multiple ones are desired for a systematic evaluation in large river basins such as the ARB in the verification process.

In addition, a comprehensive comparison of the multi-dimensional performances of the ReMPMID approach in the verification process reveals that the accuracies of this approach in reproducing high-resolution local climate over the ARB are relatively high for both $T_{min}$ and $T_{max}$ and acceptable for Prec. For instance, the $PC$ ranges from 0.9633 to
0.9785 for $T_{min}$, from 0.9600 to 0.9804 for $T_{max}$, and from 0.4643 to 0.7796 for $Prec$. The grids where the accuracies of $ReMPMID$ in modeling $Prec$ are relatively low only account for a small proportion of the entire river basin. These grids concentrate in the south or north end where the precipitation intensity is the highest or the lowest compared with other grids (Chapter 3) and where the intensity of socio-economic activities is relatively low. On the contrary, the $ReMPMID$ approach shows higher accuracies for modeling $T_{min}$ and $T_{max}$ over these grids than the grids around High Prairie. Furthermore, the $ReMPMID$ approach shows the highest accuracies in modeling $T_{max}$ and the lowest for $Prec$, which is a common practice in existing downscaling studies and may be related with the relatively high complexity of precipitation processes.

The $MAE$ is introduced to quantify the multi-year-averaged deviation of the downscaling results in comparison with observations. Figure 4.6c, Figure 4.6f and Figure 4.6i illustrate that the $ReMPMID$ approach tends to overestimate the multi-year-averaged magnitudes of $T_{min}$, $T_{max}$ and $Prec$ over almost all grids except the ones around High Prairie where $Prec$ is underestimated. In general, the extent to which the multi-year-averaged climate magnitude is overestimated decreases from $T_{min}$, $T_{max}$ to $Prec$. The spatial distributions of the multi-year-averaged climate-simulation deviations are of significant dissimilarity for the selected predictands. For instance, the deviations are relatively high over the grids around Jasper and Hinton for every predictand, while being high for $T_{max}$, moderate for $T_{min}$ and low for $Prec$ over the grids at the south of Fort McMurray. When the $ReMPMID$ approach is applied to guide impact studies or adaptation practices, these deviations should be taken into account for avoiding unreliability of scientific support.
Figure 4.6. Verification Accuracies of ReMPMID in Simulating Multi-Year Climate Variability and Magnitudes.
4.4.3 Verification: Monthly or Seasonal Climate Variability and Magnitudes

The multi-dimensional performances of the ReMPMID approach in reproducing the monthly or seasonal variability and magnitudes of regional climate over the ARB at a high spatial resolution in the verification process are examined. The related result are presented in Figure 4.7 to Figure 4.10. A comparison of these figures reveals that there are significant differences among the modeling accuracies of ReMPMID among seasons. In terms of simulating the temporal variability and relative magnitudes of regional climate, the overall accuracies of ReMPMID decrease from springs, autumns, summers to winters for \(T_{min}\) and \(T_{max}\) and from autumns, winters to summers and springs for \(Prec\) in consideration of all grids in the ARB. As for the simulation of the multi-year-averaged seasonal absolute magnitudes of predictands, the overall accuracies of ReMPMID decrease from autumns, summers, winters to springs for \(T_{min}\), from autumns, summers, springs to winters for \(T_{max}\), and from winters, springs, autumns to summers for \(Prec\).

In addition, these figures also illustrate that the ReMPMID approach tends to over-estimate \(T_{min}\) and \(T_{max}\) in springs and winters and \(Prec\) in summers and autumns over all or almost all grids while underestimating \(T_{min}\) and \(T_{max}\) in summers and autumns and \(Prec\) in springs and winters. Meanwhile, the spatial variability of the modeling accuracies of ReMPMID decreases from \(Prec\) to \(T_{min}\) and \(T_{max}\) due to relatively high spatial heterogeneity of \(Prec\). The distributed modeling accuracies of ReMPMID are of significant similarity for \(T_{min}\) and \(T_{max}\) and significant dissimilarity between \(Prec\) and \(T_{min}/T_{max}\). For instance, the ReMPMID approach shows relatively high \(PC\) values for \(T_{min}\), moderate for \(T_{max}\), and low for \(Prec\) over the grids around High Prairie in summers.
In the meantime, the spatial patterns of the three selected modeling-accuracy indicators remain for most grids in the ARB; one of exceptions is the grids at the north of the ARB on which, while the accuracies of the temporal variability of $T_{max}$ in springs are lower than those on other grids, the absolute magnitude is well captured.

Furthermore, a cross comparison of the modeling accuracies of ReMPMID in modeling the monthly variability and magnitudes of the selected predictands over all grids in the ARB reveals that there are significant differences among the accuracies of ReMPMID under various combinations of predictands, months, statistical metrics, and geographic locations. The overall accuracies of ReMPMID at the monthly scale are significantly decreased and of increased temporal heterogeneities compared with those at the seasonal scale. In comparison with Prec, the ReMPMID shows similar accuracies in modeling $T_{min}$ and $T_{max}$ which are of inherent correlations. In consideration of all grids in this river basin, the highest modeling accuracies of ReMPMID for $T_{min}/T_{max}$ and Prec are in December and August, while the lowest in May and July, respectively. The spatial heterogeneities of the accuracies of ReMPMID are higher for Prec, especially in January, March and July, than those for $T_{min}$ and $T_{max}$. The overall accuracies of ReMPMID for Prec are higher than those for $T_{min}$ and $T_{max}$ in January, February, March, May, August, September and November, while lower in April, June, July and December; in October and in comparison with $T_{min}$ and $T_{max}$, the temporal trend of Prec is better captured and the monthly relative magnitude is worse captured. When the ReMPMID approach is applied to the related studies or practices, these complicated differences of this approach’s
modeling accuracies among various months, climate variables, geographic locations and statistical metrics should be paid high attentions.
Figure 4.7. Verification Accuracies of ReMPMID in Simulating Climate Variability and Magnitudes in Winter.
Figure 4.8. Verification Accuracies of ReMPMID in Simulating Climate Variability and Magnitudes in Spring.
Figure 4.9. Verification Accuracies of ReMPMID in Simulating Climate Variability and Magnitudes in Summer.
Figure 4.10. Verification Accuracies of ReMPMID in Simulating Climate Variability and Magnitudes in Autumn.
4.4.4 Optimal Values of Nmin

In the section of 3.4, it has been disclosed that the parameter of the statistical significance level ($\alpha$) had no significant impact on the modeling accuracies of the developed ReMPMID approach; however, the parameter of the minimum partition row number ($N_{min}$) was significant for ReMPMID’s performances. One remaining problem is what the optimal value of $N_{min}$ is in downscaling studies. In this case study, the optimal value of $N_{min}$ for every predictand and representative grid is identified and the corresponding accuracies of ReMPMID in the calibration and verification periods are calculated. The related results in which the modeling accuracy is quantified by the Nash coefficient are presented in Figure 4.11. It is illustrated again that the calibration accuracies decrease with the climbing of the optimal $N_{mins}$ and that there is not a significant monotonic relationship between the optimal $N_{mins}$ and the verification accuracies. In addition, an analysis of the frequencies of the optimal $N_{mins}$ reveals that the optimal selection of $N_{min}$ varies with grids and predictands and shows higher uncertainties for $T_{min}$ and $T_{max}$ than $Prec$. If one initial value of $N_{min}$ is desired for testing the reliability of the ReMPMID approach in a similar downscaling study, it would be 15 for $T_{min}$, 18 for $T_{max}$, and 3 for $Prec$. This scheme may not directly help identify the truly optimal values of $N_{min}$ for any downscaling practice, but at least provides an initial value of $N_{min}$ that deserves a test.
Figure 4.11. Optimal Values of $N_{\text{min}}$. 
4.4.5 Uncertainty Analyses

One advantage of the developed ReMPMID approach is that it can mitigate data uncertainties, irregular nonlinearities, and multivariate dependencies in the correspondence between coarser-resolution GCM outputs and finer-resolution climate observations. These complexities are incorporated into the ReMPMID process and results. As a result, the ReMPMID approach enables climate simulation under uncertainties. For any given combination of predictor values, the corresponding values of predictands are the multi-dimensional samples in an end node. These various samples represent the potential values of predictands under system complexities. Corresponding to four seasons and the three selected predictands, the value ranges of predictand simulation over all grids in the ARB are calculated as presented in Figure 4.12. It is illustrated that the uncertainties of high-resolution climate simulation gradually increase from summers, autumns, springs to winters for Tmin and Tmax and from winters, springs, autumns to summers for Prec. In addition, there are significant spatial heterogeneities for the simulation uncertainties of every predictand among all grids in the ARB. When the ReMPMID approach is applied to downscaling practices, the users are not suggested to only focus on the averaged climate conditions that are just a simplification of abundant simulation results under uncertainties and other complexities. Instead, the potential ranges or even distributions of simulation results should be paid attentions as well to avoid provision of dogmatic climate simulation for impact or adaptation practices and to facilitate various risk assessment and management (Dong et al., 2011, 2012, 2013, 2014a,b,c, 2015; Cheng et al., 2009, 2015a,b,c, 2016a,b) under climate change.
Figure 4.12. Uncertainties in the Results of ReMPMID.
4.5 SUMMARY

In this study, an advanced approach \((ReMPMID)\) was proposed for supporting climate downscaling under complexities such as data uncertainties, nonlinear predictors-predictands correspondences, predictands’ interactions, non-normal distributions, spatial homogeneities, and temporal nonstationarities. The principle, innovations and technical details of the \(ReMPMID\) approach were discussed. Subsequently, this approach was applied to the \(ARB\), a large river basin on the Canadian prairies that was closely connected with climate change, to verify methodological effectiveness and facilitate local impact or adaptation practices. Specifically, based on predictor and predictand (i.e. \(T_{min}\), \(T_{max}\) and \(Prec\)) selection, data collection and processing, a grids-similarity analysis, a sensitivity analysis, and parameter calibration, a \(ReMPMID\) model was constructed for every selected grid in the \(ARB\). The multi-dimensional modeling accuracies of the \(ReMPMID\) approach were verified and examined in various aspects. The optimal parameter values and the uncertainties in high-resolution climate simulation were analyzed through a series of comparisons.

From the viewpoint of methodology, this study contributed an additional reliable approach or framework, as an integration of statistical inferential methods, to enable downscaling under data uncertainties, nonlinear predictors-predictands correspondences, predictands’ dependencies, non-normal distributions, spatial homogeneities, and temporal nonstationarities. A similarity-analysis-based grid selection method was developed to facilitate identification of the grids that can effectively represent the regional climate, avoiding redundant computations and improving computational efficiencies. A bi-period calibration strategy was proposed to mitigate the challenge of over-parameterization in
the calibration process for statistical downscaling approaches including ReMPMID. Multiple statistical metrics were employed to quantify the modeling accuracies of ReMPMID in various aspects that may be independent in some cases, which was helpful for avoiding the partiality in evaluating a downscaling approach’s performance through a single measure. A series of sensitivity analyses were conducted to reveal the impacts of the parameters of ReMPMID on modeling results and accuracies, facilitating this approach being extensively applied to various impact and adaptation studies. The analysis on the uncertainties existing in the results of ReMPMID disclosed the necessity of not selectively using the averaged simulation and of paying attentions on abundant simulation results that were meaningful for risk assessment and management.

Through the case study of climate downscaling over the ARB, a range of findings were revealed as summarized below. These findings would be much helpful for gaining insights into the developed ReMPMID approach and the regional climate in the ARB or neighbouring regions.

- The spatial heterogeneity of precipitation was higher than that of temperature in the ARB. The overall accuracies of the ReMPMID approach in reproducing high-resolution local climate over the ARB were relatively high for both Tmin and Tmax and acceptable for Prec. There were significant differences among the accuracies of ReMPMID under various combinations of predictands, months, seasons, statistical metrics, and geographic locations.
- In the calibration process, the modeling accuracies of ReMPMID were of significant similarity among the selected modeling-accuracy indicators for most grids,
while showing dissimilarity for some local zones. Although a single statistical metrics was capable of evaluating the multi-dimensional performances of the ReMPMID approach over most geographic locations, multiple ones were desired for a systematic evaluation in large river basins such as the ARB in the verification process.

- The ReMPMID approach tended to overestimate the multi-year-averaged magnitudes of Tmin, Tmax and Prec over almost all grids except the ones around High Prairie where Prec was underestimated. The extent to which the multi-year-averaged climate magnitude was overestimated decreased from Tmin, Tmax to Prec. Meanwhile, this approach might over-estimate Tmin and Tmax in springs and winters and Prec in summers and autumns over all or almost all grids while underestimating Tmin and Tmax in summers and autumns and Prec in springs and winters.

- In terms of simulating the temporal variability and relative magnitudes of regional climate, the overall accuracies of ReMPMID decreased from springs, autumns, summers to winters for Tmin and Tmax and from autumns, winters to summers and springs for Prec in consideration of all grids in the ARB. As for the simulation of the multi-year-averaged seasonal absolute magnitudes of predictands, the overall accuracies of ReMPMID decreased from autumns, summers, winters to springs for Tmin, from autumns, summers, springs to winters for Tmax, and from winters, springs, autumns to summers for Prec.

- In consideration of all grids in this river basin, the highest modeling accuracies of ReMPMID for Tmin/Tmax and Prec were in December and August, while the lowest in May and July, respectively. The overall accuracies of ReMPMID for Prec were higher
than those for $T_{min}$ and $T_{max}$ in January, February, March, May, August, September and November, while lower in April, June, July and December.

- The modeling accuracies was not sensitive to the parameter of the statistical significance level ($\alpha$) for any predictands and significantly changed with another parameter: the minimum partition row number ($N_{min}$). The calibration accuracies decreased with the climbing of the optimal $N_{mins}$ and there was not a significant monotonic relationship between the optimal $N_{mins}$ and the verification accuracies. The optimal selection of $N_{min}$ varied with grids and predictands and shown higher uncertainties for $T_{min}$ and $T_{max}$ compared with $Prec$.

- The uncertainties of high-resolution climate simulation through the ReMPMID approach gradually increased from summers, autumns, springs to winters for $T_{min}$ and $T_{max}$ and from winters, springs, autumns to summers for $Prec$.

On the other hand, there are many potential opportunities of improving the ReMPMID approach, revealing the mechanism of regional climate, and enhancing the reliability of scientific support for climate-change impact or adaptation practices. The representative ones include but are not limited to the followings. A continuous function that could enable the reversible transformation between any non-normal distribution and a normal distribution may be more effective than the DDT approach used in this study. The impacts of distribution transformation on downscaling results and accuracies deserves an attempt. There may be a more reasonable scheme of nonlinearity characterization than the current one that the nonlinearity between predictors and predictands is represented by the principal monotonicity in predictor-predictand paired samples. Incorporating variance analyses into the framework of ReMPMID may facilitate a comprehensive evaluation of
the difference between any two groups of multi-dimensional predictand samples. A scientific scheme, possibly integration with physically-based downscaling models, is desired for mitigating one challenge of the ReMPMID approach that the simulation results are trapped in the historical records. Adjustment of the ReMPMID process may be helpful for eliminating the high computational loads of this data-driven approach. The methods of global sensitivity analysis can help reveal complicated interactions among the parameters, boundary conditions and modeling results of the ReMPMID approach. In-depth analyses on the causes of variation and diversity of the modeling accuracies of ReMPMID may help advance this approach’s performances. Attentions should be paid on building the connections between real-world climate systems and the ReMPMID approach which is categorized as a statistical or data-driven downscaling approach. Developing a dynamic version of the ReMPMID approach to adapt to the temporal variability of regional climate may be one of schemes that could improve this approach. Consideration of various uncertainties except data uncertainties can also be an alternative direction in the following studies. More case studies such as enabling long-term high-resolution climate projection over the ARB are required to verify the reliability of this approach in various downscaling studies or practices, reveal its limitations for further improvement, and extend its potential contributions.
CHAPTER 5
HIGH-RESOLUTION PROJECTION OF 21ST CENTURY CLIMATE OVER
ATHABASCA RIVER BASIN

5.1 BACKGROUND

General circulation models (GCMs) are a widely used tool to simulate the present and future climates under assumed greenhouse gas emission scenarios, both in space and time (Mehrotra and Sharma, 2010). Due to coarse grid resolutions, GCMs can hardly represent the features and dynamics of local climatic conditions which are desired for climate-change impact or adaptation studies. To fill this gap, a number of downscaling techniques are developed for converting GCM outputs from a coarse spatial scale to a local or regional scale. These techniques consist of dynamical downscaling which simulates near-surface atmospheric processes at a finer scale by regional climate models (e.g., Giorgi et al., 1999; Fowler et al., 2007) and statistical downscaling which builds statistical relationships between large-scale GCM outputs (i.e. predictors) and regional climate (i.e. predictands) by statistical analyses (e.g., Maran et al., 2010; Schoof, 2013). The biases and other deficiencies of GCMs are usually inherited in dynamic downscaling, and statistical downscaling is computationally inexpensive so that it can be employed to downscale the outputs of GCMs or even regional climate models (Sunyer et al., 2012). Consequently, many studies have been conducted on statistical downscaling over the past years, resulting in a large number of the related methods (e.g., Yarnal et al. 2001; Charles et al., 2004; IPCC, 2007; Ekström et al., 2015).

It was stated in Chapter 4 that few methods were dedicated to support statistical downscaling under complexities such as data uncertainties, nonlinear predictors-
predictands correspondences, predictands’ interactions, non-normal distributions, spatial homogeneities, and temporal nonstationarities. Accordingly, Cheng et al. (2016e) developed an advanced statistical downscaling approach named as a recursive multivariate principal-monotonicity inferential downscaling (ReMPMID) approach. The key of this approach was a module of ReMPMID that could enable climate downscaling under these complexities through integrating advanced statistical inferential techniques within an ingeniously-designed framework. The most significant innovation of this approach was that, according to the effect of principal monotonicity, the complicated correspondence between atmospheric variables and climate variables was systematically discretized as a series of internally inseparable, externally divergent, mutually exclusive and collectively exhaustive end nodes. This approach also consisted of a series of other modules, e.g. spatial similarity analysis, bi-period calibration, multi-criteria analysis, sensitivity analysis and uncertainty analysis, which were helpful for enhancing methodological applicability and reliability. A case study demonstrated that the overall accuracies of this approach were relatively satisfactory, especially for temperature. This approach could be an alternative statistical downscaling tool for enabling high-resolution climate projection.

Meanwhile, reliable high-resolution climate projection are highly desired for impact or adaptation studies over the ARB, Canada. The main stream of this river basin, the Athabasca River, originates from the Columbia Icefield on the eastern slopes of the Rocky Mountains around Jasper in Alberta, Canada, flows approximately 1500 km northeast, and drains into the Peace-Athabasca Delta which is the largest freshwater continental river delta in North America and the home to numerous migratory birds. The ARB covers an area of approximately 138,000 km² in which around 0.15 to 0.17 million residents are
living. Many issues in the $ARB$, e.g. water availabilities, wild fires, flooding, droughts and energy, are associated with climate change. $GCM$s-based climate projection for addressing these issues are challenged by the diverse finer-scale climatic regimes over this region. Previously, very few studies were conducted to enhance the resolution of climate projection over the $ARB$. In (Erler et al., 2015), the Weather Research and Forecasting Model (Skamarock and Klemp, 2008) was applied to support dynamical downscaling of climatic conditions over the western Canada (including the $ARB$). That study was associated with multi-aspect simplifications. Only one $GCM$, i.e. the Community Earth System Model (Gent et al., 2011), was used to drive the $WRF$ model without a comprehensive evaluation of the multi-dimensional performances of multiple $GCM$s. The projection period was only from 2045 to 2060; the future status of climate over the $ARB$ in the 21st century could not be provided. Only one emission scenario, i.e. $RCP$ 8.5, which was just one of multiple potential scenarios and relatively extreme was considered; this could hardly guarantee the systematization of projection results. These simplifications would not be helpful for increasing the reliability of long-term planning of impact or adaptation practices.

Therefore, this study aims to enable high-resolution projection of 21st century climate over the $ARB$, Canada under the four $RCP$ scenarios using the $ReMPMID$ approach (Chapter 4). Based on previous studies such as $GCM$ evaluation (Chapter 2), climate classification (Chapter 3) and model assessment (Chapter 4), we will systematically examine the future status of climatic conditions over the $ARB$ at a finer spatial resolution. The framework of this study is displayed in Figure 5.1 for readers’ convenience. Specifically, section 5.2 will focus on presenting the results related with
system analyses, data collection, method application, and climate characterization. In section 5.3, the future climate conditions over the *ARB* will be analyzed from multiple dimensions, e.g. spatial averages, spatial variability, bi-decadal variability, and intra-annual variability. A series of issues regarding uncertainties, modeling biases, climate-change impacts, and potential extensions of this study will be examined in section 5.4.
Figure 5.1. Framework of Chapter 5.
5.2 DATA AND METHOD

5.2.1 Study Area

The Athabasca River originates from the Columbia Icefield on the eastern slopes of the Rocky Mountains around Jasper in Alberta, Canada. It flows approximately 1500 km northeast before entering the Peace-Athabasca Delta, the largest freshwater continental river delta in North America and the home to numerous migratory birds, and draining into Lake Athabasca. This river is the longest undammed river in the Canadian prairies. The ARB covers an area of approximately 138,000 km² over 52° ~ 59° N and -119° ~ -107° W (Figure 5.2a) in the provinces of Alberta and Saskatchewan. Around 0.15 to 0.17 million residents distribute over 22 rural or regional municipalities, 1 city, 12 towns, and 14 Aboriginal settlements in this region.

