Regional Frequency Analysis Estimates of Extreme Rainfall Events under Climate Change

by

Zhe Yang

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Civil Engineering (Water)

Waterloo, Ontario, Canada, 2019

Examining Committee Membership

The following are the members who served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Taha B.M.J Ouarda Professor
Supervisor	Donald H. Burn Professor
Internal Member	William K. Annable Associate Professor
Internal Member	Liping Fu Professor
Internal-External Member	Kumaraswamy Ponnambalam Professor

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Chapter 2 (Yang & Burn, 2019) was completed by Zhe Yang in collaboration with Dr. Donald H. Burn. Zhe Yang conceived the proposed idea, constructed the proposed framework, designed the experiment, and conducted the calculations under the supervision of Dr. Donald H. Burn. Donald Burn contributed to the analysis and interpretation of the results and acted as the editor of the written manuscript.

Chapter 3 was a collaborative work between Zhe Yang and Dr. Donald H. Burn. Zhe Yang conceived the proposed idea, constructed the proposed framework, designed the experiment, conducted the calculations and prepared the manuscript. Dr. Donald H. Burn supervised the experiment, provided feedback on the research and edited the manuscript.

Chapter 4 was completed by Zhe Yang. Zhe Yang conceived the proposed idea, constructed the proposed framework, designed the simulation, conducted the calculations and prepared the manuscript. Dr. Donald H. Burn provided feedback on the research and edited the manuscript

Abstract

Extreme rainfall events have a long history of causing large economic damages in urban areas and even loss of human life. Reliable estimates of extreme rainfall intensities for different rainfall durations are essential for the effective planning of drainage systems under climate change to balance the construction costs and potential damages caused by future extreme rainfall events. The information required for design rainfall events can be obtained through frequency analysis of extreme rainfall. However, extreme rainfall quantiles obtained from the traditional approach of frequency analysis have become increasingly unreliable under climate change. With increasing global temperatures and the uneven distribution of atmosphere moisture, the frequency and magnitude of extreme rainfall events can experience accelerated changes. Thus, urban drainage systems designed based on extreme rainfall quantiles obtained from historical records are becoming increasingly ineffective. Under the impacts of climate change, extreme rainfall events are becoming one of the most destructive natural hazards in the world.

Frequency analysis of the extreme rainfall events used to estimate the probability of exceedance of extreme rainfall events of a given magnitude in the future context can generate unreliable estimates under climate change because of two issues. Firstly, there are often insufficient data records available for the quantification of extreme rainfall events of interest from a design perspective. Since extreme rainfall events are rare, there is large uncertainty in quantile estimates obtained from using only the information from the site of interest. Thus, regional frequency analysis, which expands the data records through gathering information from sites sharing similar rainfall patterns, is widely used and is applied in this research. Secondly, the traditional assumption that there is a repetitive pattern in the occurrences of extreme rainfall events has become invalid in a nonstationary environment. Since extreme rainfall patterns can be altered in the future, estimates for rainfall quantiles obtained from using frequency analysis in a historical stationary environment can be unreliable when applied for future conditions. Further research is required into applying the regional frequency analysis approach for the estimation of extreme rainfall quantiles under climate change.

To provide reliable regional estimates of rainfall quantiles for different rainfall durations under climate change, this research improves regional frequency analysis through exploring the following issues: 1) An improved procedure for homogeneous group formation for historical

stationary periods. Extreme rainfall events have been affected by climate change. A three-layer searching algorithm is proposed for homogeneous group formation in a stationary environment for the consideration of climate change impacts on the spatial distribution of extreme rainfall events. 2) An adjustment procedure for homogeneous group formation in the future stationary environment. Under the assumption that extreme rainfall patterns remain stationary within a 30-year period, a procedure is proposed to adjust the optimal homogeneous group formation from the previous temporal periods to reflect conditions in future 30-year periods. 3) A procedure used for rainfall quantile estimation in a future nonstationary environment. Under the assumption that the extreme rainfall series exhibit nonstationary behavior during the whole future period, a one-step forward procedure is constructed based on the unscented Kalman filter to consider the potential non-monotonic change behavior of extreme rainfall events at different return periods. In this approach, the homogeneous groups are formed using a trend centered pooling approach.

The proposed methodology fills the gaps of considering climate change impacts on homogeneous group formation in both historical and future stationary environments and challenges the assumption of monotonic change behavior of extreme rainfall quantiles used in the traditional regional frequency analysis for stations exhibiting nonstationary behavior. The proposed procedures have been extensively tested using large sets of climate data in both historical and future contexts and have been shown to improve the extreme rainfall quantile estimates in both historical and future contexts.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. Donald H. Burn, for his advice and academic guidance throughout my program. I am very fortunate to be able to learn from such an accomplished researcher and distinguished supervisor. I am always grateful for all his contributions of time and support through my staying at University of Waterloo.

I am also thankful for my amazing research collaborators, Ana Requena Rodriguez and Martin Durocher, for their insightful advice and friendship during my exploration of research.

I am also very grateful to the FloodNet program for the funding, which provided me with the opportunity to advance my understanding in the statistical aspects of hydrology.

I would like to express my sincere appreciation to my defence committee members, Dr. Taha B.M.J Ouarda from INRS, Dr. William Annable, Dr. Liping Fu, and Dr. Kumaraswamy Ponnambalam from the University of Waterloo.

During my PhD program at University of Waterloo, I was honored to receive a Collaborative Water Program RBC Water Scholars Graduate Entrance Scholarship, Provost Doctoral Entrance Award (PDEA) for Women, University of Waterloo International Doctoral Student Awards and Inch Engineering Research Award, which are hereby acknowledged. I also thank the Department of Civil and Environmental Engineering staff who have helped me in many ways during my studies.

My warmest gratitude goes to my family for providing unconditional support during my pursuing of a higher degree. I would never be here today without them believing in me.

Dedication

I dedicate this thesis to my loving parents, for their unconditional supports and encouragements to me throughout my life.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

Extreme rainfall events, which can lead to extreme floods, have a long history of causing devastating impacts on human society. In the last two decades, more than 3,000 flooding events, including rainfall-induced and snowmelt floods, have been recorded worldwide, affected the lives of 1.5 billion people and caused direct economic damages of $ 656 billion (CRED and UNISDR, 2018). In Canada, extreme floods are the most common and largely distributed natural hazards to modern society. Many of these floods were caused by extreme rainfall events including the historical events in southern Alberta 1929, Toronto region 1954 and 2005, and Saguenay 1996 (Gifford, 2004; Library and Archives Canada, 2006). Under the impacts of climate change, increasingly frequent extreme rainfall events have caused staggering economic losses in the last decades. In 2012, the year of the urban flood, 206 mm of rainfall accumulated over the Thunder Bay area in May and caused extensive damages to both public infrastructure and private properties with initial costs exceeding $100 million (Environment and Climate Change Canada, 2017). In 2017, major flooding that was caused by low-pressure systems in April led to 155 mm of rain accumulation in May resulting in more than $223 million in insured damages in Ontario and Quebec (Insurance Bureau of Canada, 2017). In 2018, 72 mm of rain on August 7th over Toronto caused over $80 million in insured damage (Insurance Bureau of Canada, 2018). Furthermore, based on the projected future trends of temperature and precipitation in Canada, the frequency of extreme rainfall events of large magnitude are expected to increase based on different emission scenarios (Bush et al., 2019).

To reduce the devastating impacts caused by extreme rainfall events, drainage systems are implemented in urban areas for flood mitigation and the minimization of potential economic damages. To design an effective urban drainage system for prevention of future flooding, accurate estimates of Intensity-Duration-Frequency (IDF) curves are needed. IDF curves are used to describe the frequency of occurrence and the intensity of extreme rainfall events for a specified rainfall duration and can be obtained through frequency analysis. Rainfall frequency analysis, which is conducted under the assumption that historical extreme events are representative of future events in a stationary environment, can determine the probability of future occurrence of these events based on the patterns exhibited in the historical extreme observations (Hosking & Wallis, 1997).

At-site frequency analysis is the approach of using the historical extreme rainfall observations from the site of interest for the characterization of the extreme rainfall patterns at a specific geographic location and is widely used for the estimation of IDF curves. However, a problem associated with this approach is that sufficiently long data records are generally unavailable for obtaining reliable estimates of the extreme rainfall quantiles that are required for the design of urban drainage systems. Regional frequency analysis has often been applied to improve the accuracy and reduce the uncertainty in extreme rainfall quantile estimation. Under the assumption that extreme rainfall events do not only appear over a single geographic location, the rainfall records from sites sharing similar extreme rainfall patterns with the site of interest can be collected for the purpose of extreme rainfall characterization at a target site. The regional approach of frequency analysis is the focus in this thesis.