Many issues in the ARB, e.g. water availabilities, wild fires, flooding frequencies, drought durations and energy demands, are associated with climate change. For instance, the Athabasca Oil Sands at the downstream of the ARB is one of the largest remaining reserves of petroleum on the planet and a central point of friction in Canadian, American and global climate politics and policy (Leong and Donner, 2015). Surface water use by oil sands mining operations accounts for the largest sectoral water allocations (62 %) and actual water use (57 %) in the ARB (AMEC Earth and Environmental, 2007). As mining activity expands, surface water use demand is projected to rapidly increase, adding pressure to water availability in the ARB (Natural Resources Canada, 2009). A relatively large amount of water use for mining in winters is threatening the health of habitats in the lower-stream wetlands. In addition, it appears that climate warming may be an important influencing factor for wild fires in Canada. Increased burned areas are associated with
warmer temperatures (Flannigan et al., 2005). In 2016, a fire in Fort McMurray, a city at the downstream of the ARB in Alberta, forced more than 80,000 people to flee and is recognized as one of the most devastating disasters in Alberta’s history.

A reliable projection of the future climatic conditions is much desired for guiding local socio-economic and eco-environmental activities and eliminating occurrences of related losses under climate change. Currently, GCMs in the CMIP5 are a widely used tool for climatic projection. On the other hand, the ARB includes diverse hydro-climatic regimes due to physiographical heterogeneity; snowcapped mountains, coniferous forest, mixed wood and deciduous forest are found in the uplands, whereas willow brush, shrubs, black spruce and sphagnum moss dominate the lowlands (Kerkhoven and Gan, 2006). The region can be divided into nine sub-catchments (Figure 5.2b) and twenty climate zones (Figure 5.2d) (Chapter 3). There are significant differences among the local climatic conditions over the whole region (Chapter 2). In comparison, the spatial resolution of GCMs is relatively too coarse to capture the finer-resolution regional climate. This may lead to a decrease of the reliability and applicability of GCMs-based projection for the related impact studies over the ARB.

Recently, an advanced statistical downscaling approach, i.e. recursive multivariate principal-monotonicity inferential downscaling (ReMPMID) (Chapter 4), was developed for enhancing the spatial resolution of climate simulation. As indicated in Chapter 4, the overall modeling accuracy of ReMPMID was relatively satisfactory although it varied with multiple factors such as climate variables, geographic locations, temporal scales, and accuracy indicators. In this study, this approach is applied to enable high-resolution projection of climatic conditions in the 21st century over the ARB, building a solid
foundation for long-term planning of climate-change impact or adaptation practices in this region.
Figure 5.2. Study Area and Data Collection.
5.2.2 Data

Three climate variables, i.e. daily minimum near-surface air temperature \((T_{min})\), daily maximum near-surface air temperature \((T_{max})\), and daily precipitation \((Prec)\), which are of relatively high interests for impact studies are selected as predictands to represent the regional climate over the \(ARB\). The corresponding predictors are mainly the outputs of \(CMIP5\) GCMs. As revealed in Chapter 2, the multi-model ensemble shown relatively high accuracies, although the accuracies of any \(CMIP5\) GCM or their ensemble varied with climate variables, geographical locations, statistical metrics, and temporal or spatial scales. Meanwhile, the \(CMIP5\) GCMs for which all atmospheric variables of potentially significant impacts for regional climate in continental regions such as the \(ARB\) are available are IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC-ESM-CHEM, MIROC5, GFDL-ESM2G and GFDL-ESM2M (Table 1). Hence, the ensemble mean of the six \(CMIP5\) GCMs is used in this study to support projection of predictors. In consideration of the advantage of \(ReMPMID\) in automatically identifying the atmospheric variables of significant impacts on climate variables, a total of 31 GCM outputs, as listed in the appendix, at different heights such as surface, 200 hpa, 300 hpa, 500 hpa, 700 hpa, 850 hpa, or 1000 hpa are selected as predictors in this study. The four future socioeconomic scenarios named as \(RCPs\) 2.6, 4.5, 6.0, and 8.5 in the \(CMIP5\) simulation (Moss et al. 2010; van Vuuren et al. 2011) are considered in this study.

The monthly averages of the selected predictands from 1961 to 1990 over 1615 10-km grids in the \(ARB\) (Figure 5.2c) are derived from a raster-gridded climate dataset. These data will be used for analyzing climate anomalies in the following sections. To drive the constructed \(ReMPMID\) model for high-resolution projection in this study, the projection
of the selected predictors from 2020 to 2100 under the four emission scenarios and by the six selected CMIP5 GCMs over the ARB are obtained from the website of World Data Center for Climate (http://cera-www.dkrz.de/WDCC/ui/). Subsequently, these datasets are used to calculate the multi-model ensemble means of the selected predictors in the 21st century, providing inputs for ReMPMID-based high-resolution climate projection.
Table 5.1. *CMIP5* Models Used in Chapter 5.

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Resolution (Lon × Lat)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPSL-CM5A-LR</td>
<td>3.75° × 1.875°</td>
<td>Institut Pierre-Simon Laplace, France</td>
</tr>
<tr>
<td>IPSL-CM5A-MR</td>
<td>2.5° × 1.259°</td>
<td></td>
</tr>
<tr>
<td>MIROC5</td>
<td>1.406° × 1.406°</td>
<td>Atmosphere and Ocean Research Institute, Japan</td>
</tr>
<tr>
<td>MIROC-ESM-CHEM</td>
<td>2.812° × 2.812°</td>
<td></td>
</tr>
<tr>
<td>GFDL-ESM2M</td>
<td>2.5° × 2.0°</td>
<td>Geophysical Fluid Dynamics Laboratory, NOAA, USA</td>
</tr>
<tr>
<td>GFDL-ESM2G</td>
<td>2.5° × 2.0°</td>
<td></td>
</tr>
</tbody>
</table>
5.2.3 Method

The foundation of high-resolution climate projection in this chapter is the ReMPMID framework which was developed in Chapter 4. This framework is an integration of an ReMPMID approach, a grids-similarity analysis, a sensitivity analysis, a bi-period calibration strategy, a multi-criteria analysis, and an uncertainty analysis. As the key of this framework, the ReMPMID approach enables recursive analyses of the dissimilarity and similarity among the paired samples of predictors and predictands through an ingenious incorporation of advanced statistical inferential techniques: the discrete distribution transformation (Cheng et al., 2016a,b), the ZS test (Zhou and Shao, 2014) and the MNV test (Krishnamoorthy and Yu, 2004). Through this approach, the complicated responsive relationship between large-scale atmospheric variables and regional- or local-scale climate variables is systematically characterized as a discrete correspondence simulation model. This model is composed of a series of groups of the predictors-predictands paired samples that are internally inseparable, externally divergent, mutually exclusive and collectively exhaustive. In addition, the module of grids-similarity analysis is incorporated into the framework of ReMPMID to analyze the similarity among all grids in a large study area based on climate classification (Chapter 3), geographical conditions and a similarity measure, which is helpful for eliminating redundant computations in downscaling practices. The most representative advantage of ReMPMID is the skill of enabling climate downscaling under complexities of data uncertainties (Cheng et al., 2015a,b,c), nonlinear predictors-predictands correspondences, predictands’ dependencies, non-normal distributions, spatial homogeneities, and temporal nonstationarities.
It was revealed in Chapter 4 that in general the modeling accuracies of the ReMPMID approach are high for temperature and acceptable for precipitation. An ReMPMID model was constructed for simulating the correspondence between the finer-resolution observations of every predictand and the multi-model ensemble simulation of the predictors for every 10-km grid over the ARB. This model is equivalent to a systematic discretization of the multi-dimensional space of the predictors. The space is deconstructed as a range of sub-spaces that correspond to particular value ranges of every predictor. Every sub-space also corresponds to a series of multi-dimensional samples of the predictands. When this model is used for projecting the future climatic conditions, any given multi-dimensional sample of predictors is taken as the input of the model. The corresponding sub-space of predictors is located based on a discriminant analysis. Accordingly, the corresponding multi-dimensional predictands samples can be identified and used for estimating the mean values or distributions of the predictands for the given combination of the values of the predictors. Through this projection scheme, the mean value and the discrete distribution (expressed as percentiles) of every predictand in every temporal and spatial unit under every emission scenario are generated to support climate-change analyses in the 21st century over the ARB.

5.2.4 Climate Indices

The projection result consists of the future climatic conditions over the ARB in multiple dimensions: 4 emission scenarios, 3 climate variables, 1615 grids, 80 years (or 4 bi-decades: [2021, 2040], [2041, 2060], [2061, 20880], [2081, 2100]), 12 months, and 101 percentiles. This study focuses on analyzing the monthly averaged climatic conditions. For every emission scenario and every climate variable, the following indices are
calculated: a) the bi-decadal anomalies, b) the bi-decadal-averaged annual anomalies, the
c) seasonal and d) monthly averages in four bi-decades, the entire projection period, and
the baseline (1961~1990), and e) the climate anomalies over every 10-km grid in four bi-
decades and the 80-year period. The differences of climate anomalies among emission
scenarios, climate variables, bi-decades, seasons, and geographical locations are also
examined based on the corresponding indices. One representative characteristics of the
ReMPMID approach is that the uncertainties in the correspondence between the
simulation of large-scale atmospheric variables and the observations of regional- or local-
scale climate variables are incorporated into the projection result which is of uncertainties
accordingly. To analyze these uncertainties, the median and the quartiles of the bi-decadal-
averaged monthly climate anomalies are estimated for every emission scenario, climate
variable, bi-decade, and month. The related results are presented in the following figures.

5.3 RESULTS ANALYSIS

5.3.1 Spatial Averages: Bi-Decadal Anomalies

The bi-decadal climate anomalies over the ARB under the four RCP scenarios are
presented in Figure 5.3. In general, the $T_{min}$, $T_{max}$ and $Prec$ averaged over all temporal
and spatial units would increase during 2021 and 2100 in comparison with historical
averages under every RCP scenario; corresponding to RCPs 2.6, 4.5, 6.0 and 8.5, the
increment would be 1.742, 2.261, 2.107 and 2.975 °C for $T_{min}$, 1.353, 1.796, 1.612 and
2.217 °C for $T_{max}$, and 0.030, 0.074, 0.080 and 0.160 mm for $Prec$, respectively. In
addition, $T_{max}$ and $T_{min}$ follow the same pattern of differences among bi-decades. On
average, the increment of temperature or precipitation would decrease from [2021, 2040]
to [2061, 2080] and increase from [2061, 2080] to [2081, 2100] under RCP 2.6. This
pattern would be reversed for temperature under RCP 4.5, i.e. increasing from [2021, 2040] to [2061, 2080] and decreasing from [2061, 2080] to [2081, 2100]. In contrast, the increment of precipitation would gradually increase from [2021, 2040] to [2081, 2100]. This trend of gradually increasing is also followed for the increment of either temperature or precipitation under RCPs 6.0 and 8.5. The difference among the variations of climate anomalies among bi-decadal periods under RCP scenarios is closely associated with the overall emission control strategies under these scenarios. Meanwhile, the ranges of climate-variable increments would be 4.223, 3.444, 4.821 and 6.164 °C for Tmin, 3.503, 2.588, 3.504 and 4.968 °C for Tmax, and 0.282, 0.322, 0.362 and 0.486 mm for Prec, respectively under RCPs 2.6, 4.5, 6.0 and 8.5. The implication is that the variability of bi-decadal averaged climate anomalies is maximized for either temperature or precipitation under RCP 8.5 and minimized for temperature under RCP 4.5 and for precipitation under RCP 2.6. This is closely related with the difference among the four RCP scenarios and the particular characteristics of regional climate over the ARB.
Figure 5.3. Bi-decadal Anomalies.
5.3.2 Spatial Averages: Annual Anomalies

The annual anomalies of every selected climate variable over the ARB under RCP scenarios are illustrated in Figure 5.4. Generally, the annual anomalies would be positive for most years and all RCP scenarios. Namely, the annually-averaged temperature and precipitation would likely climb in most cases. The overall increasing trends of the increments of annual temperature or precipitation would increase from RCP 2.6, RCP 4.5, RCP 6.0 to RCP 8.5. It is against the overall increasing trend of annually-averaged climatic conditions that in some years the average temperature or precipitation would be lower than the historical averages. The ratio of the annual anomaly of a climate variable being lower than zero in the 80-year period of the 21st century would be 0.013, 0, 0.013 and 0 for Tmin, 0.025, 0, 0 and 0 for Tmax, and 0.325, 0.113, 0.163 and 0.125 for Prec corresponding to RCPs 2.6, 4.5, 6.0 and 8.5, respectively. It is indicated that in consideration of inter-annual variations the increasing climate trend is of higher universality for temperature compared with precipitation and its significance would increase under the scenarios of from RCP 2.6, RCP 6.0, RCP 4.5 to RCP 8.5. Besides, there would be significant differences among the annual anomalies of climate variables. Corresponding to RCPs 2.6, 4.5, 6.0 and 8.5, their ranges would be 4.223, 3.444, 4.821 and 6.164 °C for Tmin, 3.503, 2.588, 3.504 and 4.968 °C for Tmax, and 0.282, 0.322, 0.362 and 0.486 mm for Prec, respectively. This implies that the inter-annual variability of climate anomalies would increase from RCP 4.5, RCP 2.6, RCP 6.0 to RCP 8.5 for temperature while increasing from RCP 2.6, RCP 4.5, RCP 6.0 to RCP 8.5 for precipitation.
Figure 5.4. Annual Anomalies. (Orange lines: no temperature change; pink lines: no precipitation change).
5.3.3 Spatial Averages: Intra-Annual Variability

The intra-annual variability of climate anomalies over the ARB is displayed in Figure 5.5. As a general rule, either temperature or precipitation would increase for most combinations of RCP scenarios, bi-decadal periods, and seasons or months except: $T_{min}$ in July and October, $T_{max}$ in October, and $Prec$ in September and from Nov to March. The increase of temperature is mainly contributed by warming in winters, while that of precipitation is dominated by heavy rainfall from May to August. For instance, under RCP 2.6, the increment of climate variables would gradually decrease from 4.17 °C in DJF, 1.75 °C in MAM, 0.63 °C in JJA SON to 0.43 °C in JJA for $T_{min}$, from 3.46 °C in DJF, 0.85 °C in JJA, 0.58 °C in SON to 0.53 °C in MAM, and from 0.20 mm in JJA to 0.67 mm in MAM for Prec; on the contrary, there would be less precipitation in DJF (-0.085 mm) and SON (-0.065 mm) on average. As for $T_{min}$ and $T_{max}$, they would share the nearly same changing trends among bi-decadal periods for almost all seasons, months and RCP scenarios due to the inherent correlations. The overall variability of either temperature or precipitation averaged over various bi-decades would be the highest for RCP 8.5 and gradually decrease for RCP 6.0, RCP 4.5 and RCP 2.6. This variability would vary with months and seasons and is the most significant in MAM for temperature and in JJA for precipitation. In most seasons or months, the variability of every selected climate variable among the four bi-decadal periods or the four RCP scenarios would be of high similarity with that of the bi-decadal climate anomalies.
Figure 5.5. Intra-Annual Variability of Climate Anomalies.
5.3.4 Spatial Variability: Bi-Decadal Averages

The spatial variability of octo- or bi-decadal climate anomalies is of high similarity among the four RCP scenarios although some differences exist. RCP 2.6 is taken as an example to present such variability. The related results are illustrated in Figure 5.6 and the spatial variability of climate anomalies under the other RCP scenarios are presented in Figure 5.7 to Figure 5.9. In general, the whole river basin would experience warmer nights. On the contrary, people in the north end of the ARB would have colder days, while those in the other grids would experience warmer days. The upstream would become dryer and the other regions would be wetter than the baseline period. Besides, the octo- or bi-decadal averaged climate anomalies would decrease from the upstream to the lower one for temperature while following the reverse order for precipitation. These patterns would be relatively consistent for every bi-decadal period, but from the viewpoint of magnitudes, the climate anomalies would decrease from [2021, 2040] to [2061, 2080] and increase from [2061, 2080] to [2081, 2100] for every selected climate variable and most grids over the ARB. The historical spatial variability of climatic conditions over the ARB is shown in Figure 3.2. A comparison of Figure 5.6 and Figure 3.2 reveals that the spatial climatic variability over the ARB would not change significantly and that the climatic anomalies are highly correlated with the absolute magnitudes of climatic conditions. The increment of temperature would be the highest for highlands that only account for a small proportion of the entire river basin, while that of precipitation would be maximized over the downstream. In addition, the differences among the spatial variability of climatic anomalies over the ARB under various RCP scenarios are concentrated in the following aspects. In comparison with RCP 2.6, the range of climatic anomalies would be different
under the other $RCP$ scenarios, which is synchronic with the overall climate change under these scenarios. For instance, under $RCP$s 2.6 and 4.5, the range for $T_{max}$ would be $[-0.13, 6.74]$ and $[0.33, 7.23] °C$, while the octo-decadal averaged increment of $T_{max}$ would be 1.35 and 1.79 °C, respectively. Due to the change of climate-anomaly ranges, the north end of the $ARB$ would experience warmer days under $RCP$s 4.5, 6.0 and 8.5 instead of colder days under $RCP$ 2.6.
Figure 5.6. Spatial Variability of Climate Anomalies under RCP 2.6.
Figure 5.7. Spatial Variability of Climate Anomalies under RCP 4.5.
Figure 5.8. Spatial Variability of Climate Anomalies under RCP 6.0.
Figure 5.9. Spatial Variability of Climate Anomalies under RCP 8.5.
5.3.5. Spatial Variability: Inter-Seasonal Differences

Figure 5.10 shows the seasonal-scale differences of the spatial variability of climate anomalies over the ARB among RCP scenarios, climate variables, and bi-decadal periods. It is illustrated that the differences would be very significant for varying seasons. In general, the bi-decadal averaged climate anomalies would decrease from DJF, MAM, SON to JJA for night temperature, from DJF, SON, MAM to JJA for day temperature, and from JJA, MAM, SON to DJF for precipitation under every RCP scenario and bi-decadal periods. The increment of temperature would be dominated by warming in the upstream of the ARB in winters, while that of precipitation by increased rainfall in the downstream in summers. For almost every combination of seasons, bi-decades and RCP scenarios, the climate anomalies would decrease from the upstream to the downstream for either Tmin or Tmax, while this spatial pattern would be reversed for precipitation.

For every season and climate variable, the seasonal averaged climate anomalies would vary with bi-decadal periods and RCP scenarios. This would lead to complicated fluctuation of seasonal climate anomalies. For instance, the increment of night temperature averaged over almost all grids of the ARB in springs of [2041, 2060] would increase from RCP 2.6 to RCP 4.5, decrease from RCP 4.5 to RCP 6.0, and increase from RCP 6.0 to RCP 8.5. Meanwhile, the variation pattern of seasonal climate anomalies would be of high similarity with that among bi-decades and RCP scenarios for most combinations of seasons, geographic locations, bi-decadal periods, RCP scenarios, and climate variables. There are also a few of exclusions. For example, the increment of day temperature in SON on the upstream of the ARB would increase from [2061, 2080] to
[2081, 2100], while the bi-decadal increment would decrease from [2061, 2080] to [2081, 2100] as shown in Figure 5.3.

Increases of temperature or precipitation would be experienced by the residents distributed over the entire river basin under most combinations of bi-decadal periods, RCP scenarios, and geographic locations. The exclusions consist of the followings. Night temperature would decrease at the downstream of the ARB in summers for every RCP scenario and in autumns for RCP 2.6 and [2021, 2040]. As for day temperature, it would decrease in summers at the downstream of the ARB, tend to decrease from [2081, 2100] to [2021, 2040] and from RCP 8.5, RCP 4.5, RCP 6.0 to RCP 2.6, and likely decrease in springs, summers and autumns at the downstream. In winters and autumns, precipitation might decrease for grids at the upstream, which would also vary with RCP scenarios and bi-decadal periods.
Figure 5.10. Inter-Seasonal Differences of Spatial Climate-Anomaly Variability.
5.4 DISCUSSIONS

5.4.1 Uncertainty Analysis

The above revelations are achieved through analyses of the averaged high-resolution climatic projection results based on the ReMPMID approach. However, long-term projection of climatic conditions are challenged by the existence of uncertainties resulting from multiple sources such as data uncertainty and the complicated correspondence between coarser-resolution atmospheric variables and finer-resolution climate variables. As a representative advantage of ReMPMID, these uncertainties can be incorporated into the downscaling process and results, enabling climate projection under uncertainties. For any given combination of the values of atmospheric variables, the potential discrete distributions of climate variables can be obtained. In this case study, the quartiles of the potential monthly averaged anomalies of $T_{min}$, $T_{max}$ or Prec in every bi-decade under every RCP scenario are calculated to present the uncertainties in climate change over the ARB. The related results are displayed in Figure 5.11. It is revealed that in general there are relatively significant uncertainties in the projection results of climate anomalies. For every selected climate variable, the uncertainties in the projected climate anomalies would significantly vary with months. These uncertainties can be represented as the fluctuation ranges of monthly climate anomalies (Dong et al., 2011, 2012, 2013, 2014a,b,c, 2015). Accordingly, these uncertainties would be relatively low in May to September (about [0.7, 1.2] °C) and high in other months (about [1.5, 3.0] °C) for temperature, while they would be lower in November to April ([0.1, 0.13] mm) than those in the other months ([0.17, 0.38] mm) for precipitation. Meanwhile, these uncertainties would monotonically increase with the increase of climate anomalies in various bi-decadal periods and RCP scenarios.
for every selected climate variable. When the projected climate anomalies are applied to
the related studies, these uncertainties should be taken into account to avoid provision of
dogmatic climate projection for risk assessment and management under uncertainties.
Figure 5.11. Quartiles of Monthly Climatic Anomalies. (Red: maximum; orange: 75%; black: median; blue: 25%; green: minimum).
5.4.2 Bias Analysis

In the process of verification, it was revealed that the ReMPMID approach tended to over-estimate $T_{min}$ and $T_{max}$ in springs and winters and $Prec$ in summers and autumns while underestimating $T_{min}$ and $T_{max}$ in summers and autumns and $Prec$ in springs and winters over all or almost all grids in the ARB (Chapter 4). If these biases are taken into account, the climate projection results would be changed in some aspects. For example, as discussed in the former section, warming would mainly occur in winters and springs while cooling in summers and autumns for every $RCP$ scenario and bi-decadal period; meanwhile, the increases of precipitation would concentrate in summers and springs while the decreases in winters and autumns. If the overestimation or underestimation of climatic conditions is removed, the significance of the differences of climate anomalies among seasons as well as the overall intensity of climate change would be decreased for both temperature and precipitation. In general, the increments of temperature and precipitation would shrink by approximately 1.09 °C and 0.12 mm, respectively. Accordingly, corresponding to $RCP$s 2.6, 4.5, 6.0 and 8.5, the octo-decadal averaged increment of climate variables would be 0.652, 1.171, 1.017 and 1.885 °C for $T_{min}$, 0.263, 0.706, 0.522 and 1.127 °C for $T_{max}$, and -0.090, -0.046, -0.040 and 0.040 mm for $Prec$, respectively. It is implied that, if the modeling biases are removed, the octo-decadal averaged temperature would still increase for every $RCP$ scenario with a decreased increment, while the octo-decadal averaged precipitation would only increase under $RCP$ 8.5 and decrease under the other $RCP$ scenarios.