Regional frequency analysis is a well-established method with numerous applications for extreme rainfall frequency analysis in the last two decades. The early example for the regional approach is the index-event model introduced by Dalrymple (1960), which has become the foundation for subsequent developments. In a stationary environment, efforts for improving extreme rainfall quantile estimates from using the regional approach can be summarized into the following aspects: 1) Increasing the length of input data records by using peaks-over-threshold records instead of annual maximum series (Mailhot, et al., 2013). 2) Developing improved procedures for identifying rainfall similar stations (Satyanarayana & Srinivas, 2008; Yang et al., 2010; Gabriele & Chiaravalloti, 2013; Asong et al., 2015). 3) Choosing the best distribution model for the calculation of rainfall quantiles (Forestieri et al., 2018). In contrast to the numerous developments of the regional approach in a stationary environment, fewer efforts have been made in applying a regional approach in a nonstationary environment.

The exploration of regional frequency analysis in a nonstationary environment is essential for the successful design of urban drainage systems under climate change for two reasons: 1) Under the impacts of climate change, the frequency and intensity of extreme rainfall events are projected to change across Canada in the future (Simonovic, et al., 2017). The current drainage systems in urban areas that are designed based on rainfall intensity estimated using historical rainfall records are becoming increasingly ineffective in serving the traditional function of flood mitigation (D. Sandink, 2016). 2) The current urban drainage systems, which have traditionally

only been used for the mitigation of floods in the urban area, serve other functions including climate change mitigation, biodiversity enhancement, stormwater management and public health improvement (Fletcher et al., 2015). To effectively manage the urban drainage systems to achieve an optimal balance amongst different functions and to reduce their construction costs, accurate estimates for extreme rainfall quantiles at each temporal period are important.

This research explores the application of regional frequency analysis of extreme rainfall events under climate change and develops a system for extreme rainfall quantile estimation in both stationary and nonstationary environments. This system is applied to a large number of rainfall stations in Canada. This research was funded by the Canadian research network called NSERC FloodNet (FloodNet, 2015). FloodNet endeavours to improve the flood forecasting and management capacity in Canada and reduce the negative impacts of flood-related disasters on both human society and the natural environment. Through using annual maximum rainfall series (AMS), the system developed in this research contributes to advancing the understanding of extreme rainfall patterns in Canada with focuses on the development of an automatic procedure for collecting similar extreme rainfall stations in both historical (stationary) and future (stationary) contexts, and dynamic rainfall quantile estimation in a future nonstationary environment.

The system for regional frequency analysis of extreme rainfall events under climate change is developed to advance understanding in the following aspects:

1) The effective selection of similarity indicators for the identification of similar rainfall stations for the site of interest in the historical context. Considering the spatial differences among the extreme rainfall patterns at various sites of interest, an automatic procedure for the selection of effective similarity indicators for the homogeneous group (HG) formation (a group of stations that share similar extreme rainfall patterns) was developed. A three-layer design was proposed based on the three stages in extreme rainfall formation for distinguishing the temporal differences among potential similarity indicators used in HG formation. Then to consider the spatial difference of similarity indicators for HG formation at various target sites, an automatic feature selection and weighting algorithm, specifically the hybrid searching algorithm of Tabu Search, Lagrange Multiplier and Fuzzy C-means clustering, was

used for the search for an optimal feature combination at each site. The effectiveness of this approach was tested at target sites from eight climate regions.

2) The improved procedure for the HG formation in historical and future contexts. This work focused on improving the automatic procedure used for the HG formation in the historical context and adjusting the previous historical homogeneous group for the future context. To increase the efficiency of reducing the impacts of feature correlation on the similar rainfall stations identification process, the feature weighting method in the previous version of the algorithm was replaced by feature extraction techniques. Furthermore, to improve the HG formation results, the Region of Influence (ROI) procedure was added as an additional procedure for gathering similar rainfall stations. Spatial alterations in future rainfall characteristics can directly cause heterogeneity in the previous HGs in the future context. The second part of this work focused on adjusting the previous HGs to the new optimality in the future context, during which the rainfall statistics from the projected rainfall series in the previous HGs served as the new similarity indicators at the target site. The superiority of this work was tested at 153 target stations from four climate regions.
3) The dynamic state estimation for rainfall quantiles in a nonstationary environment. This work focuses on the estimation of non-monotonic changes in extreme rainfall quantiles in a nonstationary environment. To calculate the extreme rainfall quantiles for the stations exhibiting non-stationarity in the future context, the parameters for a regional growth curve at sequential temporal steps can be calculated using an unscented Kalman filter based on the structure established in the index-event model. A one-step forward procedure using the noise covariance estimated from using a Monto Carlo-simulation approach was proposed to successfully apply the unscented Kalman filter in regional frequency analysis and reduce the possibility of performance degradation caused by using insufficient samples for state characterization and inaccurate noise covariance for uncertainty quantification. During the process, large sets of resamples for the rainfall records at each temporal step were used in the one-step forward procedure to increase the accuracy and reduce the uncertainty for the parameter estimates. The effectiveness of this proposed methodology was tested for three homogeneous groups with positive or increasing trends in four climate regions.