5.4.3 Impacts Analysis
The potential changes of 21st-century climatic conditions over the ARB might raise a series of concerns such as ones regarding hydro-regime variation, socio-economic development, and eco-environmental conservation. For instance, the climbing of temperature may speed up the melting of snow, glacier or frozen soil, increase the intensity of evaporation and transpiration, enhance the availability of solar power and the occurrence of wild fires, ascend the demands for irrigation waters and air conditioning electricity, and alter the habitat of ecological systems. The rising of precipitation may change the regional or local hydrological regimes, increase the availability of waters for socio-economic activities and natural eco-systems, and facilitate addressing the potential challenge of water crisis. Furthermore, due to the variation of climate anomalies with multiple factors such as geographic locations, temporal units, RCP scenarios and climate variables, the impacts of climate change would be diversified over the entire river basin. For example, the glacier field on the Rocky Mountains would tend to experience both increased temperature and decreased precipitation of which the combined effect might lead to acceleration of glacier and snow melting. Warmer and dryer winters might deteriorate the conflict of water consumptions between oil-sands mining operations, socio-economic development and eco-environmental conservation and lead to increased occurrence intensities, durations and frequencies of droughts in springs.

5.4.4 Potential Extensions

This study is the first attempt of applying the ReMPMID approach to a case study of high-resolution climate projection. A series of interesting findings related with the future potential status of the regional climate at a high spatial resolution over the ARB, Canada are revealed from this study. These findings would be helpful for providing
scientific support for guiding climate-change impact studies or practices. However, there are also many opportunities regarding either methodology advancement or engineering practices in the future studies. For instance, the ReMPMID approach can be improved or extended in many aspects, which is comprehensively discussed in Chapter 4. Representative ones consist of, but are not limited to, the followings: enabling the reversible transformation between any non-normal distribution and a normal distribution by a continuous function, identifying a more reasonable scheme of characterizing the complicated nonlinearity between atmospheric variables and climate variables, and developing a dynamic version of the ReMPMID approach to address the temporal variability of regional climate. In addition, it deserves further efforts that the temporal resolution is enhanced from monthly to daily based on advancement of the modeling accuracies of ReMPMID. Extreme climatic conditions that may pose severe threats on socio-economic development and eco-environmental health would be examined by the ReMPMID approach. It may be helpful for improving modeling accuracies that ReMPMID is used as an ensemble tool to integrate the advantages of various dynamic or statistical downscaling techniques or taken as an element in an ensemble downscaling framework under the uncertainty of climate simulation or downscaling models. On the other hand, the long-term high-resolution projection of climatic conditions over the ARB could be used for various studies or practices related with climate-change impacts or adaptations. For example, these projection could be incorporated into hydrological models to analyze the variation or status of hydrological regimes in the future and to guide water resources management. Calculation of climate indices such as the Palmer drought severity index (Palmer, 1965) and the standardised precipitation index (McKee et al., 1993) would
facilitate examining the long-term impacts of climate change on various receptors such as ecosystem, environment, energy, society, and economy. In addition, the ReMPMID approach could be applied to more case studies of high-resolution projection of climate in other river basins, enhancing the reliability of scientific support for climate-change impact or adaptation practices over these regions.

5.5 SUMMARY

In this chapter, high-resolution projection of 21st century climate over the ARB, Canada were enabled through ReMPMID (Chapter 4), an advanced statistical approach dedicated in climate downscaling under complexities of data uncertainties, nonlinear correspondences, multivariate dependencies, non-normal distributions, spatial homogeneities, and temporal nonstationarities. The spatial resolution of CMIP5 GCMs was effectively enhanced to 10 kilometres. These projection provided rich information of the climatic conditions ($T_{min}$, $T_{max}$ and Prec) over this river basin in the 21st century under four RCP scenarios: RCPs 2.6, 4.5, 6.0 and 8.5. A series of findings were revealed from in-depth analyses of these projection.

In general, temperature and precipitation would increase during 2021 and 2100. Corresponding to RCPs 2.6, 4.5, 6.0 and 8.5, the octo-decadal averaged increment would be 1.742, 2.261, 2.107 and 2.975 °C for $T_{min}$, 1.353, 1.796, 1.612 and 2.217 °C for $T_{max}$, and 0.030, 0.074, 0.080 and 0.160 mm for Prec, respectively. If the modeling biases are removed, these increments would shrink by approximately 1.09 °C for temperature and 0.12 mm for precipitation; temperature would still increase for every RCP scenario with a decreased increment, while precipitation would only increase under RCP 8.5 and
decrease under the other RCP scenarios. The variability of bi-decadal averaged climate anomalies is maximized under RCP 8.5 and minimized for temperature under RCP 4.5 and for precipitation under RCP 2.6. The inter-annual variability of climate anomalies would increase from RCP 4.5, RCP 2.6, RCP 6.0 to RCP 8.5 for temperature while increasing from RCP 2.6, RCP 4.5, RCP 6.0 to RCP 8.5 for precipitation.

The bi-decadal averaged climate anomalies would decrease from DJF, MAM, SON to JJA for night temperature, from DJF, SON, MAM to JJA for day temperature, and from JJA, MAM, SON to DJF for precipitation. Either temperature or precipitation would increase in most months except decreases of Tmin in July and October, Tmax in October, and Prec in September and from November to March. The increase of temperature is mainly contributed by warming in winters, while that of precipitation is dominated by heavy rainfall from May to August. The uncertainties in the projection results would be relatively low in May to September and high in other months for temperature, while they would be lower in November to April than those in the other months for precipitation.

The spatial climatic variability over the ARB would not significantly change with RCP scenarios and the climatic anomalies are highly correlated with the absolute magnitudes of climatic conditions. This river basin would experience warmer nights; people in the north end of the ARB would have colder days, while those in the other grids would experience warmer days; the upstream would become dryer and the other regions would be wetter than the baseline period. The climate anomalies would decrease from the upstream to the downstream for either Tmin or Tmax, while this spatial pattern would be reversed for precipitation. The increment of temperature would be dominated by warming
in the upstream of the ARB in winters, while that of precipitation by increased rainfall in the downstream in summers.

The potential changes of 21st-century climatic conditions over the ARB might raise a series of concerns. Representative ones include but are not limited to the followings: acceleration of glacier and snow melting, alternation of the hydrological regime, deterioration of the water conflict between oil-sands mining operations and eco-environmental health, and increased intensities, durations and frequencies of droughts in springs. Additional attentions should be paid on addressing these concerns through scientific climate-change adaptation strategies.

On the other hand, further efforts would be made to deepen understandings of the ReMPMID approach and regional climate, facilitate proposition of more advanced downscaling techniques, reveal the complicated correspondence between large-scale atmospheric variables and regional or local scale climatic variables, and enable provision of scientific support for climate-change impacts or adaptation practices. For instance, the temporal resolution would be enhanced from monthly to daily. Extreme climatic conditions would be examined. The ReMPMID approach would be used as an ensemble tool to integrate the advantages of various dynamic or statistical downscaling techniques or be taken as an element in an ensemble downscaling framework under the uncertainty of climate simulation or downscaling models. The climate projection results would be incorporated into hydrological models to analyze the variation or status of hydrological regimes in the future and to guide water resources management.
CHAPTER 6

BAYESIAN PRINCIPAL-MONOTONICITY INFERENCE: METHODOLOGY
DEVELOPMENT AND AN APPLICATION TO A HYDROMEΤEOROLOGICAL SYSTEMS ANALYSIS OVER THE ATHABASCA
RIVER BASIN

6.1 BACKGROUND

Hydrological systems are intrinsically interrelated with many other systems, e.g. global climate, local weather, human society, economic development, and ecological environment. Decisions relating to water resources management and flood control rely on hydro-system analysis, where the temporal-spatial interrelationships within hydrological systems are modelled. A critical task is to effectively model hydrological processes, e.g. soil moisture, groundwater recharge and surface runoff, and understand their driving mechanisms. However, the task is challenged by many complexities in hydrological systems: e.g. multiplicity, interactions, uncertainties and weak predictability of influencing factors; multiplicity, interactions, nonlinearities and tempo-spatial variations of hydrological processes; diversity and interrelations of uncertainties; unavailability or lack of high-quality data; unidentifiability and non-stationarity of processes and influencing factors; and, heavy computational loads for alternative analysis methods. Nonlinearities exist extensively in hydrological systems and significantly affect the robustness of hydro-systems analysis.

Previously, a number of studies have been conducted on this issue (Izzard, 1966; Dooge, 1967). These may be categorized as metric (also called data-based, empirical or black box) models, parametric (also called conceptual or grey box) models, and
mechanistic (also called physically-based or white box) models. Representative ones include, but were not limited to, artificial neural networks (Stelling, 2000; Govindaraju, 2000), recurrent neural networks (Price et al., 1998; Proaño et al., 1998; Van den Boogaard et al., 1998), nonlinear stochastic simulation (Kavvas, 2003), response surface analysis (Box and Wilson, 1951), nonlinear dynamics and chaos (Sivakumar, 2000), supportive vector machines (Cristianini and Shawe-Taylor, 2000), lumped models (Rockwood, 1966; Anderson, 1967), distributed models (Huggins and Monke, 1968; Beasley et. al., 1980; Quick, 1995), semi-distributed models (Hughes and Sami, 1994; Arnold et al., 1998; Saleh et al., 2000; Gassman et al., 2001; Nasr et al., 2005), the Systeme Hydrologique Européen model (Abbott et al., 1986a, b), and the Data-based Mechanistic model (Young and Beven, 1994).

Reflection of nonlinearities in hydrological systems vary with the hydro-system analysis methods. In metric models, this is commonly derived from analyses of available data series (Wüst, 1995; Minns, 1996; Scardi, 1996; Recknagel et al., 1997; Sanchez et al., 1998;See and Openshaw, 1999; Abrahart et al., 2004; Dong et al., 2015). Especially for classification methods (De Bruin and Stein, 1998; Hong et al., 2004; Minglei et al., 2010; Olden et al., 2012), a particular type of metric models, the analyses are focused on either dissimilarity or similarity of data series based on an artificially designated threshold or criterion. When parametric models are employed, hydrological systems are modeled through analyses of balances among fluxes such as rainfall, infiltration, percolation, evapotranspiration, runoff and drainage. In contrast to metric models, parameter relationships that relate to nonlinearities are defined by the modeller’s understanding. As for mechanistic models, nonlinearities are formulated as equations based on an
understanding of the physical processes and spatial discretization of hydrological systems. In general, modeling of nonlinearities is achieved through either one-way empirical evaluation in metric models or predefined functions in parametric or mechanistic models.

Nevertheless, nonlinearities in hydrological systems may be temporally or spatially heterogeneous, leading to irregularity of the related nonlinear relationships. A particular function generalized over the entire temporal-spatial horizon, which is commonly used in existing parametric or mechanistic methods, may not reasonably simulate the irregularity. In addition, a hydrological system analysis may be further challenged by the coexistence of irregular nonlinearities and other system complexities. For instance, uncertainties may appear in the observations of both independent and dependent factors such as daily precipitation and surface runoff. The same values of independent factors may correspond to significantly different values of the dependent variable. Besides, a hydrological variable of interest may be nonlinear with multiple influencing factors, and these nonlinear relationships are dependent upon each other. The significance of a factor for the hydrological variable may vary with value levels, e.g. low significance of the factor for high values of the variable and high significance for moderate values. It is possible that a factor is significant for global variation of the hydrological variable and that another factor is only significant for local variation. Furthermore, massive computational loads are a common problem for hydrological modeling, especially for large-scale finer-resolution problems. For existing hydrological systems analysis methods such as classification methods, all of these complexities can hardly be incorporated into the modeling process. Failure to address these integrated complexities in hydro-system analyses may lead to a misrepresentation of the hydrological system, distortion of analysis processes and results,
unreliability of resultant decision alternatives, and resultant disasters in socio-economic development or eco-environmental qualities.

To mitigate these challenges, a Bayesian principal-monotonicity inference (BaPMI) method is proposed in this chapter. In BaPMI, the responsive relationship from influencing factors (named as predictors) to the hydrological variable of interest (named as the predictand) will be discretized as interrelated end nodes, i.e. groups of paired samples of predictors and the predictand, under irregular nonlinearities, data uncertainties, and multivariate dependencies. In detail, a discrete distribution transformation approach will be developed to enable transformation of non-normally distributed predictand samples as a normal distribution and invertible restoration of the simulated predictand values as the original non-normal distribution. To eliminate data uncertainties, statistical inference will be employed to assess the significance of differences among groups of predictand samples. The nonlinear responsive relationship between predictors and the predictand will be interpreted as piecewise monotonicity, similar to piecewise linearization of a nonlinear function. The piecewise monotonicity will be further represented as principal monotonicity, representing the most significant monotonicity between the predictand and one of all predictors, for dealing with existing correlations of predictors. The process of identifying the principal monotonicity, which is equivalent to an integer optimization problem, will be accelerated by an incorporation of Gaussian process analyses and Bayesian optimization. Based on a recursive classification and cluster analysis, all paired samples of predictors and the predictand that represent the responsive relationship between them will be discretized as a series of end nodes that are internally inseparable, externally divergent, mutually exclusive and collectively
exhaustive. Given any combination of predictors, the corresponding predictand value will be estimated through a BaPMI prediction scheme. The properties, strengths, extensions and shortcomings of BaPMI will be also examined through an application to streamflow simulation in the Athabasca River Basin, Canada.

6.2 METHODOLOGY DEVELOPMENT

6.2.1 Overview

One challenging task in hydro-meteorological systems analyses is to effectively quantify the correspondence between a dependent variable of interests (i.e. the predictand), e.g. the discharge in this study, and multiple potential influencing factors (i.e. the predictors), e.g. temperature and precipitation. This task is challenged by the complexities of data uncertainties, non-normal distributions, nonlinear hydro-meteorological correspondences, multivariate dependencies, and massive computations in large-scale real case studies. To enable hydro-meteorological systems analyses under these complexities, advanced statistical inferential techniques, Gaussian process models, and Bayesian optimization are integrated into an ingeniously design framework, leading to development of the BaPMI approach. The principle, structure and significance of this approach are presented in Figure 6.1. The essence of BaPMI is to systematically discretize the correspondence between the predictand and the predictors as a series of interrelated groups. The key procedures of BaPMI consist of the followings: normality analyses, recursive dissimilarity inferences, Bayesian principal monotonicity analyses, and recursive similarity and dissimilarity inferences.
Figure 6.1. Principle, Structure and Significance of BaPmI.
6.2.2 Normality Analyses

Let the predictand and the predictors be denoted as $y$ and $\{x_j\}_{j=1}^n$ or $X = (x_1, x_2, \ldots, x_n)$, respectively where $n$ is the number of the predictors. The corresponding datasets of the predictand and the predictors are $\{ y_t \}_{t=1}^T$ and $\{ X_t \}_{t=1}^T$, respectively which constitute the multi-dimensional paired samples of the predictors and the predictand:

$$\{ X_{yt} \}_{t=1}^T = \{ (X_t, y_t) \}_{t=1}^T. \quad (6.1)$$

One prerequisite of $BaPMI$ is that the population of predictand samples should be normally distributed. Approximately forty methods are available for a uni-variate normality test (Kolmogorov, 1933; Massey Jr, 1951; Lilliefors, 1967; Dufour et al., 1998; Oztuna et al., 2006; Peat and Barton, 2005). Among them, the Shapiro-Wilk ($SW$) test (Shapiro and Wilk, 1965; Royston, 1995) is recommended by many researchers or institutions (Mendes and Pala, 2003; Keskin, 2006; Razali and Wah, 2011; USEPA, 2006). The $SW$ test has a competitive power performance in a univariate normality test and can be adopted for both symmetric non-normal and asymmetric distributions (Royston, 1995; Razali and Wah, 2011). The test is achieved through an analysis of the correlation between the quantiles of the standard normal distribution and the ordered samples of the predictand. The related technical details are presented in the appendix.

If the normality prerequisite does not hold, a transformation from an abnormal distribution to a normal distribution is required to guarantee the reliability of the $BaPMI$ results. The process must be reversible so that the original values of the hydrological variables can be restored, avoiding distortion of the similarity and dissimilarity among samples of hydrological variables. To achieve this, a discrete distribution transformation
(DDT) approach (Cheng et al., 2016b) which was developed based on the reversible convertibility of any distribution to the 0-1 uniform distribution is incorporated into the framework of \textit{BaPMI}. The related details are presented in Chapter 3. Through this approach, the original predictand samples are replaced with normally-distributed ones. At the end of \textit{BaPMI}, the predictand values in the constructed model are restored to the original values.

6.2.3 Recursive Dissimilarity Inferences

All paired samples of the predictors and the predictand constitute an initial node \((N_1)\) of which the sample size is \(T\). This node is partitioned as many child nodes in the process of \textit{BaPMI} step by step. For any node, a recursive dissimilarity analysis (RDA) is carried out to reveal the dissimilarity of the complicated correspondence between the predictors and the predictand. The initial node \((N_1)\) is taken as an example to present the procedures of RDA.

For any predictor, e.g. the \(j\)th predictor \((x_j)\) where \(j \in \{1, 2, \ldots, n\}\), its samples in \(N_1\), i.e. \(\{x_j\}_{t=1}^T\), are sequenced from the lowest value to the highest. Let the samples of \(x_j\) after sequencing be denoted as \(\{sjx_j\}_{t=1}^T\), the sequence numbers of \(\{x_j\}_{t=1}^T\) corresponding to \(\{sjx_j\}_{t=1}^T\) as \(\{s(j, 1), s(j, 2), \ldots, s(j, T)\}\), and the sequence numbers of \(\{sjx_j\}_{t=1}^T\) corresponding to \(\{x_j\}_{t=1}^T\) as \(\{\hat{s}(j, 1), \hat{s}(j, 2), \ldots, \hat{s}(j, T)\}\). We have \(x_{jt} = sjx_{s(j, t)}\) and \(sjx_{jt} = \hat{x}_{\hat{s}(j, t)}\) for any \(t \in \{1, \ldots, T\}\). According to the sequences of \(x_j\) samples after sequencing, the paired samples of the predictors and the predictand are sequenced as \(\{X_{Y_{\hat{s}(j, t)}}\}_{t=1}^T\). The sequenced paired samples are partitioned as two groups in two child nodes \(\{N_{1u_1}\} \text{ and } \{N_{1u_2}\}\) from a row. We denote the row number of the partition row as \(u\). Due to the requirement of degrees of freedom in most statistical analyses, a parameter \((N\text{min})\) named as the
minimum partition row number is introduced to represent the minimum value of \( u \). The partition row numbers range from \( N_{min} \) to \( T - N_{min} \). As the result of partitioning, we have

\[
N_{1_{\mu_1}} = \{X_{y(j, n)}\}_{t=1}^{u}, \quad (6.2)
\]

\[
N_{1_{\mu_2}} = \{X_{y(j, n)}\}_{t=u+1}^{T}, \quad (6.3)
\]

\[
N_1 = N_{1_{\mu_1}} \cup N_{1_{\mu_2}}, \quad (6.4)
\]

and

\[
N_{1_{\mu_1}} \cap N_{1_{\mu_2}} = \emptyset. \quad (6.5)
\]

The predictand samples in \( N_{1_{\mu_1}} \) and \( N_{1_{\mu_2}} \) are \( \{y_{(j, n)}\}_{t=1}^{u} \) and \( \{y_{(j, n)}\}_{t=u+1}^{T} \), respectively. The Aspin-Welch-Satterthwaite (AWS) \( t \)-test (Jácome et al., 2007) is employed to evaluate the difference of the populations of \( \{y_{(j, n)}\}_{t=1}^{u} \) and \( \{y_{(j, n)}\}_{t=u+1}^{T} \). The related details are stated in the appendix. An indicator \( w_{j\mu}(\alpha) \) named as the marginal monotonicity significance is introduced to quantify the significance of the local marginal monotonicity of the predictand with the \( j \)th predictor between \( N_{1_{\mu_1}} \) and \( N_{1_{\mu_2}} \). This indicator is formulated as

\[
w_{j\mu}(\alpha) = \alpha/p_{ju} \quad (6.6)
\]

where \( \alpha \) is an acceptable level of statistical significance and \( p_{ju} \) is the \( p \) value in the AWS \( t \) test. From the viewpoint of statistical inference, the predictand samples in \( N_{1_{\mu_1}} \) and \( N_{1_{\mu_2}} \) are significantly different if \( w_{j\mu}(\alpha) > 1 \).

Through the sequencing operation, there must be a significant difference between the samples of the \( j \)th predictor in \( N_{1_{\mu_1}} \) and \( N_{1_{\mu_2}} \). If the \( j \)th predictor is of significant
impacts on the predictand, it is highly possible that the difference between the predictand samples in $N_{i1}$ and $N_{i2}$ is also significant. The impact of the $j$th predictor on the predictand can be quantified as the *marginal monotonicity significance* ($w_{ju}(\alpha)$). Furthermore, the most significant impact of the $j$th predictor on the predictand corresponds to the highest value of $w_{ju}(\alpha)$, $w^*_{ju}(\alpha)$, which represents the *principal local marginal monotonicity* of the predictand with a predictor. This indicator named as the *principal local marginal monotonicity significance* is incorporated into the framework of RDA to support partition of a node such as the initial node.

### 6.2.4 Bayesian Principal Monotonicity Analyses

Due to the variation of $w^*_{ju}(\alpha)$ with $j \in \{1, 2, \ldots, n\}$ and $u \in \{Nmin, Nmin + 1, \ldots, T - Nmin\}$, the process of identifying the *principal local marginal monotonicity significance* from $n \cdot (T - 2 \cdot Nmin + 1)$ alternatives is equivalent to an integer programming problem:

$$\max \{w_{ju}(\alpha) \mid j \in \{1, 2, \ldots, n\}; u \in \{Nmin, Nmin + 1, \ldots, T - Nmin\}\}. \quad (6.7)$$

Computational loads geometrically increase with temporal resolutions and ranges, the number of predictors, the length of lead times, the number of spatially correlated grids, and other potential factors. Hence, the loads can be very heavy when the *BaPMI* approach is applied to large-scale hydrological systems analysis problems. An example is presented in the section of discussions. It is desired that an effective discrete optimization algorithm is incorporated into the framework of *BaPMI* to facilitate identifying the *principal local marginal monotonicity significance*. The algorithm employed in this study is Bayesian Gaussian process optimization (*BaGPO*) (Snoek et al., 2012).
The *BaGPO* algorithm, as an integration of the Gaussian process (*GP*) (Sacks et al., 1989) and Bayesian optimization (*BO*) (Mockus et al., 2012; Brochu et al., 2010), has been shown to outperform other state of the art global optimization algorithms on a number of challenging optimization benchmark functions (Snoek et al., 2012). *GP* models are a currently popular choice for use as a surrogate for simulation models because they are flexible for a wide variety of nonlinearities and parsimonious in the number of parameters (Jones and Johnson, 2009). Meanwhile, *BO* is verified as an effective tool in terms of systematically and intelligently optimizing the hyperparameters of a complicated system simulation model (Swersky et al., 2014). Based on a principled characterization of uncertainties in a model, *BO* is capable of identifying the best combination of hyperparameters with as few runs as possible. In the process of *BaGPO*, *GP* is used as a surrogate model to approximate the relationship between \( j (j \in \{1, 2, \ldots, n\}) \), \( u (u \in \{N_{\text{min}}, N_{\text{min}} + 1, \ldots, T - N_{\text{min}}\}) \) and \( w_{j\alpha} \) based on sampling, and *BO* is employed as an optimizer to find out the maximal value of \( w_{j\alpha} \). For the related technical details, see (Swersky et al., 2014).

Through the *BaGPO* method, \( w^*_{j\alpha} \) can be identified with a relatively high accuracy, which will be demonstrated by an example in the section of discussions. Let the corresponding predictor and row number be named as the *criterion predictor* and the *criterion partition row number* and be denoted as \( x_j \) and \( u_0 \), respectively. In comparison with other predictors, this predictor is of the most significant impact on the variation of the predictand within the initial node. Accordingly, all paired samples in the initial node are classified as two child nodes: \( N_2 \) and \( N_3 \) where
In the process of BaPMI, every node is examined in terms of the dissimilarity among the paired samples of the predictors and the predictand. Any node is partitioned as two child nodes if the maximal value of $w_{ju}(\alpha)$ is higher than 1. As the result of RDA, all nodes are internally inseparable; namely, every node cannot be further classified in accordance with the principal local marginal monotonicity between any predictor and the predictand. The number of nodes after RDA can be assumed as an integer $\mathcal{D}$. These nodes constitute a node list of which the length is $\mathcal{D}$. Let the nodes be denoted as $\mathcal{N}^d$ where $d = 1, 2, \ldots, \mathcal{D}$. We have

$$\mathcal{N}^d = \{X_{\mathcal{U}(d, p)}\}_{p=1}^{\mathcal{T}(d)}$$

(6.10)

for any $d \in \{1, 2, \ldots, \mathcal{D}\}$ where $\mathcal{T}(d)$ is the number of included paired samples, $p$ is the index for paired, and $\mathcal{U}(d, p)$ is the original sequencing numbers of paired samples in the initial node.

### 6.2.5 Recursive Similarity and Dissimilarity Inferences

To reveal the nonlinear correspondence between the predictors and the predictand and avoid redundant computation, a recursive similarity analysis (RSA) is conducted on every pair of nodes obtained from RDA. The statistical significance of the difference between predictand samples in any two nodes is also quantified by the marginal monotonicity significance. If the difference is statistically insignificant, the two nodes are
merged as a new one. The difference of predictand samples in every pair of nodes is evaluated until any pair of nodes cannot be merged. The result of RSA is a series of nodes for which the samples of the predictand in any two nodes are not significantly different. For any of them, however, it is possible that there is significant local marginal monotonicity between a predictor and the predictand. To completely reveal the dissimilarity and similarity of the correspondence between the predictors and the predictand, the modules of RSA and RDA are alternately conducted on all nodes in the node list until the list does not change any more. The nodes in the final node list are named as end nodes.

In the process of RSA, there is a gap between the boundaries of the criterion predictor samples in two targeted child nodes before partitioning. To fill this gap, the boundaries of the criterion predictor samples in two child nodes should be extended. Through this correction, the end nodes obtained from BaPMI are continuously inclusive, the predictand corresponding to any combination of predictors can be projected, and an infinite loop of the BaPMI program can be avoided. Subsequently, all paired samples \( \{X_{yt}\}_{t=1}^T \) representing the response from the predictors to the predictand are discretized as a series of end nodes. These end nodes are internally inseparable, externally divergent, mutually exclusive and collectively exhaustive. These characteristics of end nodes indicate that a systematic discretization of the complicated correspondence between predictors and predictand, which was hardly achieved by existing statistical downscaling methods, can be effectively enabled by the BaPMI approach. After construction of a BaPMI model, the corresponding hydrological conditions of any given combination of the values of the
hydro-system influencing factors can be obtained from a clustering analysis based on various statistical metrics such as the mean value or the quartiles.

6.3 CASE STUDY: ATHABASCA RIVER BASIN

6.3.1 Study Area

The ARB covers an area of approximately 138,000 km² over 52° ~ 59° N and -119° ~ -107° W. Its origin is the Columbia Icefield on the eastern slopes of the Rocky Mountains around Jasper in Alberta, Canada, leading to a close connection with climate change. This river is the longest undammed river in the Canadian prairies, which is another particular feature of it. This river flows approximately 1500 km northeast before entering the Peace-Athabasca Delta, the largest freshwater continental river delta in North America and the home to numerous migratory birds, and draining into Lake Athabasca. The evaluation drops from approximately 3715 m at Jasper to 211 m at the outlet in the Peace-Athabasca Delta (Figure 6.2a), resulting in a large variation of discharges across the ARB. The ARB includes diverse hydro-climatic regimes due to physiographical heterogeneity; snowcapped mountains, coniferous forest, mixed wood and deciduous forest are found in the uplands, whereas willow brush, shrubs, black spruce and sphagnum moss dominate the lowlands (Kerkhoven and Gan, 2006). Around 0.15 to 0.17 million residents distribute across this river basin. Many issues regarding regional socio-economic development and eco-environmental conservation, e.g. the conflict between water consumptions of oil-sands mining and ecosystem health at the downstream, is closely related with the temporal and spatial variability of the streamflow over the ARB. Provision of scientific support for guiding addressing these issues desires a robust model to simulate the temporal variability and spatial distribution of the discharges. Meanwhile, this task is challenged by the
complexities of uncertainties, non-normality, nonlinearity, dependencies, and enormousness which could be mitigated by the \textit{BaPMI} approach. Therefore, the \textit{BaPMI} approach is applied to simulate the streamflow over this river basin, which is helpful for verifying methodological reliability, revealing the regional hydrological regime of the \textit{ARB}, and supporting socio-economic and eco-environmental management.
Figure 6.2. Study Area and Data Collection.
6.3.2 Data

In this case study, a *BaPMI* model is constructed for examining the correspondence between discharges and the related influencing factors for every sub-catchment. First of all, a hydrometric database (HYDAT) is derived from Environment and Climate Change Canada (https://ec.gc.ca/). This database is compiled by Water Survey of Canada. The historical daily or monthly flow observations at a total of 190 gauge stations (Figure 6.2b) are included. The records of discharges at the temporal resolution of daily for some gauge stations are upcaled to monthly averages in consideration of the potentiality of long-term flow projection through connecting the constructed *BaPMI* models with global climate model outputs or dynamic/statistical downscaling results. Based on a comprehensive analysis of the temporal resolutions and ranges of records at these stations, ten representative gauge stations (Figure 6.2b) are selected. The related details are presented in Table 6.1. It is illustrated that there is a significant difference among the magnitudes and the temporal variability of the average discharges at the selected gauge stations. According to the ten gauge stations, the *ARB* is divided as ten sub-catchments (Figure 6.2b).

The variation of discharges (i.e. the predictand) across the *ARB* is related with multiple categories of influencing factors (i.e. predictors), e.g. human activities, land cover, soil types, and climatic conditions. Among them, the ones of the most significant impacts on streamflow are climatic conditions, while the others are relatively stable in terms of inter-annual variability. As a statistical hydrological simulation approach, the *BaPMI* approach focuses on revealing the causes of the long-term variation of discharges. Hence, the predictors selected in this case study are three climate variables, i.e. daily minimum
near-surface air temperature (Tmin), daily maximum near-surface air temperature (Tmax), and daily precipitation (Prec), which are widely used to represent the regional climate over a region. Accordingly, the monthly averages of the selected predictors from 1961 to 2003 over 1615 10-km grids in the ARB (Figure 6.2c) are derived from a raster-gridded climate dataset. Furthermore, 71 representative grids (Figure 6.2d) are selected to represent the regional climate of spatial heterogeneities in various sub-catchments and climate zones for every climate variable based on climate classification and similarity analyses (Chapter 3). Otherwise, the number of predictors for the gauge station in the lowest stream would be enormous, reaching tens of thousands and resulting in a disaster for computations.
Table 6.1. Selected Gauge Stations around the Outlets of Sub-Catchments.

<table>
<thead>
<tr>
<th>Station ID</th>
<th>07AA002</th>
<th>07AD002</th>
<th>07AE001</th>
<th>07AG004</th>
<th>07BC002</th>
<th>07BE001</th>
<th>07BK006</th>
<th>07CD001</th>
<th>07DA001</th>
<th>07DD001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude</td>
<td>52.91</td>
<td>53.42</td>
<td>54.21</td>
<td>54.01</td>
<td>54.45</td>
<td>54.72</td>
<td>55.29</td>
<td>56.69</td>
<td>56.78</td>
<td>58.21</td>
</tr>
<tr>
<td>Longitude</td>
<td>-118.06</td>
<td>-117.57</td>
<td>-116.06</td>
<td>-115.84</td>
<td>-113.99</td>
<td>-113.29</td>
<td>-114.59</td>
<td>-111.26</td>
<td>-111.40</td>
<td>-111.39</td>
</tr>
<tr>
<td>Drainage Area (km²)</td>
<td>3.87E+03</td>
<td>9.76E+03</td>
<td>1.96E+04</td>
<td>9.11E+03</td>
<td>1.31E+04</td>
<td>7.46E+04</td>
<td>1.44E+04</td>
<td>3.08E+04</td>
<td>1.33E+05</td>
<td>1.55E+05</td>
</tr>
<tr>
<td>Multi-Year Average Flow Rate (m³/s)</td>
<td>86</td>
<td>172</td>
<td>267</td>
<td>59</td>
<td>34</td>
<td>426</td>
<td>51</td>
<td>115</td>
<td>630</td>
<td>922</td>
</tr>
<tr>
<td>DJF</td>
<td>14</td>
<td>37</td>
<td>55</td>
<td>8</td>
<td>5</td>
<td>103</td>
<td>31</td>
<td>57</td>
<td>178</td>
<td>224</td>
</tr>
<tr>
<td>MAM</td>
<td>38</td>
<td>83</td>
<td>149</td>
<td>54</td>
<td>51</td>
<td>362</td>
<td>47</td>
<td>130</td>
<td>559</td>
<td>644</td>
</tr>
<tr>
<td>JJA</td>
<td>231</td>
<td>438</td>
<td>566</td>
<td>95</td>
<td>62</td>
<td>917</td>
<td>69</td>
<td>163</td>
<td>1234</td>
<td>1400</td>
</tr>
<tr>
<td>SON</td>
<td>60</td>
<td>128</td>
<td>182</td>
<td>27</td>
<td>19</td>
<td>321</td>
<td>56</td>
<td>112</td>
<td>537</td>
<td>755</td>
</tr>
<tr>
<td>Jan</td>
<td>13</td>
<td>36</td>
<td>53</td>
<td>5</td>
<td>5</td>
<td>100</td>
<td>29</td>
<td>55</td>
<td>175</td>
<td>224</td>
</tr>
<tr>
<td>Feb</td>
<td>11</td>
<td>32</td>
<td>50</td>
<td>7</td>
<td>4</td>
<td>93</td>
<td>31</td>
<td>51</td>
<td>161</td>
<td>198</td>
</tr>
<tr>
<td>Mar</td>
<td>11</td>
<td>32</td>
<td>50</td>
<td>10</td>
<td>9</td>
<td>101</td>
<td>35</td>
<td>50</td>
<td>168</td>
<td>208</td>
</tr>
<tr>
<td>Apr</td>
<td>16</td>
<td>44</td>
<td>102</td>
<td>51</td>
<td>71</td>
<td>336</td>
<td>47</td>
<td>123</td>
<td>492</td>
<td>627</td>
</tr>
<tr>
<td>May</td>
<td>84</td>
<td>175</td>
<td>294</td>
<td>100</td>
<td>72</td>
<td>649</td>
<td>62</td>
<td>216</td>
<td>1018</td>
<td>1186</td>
</tr>
<tr>
<td>Jun</td>
<td>242</td>
<td>491</td>
<td>631</td>
<td>111</td>
<td>63</td>
<td>991</td>
<td>67</td>
<td>179</td>
<td>1304</td>
<td>1501</td>
</tr>
<tr>
<td>Jul</td>
<td>253</td>
<td>472</td>
<td>617</td>
<td>111</td>
<td>80</td>
<td>1047</td>
<td>72</td>
<td>170</td>
<td>1402</td>
<td>1661</td>
</tr>
<tr>
<td>Aug</td>
<td>199</td>
<td>351</td>
<td>450</td>
<td>64</td>
<td>42</td>
<td>712</td>
<td>68</td>
<td>138</td>
<td>995</td>
<td>1127</td>
</tr>
<tr>
<td>Sep</td>
<td>104</td>
<td>207</td>
<td>280</td>
<td>42</td>
<td>27</td>
<td>468</td>
<td>65</td>
<td>129</td>
<td>733</td>
<td>943</td>
</tr>
<tr>
<td>Oct</td>
<td>50</td>
<td>117</td>
<td>173</td>
<td>33</td>
<td>20</td>
<td>318</td>
<td>58</td>
<td>122</td>
<td>556</td>
<td>711</td>
</tr>
<tr>
<td>Nov</td>
<td>25</td>
<td>59</td>
<td>106</td>
<td>24</td>
<td>11</td>
<td>177</td>
<td>46</td>
<td>86</td>
<td>322</td>
<td>482</td>
</tr>
<tr>
<td>Dec</td>
<td>17</td>
<td>42</td>
<td>64</td>
<td>10</td>
<td>7</td>
<td>117</td>
<td>32</td>
<td>65</td>
<td>197</td>
<td>250</td>
</tr>
</tbody>
</table>
6.3.3 Correlation Analysis

There is a lead time between climatic conditions and discharges. The length of the lead time can range from hours to months. To support the selection of predictors in discharge simulation, a correlation analysis is carried out on the gauge station of 07DD001. This gauge station is located at the outlet of the ARB, thus tends to correspond to the longest lead time. The longest lead time is assumed as three months in this experiment. Thus, there are a total of 3*3*71 alternative predictors for the discharge in any month. Specifically, they are the three climate variables in the current and the last two months over 71 representative grids. Both the Pearson product-moment correlation coefficient (PC) (Pearson, 1895) and the Spearman’s rank correlation coefficient (SC) (Spearman, 1904) are employed to quantify the correlation between the discharge and any predictor. The related results are shown in Figure 6.3. The p values of PC and SC are much lower than any alternative statistical significance level (e.g. 0.01, 0.05 or 0.10) due to a relatively large sample size, implying that the correlation analysis result is credible.

The SC is capable of reflecting a monotonic relationship that can be nonlinear, while the PC is just a measure of the linear correlation between two variables. The synergistic variation of SCs and PCs reveals that the monotonicity between climatic conditions and streamflow may be dominated by linearity. In comparison with Tmin and Tmax, precipitation has a relatively week correlation with streamflow in the ARB. The temporal variation of streamflow is of more significant synchronism with that of temperature than that of precipitation. Meanwhile, it is illustrated that the climatic conditions in the current month are of more significant impacts on streamflow than those in the last two months and that there is significant colinearity between climatic conditions in three continuous
months. It is implied that the lead time from climate to streamflow may be one month and that the climatic conditions in the current month can be used to represent those in the three selected months. This implication can be used to facilitate screening of predictors in statistical hydrological modeling. However, prescreening of predictors is not required for the BaPMI approach which is capable of automatically identifying the predictors of the most significant impacts on streamflow.

The grids that are relatively close to the gauge station of 07DD001 correspond to the plots in the green rectangulars in Figure 6.3. It is shown that the values of either $PC$ or $SC$ are relatively high for these grids. An implication is that the climatic conditions over these grids have the most significant impacts on streamflow at the outlet of the ARB. Furthermore, the values of $PC$ or $SC$ for every climate variable, especially for precipitation, tend to increase from the downstream to the upper, excluding the grids nearby the gauge station. A guess is that the contributions of climate variables to streamflow decrease from the nearby grids, the upstream grids to the downstream grids excluding these nearby ones. In addition, the spatial variability of the correlations between precipitation and streamflow is higher than that between temperature and streamflow. It is implied that the spatial heterogeneity of precipitation is more significant than that of temperature.
Figure 6.3. Correlations between Climatic Conditions and Discharges. (PC: the Pearson product-moment correlation coefficient; SC: the Spearman’s rank correlation coefficient; PCP: the P value of PC; SCP: the P value of SC).
6.3.4 Model Configuration

The *BaPMI* approach is applied to simulate the correspondence between discharges at every selected gauge station (Figure 6.2b) and the three climate variables in the current and the last two months at the upstream representative grids (Figure 6.2d). The process of maximizing the *principal local marginal monotonicity significance* from \( n \cdot (T - 2 \cdot N_{min} + 1) \) alternatives is compiled as a function. In this function, the independent variables are \( j (j \in \{1, 2, \ldots, n\}) \) and \( u (u \in \{N_{min}, N_{min} + 1, \ldots, T - N_{min}\}) \), the dependent variable is \( w_{ju}(\alpha) \), and the parameters consist of (a) the number of randomly chosen points to sample the target function before Bayesian Optimization fitting the Gaussian Process and (b) the total number of times the Bayesian Optimization is to be repeated. The two parameters are denoted as *InitPnt* and *NIter*, respectively. Based on an analysis of the influences of these parameters, which is presented in the section of discussions, *InitPnt* and *NIter* are set as 6 and 2, respectively in the process of *BaPMI*.

In addition, the *BaPMI* approach has another two parameters: \( \alpha \) (the statistical significance level) and \( N_{min} \) (the minimum partition row number). Normally, they fluctuate within \{0.001, 0.01, 0.05, 0.10\} and \{3, 4, \ldots, 30\}, respectively. Meanwhile, as revealed in (Cheng et al., 2016b), the conventional parameter calibration strategy that the parameter combination of the highest accuracy in one period is selected as the most desired one is unreasonable for statistical hydrological simulation methods. Instead, a bi-period calibration strategy is desired for the *BaPMI* approach. A *BaPMI* model is built for the first calibration period under every parameter combinations, its effectiveness is estimated as the simulation accuracy over the second calibration period, and the optimal parameter combination is the one showing the highest accuracy in the latter period.
In the process of calibration, the Nash-Sutcliffe coefficient \((Nash)\) (Nash and Sutcliffe, 1970) is employed as the principal indicator of the modeling accuracy of \(BaPMI\). Another nine statistical metrics are used to evaluate the multi-dimensional performances of the \(BaPMI\) approach in either calibration or verification. They are: (a) the Pearson product-moment correlation coefficient \((PC)\) (Pearson, 1895), (b) the Kendall rank correlation coefficient \((KC)\) (Kendall, 1938), (c) the Spearman’s rank correlation coefficient \((SC)\) (Spearman, 1904), (d) the root relative squared error \((RRSE)\) (Diaz and Jones, 2004), (e) the root mean square error \((RMSE)\) (Hyndman et al., 2006), (f) the normalized root-mean-square error \((NRMSE)\) (Hyndman et al., 2006), (g) the scatter index \((SI)\) (Hanna and Heinold, 1985), (h) the relative absolute error \((RAE)\) (Armstrong and Collopy, 1992), and (i) the mean absolute error \((MAE)\) (Willmott and Matsuura, 2005).

For any given paired data of hydrological observations and simulation, metrics (a) to (c) are measures of their linear or nonlinear correlation, while metrics (d) to (j) and \(Nash\) are measures of their absolute or relative difference. For any particular metrics, its detailed formulation is stated in the corresponding reference.

6.4 RESULTS ANALYSIS

6.4.1 Calibration

The calibration results are illustrated in Figure 6.4. It is shown that the correlation coefficients, i.e. \(PC\), \(KC\) and \(SC\), are relatively satisfactory for every gauge station, implying the effectiveness of the \(BaPMI\) approach at capturing the temporal variability of streamflow. The \(Nash\) coefficient is higher than 0.50 for the gauge stations of 07AA002, 07AD002, 07AE001, 07AG004, 07BE001, and 07DA001, while relatively low for the other stations especially for 07BK006. The main cause of the unsatisfactory performances
of the *BaPMI* approach for the latter stations excluding 07BK006 is data unrepresentativeness due to data missing. As for 07BK006, the cause may be the existence of a lake, Lesser Slave Lake, which covers 1,160 km$^2$ and is located at the upstream of this gauge station (Figure 6.2b). This lake may significantly alter the correspondence between streamflow and climatic conditions. The optimal values of parameters $\alpha$ and $N_{\text{min}}$ are also displayed in Figure 6.4. It is revealed that there is a significant difference among the optimal parameter values for various gauge stations. An implication is that it is hardly feasible to use the identical parameter values for every gauge stations under spatial heterogeneities.
Figure 6.4. Calibration and Verification Results.
6.4.2 Verification

Subsequently, the effectiveness of the \textit{BaPMI} approach is verified through another period. The overall performances of \textit{BaPMI} are shown in Figure 6.4. Due to the bi-period calibration strategy used in this study, the performances of \textit{BaPMI} in verification are relatively identical with those in calibration. For instance, the correlation coefficients are relatively satisfactory for every gauge stations. The \textit{BaPMI} approach is effective at reproducing the temporal variability of streamflow. The modeling accuracies are the worst for the station of 07BK006 due to the significant influences of Lesser Slave Lake. The seasonal averages of discharges at three selected gauge stations, 07AD002, 07AE001, and 07BE001, are also presented in Figure 6.4 to elaborate the capability of \textit{BaPMI}. It is indicated that the temporal variability of streamflow is well captured by the \textit{BaPMI} approach. There would be a slight difference between the observation and simulation of streamflow which fluctuates within a relatively small range. In addition, the \textit{BaPMI} approach is capable of enabling streamflow simulation under uncertainties. For any given observation of streamflow, a distribution of simulation can be generated. To evaluate this capability of \textit{BaPMI}, the relative location of the observation in simulation is quantified as a percentile ranging from 0 to 1. It is shown in Figure 6.4(c) to Figure 6.4(e) that the percentiles of streamflow observations are distributed over \([0, 1]\) without significant skewness for these stations. This implies that the \textit{BaPMI} approach is effective at reflecting the uncertainties in the correspondence between climatic conditions and streamflow at these stations. Consequently, this approach can be used to support the related risk assessment and management.

6.4.3 Bias Analysis
As revealed in the calibration and verification process, there are biases between streamflow observations and simulation. When the simulation results of the *BaPMI* approach are applied to guide water resources management or other related human activities, these biases should be taken into account to avoid unreasonable decisions. A detailed analysis of these biases is presented in Figure 6.5. Figure 6.5(a) and Figure 6.5(b) show the absolute bias, i.e. the average of (simulation – observation), and the relative bias, i.e. the average of (simulation – observation)/observation, at various temporal scales, respectively. It is illustrated in Figure 6.5(b) that the multi-year averaged relative bias is approximately zero for every gauge stations. Besides, the *BaPMI* approach tends to overestimate streamflow in January, February, March, August, September and October, while underestimating it in April to July and November to December. The relative bias is more significant for 07AG004, 07BC002 and 07DD001 than for the other stations. Figure 6.5 can be helpful for correcting the modeling biases of the *BaPMI* approach when the simulation results are directly used for guiding socio-economic development or eco-environmental conservation practices over the *ARB*. 
Figure 6.5. Modeling Biases.
6.5 DISCUSSIONS

6.5.1 Bayesian Gaussian Process Optimization

One of representative advantages of the BaPMI approach is that the BaGPO algorithm is integrated with advanced statistical inferential techniques to enable hydro-system analyses under complexities such as data uncertainties, non-normal distributions, nonlinear correspondences, and heavy computational loads. The effectiveness of this integration is evaluated through a comparison of the BaPMI results with or without the module of BaGPO. The gauge station of 07DD001 around the outlet of the ARB is selected as a representative case because the computational load is highest at this station. The number of alternative predictors reaches 639 for this station. The \( p \) values corresponding to four selected predictors are shown in Figure 6.6(a). If the BaGPO algorithm is not used, the process of identifying the minimal \( p \) value which corresponds to the principal local marginal monotonicity significance would cost 162.44 seconds. The minimal \( p \) value equals to \( 1 \times 10^{-85} \).

The BaGPO algorithm involves two parameters: the number of randomly chosen points to sample the target function before Bayesian Optimization fitting the Gaussian Process (\( \text{InitPnt} \)) and the total number of times the Bayesian Optimization is to repeated (\( \text{NIter} \)). Prior to incorporating this algorithm into the framework of BaPMI, a sensitivity analysis is carried out to examine the impacts of the two parameters. The parameter value ranges are set as \{6, 7, \ldots, 10\} and \{1, 2, \ldots, 5\} for \( \text{InitPnt} \) and \( \text{NIter} \), respectively. The computational time and the minimal \( p \) value which corresponds to the principal local marginal monotonicity significance for every combination of \( \text{InitPnt} \) and \( \text{NIter} \) are illustrated in Figure 6.6(b). It is revealed from the distributions of \( p \) values that in this case...
the minimal $p$ value obtained from $BaGPO$ or the estimated value of the principal local marginal monotonicity significance is not sensitive to either $InitPnt$ or $NIter$.

In addition, there is not a significant monotonic relationship between the minimal $p$ values and the computational time, and the computational time remains around 6 seconds in most runs. This implies that the computational efficiency of $BaGPO$ is insensitive to the two parameters. The averaged computational time is 6.05 seconds which is much lower than that in the case of not using $BaGPO$. Meanwhile, the estimation of the minimal $p$ value ranges from $1 \times 10^{-5}$ to $1 \times 10^{-85}$ that are very close to the real value ($1 \times 10^{-85}$). From the viewpoint of statistical inference, the difference between the estimated $p$ values and the real value is negligible. It is indicated that, while guaranteeing the reliability of the obtained minimal $p$ value, the $BaGPO$ algorithm is advantageous at highly reducing computational time. Incorporating this algorithm into the framework of $BaPMI$ is much helpful for improving the applicability of $BaPMI$ in large-scale hydro-system analysis problems.
Figure 6.6. Bayesian Optimization.
6.5.2 Correspondence Analysis

Another advantage of the $BaPMI$ approach is that the pre-screening of alternative predictors is not required. The predictors of significant impacts on streamflow can be automatically detected in the process of $BaPMI$. These predictors are equivalent to the criterion predictors on which the grouping of the paired samples of the predictors and the predictand relies and that represent the most significant marginal monotonicity between the predictors and the predictand within a node. An analysis of the criterion predictors in this case study is helpful for revealing the correspondence between climatic conditions and streamflow. The criterion predictors for streamflow at every gauge station are presented in Figure 6.7. The key finding from a comparison of the criterion predictors for all gauge stations is that the daily minimum temperature, followed by the daily maximum temperature and the daily cumulative precipitation, around the glacier field at the upstream is of very significant impacts on the streamflow in the main channel over the $ARB$. In addition, the streamflow at a gauge station is close related with the nearby climatic conditions, which is a common law in hydro-meteorological systems. Besides, the ratios of the three climate variables, i.e. $T_{min}$, $T_{max}$ and $Prec$, being taken as the criterion predictors for all gauge stations equal to 0.40, 0.38 and 0.22, respectively. It is implied that the overall impact of climate variables on streamflow over the $ARB$ decreases from $T_{min}$, $T_{max}$ to $Prec$. 

228
Figure 6.7. Criterion Predictors for Every Gauge Station.
6.5.3 Sensitivity Analysis

The BaPMI approach includes two parameters: \( \alpha \) (the statistical significance level) and \( N_{min} \) (the minimum partition row number). Normally, their alternative values are \{0.001, 0.01, 0.05, 0.10\} and \{3, 4, ..., 30\}, respectively. In large-scale hydrosystem analysis problems, the upper bound of \( N_{min} \) can be extended to a larger integer, while the alternative values of \( \alpha \) remain from the viewpoint of conventional statistical analysis. To examine the impacts of these parameters on the modeling accuracy, the BaPMI approach is applied to streamflow simulation for every gauge station and every combination of \( \alpha \) and \( N_{min} \). The correspondence among modeling accuracies (quantified by the Nash coefficient) and parameter values is illustrated in Figure 6.8. The variation of modeling accuracies with any parameter at every gauge station is fitted by a high-order polynomial function for better visualizing the correspondence. It is revealed that there is the sensitivities of \( \alpha \) and \( N_{min} \) significantly vary with gauge stations. The parameter sensitivities decrease with the modeling accuracies of the BaPMI approach at various gauge stations. Under the same variation of the two parameters, the modeling accuracies do not significant change for the gauge stations for which the modeling accuracies are relatively high, while showing significant changes for the other gauge stations such as 07BK006, 07DD001, and 07BC002. It is also shown in Figure 6.8 that in general the modeling accuracies of BaPMI tend to decrease with \( \alpha \) and increase with \( N_{min} \). Hence, the suggested values of \( \alpha \) and \( N_{min} \) are 0.001 and 30, respectively when the BaPMI approach is applied to other case studies of hydro-system analyses. This suggestion on parameter setting may not directly help identify the real optimal values of the two parameters, but at least provides an initial parameter combination that deserves a test. The
upper bound of $N_{min}$ is set as 30 in this sensitivity analysis. This may lead to a potentially more optimal value being neglected, which is implied in the monotonically increasing relationship between $N_{min}$ and the modeling accuracy. Extending the upper bound of $N_{min}$ to a larger integer may be helpful for improving the modeling accuracies of the BaPMI approach.
Figure 6.8. Impacts of Parameters in BaPMI on the Verification Results.
6.6 SUMMARY

In this chapter, an advanced statistical hydro-system analysis approach abbreviated as *BaPMI* was proposed for supporting hydro-meteorological system analyses under complexities of data uncertainties, non-normal distributions, nonlinear correspondences, multivariate dependencies, and massive computations. In essence, the *BaPMI* approach was an integration of statistical inference, Gaussian process analysis, and Bayesian optimization. These techniques were integrated into an ingeniously design framework that could mitigate the challenges of these complexities in hydro-meteorological systems. The reliability of this approach was verified through a case study of streamflow simulation in the *ARB*. A series of analyses such as a correlation analysis, a bias analysis, a correspondence analysis, and two sensitivity analyses helped deepen the understandings of this approach, reveal the mechanisms of the hydrological regime over the *ARB*, disclose the limitations of this approach, and divulge the potential opportunities in the subsequent studies.

A few of findings were revealed from these analyses. For instance, the lead time from climatic conditions to streamflow may be one month for the *ARB*. The contributions of climate variables to streamflow generation decrease from the nearby grids, the upstream grids to the downstream grids excluding these nearby ones. The spatial heterogeneity of precipitation is more significant than that of temperature. The *BaPMI* approach is effective at capturing the temporal variability and the multi-year averages of streamflow for every gauge station and reflecting the uncertainties in the correspondence between climatic conditions and streamflow at part of stations. Meanwhile, this approach tends to overestimate streamflow in January, February, March, August, September and October,
while underestimating it in the other months. The \textit{BaGPO} algorithm is capable of highly reducing computational time while guaranteeing modeling reliability, which is much helpful for enhancing the applicability of \textit{BaPMI}. The daily minimum temperature, followed by the daily maximum temperature and the daily cumulative precipitation, around the glacier field at the upstream is of very significant impacts on the streamflow in the main channel over the \textit{ARB}. In addition, the overall impact of climate variables on streamflow over the \textit{ARB} decreases from \textit{Tmin}, \textit{Tmax} to \textit{Prec}. The parameter sensitivities decrease with the modeling accuracies of the \textit{BaPMI} approach at various gauge stations, and the modeling accuracies of \textit{BaPMI} tend to decrease with \( a \) and increase with \( N_{min} \).

On the other hand, there are also many opportunities to further improve this approach or apply it to reveal the mechanism of more hydro-meteorological systems. Among them, the representative ones include but are not limited to the followings. A continuous function that could enable the reversible transformation between any non-normal distribution and a normal distribution may be more effective than the \textit{DDT} approach used in this study. The impacts of distribution transformation on the modeling results and accuracies deserves an attempt. There may be a more reasonable scheme of nonlinearity characterization than the current one that the nonlinearity between climatic conditions and streamflow is represented by the principal monotonicity in the paired samples. Incorporating variance analyses into the framework of \textit{BaPMI} may facilitate a comprehensive evaluation of the difference between any two groups of multi-dimensional streamflow samples. A scientific scheme, possibly integration with physically-based hydrological models, is desired for mitigating one challenge of the \textit{BaPMI} approach that the simulation results tend to fluctuate within the historical records. The methods of global
sensitivity analysis can help reveal complicated interactions among the parameters, the boundary conditions, and the modeling results of the BaPMI approach. In-depth analyses on the causes of variation and diversity of the modeling accuracies of BaPMI may help advance this approach’s performances. Attentions should also be paid on building the connections between real-world hydrological systems and the BaPMI approach which is a statistical or data-driven hydro-system analysis approach. Developing a dynamic version of the BaPMI approach to adapt to the temporal variability of hydrological regimes may be one of schemes that could improve this approach. Consideration of various uncertainties except data uncertainties can also be an alternative direction in the following studies. More case studies such as enabling long-term projection of the hydrological regime over the ARB are required to verify the reliability of this approach in various studies or practices, reveal its limitations for further improvement, and extend its potential contributions.
CHAPTER 7
STREAMFLOW PROJECTION OVER THE ATHABASCA RIVER BASIN
UNDER CLIMATE CHANGE

7.1 BACKGROUND

Changes in climate can have profound effects on river systems, cause significant variation in streamflow and water availability, and influence water uses such as water withdrawals for agricultural irrigation that are highly dependent on the hydrological regime (Ravazzani et al., 2015). For instance, surface water use by oil sands mining operations at the downstream of the ARB accounts for the largest sectorial water allocations (62%) and actual water use (57%) in the ARB (AMEC Earth and Environmental, 2007). As mining activity expands, a relatively large amount of water use for mining in winters is threatening the health of habitats in the lower-stream wetlands. This conflict would deteriorate in the following decades if streamflow in the ARB would decline due to impacts of climate change, especially in winters. Hence, reliable projection of streamflow under climate change are highly desired for scientific planning of socio-economic development and eco-environmental conservation practices, e.g. reservoir operations and irrigation management, protection of the environment, and flood and drought mitigation.

Previously, many studies were carried out to achieve streamflow projection under climate change, e.g. (Gutiérrez and Dracup, 2001; Chiew et al., 2003; Chandimala and Zubair, 2007; Ionita et al., 2008; Gámiz-Fortis et al., 2010; Lamb et al., 2010; Oubeidillah et al., 2011). More recently, Sulis et al. (2011) applied a fully integrated surface-subsurface hydrological model for climate change impact analysis. Das et al. (2013) examined the
increases of flood magnitudes in California under warming climates through downscaled and hydrologically modeled projection from an ensemble of 16 GCMs. Kalra et al. (2013) used the support vector machine which was driven by annual average oceanic–atmospheric indices to predict annual streamflow volumes for multiple sites in the Gunnison River Basin and San Juan River Basin. Sittichok et al. (2014) analyzed the ability of various statistical techniques to forecast the July-August-September total rainfall and monthly streamflow in the Sirba watershed (West Africa). Musau et al. (2015) assessed the impact of climate change on the streamflow in Mt. Elgon watersheds using the 10-GCM Special Report on Emissions Scenarios in combination with a hydrological model (SWAT). To examine water resource shortage and salt-water intrusion during dry seasons in Pearl River, China, Yan et al. (2015) evaluated the variations in low flow using the Variable Infiltration Capacity model driven by bias-corrected results of five GCMs under scenarios RCP4.5 and 8.5. In (Zhang et al., 2016), climate data was projected from three CMIP5 GCMs under RCP scenarios using a statistical downscaling model (SDSM), and future streamflow was modeled through coupling climatic projection with SWAT.

Meanwhile, many issues regarding regional socio-economic development and eco-environmental conservation, e.g. the conflict between water consumptions of oil-sands mining and ecosystem health at the downstream, is closely related with the temporal and spatial variability of the streamflow over the ARB. The streamflow is associated with climate change that would significantly alter the hydrological regimes over this river basin. Long-term planning of socio-economic and eco-environmental systems desires reliable projection of the streamflow under climate change. Few studies were carried out to examine the future changes of streamflow over the ARB. Until recently, two papers
presented some related studies. In (Leong and Donner, 2015), the combination of a land surface process model and a hydrological routing model was used to evaluate the influence of water withdrawals and climate change on streamflow in the ARB. In (Sauchyn et al., 2015), the decadal-scale variability in river discharge in the ARB was examined by a generalized least-squares regression analysis of the trend and variability in gauged flow. It was claimed in that study that there would be long-term declining flows throughout the ARB.

However, the reliability of these existing studies is challenged by multiple complexities in coupled climatic and hydrological systems analyses. Some representative ones are stated as below. In particular, a comprehensive evaluation of the multi-dimensional performances of GCMs, which could build a solid foundation for providing reliable climatic projection to drive streamflow forecasting, was barely carried out in existing studies. These challenges would further propagate into streamflow projection, leading to decreased reliability of projection results and posing threats on socio-economic development and eco-environmental health.

(a) Normally, the main driver of streamflow projection is climatic projection that are achieved through GCMs, especially ones in the CMIP5. A challenging problem for climate-change impact studies is to decide which GCM should be chosen. Furthermore, the performances of GCMs vary with multiple factors such as climate variables, geographic locations, temporal units and statistical metrics. A comprehensive evaluation of the multi-dimensional performances of GCMs is desired to support GCM screening.

(b) For large regions such as the ARB, redundant computations are a key challenge for climate modeling and impact studies. Meanwhile, there may be significant similarities
among the climatic conditions over a large number of grids. Climate classification is helpful for reducing computational loads, which is further challenged by complexities such as data uncertainties, multivariate dependencies, and normal-distribution assumptions.

(c) The spatial resolution of GCMs is relatively coarse. The atmospheric mesoscale features and the land surface heterogeneity can hardly be properly resolved in GCMs. Statistical downscaling can help fill the gap between coarse-resolution GCMs and finer-resolution regional or local climate. However, data uncertainties, nonlinear correspondences, climate-variable dependencies, non-normal distributions, spatial homogeneities, and temporal nonstationarities are challenging the reliability of statistical downscaling.

(d) High-resolution climatic projection with high reliability can build a solid foundation for climate-change impact or adaptation studies. Nevertheless, very few studies were conducted to enhance the resolution of climate projection over the ARB and the existing ones rely on a few of simplifications of the downscaling process. Furthermore, the challenges in GCM selection, climate classification and statistical downscaling are propagated into climatic projection which becomes more challenging.

(e) Most importantly, streamflow projection under climate change desire a reliable model to quantify the correspondence between streamflow and climatic conditions. Nonetheless, there are a few of complexities in hydro-meteorological systems, e.g. massive computations, data uncertainties, nonlinear correspondences and multivariate dependencies, which are challenging the reliability of the hydro-meteorological model to be constructed.
Therefore, the objective of this chapter is to develop an integrated hydro-climatic systems analysis framework (IHCSA) that can mitigate these challenges, enhance the reliability of streamflow projection under climate change, and facilitate climate-change impact or adaptation studies. This framework will be applied to a case study of streamflow projection in the ARB. Specifically, the characteristics of the ARB, the key structure of IHCSA, data requirements, and the streamflow projection analysis method will be presented in section 7.2. In section 7.3, the multi-dimensional impacts of climate change on streamflow over the ARB will be examined systematically, and the advantages and shortcomings of IHCSA will also be discussed.

7.2 STUDY AREA AND METHOD

7.2.1 Study Area

The ARB covers an area of approximately 138,000 km² over 52° ~ 59° N and -119° ~ -107° W. Its origin is the Columbia Icefield on the eastern slopes of the Rocky Mountains around Jasper in Alberta, Canada, leading to a close connection with climatic warming. This river is the longest undammed river in the Canadian prairies, which is another particular feature of it. This river flows approximately 1500 km northeast before entering the Peace-Athabasca Delta, the largest freshwater continental river delta in North America and the home to numerous migratory birds, and draining into Lake Athabasca. The evaluation drops from approximately 3715 m at Jasper to 211 m at the outlet in the Peace-Athabasca Delta (Figure 7.1a), resulting in a large variation of discharges across the ARB. The ARB includes diverse hydro-climatic regimes due to physiographical heterogeneity; snowcapped mountains, coniferous forest, mixed wood and deciduous forest are found in the uplands, whereas willow brush, shrubs, black spruce and sphagnum moss dominate
the lowlands (Kerkhoven and Gan, 2006). Around 0.15 to 0.17 million residents distribute across this river basin. Many issues regarding regional socio-economic development and eco-environmental conservation, e.g. the conflict between water consumptions of oil-sands mining and ecosystem health at the downstream, is closely related with the temporal and spatial variability of the streamflow over the ARB. Meanwhile, the streamflow is associated with climate change that would significantly alter the hydrological regimes over this river basin. On the contrary, the impacts of human activities on the streamflow in the ARB are very limited, because there is not water conservancy facilities in the channels of the ARB and, as revealed in (Leong, 2014), water withdrawals for socio-economic development were relatively little. Long-term planning of socio-economic and eco-environmental systems desires reliable projection of the streamflow under climate change. Therefore, this chapter focuses on streamflow projection based on an integration of multiple achievements in hydro-meteorological system analyses.
Figure 7.1. Study Area and Data.
7.2.2 Framework Development

Streamflow projection under climate change are confronted with multiple challenges as discussed in the section of introduction. Accordingly, we developed an integrated climatic and hydro-meteorological systems analysis framework. This framework consists of several interrelated modules that could mitigate these challenges. Specifically, (a) the multi-dimensional accuracies of CMIP5 GCMs and their ensemble in reproducing the historical climatic conditions over the ARB were comprehensively evaluated (Chapter 2). It was revealed that, while the accuracies of CMIP5 GCMs varied with climate variables, geographical locations, statistical metrics, and temporal or spatial scales, the overall performance of the multi-model ensemble was relatively better than that of the other climate simulation. (b) A recursive dissimilarity and similarity inferential meteorological classification approach was developed to support classification of regional climate under various complexities (Chapter 3). Through this approach, the ARB was classified as 20 meteorological zones under a given parameter combination. (c) A recursive multivariate principal-monotonicity inferential downscaling approach was proposed for climate downscaling under complexities such as data uncertainties and temporal nonstationarities (Chapter 4). The spatial resolution of climate modeling over the ARB was enhanced to 10 kilometres with relatively high reliability. (d) Projection of the climate over the ARB from 2021 to 2100 at a resolution of 10 kilometres under RCPs 2.6, 4.5, 6.0 and 8.5 were enabled through an integration of the results in (a) to (c) (Chapter 5). Climate change over the ARB and their related complexities were examined systematically. (e) An advanced statistical hydro-system analysis approach was developed for quantifying the complicated correspondence between climatic conditions and
streamflow over the ARB (Chapter 6). The modeling accuracies of this approach were relatively satisfactory as revealed by a series of evaluations and comparisons. The related details are presented in the corresponding publications that readers may refer to. These studies build a solid foundation for streamflow projection under climate change over the ARB. In particular, high-resolution projection of climate over the ARB are provided with relatively high reliability in (d) and the correspondence between climatic conditions and streamflow is reproduced by an advanced model in (e). In this chapter, these results will be integrated within a general framework to enable streamflow projection under climate change and to assess the impacts of climate change on streamflow over the ARB.

7.2.3 Data

In this study, streamflow projection over the ARB are driven by two datasets. One is high-resolution projection of 21st century climate over the ARB which are derived from Cheng et al. (2016f). This dataset is generated through downscaling of the ensemble means of six CMIP5 GCMs: IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC-ESM-CHEM, MIROC5, GFDL-ESM2G, and GFDL-ESM2M (Table 1). Three climate variables, i.e. daily minimum near-surface air temperature ($T_{min}$), daily maximum near-surface air temperature ($T_{max}$), and daily precipitation ($Prec$), which are of relatively high interests for impact studies are involved in this dataset to represent the regional climate over the ARB. This dataset consists of the monthly averages of these climate variables from 2021 to 2100 at 1615 10-km grids (Figure 7.1c) over the ARB under four RCP scenarios: RCPs 2.6, 4.5, 6.0 and 8.5. In addition, the other one is a hydro-meteorological correspondence analysis dataset that was contributed by Cheng et al. in (Chapter 6). This dataset involves a series of Bayesian principal-monotonicity inference (BaPMI) models that were
constructed for quantifying the correspondence between the monthly averages of three climate variables \((\text{Tmin}, \text{Tmax} \text{ and } \text{Prec})\) and the discharges at ten selected gauge stations (Figure 7.1b and Table 6.1) in the \(ARB\) based on an integration of advanced statistical inference, Gaussian process analysis, and Bayesian optimization. Meanwhile, this dataset also includes many results related with various analyses such as modeling accuracies, biases, uncertainties, and parameter sensitivities that are helpful for gaining insights into the potential changes of climate over the \(ARB\). In addition, the historical discharges at the selected gauge stations are required for analyzing the changes of streamflow. These data are downloaded from the website of Environment and Climate Change Canada (https://ec.gc.ca/). The records of daily discharges are upscaled to monthly averages for matching the temporal resolution of climatic projection.
<table>
<thead>
<tr>
<th>Model Name</th>
<th>Resolution (Lon × Lat)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPSL-CM5A-LR</td>
<td>$3.75^\circ \times 1.875^\circ$</td>
<td>Institut Pierre-Simon Laplace, France</td>
</tr>
<tr>
<td>IPSL-CM5A-MR</td>
<td>$2.5^\circ \times 1.259^\circ$</td>
<td></td>
</tr>
<tr>
<td>MIROC5</td>
<td>$1.406^\circ \times 1.406^\circ$</td>
<td>Atmosphere and Ocean Research Institute, Japan</td>
</tr>
<tr>
<td>MIROC-ESM-CHEM</td>
<td>$2.812^\circ \times 2.812^\circ$</td>
<td></td>
</tr>
<tr>
<td>GFDL-ESM2M</td>
<td>$2.5^\circ \times 2.0^\circ$</td>
<td>Geophysical Fluid Dynamics Laboratory, NOAA, USA</td>
</tr>
<tr>
<td>GFDL-ESM2G</td>
<td>$2.5^\circ \times 2.0^\circ$</td>
<td></td>
</tr>
</tbody>
</table>
7.2.4 Projection

For every selected gauge station (Table 2) in the ARB, there is a BaPMI model that simulates the correspondence between the discharges at this station and a series of predictors. These predictors are mainly composed of the monthly averages of the three climate variables (i.e. \(T_{min}\), \(T_{max}\) and \(Prec\)) in the current and the last two months at representative grids (Figure 7.1d) which are identified through climate classification (Chapter 3) and a grids-similarity analysis (Chapter 4). The BaPMI model is equivalent to a systematic discretization of the multi-dimensional space of these climate variables. The space is deconstructed as a range of sub-spaces that correspond to particular value ranges of every climate variable. Every sub-space also corresponds to a series of the samples of discharges. When this model is used for projecting the future streamflow conditions, any given multi-dimensional sample of climate variables is taken as the input of the BaPMI model. The corresponding sub-space of climate variables is located based on a discriminant analysis. Accordingly, the potential values of the discharge can be identified and used for estimating the mean value or distribution of the discharge for the given combination of the values of the climate variables. Through this projection scheme, the mean value and the discrete distribution (expressed as percentiles) of discharges in every temporal and spatial unit under every RCP scenario are generated, enabling streamflow projection under climate change.

The changes of streamflow vary with temporal scales and can be quantified by various statistical metrics. A series of hydrological indices are proposed to systematically indicate the streamflow changes over the ARB. First of all, the historical averaged discharges in multiple years and 12 months are calculated for every gauge station, building
a reference for estimating streamflow changes. For every gauge station and every RCP scenario, the originally simulated discharge averages (OSDAs) at the temporal scales of octo-decades (2021 to 2100), bi-decades ([2021, 2040], [2041, 2060], [2061, 2080], and [2081, 2100]), annuals, and months are calculated, respectively. Besides, the biases in the process of building BaPMI models for every gauge station would influence the reliability of the originally simulated discharge averages. As one of means to enhance the reliability of these results, these modeling biases are removed, leading to another set of modeling results: the corrected simulated discharge averages (CSDAs). Meanwhile, there is a significant difference among the magnitudes of discharges at all selected gauge stations. To eliminate such a difference, the discharge changing ratio which equals to the average of (projection – historical averages)/(historical averages) is introduced into this study, generating another two sets of projection results: the originally simulated discharge changing ratios (OSDRs) and the corrected simulated discharge changing ratios (CSDRs). Furthermore, there are uncertainties in the correspondence between climatic conditions and discharges. These uncertainties are propagated into the streamflow projection results through the constructed BaPMI models, leading to the uncertainties in streamflow projection. To examine streamflow changes under uncertainties, two quartiles (0.25 and 0.75) of the potential values of OSDRs and CSDRs are estimated at the monthly scale. The related results are shown in the following figures.

7.3 RESULTS ANALYSIS AND DISCUSSIONS

7.3.1 Octo-Decadal Averaged Discharge Changes

The octo-decadal averaged changes of discharges at the selected gauge stations over the ARB are presented in Figure 7.2. The detailed results are listed in the appendix. It is
illustrated that the discharges would increase for the gauge stations of 07AA002, 07AD002, 07BC002, 07BE001, 07BK006 and 07CD001, while decreasing for the other stations. In comparison with the absolute magnitudes of discharges, the impacts of either RCP scenarios or modeling biases on the octo-decadal averaged discharges are relatively low for every gauge station. Furthermore, the impacts of RCP scenarios on octo-decadal averaged discharges vary with gauge stations, showing high spatial heterogeneity. The octo-decadal averaged discharges would be the highest for RCP 8.5 and gradually decrease from RCPs 4.5, 6.0 to 2.6 for most gauge stations, while being of completely different patterns for the other stations. In addition, the octo-decadal averaged discharge changing ratios would be the highest for 07BK006 and 07AA002 before and after bias corrections, while being the lowest for 07AG004 and 07DD001, respectively. This implies that the modeling biases, of which the impacts are limited compared with the absolute magnitudes of discharges, are significant for the variability of octo-decadal averaged discharge changing ratios among gauge stations. In most combinations of gauge stations and RCP scenarios, the modeling biases would not reverse the octo-decadal averaged discharge changes from decreasing to increasing or from increasing to decreasing. An implication is that the modeling biases are relatively low in comparison with the octo-decadal averaged discharge changing ratios over the ARB.
Figure 7.2. Octo-Decadal Averaged Discharges.
<table>
<thead>
<tr>
<th>StationID</th>
<th>OSDA (m³/s)</th>
<th>OSDR (m³/s)</th>
<th>CSDA (m³/s)</th>
<th>CSDR (m³/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.6</td>
<td>4.5</td>
<td>6.0</td>
<td>8.5</td>
</tr>
<tr>
<td>07AA002</td>
<td>92.3</td>
<td>94.9</td>
<td>94.7</td>
<td>97.3</td>
</tr>
<tr>
<td>07AD002</td>
<td>184.0</td>
<td>186.7</td>
<td>185.3</td>
<td>193.5</td>
</tr>
<tr>
<td>07AE001</td>
<td>241.2</td>
<td>251.0</td>
<td>249.8</td>
<td>259.3</td>
</tr>
<tr>
<td>07AG004</td>
<td>50.9</td>
<td>50.4</td>
<td>49.8</td>
<td>52.5</td>
</tr>
<tr>
<td>07BC002</td>
<td>36.6</td>
<td>37.7</td>
<td>37.4</td>
<td>38.6</td>
</tr>
<tr>
<td>07BE001</td>
<td>436.9</td>
<td>453.5</td>
<td>433.7</td>
<td>457.4</td>
</tr>
<tr>
<td>07BK006</td>
<td>56.6</td>
<td>57.5</td>
<td>57.0</td>
<td>57.8</td>
</tr>
<tr>
<td>07CD001</td>
<td>121.1</td>
<td>122.1</td>
<td>123.0</td>
<td>126.2</td>
</tr>
<tr>
<td>07DA001</td>
<td>603.8</td>
<td>587.9</td>
<td>587.7</td>
<td>603.5</td>
</tr>
<tr>
<td>07DD001</td>
<td>816.6</td>
<td>851.2</td>
<td>839.5</td>
<td>842.9</td>
</tr>
</tbody>
</table>
7.3.2 Bi-Decadal Averaged Discharge Changes

The bi-decadal averaged values of $O_{SDA}$, $C_{SDA}$, $O_{SDR}$ and $C_{SDR}$ at every gauge station under four $RCP$ scenarios are shown in Figure 7.3. It is illustrated that the bi-decadal variability of averaged discharge averages or changing ratios varies with gauge stations and $RCP$ scenarios significantly. In most cases, the bi-decadal averaged discharges would decrease from [2021, 2040] to either [2041, 2060] or [2061, 2080] and increase to [2081, 2100] under $RCP$ 2.6, while gradually increasing from [2021, 2040] to [2081, 2100] under the other $RCP$ scenarios. These patterns would not be valid for some gauge stations, especially for 07DA001 and 07DD001 that are of totally different patterns. Furthermore, the increasing rate under $RCP$ 8.5 is the highest in comparison with that under any other scenario. In addition, whether the averaged discharges would increase or not at the scale of bi-decades remains identical with that at the octo-decadal scale for almost all combinations of $RCP$ scenarios and gauge stations.
Figure 7.3. Bi-Decadal Variability of Averaged Discharges.
7.3.3 Inter-Annual Variability of Discharge Changes

The annual averages of *OSDA* and *CSDA* from 2021 to 2100 at the selected gauge stations under every *RCP* scenario are shown in Figure 7.4. High-order polynomial functions are introduced to fit the annual averages and visualize the temporal trend of discharges within the period of from 2021 to 2100. It is illustrated that, for most gauge stations, the annually averaged discharges would remain above or below the historical averages for almost all years under every *RCP* scenario. As for the station of 07BE001 or 07CD001, the annually averaged discharge may be lower than the historical average in part of years while exceeding it in most years. On the contrary, the annual discharge average may exceed the historical average in the late 21st century for the stations of 07AE001, 07DA001 and 07DD001 due to the impacts of climate change. For most gauge stations, the modeling biases would not have significant impacts on whether the annual discharge average is higher than the historical average or not; one representative exclusion is the station of 07BC002. For this station, the original simulation of the annual discharge averages would be higher than the historical average, and the corrected simulation would be lower than it. Namely, the annual discharge anomalies at some stations such as 07BC002 would be sensitive to modeling biases while those at the other stations not. In addition, there are significant differences among the annual trends of discharges among the four *RCP* scenarios for every gauge station, of which the temporal patterns are of high similarity with those at the bi-decadal scale.
Figure 7.4. Annual Flow. (Pink lines: historical averages; *OSDA*: originally simulated discharge averages; *CSDA*: corrected simulated discharge averages).
7.3.4 Monthly Variability of Discharge Changes

Figure 7.5 shows the multi-year-averaged means, the first quartiles and the third quartiles of monthly \textit{OSDR} and \textit{CSDR} at the selected gauge stations under four \textit{RCP} scenarios. It is illustrated that the monthly variability of discharge changes would significantly vary with gauge stations. For most gauge stations, the monthly discharges would tend to increase in January, February, March, August, and September, while decreasing in April to July. In the other months, whether the monthly discharges would increase or decrease varies with gauge stations. For instance, the discharges would decrease in November and December for 07AA002, 07AD002, 07AG004, 07DA001 and 07DD001, while increasing for the other stations. A comparison of the averaged discharge changes in all months reveals that the increments of monthly discharges would be the highest in March for almost all gauge stations. In general, the modeling biases as well as the \textit{RCP} scenarios have insignificant impacts on the overall trends of monthly discharges in most combinations of months and gauge stations. In most combinations of gauge stations and months, the conclusion of whether monthly discharges would increase or decrease would not vary with \textit{RCP} scenarios or bias corrections. On the contrary, the uncertainties in streamflow projection would significantly alter the overall trends of monthly discharges, especially for the gauge stations of 07AD002, 07AE001, 07BC002 and 07BK006 and the months of August to October. Furthermore, for the gauge stations where the octo-decadal averaged discharges would decrease, e.g. 07AG004 and 07DA001, the decrements of discharges would be dominated by less flow in July or \textit{JJA} when the magnitudes of streamflow are the highest.
Figure 7.5. Monthly Variability of Flow. (CSDR: corrected simulated discharge changing ratios; OSDR: originally simulated discharge changing ratios; Mean: the mean values; Q0.25: the first quartile; Q0.75: the third quartile).
7.3.5 Comparisons with Existing Studies

Previously, few studies were carried out to examine the future changes of streamflow over the ARB. Until recently, two papers presented some related studies. In (Leong and Donner, 2015), the combination of a land surface process model and a hydrological routing model was used to evaluate the influence of water withdrawals and climate change on streamflow in the ARB. It was revealed that there would be a decline in summer flows and an increase in winter flows around a gauge station that may be located between 07DA001 and 07DD001. In (Sauchyn et al., 2015), the decadal-scale variability in river discharge in the ARB was examined by a generalized least-squares regression analysis (GLSRA) of the trend and variability in gauged flow. Sauchyn et al. (2015) claimed that there would be long-term declining flows throughout the ARB. The evidence was mainly the decline of discharges at several selected gauge stations that might correspond to 07AD002, 07AG004, 07BE001, 07DA001 and 07CD001 in this study.

A comparison of the findings in this study and ones in these existing ones can help reveal several interesting viewpoints. For instance, it was disclosed in this study that the octo-decadal averages of the streamflow at 07DA001 and 07DD001 which are adjacent with each other would decline under every RCP scenario (Figure 7.2). Hence, it is highly possible that, for the gauge station used in (Leong and Donner, 2015), the streamflow would also decline. It is implied that this study may reach the same conclusion regarding how the streamflow between 07DA001 and 07DD001 would change in the future. However, many findings in this study, e.g. the spatial heterogeneity of streamflow changes, were not mentioned in (Leong and Donner, 2015), which reflects the advantages of this study.
It was illustrated in Figure 7.4 and Figure 7.5 that the discharges around 07AG004 and 07DA001 would tend to decrease while increasing for 07AD002, 07BE001 and 07CD001. There is an agreement as well as a conflict between this study and (Sauchyn et al., 2015) regarding projection of the discharges around these gauge stations. The discharges at 07AG004 and 07DA001 would decline for both the *BaPMI* approach in this study and the GLSRA method in (Sauchyn et al., 2015). An implication is that the historical temporal trend of discharges at the two stations might remain in the 21st century. On the other hand, the opposite conclusion regarding how the discharges around 07AD002, 07BE001 and 07CD001 would change in this study and (Sauchyn et al., 2015) reveals that due to the impacts of climate change the historical trend of discharges at these stations might not remain in the 21st century. Climate change may alter the historical or current hydrological regimes around some gauge stations and result in a challenge of nonstationarity on conventional trend or frequency analysis techniques.

### 7.3.6 Potential Extensions

From the viewpoint of either methodology or application, there are a few of opportunities to improve this study and extend it for the related studies or practices. The representative ones include but are not limited to the followings. It was stated in Chapter 4 that the method of generating the climatic projection results might over-estimate temperature in springs and winters and precipitation in summers and autumns while underestimating temperature in summers and autumns and precipitation in springs and winters. A test is required to examine the impacts of these biases on the projection of streamflow over the *ARB*. The temporal scale of streamflow projection in this study is monthly which may be unsatisfactory for climate-change impact or adaptation studies.
such as flooding control. Although streamflow projection at the scale of daily or finer are challenged by the modeling accuracies of the desirable climate projection and hydrological models, it deserves some efforts that the streamflow projection framework in this study is improved for enabling reliable projection of daily streamflow through coupling with advanced climatic and hydro-meteorological systems analysis techniques such as ensembles. The streamflow projection results and the corresponding findings could be employed to guide climate-change impact studies or practices, e.g. water resources planning, socio-economic management, and eco-environmental protections. Furthermore, streamflow projection under uncertainties are achieved through the BaPMI method used in this study. The projection results of uncertainties could eliminate unreasonable decision making resulting from simplifications of these uncertainties and facilitate risk assessment and management.

7.4 SUMMARY

In this chapter, the projection of streamflow over the ARB under multiple RCP scenarios were enabled based on an integrated climatic and hydro-meteorological systems analysis framework. This framework consisted of multiple interrelated modules. (a) GCMs screening was developed to identify the GCM simulation of the best overall modeling accuracy through a comprehensive evaluation of the multi-dimensional performances of CMIP5 GCMs and their ensemble over the ARB. (b) Climate classification was conducted to avoid redundant computations in downscaling through a systematic analysis of the similarity and dissimilarity among climate observation grids over the ARB. (c) Based on modules (a) and (b), climate downscaling was carried out to enhance the spatial resolution of GCM simulation while correcting its biases in
reproducing the regional climate over the ARB. (d) Coupling modules (a) to (c) helped provide high-resolution projection of the climatic conditions over the ARB in the 21st century which were the boundary inputs of streamflow projection in this study. (e) The complicated correspondences between climatic conditions and streamflow at selected gauge stations over the ARB were simulated by an advanced statistical hydro-meteorological systems analysis model. (f) An integration of the climatic projection in (d) and the streamflow simulation in (e) enabled streamflow projection over the ARB under climate change. A few of interesting findings as summarized below were revealed from a series of analyses and comparisons in this study.

At the scale of octo-decades, the discharges would increase for the gauge stations of 07AA002, 07AD002, 07BC002, 07BE001, 07BK006 and 07CD001 and decrease for the other stations, showing significant spatial heterogeneities. The bi-decadal variability of averaged discharge averages or changing ratios might vary with gauge stations and RCP scenarios significantly. In most cases, the bi-decadal averaged discharges would decrease from [2021, 2040] to either [2041, 2060] or [2061, 2080] and increase to [2081, 2100] under RCP 2.6, while gradually increasing from [2021, 2040] to [2081, 2100] under the other RCP scenarios. The increasing rate of bi-decadal discharge averages would be the highest under RCP 8.5 in comparison with other scenarios.

For most gauge stations, the annual discharge averages would remain above or below the historical averages for almost all years under every RCP scenario. Meanwhile, the monthly discharge averages would tend to increase in January, February, March, August, and September, while decreasing in April to July. The increments of monthly discharges would be the highest in March for almost all gauge stations. For the gauge
stations where the octo-decadal averaged discharges would decrease, e.g. 07AG004 and 07DA001, the decrements of discharges would be dominated by less flow in July or $JJA$ when the magnitudes of streamflow are the highest.

In general, either $RCP$ scenarios or modeling biases were of significant impacts on the temporal variability and trend of discharges, while they would not significantly alter the overall magnitudes of discharges. The modeling biases were relatively significant for the variability of octo-decadal averaged discharge changing ratios among gauge stations. The annual discharge anomalies at some stations such as 07BC002 would be sensitive to modeling biases while those at the other stations not. In addition, the modeling biases as well as the $RCP$ scenarios had insignificant impacts on monthly averaged discharges in most combinations of months and gauge stations.

As revealed in a comparison of this study and existing studies on streamflow projection over the $ARB$, the impacts of climate change on streamflow might vary with geographic locations. The historical temporal trend of discharges might remain at two gauge stations in the 21st century, while not for three other stations. In addition, this comparison also disclosed one advantage of this study over these existing related ones. The impacts of climate change on steamflow and the spatial heterogeneity of these impacts could be evaluated in this study.

These findings would be much helpful for guiding the regional socio-economic development and eco-environmental conservation over the $ARB$, gaining insights into the changes of streamflow in the $ARB$ under climate change, and promoting the development of more advanced climatic and hydro-meteorological systems analysis methods. On the other hand, there are also a few of opportunities to improve this study and extend it for the
related studies or practices. For instance, an analysis of the impacts of modeling biases in
downscaling on the projection of streamflow over the ARB deserves an attempt. The
streamflow projection framework in this study would be improved to enable reliable
streamflow projection at the scale of daily. The streamflow projection results and the
corresponding findings could be applied to guide climate-change impact studies or
practices, e.g. water resources planning, socio-economic management, and eco-
environmental protections. The streamflow projection results of uncertainties could
eliminate unreasonable decision making resulting from simplifications of these
uncertainties and facilitate risk assessment and management under climate change.
CHAPTER 8

CONCLUSIONS

8.1 SUMMARY

(1) The accuracies of CMIP5 GCMs and their ensemble in reproducing historical climate conditions in the ARB were evaluated comprehensively. In section 2.2, thirty representative outputs of CMIP5 GCMs were identified through a review of existing continental climate-change impact studies; in consideration of data availabilities in these representative GCM outputs and four emission scenarios of CMIP5, six CMIP5 GCMs and their ensemble were selected as alternatives for evaluation; based on ten statistical metrics, ten groups of indicators for the simulation accuracies of CMIP5 GCMs in various aspects were proposed. The evaluation results were presented in section 2.3; the CMIP5 GCMs of the highest accuracies for diverse impact studies in the ARB was identified through a series of systematic comparisons. In section 2.4, scaling effects and statistical-metrics interactions were revealed, a few of suggestions on CMIP5 GCM selection for impact studies in the ARB were specified, and the potential extensions of this study were discussed.

(2) The ReDSICC approach was developed to provide an alternative for enabling effective classification of climate under data uncertainties and multivariate dependencies. Based on development of a discrete distribution transformation method and integration of advanced statistical inferential methods, a recursive framework of dissimilarity and similarity inferences was proposed for gradually grouping multi-dimensional climate-variables observations. In comparison with existing climate classification methods, the developed ReDSICC approach was advantageous at multiple aspects. For instance, the
module of $DDT$ eliminated the restriction of samples being normally distributed. Statistical inferential methods enabled classification of regional climate under data uncertainties and multivariate dependencies. The recursive process of dissimilarity and similarity inferences facilitated identifying the most desired climate classification result in which climatic conditions were significantly different for any two climate zones and were not of significant differences for grids in the same climate zone. The whole process of $ReDSICC$ was independent with subjective judgement which were replaced with statistical inferences. To verify methodological effectiveness and facilitate local studies such as downscaling and hydrological simulation, the $ReDSICC$ approach was applied to a case study of climate classification in the $ARB$, Canada.

(3) An advanced approach ($ReMPMID$) was proposed for supporting climate downscaling under complexities such as data uncertainties, nonlinear predictors-predictands correspondences, predictands’ interactions, non-normal distributions, spatial homogeneities, and temporal nonstationarities. The principle, innovations and technical details of the $ReMPMID$ approach were discussed. Subsequently, this approach was applied to the $ARB$, a large river basin on the Canadian prairies that was closely connected with climate change, to verify methodological effectiveness and facilitate local impact or adaptation practices. Specifically, based on predictor and predictand (i.e. $T_{min}$, $T_{max}$ and $Prec$) selection, data collection and processing, a grids-similarity analysis, a sensitivity analysis, and parameter calibration, a $ReMPMID$ model was constructed for every selected grid in the $ARB$. The multi-dimensional modeling accuracies of the $ReMPMID$ approach were verified and examined in various aspects. The optimal parameter values and the
uncertainties in high-resolution climate simulation were analyzed through a series of comparisons.

(4) High-resolution projection of 21st century climate over the ARB, Canada were enabled through ReMPMID (Chapter 4), an advanced statistical approach dedicated in climate downscaling under complexities of data uncertainties, nonlinear correspondences, multivariate dependencies, non-normal distributions, spatial homogeneities, and temporal nonstationarities. The spatial resolution of CMIP5 GCMs was effectively enhanced to 10 kilometres. These projection provided rich information of the climatic conditions (Tmin, Tmax and Prec) over this river basin in the 21st century under four RCP scenarios: RCPs 2.6, 4.5, 6.0 and 8.5. A series of findings were revealed from in-depth analyses of these projection.

(5) An advanced statistical hydro-system analysis approach abbreviated as BaPMI was proposed for supporting hydro-meteorological system analyses under complexities of data uncertainties, non-normal distributions, nonlinear correspondences, multivariate dependencies, and massive computations. In essence, the BaPMI approach was an integration of statistical inference, Gaussian process analysis, and Bayesian optimization. These techniques were integrated into an ingeniously design framework that could mitigate the challenges of these complexities in hydro-meteorological systems. The reliability of this approach was verified through a case study of streamflow simulation in the ARB. A series of analyses such as a correlation analysis, a bias analysis, a correspondence analysis, and two sensitivity analyses helped deepen the understandings of this approach, reveal the mechanisms of the hydrological regime over the ARB, disclose
the limitations of this approach, and divulge the potential opportunities in the subsequent studies.

(6) The projection of streamflow over the ARB under multiple RCP scenarios were enabled based on an integrated climatic and hydro-meteorological systems analysis framework. This framework consisted of multiple interrelated modules. (a) GCMs screening was developed to identify the GCM simulation of the best overall modeling accuracy through a comprehensive evaluation of the multi-dimensional performances of CMIP5 GCMs and their ensemble over the ARB. (b) Climate classification was conducted to avoid redundant computations in downscaling through a systematic analysis of the similarity and dissimilarity among climate observation grids over the ARB. (c) Based on modules (a) and (b), climate downscaling was carried out to enhance the spatial resolution of GCM simulation while correcting its biases in reproducing the regional climate over the ARB. (d) Coupling modules (a) to (c) helped provide high-resolution projection of the climatic conditions over the ARB in the 21st century which were the boundary inputs of streamflow projection in this study. (e) The complicated correspondences between climatic conditions and streamflow at selected gauge stations over the ARB were simulated by an advanced statistical hydro-meteorological systems analysis model. (f) An integration of the climatic projection in (d) and the streamflow simulation in (e) enabled streamflow projection over the ARB under climate change. A few of interesting findings as summarized below were revealed from a series of analyses and comparisons in this study.

8.2 RESEARCH ACHIEVEMENTS

(1) In brief, the most reasonable scheme of selecting CMIP5 GCMs for impact
studies in the *ARB* was to integrate them into a general framework because their accuracies varied with climate variables, geographical locations, temporal windows, statistical metrics, and temporal and spatial scales. If only one set of climate simulation was demanded, the multi-model ensemble was recommended. If only one *CMIP5 GCM* was preferred in some cases, ESM2G would be the desired one. In addition, a *CMIP5 GCM* of the highest accuracies in modeling climate conditions of the *ARB* at a higher spatial or temporal scale might not be the *CMIP5 GCM* which was the most accurate at a coarser scale, and vice versa. This scaling effect mainly occurred when the most accurate *CMIP5 GCMs* were significantly heterogeneous at the finer scale. In consideration of the similarity and dissimilarity of the selected statistical metrics, *SC*, *RMSE* and *Nash* were recommended for evaluation of *CMIP5 GCMs* in multiple aspects such as correlations and magnitudes. These findings would be much helpful for guiding researchers or practisers to select the desired *CMIP5 GCM* for climate-change impact or adaptation studies in the *ARB* or neighbouring regions.

(2) It was revealed that the complicated dissimilarities and similarities of climatic conditions among all grids over the *ARB* were effectively reflected in the climate classification results obtained from *ReDSICC*. The objective of enabling reversible transformation between an abnormal distribution and a normal distribution was effectively achieved by the *DDT* approach. The effectiveness of the climate classification result at reflecting the spatial dissimilarity and similarity of regional climate over the *ARB* was decreased if the *DDT* approach was not employed. In addition, a series of comparisons helped gain insights into the regional climate in the *ARB* and the mechanism of *ReDSICC*. For instance, in comparison with *Tmin*, the spatial heterogeneity of *Tmax* was higher while
that of Prec was lower over the ARB. The classification result of ReDSICC varied with changes of representative climate variables. An increase of Nmin might lead to a decrease of the obtained climate zones, which mainly occurred for zones of which the frequency of the corresponding climate-variable samples was relatively low. As $\alpha$ decreased, an increased number of climate zones would be merged. These advantages of ReDSICC and revelations from this case study were much helpful for enhancing the reliability of climate classification results, improving the effectiveness of existing climate classification methods, and providing scientific support for climate-change impact studies or the other related ones.

(3) From the viewpoint of methodology, chapter 4 contributed an additional reliable approach or framework, as an integration of statistical inferential methods, to enable downscaling under data uncertainties, nonlinear predictors-predictands correspondences, predictands’ dependencies, non-normal distributions, spatial homogeneities, and temporal nonstationarities. A similarity-analysis-based grid selection method was developed to facilitate identification of the grids that can effectively represent the regional climate, avoiding redundant computations and improving computational efficiencies. A bi-period calibration strategy was proposed to mitigate the challenge of over-parameterization in the calibration process for statistical downscaling approaches including ReMPMID. Multiple statistical metrics were employed to quantify the modeling accuracies of ReMPMID in various aspects that may be independent in some cases, which was helpful for avoiding the partiality in evaluating a downscaling approach’s performance through a single measure. A series of sensitivity analyses were conducted to reveal the impacts of the parameters of ReMPMID on modeling results and accuracies, facilitating this approach
being extensively applied to various impact and adaptation studies. The analysis on the uncertainties existing in the results of ReMPMID disclosed the necessity of not selectively using the averaged simulation and of paying attentions on abundant simulation results that were meaningful for risk assessment and management.

Through the case study of climate downscaling over the ARB, a range of findings were revealed as summarized below. These findings would be much helpful for gaining insights into the developed ReMPMID approach and the regional climate in the ARB or neighbouring regions.

The spatial heterogeneity of precipitation was higher than that of temperature in the ARB. The overall accuracies of the ReMPMID approach in reproducing high-resolution local climate over the ARB were relatively high for both Tmin and Tmax and acceptable for Prec. There were significant differences among the accuracies of ReMPMID under various combinations of predictands, months, seasons, statistical metrics, and geographic locations.

In the calibration process, the modeling accuracies of ReMPMID were of significant similarity among the selected modeling-accuracy indicators for most grids, while showing dissimilarity for some local zones. Although a single statistical metrics was capable of evaluating the multi-dimensional performances of the ReMPMID approach over most geographic locations, multiple ones were desired for a systematic evaluation in large river basins such as the ARB in the verification process.

The ReMPMID approach tended to overestimate the multi-year-averaged magnitudes of Tmin, Tmax and Prec over almost all grids except the ones around High
Prairie where *Prec* was underestimated. The extent to which the multi-year-averaged climate magnitude was overestimated decreased from *Tmin*, *Tmax* to *Prec*. Meanwhile, this approach might over-estimate *Tmin* and *Tmax* in springs and winters and *Prec* in summers and autumns over all or almost all grids while underestimating *Tmin* and *Tmax* in summers and autumns and *Prec* in springs and winters.

In terms of simulating the temporal variability and relative magnitudes of regional climate, the overall accuracies of ReMPMID decreased from springs, autumns, summers to winters for *Tmin* and *Tmax* and from autumns, winters to summers and springs for *Prec* in consideration of all grids in the ARB. As for the simulation of the multi-year-averaged seasonal absolute magnitudes of predictands, the overall accuracies of ReMPMID decreased from autumns, summers, winters to springs for *Tmin*, from autumns, summers, springs to winters for *Tmax*, and from winters, springs, autumns to summers for *Prec*.

In consideration of all grids in this river basin, the highest modeling accuracies of ReMPMID for *Tmin/Tmax* and *Prec* were in December and August, while the lowest in May and July, respectively. The overall accuracies of ReMPMID for *Prec* were higher than those for *Tmin* and *Tmax* in January, February, March, May, August, September and November, while lower in April, June, July and December.

The modeling accuracies was not sensitive to the parameter of the statistical significance level (*α*) for any predictands and significantly changed with another parameter: the minimum partition row number (*Nmin*). The calibration accuracies decreased with the climbing of the optimal *Nmin* and there was not a significant monotonic relationship between the optimal *Nmin* and the verification accuracies. The
optimal selection of $N_{min}$ varied with grids and predictands and shown higher uncertainties for $T_{min}$ and $T_{max}$ compared with $Prec$. The uncertainties of high-resolution climate simulation through the ReMPMID approach gradually increased from summers, autumns, springs to winters for $T_{min}$ and $T_{max}$ and from winters, springs, autumns to summers for $Prec$.

(4) In general, temperature and precipitation would increase during 2021 and 2100. Corresponding to $RCP$s 2.6, 4.5, 6.0 and 8.5, the octo-decadal averaged increment would be 1.742, 2.261, 2.107 and 2.975 °C for $T_{min}$, 1.353, 1.796, 1.612 and 2.217 °C for $T_{max}$, and 0.030, 0.074, 0.080 and 0.160 mm for $Prec$, respectively. If the modeling biases are removed, these increments would shrink by approximately 1.09 °C for temperature and 0.12 mm for precipitation; temperature would still increase for every $RCP$ scenario with a decreased increment, while precipitation would only increase under $RCP$ 8.5 and decrease under the other $RCP$ scenarios. The variability of bi-decadal averaged climate anomalies is maximized under $RCP$ 8.5 and minimized for temperature under $RCP$ 4.5 and for precipitation under $RCP$ 2.6. The inter-annual variability of climate anomalies would increase from $RCP$ 4.5, $RCP$ 2.6, $RCP$ 6.0 to $RCP$ 8.5 for temperature while increasing from $RCP$ 2.6, $RCP$ 4.5, $RCP$ 6.0 to $RCP$ 8.5 for precipitation.

The bi-decadal averaged climate anomalies would decrease from $DJF, MAM, SON$ to $JJA$ for night temperature, from $DJF, SON, MAM$ to $JJA$ for day temperature, and from $JJA, MAM, SON$ to $DJF$ for precipitation. Either temperature or precipitation would increase in most months except decreases of $T_{min}$ in July and October, $T_{max}$ in October, and $Prec$ in September and from Nov to March. The increase of temperature is mainly
contributed by warming in winters, while that of precipitation is dominated by heavy rainfall from May to August. The uncertainties in the projection results would be relatively low in May to September and high in other months for temperature, while they would be lower in November to April than those in the other months for precipitation.

The spatial climatic variability over the ARB would not significantly change with RCP scenarios and the climatic anomalies are highly correlated with the absolute magnitudes of climatic conditions. This river basin would experience warmer nights; people in the north end of the ARB would have colder days, while those in the other grids would experience warmer days; the upstream would become dryer and the other regions would be wetter than the baseline period. The climate anomalies would decrease from the upstream to the downstream for either \( T_{min} \) or \( T_{max} \), while this spatial pattern would be reversed for precipitation. The increment of temperature would be dominated by warming in the upstream of the ARB in winters, while that of precipitation by increased rainfall in the downstream in summers.

The potential changes of 21st-century climatic conditions over the ARB might raise a series of concerns. Representative ones include but are not limited to the followings: acceleration of glacier and snow melting, alternation of the hydrological regime, deterioration of the water conflict between oil-sands mining operations and eco-environmental health, and increased intensities, durations and frequencies of droughts in springs. Additional attentions should be paid on addressing these concerns through scientific climate-change adaptation strategies.
(5) The lead time from climatic conditions to streamflow may be one month for the ARB. The contributions of climate variables to streamflow generation decrease from the nearby grids, the upstream grids to the downstream grids excluding these nearby ones. The spatial heterogeneity of precipitation is more significant than that of temperature. The \textit{BaPMI} approach is effective at capturing the temporal variability and the multi-year averages of streamflow for every gauge station and reflecting the uncertainties in the correspondence between climatic conditions and streamflow at part of stations. Meanwhile, this approach tends to overestimate streamflow in January, February, March, August, September and October, while underestimating it in the other months. The \textit{BaGPO} algorithm is capable of highly reducing computational time while guaranteeing modeling reliability, which is much helpful for enhancing the applicability of \textit{BaPMI}. The daily minimum temperature, followed by the daily maximum temperature and the daily cumulative precipitation, around the glacier field at the upstream is of very significant impacts on the streamflow in the main channel over the ARB. In addition, the overall impact of climate variables on streamflow over the ARB decreases from $T_{min}$, $T_{max}$ to $Prec$. The parameter sensitivities decrease with the modeling accuracies of the \textit{BaPMI} approach at various gauge stations, and the modeling accuracies of \textit{BaPMI} tend to decrease with $\alpha$ and increase with $N_{min}$.

(6) At the scale of octo-decades, the discharges would increase for the gauge stations of 07AA002, 07AD002, 07BC002, 07BE001, 07BK006 and 07CD001 and decrease for the other stations, showing significant spatial heterogeneities. The bi-decadal variability of averaged discharge averages or changing ratios might vary with gauge stations and RCP scenarios significantly. In most cases, the bi-decadal averaged discharges would decrease
from [2021, 2040] to either [2041, 2060] or [2061, 2080] and increase to [2081, 2100] under *RCP* 2.6, while gradually increasing from [2021, 2040] to [2081, 2100] under the other *RCP* scenarios. The increasing rate of bi-decadal discharge averages would be the highest under *RCP* 8.5 in comparison with other scenarios.

For most gauge stations, the annual discharge averages would remain above or below the historical averages for almost all years under every *RCP* scenario. Meanwhile, the monthly discharge averages would tend to increase in January, February, March, August, and September, while decreasing in April to July. The increments of monthly discharges would be the highest in March for almost all gauge stations. For the gauge stations where the octo-decadal averaged discharges would decrease, e.g. 07AG004 and 07DA001, the decrements of discharges would be dominated by less flow in July or *JJA* when the magnitudes of streamflow are the highest.

In general, either *RCP* scenarios or modeling biases were of significant impacts on the temporal variability and trend of discharges, while they would not significantly alter the overall magnitudes of discharges. The modeling biases were relatively significant for the variability of octo-decadal averaged discharge changing ratios among gauge stations. The annual discharge anomalies at some stations such as 07BC002 would be sensitive to modeling biases while those at the other stations not. In addition, the modeling biases as well as the *RCP* scenarios had insignificant impacts on monthly averaged discharges in most combinations of months and gauge stations.

As revealed in a comparison of this study and existing studies on streamflow projection over the *ARB*, the impacts of climate change on streamflow might vary with geographic locations. The historical temporal trend of discharges might remain at two
gauge stations in the 21st century, while not for three other stations. In addition, this comparison also disclosed one advantage of this study over these existing related ones. The impacts of climate change on streamflow and the spatial heterogeneity of these impacts could be evaluated in this study.

These findings would be much helpful for guiding the regional socio-economic development and eco-environmental conservation over the ARB, gaining insights into the changes of streamflow in the ARB under climate change, and promoting the developaton of more advanced climatic and hydro-meteorological systems analysis methods.

8.3 RECOMMENDATIONS FOR FUTURE RESEARCH

Based on the research presented in this dissertation, further studies are desired in the following aspects:

(1) In the future, the framework of CMIP5 GCM evaluation employed in this study can be improved in some aspects and provides scientific support for climate-change impact or adaptation studies in other regions. For example, analyses of climatic conditions projected by the selected CMIP5 GCMs under different emission scenarios as well as of the impacts of various CMIP5 GCM selection schemes on future climatic projection results deserve further efforts. The identification of relatively accurate CMIP5 GCMs for the ARB can be used for facilitating socio-economic development and eco-environmental conservation. An extension of this evaluation to extreme climatic conditions which are desired for some particular impact studies such as flooding control and drought elimination is helpful for enhancing the applicability of this study. The developed framework can be applied to the evaluation of CMIP5 GCMs for diverse impact studies in other river basins or regions worldwide, providing specified suggestions on selection
of CMIP5 GCMs according to particular requirements of climatic conditions.

(2) The developed ReDSICC approach was challenged by multiple issues that deserved further efforts to mitigate in the subsequent research. Representative ones included but were not limited to the followings: developing continuous distribution transformation functions, adding a module of covariance-matrices inequality or equality tests, improving the scheme of identifying the most significant dissimilarity, evaluating the necessity of removing the different magnitudes of climate-variable samples, developing a dynamic version of ReDSICC, optimizing the values of the parameters ($\alpha$ and $N_{min}$), and disclosing other opportunities of improving ReDSICC through extensive applications.

(3) A continuous function that could enable the reversible transformation between any non-normal distribution and a normal distribution may be more effective than the DDT approach used in this study. The impacts of distribution transformation on downscaling results and accuracies deserves an attempt. There may be a more reasonable scheme of nonlinearity characterization than the current one that the nonlinearity between predictors and predictands is represented by the principal monotonicity in predictor-predictand paired samples. Incorporating variance analyses into the framework of ReMPMID may facilitate a comprehensive evaluation of the difference between any two groups of multi-dimensional predictand samples. A scientific scheme, possibly integration with physically-based downscaling models, is desired for mitigating one challenge of the ReMPMID approach that the simulation results are trapped in the historical records. Adjustment of the ReMPMID process may be helpful for eliminating the high computational loads of this data-driven approach. The methods of global sensitivity
analysis can help reveal complicated interactions among the parameters, boundary conditions and modeling results of the ReMPMID approach. In-depth analyses on the causes of variation and diversity of the modeling accuracies of ReMPMID may help advance this approach’s performances. Attention should be paid on building the connections between real-world climate systems and the ReMPMID approach which is categorized as a statistical or data-driven downscaling approach. Developing a dynamic version of the ReMPMID approach to adapt to the temporal variability of regional climate may be one of schemes that could improve this approach. Consideration of various uncertainties except data uncertainties can also be an alternative direction in the following studies. More case studies such as enabling long-term high-resolution climate projection over the ARB are required to verify the reliability of this approach in various downscaling studies or practices, reveal its limitations for further improvement, and extend its potential contributions.

(4) Further efforts would also be made to deepen understandings of the ReMPMID approach and regional climate, facilitate proposition of more advanced downscaling techniques, reveal the complicated correspondence between large-scale atmospheric variables and regional or local scale climatic variables, and enable provision of scientific support for climate-change impacts or adaptation practices. For instance, the temporal resolution would be enhanced from monthly to daily. Extreme climatic conditions would be examined. The ReMPMID approach would be used as an ensemble tool to integrate the advantages of various dynamic or statistical downscaling techniques or be taken as an element in an ensemble downscaling framework under the uncertainty of climate simulation or downscaling models. The climate projection results would be incorporated
into hydrological models to analyze the variation or status of hydrological regimes in the future and to guide water resources management.

(5) A continuous function that could enable the reversible transformation between any non-normal distribution and a normal distribution may be more effective than the DDT approach used in this study. The impacts of distribution transformation on the modeling results and accuracies deserves an attempt. There may be a more reasonable scheme of nonlinearity characterization than the current one that the nonlinearity between climatic conditions and streamflow is represented by the principal monotonicity in the paired samples. Incorporating variance analyses into the framework of BaPMI may facilitate a comprehensive evaluation of the difference between any two groups of multi-dimensional streamflow samples. A scientific scheme, possibly integration with physically-based hydrological models, is desired for mitigating one challenge of the BaPMI approach that the simulation results tend to fluctuate within the historical records. The methods of global sensitivity analysis can help reveal complicated interactions among the parameters, the boundary conditions, and the modeling results of the BaPMI approach. In-depth analyses on the causes of variation and diversity of the modeling accuracies of BaPMI may help advance this approach’s performances. Attentions should also be paid on building the connections between real-world hydrological systems and the BaPMI approach which is a statistical or data-driven hydro-system analysis approach. Developing a dynamic version of the BaPMI approach to adapt to the temporal variability of hydrological regimes may be one of schemes that could improve this approach. Consideration of various uncertainties except data uncertainties can also be an alternative direction in the following studies. More case studies such as enabling long-term projection
of the hydrological regime over the ARB are required to verify the reliability of this approach in various studies or practices, reveal its limitations for further improvement, and extend its potential contributions.

(6) There are also a few of opportunities to improve the streamflow forecasting in this study and extend it for the related studies or practices. For instance, an analysis of the impacts of modeling biases in downscaling on the projection of streamflow over the ARB deserves an attempt. The streamflow projection framework in this study would be improved to enable reliable streamflow projection at the scale of daily. The streamflow projection results and the corresponding findings could be applied to guide climate-change impact studies or practices, e.g. water resources planning, socio-economic management, and eco-environmental protections. The streamflow projection results of uncertainties could eliminate unreasonable decision making resulting from simplifications of these uncertainties and facilitate risk assessment and management under climate change.
REFERENCES
daily circulation patterns for precipitation and temperature downscaling based on
optimized fuzzy rules. Climate Research, 23, 11-22.


Bad Homburg, Germany.

watershed planning. Transactions of the American Society of Agricultural
Engineers, 23 (4), 938-944.

screening, application and added value of statistical downscaling of a CMIP5
ensemble for single-site projection in Distrito Federal, Brazil. International Journal
of Climatology.

Box, G. E. P., and Wilson, K. B. 1951. On the Experimental Attainment of Optimum

components to surface weather conditions: Lake Superior Basin winter temperatures.
Theoretical and Applied Climatology, 63, 41-56.

Brochu, E., Cora, V. M., and De Freitas, N. 2010. A tutorial on Bayesian optimization of
expensive cost functions, with application to active user modeling and hierarchical


Dong, C., Huang, G. H., Cai, Y. P., and Liu, Y. 2013a. Robust planning of energy management systems with environmental and constraint-conservative


Potential predictability of an Iberian river flow based on its relationship with previous winter global SST. Journal of Hydrology, 385(1-4), 143-149.


Downscaling Daily Temperature, Precipitation and Wind Speed Parameters.
International Journal of Environmental Research, 8(4), 1223-1230.
Hughes, D. A., and Sami, K. 1994. A semi-distributed, variable time interval model of


Oztuna, D, Elhan, A. H., and Tuccar, E. 2006. Investigation of four different normality tests in terms of type 1 error rate and power under different distributions. Turkish Journal of Medical Sciences, 36(3), 171-176.


appraisal, Blackwell Publishing.


Meyerson, J. E., Neelin, J. D., Nigam, S., Pan, Z., Ruiz-Barradas, A., Seager, R.,
Serra, Y. L., Sun, D. Z., Wang, C., Xie, S. P., Yu, J. Y., Zhang, T., and Zhao, M.
2013. North American climate in CMIP5 experiments. Part II: evaluation of
historical simulation of intraseasonal to decadal variability. Journal of Climate,
26(23), 9247-9290.

Singh, D., Jain, S. K., Gupta, R. D., Kumar, S., Rai, S. P., and Jain, N. 2016. Analyses of
Observed and Anticipated Changes in Extreme Climate Events in the Northwest

Statistical seasonal rainfall and streamflow forecasting for the Sirba watershed,
West Africa, using sea surface temperatures. Hydrological Sciences Journal, 61(5),
805-815.


for weather research and forecasting applications. Journal of Computational Physics,
227, 3465-3485.

Snoek, J., Larochelle, H., and Adams, R. P. 2012. Practical bayesian optimization of
machine learning algorithms. In Advances in neural information processing systems,
2951-2959.

Spearman, C. 1904. The proof and measurement of association between two things. The

Stahl, K., Moore, R. D., and McKendry, I. G. 2006. The role of synoptic-scale circulation
in the linkage between large-scale ocean-atmosphere indices and winter surface climate in British Columbia, Canada. International Journal of Climatology, 26, 541-560.


over the root mean square error (RMSE) in assessing average model performance. Climate Research, 30(1), 79-82.


APPENDICES

APPENDIX A. THE MODIFIED NEL AND VAN DER MERWE TEST
(KRISHNAMOORTHY AND YU, 2004)

Suppose \( \{X_t\}_{t=1}^u \) and \( \{X_t\}_{t=u+1}^T \) are two groups of \( n \)-dimensional predictand samples where \( X_t \) is an \( n \)-dimensional vector for any \( t \in \{1, \ldots, T\} \). Let the mean vectors of the populations of the two groups of samples be denoted as \( \mu_1 \) and \( \mu_2 \). The modified Nel and van der Merwe (MNV) test (Krishnamoorthy and Yu, 2004) is used to analyze the statistical difference of the equality of mean vectors of the two groups of samples. Accordingly, the null hypothesis is

\[
H_0: \mu_1 = \mu_2, \tag{A.1}
\]

while the alternative one is

\[
H_1: \mu_1 \neq \mu_2. \tag{A.2}
\]

The test statistic is defined as

\[
T_u^2 = (\bar{X}_1 - \bar{X}_2)\hat{S}^{-1}(\bar{X}_1 - \bar{X}_2) \tag{A.3}
\]

where

\[
\hat{S} = \hat{S}_1 + \hat{S}_2, \tag{A.4}
\]

\[
\hat{S}_1 = S_1/u, \tag{A.5}
\]

\[
\hat{S}_2 = S_2/(T - u), \tag{A.6}
\]

\[
\bar{X}_1 = \sum_{t=1}^{u}X_t/u, \tag{A.7}
\]

\[
\bar{X}_2 = \sum_{t=u+1}^{T}X_t/(T - u), \tag{A.8}
\]
\[ S_1 = \sum_{t=1}^{u} (X_t - \hat{X}_1)(X_t - \hat{X}_1)'/(u - 1), \quad (A.9) \]

and

\[ S_2 = \sum_{t=u+1}^{T} (X_t - \hat{X}_2)(X_t - \hat{X}_2)'/(T - u - 1). \quad (A.10) \]

This statistic tends to obey an \( F \) distribution, i.e.

\[ T_u^2 \sim (v \cdot n/(v - n + 1))F_{n, v-n+1} \quad (A.11) \]

where

\[ v = (n + n^2)/(\text{tr}[(\hat{S}_1\hat{S}^{-1})^2]/u + \text{tr}[(\hat{S}_1\hat{S}^{-1})^2]/u + \text{tr}[(\hat{S}_2\hat{S}^{-1})^2]/(T - u) + \text{tr}[(\hat{S}_2\hat{S}^{-1})^2]/(T - u)) \]

and \( \text{tr}[\cdot] \) is the trace of a matrix.

For a given observed value \( (T_{uo}^2) \) of \( T_u^2 \), the null hypothesis \((H_0)\) should be rejected, if

\[ P\left\{(v \cdot n/(v - n + 1))F_{n, v-n+1} > T_{uo}^2\right\} < \alpha \quad (A.13) \]

where \( \alpha \) is a given level of statistical significance. In this case, the mean vectors of the two groups of multi-dimensional predictand samples are significantly different from the viewpoint of statistics.
APPENDIX B. THE ZS TEST (ZHOU AND SHAO, 2014)

All $m$-dimensional predictand data series, i.e. $\{Y_t\}_{t=1}^T$, are used as the inputs of the ZS test in order to maximize reliability of the multivariate normality test conclusion. Let superscripts ‘$\prime$’ and ‘$^{-1}$’ represent the transpose and the inverse matrix of a matrix, respectively.

1) Initialize hypothesis $H_0$ : $\{Y_t\}_{t=1}^T$ are normally distributed;

2) Calculate the value of $\hat{Y}_t$:

\[
\hat{Y}_t = S^{-1/2}(Y_t - \sum_{i=1}^T (Y_t)/T) \tag{B.1}
\]

for any $t$ ($S^{-1/2}$ is the symmetric square root of the inverse covariance matrix of $\{Y_t\}_{t=1}^T$);

3) Estimate the value of $I_A$:

\[
I_A = 1 \text{ (if } CDF^{-1}(\alpha) \leq (T/(8 \cdot m \cdot (m + 2)))^{1/2} \cdot ([\sum||\hat{Y}_t||^4]/T - (m \cdot (m + 2) \cdot (T - 1))/T + 1) \leq CDF^{-1}(1 - \alpha)) \text{ or } 0 \text{ (otherwise)} \tag{B.2}
\]

where $\alpha$ is a significance level and $CDF^{-1}()$ is the standard normal quantile function;

4) Calculate the values of $a(T, n)$:

\[
a(T, n) = (s'V^{-1}s)^{-1/2}s'V^{-1} (s = (s_1, s_2, \ldots, s_T)', \tag{B.3}
\]

\[
s_t = \{\sum_{i=1}^T (s(i, t))\}/T \text{ for any } t, \tag{B.4}
\]

\[
V = \{v_{ij} \mid i \text{ or } j = 1, 2, \ldots, T; v_{ij} = cov(\{s(i, t)\}, \{s(j, t)\})\}, \tag{B.5}
\]

where $\{s(i, t)\}_{t=1}^T$ are order statistics of $T$-size samples of a standard normal distribution for any $t$, and $cov(\cdot, \cdot)$ is the covariance of two sample sets);
5) Calculate the value of $G(\theta)$:

$$G(\theta) = W(\theta' \cdot \bar{\hat{Y}}_1, \theta' \cdot \bar{\hat{Y}}_2, ..., \theta' \cdot \bar{\hat{Y}}_T),$$  
(B.6)

$$\theta = (\theta_1, \theta_2, ..., \theta_m)',$$  
(B.7)

$$\sum_{j=1}^m (\theta_j)^2 = 1,$$  
(B.8)

$$\hat{y}_t = \theta' \cdot \bar{\hat{Y}}_t \text{ for any } t,$$  
(B.9)

$$\hat{y} = \{\sum_t (\hat{y}_t)/T, \hat{y}_t \text{ is the } t-\text{th smallest sample in } \{\hat{y}_t\}_{t=1}^T \text{ for any } t\},$$  
(B.10)

and

$$W(\hat{y}_1, \hat{y}_2, ..., \hat{y}_T) = \{\sum_t (a_{t,\theta} \cdot \hat{y}_t)\}^2/\{\sum_t (\hat{y}_t - \hat{y})^2\};$$  
(B.11)

6) $\Theta_1$ is a set of the $m$ most “extreme” directions corresponding to $m$ smallest $G$ values evaluated at random directions $\{||\bar{\hat{Y}}_t||^{-1} \bar{\hat{Y}}_t\}_{t=1}^T$ and $\Theta_2$ is a set of the $m$ marginal variates $\{(e_{j1}, e_{j2}, ..., e_{jm})'\}_{j=1}^m$ ($e_{jk} = 1$ (if $j = k$) or 0 (if $j \neq k$) for any $j$ and $k$);

7) Estimate:

$$f(\{\hat{Y}_t\}_{t=1}^T) = 1 - (\sum_{\theta \in \Theta_1 \cup \Theta_2} (I_A \cdot G(\theta)))/(2 \cdot m);$$  
(B.12)

8) Further calculate the value of $f_c$:

$$f_c = f(\{\hat{Y}_{ct}\}_{t=1}^T)$$  
(B.13)

for any $c \in \{1, 2, ..., C\}$ ($\{\hat{Y}_{ct}\}_{t=1}^T$ are samples of an $m$-variate standard normal distribution through Monte Carlo simulation and $C$ is the sampling times); and

9) $H_0$ is accepted iff $\alpha \leq p$ ($p$ = the relative quantile of $f$ in $\{f_c\}_{c=1}^C$).
Given a statistical significance level ($\alpha = 0.001, 0.01, \text{ or } 0.05$), we can have predictand samples $\{Y_t\}_{t=1}^T$ tend to be normally distributed if $\alpha \leq p$; otherwise, the assumption of multivariate normality of $\{Y_t\}_{t=1}^T$ can hardly hold.
APPENDIX C. THE SHAPIRO-WILKS TEST (SHAPIRO AND WILK, 1965)

The test statistic is

\[ W = \frac{\left( \sum_{t=1}^{T} a_t \tilde{y}_t \right)^2}{\left( \sum_{t=1}^{T} (\tilde{y}_t - \tilde{y})^2 \right)} \]  \hspace{1cm} (C.1)

where \( \tilde{y}_t \) is the \( t \)th order statistic for any \( t \in \{1, 2, \ldots, T\} \), i.e.

\[ \tilde{y}_t \leq \tilde{y}_{t+s} \text{ for any } t \in \{1, 2, \ldots, T-1\} \text{ and any } s \in \{1, 2, \ldots, T-t\} \];  \hspace{1cm} (C.2)

\[ \tilde{y} = \frac{\left( \sum_{t=1}^{T} \tilde{y}_t \right)}{T} \]  \hspace{1cm} (C.3)

or

\[ \tilde{y} = \frac{\left( \sum_{t=1}^{T} y_t \right)}{T}; \]  \hspace{1cm} (C.4)

\[ a = (a_1, a_2, \ldots, a_T) = m^T V^{-1} (m^T V^1 V^0 m)^{0.5}; \]  \hspace{1cm} (C.5)

\[ m = (m_1, m_2, \ldots, m_T)^T; \]  \hspace{1cm} (C.6)

and \( m \) and \( V \) are the mean vector and the covariance matrix of order statistics of standard normal samples. The null hypothesis is that samples \( \{y_t\}_{t=1}^{T} \) originate from a normally distributed random variable \( y(\omega) \). The value range of statistic \( W \) is \( (0, 1] \). Small values of \( W \) may represent rejection of normality, and its equivalence with one may indicate normality of the samples.

It is a more reliable means to evaluate normality of samples via significance levels, i.e. probabilities of rejecting the null hypothesis given that it is true. The statistic \( W \) can be equivalently transformed to a standard normal distribution, which enables the provision of a \( p \) value. The null hypothesis should be rejected if the \( p \) value is lower than the given significance level (e.g. 0.10, 0.05, or 0.01). Insufficient evidence supports rejection of the
null hypothesis if the $p$ value is higher than or equal to the significance level. In the latter case, the null hypothesis should be accepted, i.e. the population of samples $\{y_t\}_{t=1}^T$ is from a normal distribution.
APPENDIX D. THE ASPIN-WELCH-SATTERTHWAITE T-TEST

(JÁCOME ET AL., 2007)

Let the populations of samples \( \{y_{ij}, o\}_{t=1}^{u} \) and \( \{y_{ij}, o\}_{t=U+1}^{T} \) be flagged as 1 and 2, respectively. We denote the means of the two sample sets as \( \bar{x}_1 \) and \( \bar{x}_2 \), and the unbiased variance estimations as \( s_1^2 \) and \( s_2^2 \). We have:

\[
\bar{x}_1 = \frac{\sum_{t=1}^{u} y_{ij,o}}{u}, \quad (D.1)
\]

\[
\bar{x}_2 = \frac{\sum_{t=U+1}^{T} y_{ij,o}}{(T-u)}, \quad (D.2)
\]

\[
s_1^2 = \frac{\sum_{t=1}^{u} (y_{ij,o} - \bar{x}_1)^2}{(u-1)}, \quad (D.3)
\]

and

\[
s_2^2 = \frac{\sum_{t=U+1}^{T} (y_{ij,o} - \bar{x}_2)^2}{(T-u-1)}. \quad (D.4)
\]

A \( t \)-statistic for testing \( H_0 \) is formulated as

\[
t_s = \left[ (a \cdot x_1 - a \cdot x_2)(T-2)^{0.5} \right] / [(1/u + 1/(T-u))((u-1)s_1^2 + (T-u-1)s_2^2)]^{0.5}. \quad (D.5)
\]

The degrees of freedom of the Student \( t \)-distribution that statistic \( t_s \) approximately follows is equal to:

\[
E_{df} = [(s_1^2/u + s_2^2/(T-u))^2] / [(s_1^2/u)^2/(u-1) + (s_2^2/(T-u))^2/(T-u-1)]. \quad (D.6)
\]

Given an allowable significance level \( \alpha \), the null hypothesis is rejected if

\[
|t_s| > f(1 - \alpha/2, E_{df}) \quad (D.7)
\]
where $f(\cdot, \cdot)$ is the inverse function of the cumulative distribution function of the Student $t$-distribution. Suppose

$$w_{tt}(\alpha) = |t_s|/f(1 - \alpha/2, E_{tdf}).$$  \hfill (D.8)

Mean values of two sample sets ($\{y_{j(t, o)}\}_{t=1}^u$ and $\{y_{s(t, o)}\}_{t=u+1}^T$) are significantly different if

$$w_{tt}(\alpha) > 1,$$  \hfill (D.9)

while the difference is not statistically significant if

$$w_{tt}(\alpha) \leq 1.$$  \hfill (D.10)