This research is an ongoing effort towards a wider use of the regional approach to frequency analysis under climate change and provides automatic procedures for HG formation in the

historical stationary context and an adjustment procedure for the previous HGs reaching new optimality in the future stationary context, plus the framework of nonstationary rainfall quantile calculation using the dynamic state estimation method. All three segments from the proposed system were verified through the comparison with procedures used in a traditional regional frequency analysis.

1.1 Objectives

The overall objective of this thesis is to develop methodologies for the application of regional frequency analysis of extreme rainfall events in a nonstationary environment; the specific objectives of this research include:

1) Chapter 2: The development of a searching algorithm for the selection of similarity indicators used for HG formation at various climate regions in the historical stationary context.
2) Chapter 3: The application of feature extraction techniques and region-of-influence approach in the searching algorithm based on the previous proposed framework for improving the HG formation in the historical stationary context.
3) Chapter 3: The development of an adjustment procedure for HG formation in the future stationary context.
4) Chapter 4: The development of a dynamic extreme rainfall quantile estimation procedure for regional frequency analysis of extreme rainfall events in a future nonstationary environment.

1.2 Thesis Organization

Chapters 2 to 4 of this thesis are provided in the form of manuscripts that have been published or submitted to scientific journals. Chapter 2 has been published in Journal of Hydrology (Yang & Burn, 2019). Chapter 3 is presented as a manuscript submitted to Advances in Water Resources. Chapter 4 is presented as a manuscript submitted to Journal of Hydrology. Transition paragraphs are added at the end of each chapter for the transition from each chapter to the next and to improve the readability of the thesis. The overall conclusions from this thesis, and potential future research, are presented in Chapter 6. The list of references follows Chapter 6.

Chapter 2 Automatic Feature Selection and Weighting for the Formation of Homogeneous Groups for Regional IDF Estimation

This chapter is built upon the published article with the same title in the Journal of Hydrology. Minor differences between the paper and the chapter have been made to facilitate consistency and coherence.

Yang, Z., & Burn, D. H. (2019). Automatic Feature Selection and Weighting for the Formation of Homogeneous Groups for Regional IDF Estimation. Journal of Hydrology, 575(May), 292–307. https://doi.org/10.1016/j.jhydrol.2019.05.015

Summary

The intensity-duration-frequency (IDF) curve has been used as an effective tool to quantify the risk associated with the impact of extreme rainfall on civil infrastructure. However, recent changes in the rainfall climatology caused by climate change and urbanization have made estimates in the stationary environment provided by the traditional regional IDF approach increasingly inaccurate. This inaccuracy is mainly caused by the lack of consideration for the temporal and spatial difference in the selection of similarity indicators (attributes that are used to measure similarity of extreme rainfall patterns among different stations), resulting in ineffective formation of a homogeneous group (a group of stations that share similar extreme rainfall patterns) at various regions. To consider the temporal differences of similarity indicators, including meteorological factors, topographic features and urban impact indicators, a three-layer design is proposed based on the three stages in extreme rainfall formation: cloud formation, rainfall generation and change of rainfall intensity over an urban surface. During the process, the impacts from climate change and urbanization on extreme rainfall patterns are considered through the inclusion of potential features that relate to the rainfall mechanism at each layer. The spatial differences of similarity indicators for Homogeneous Group Formation (HGF) at various regions is resolved by using an automatic feature selection and weighting algorithm, specifically the hybrid searching algorithm of Tabu Search, Lagrange Multiplier and Fuzzy C-means clustering, to select the optimal combination of features for HGF based on the uncertainty in the regional estimates of the rainfall quantiles for a specific site. The proposed methodology fills the gap of including the urbanization impacts on the extreme rainfall patterns during HGF process and challenges the traditional assumption that the same set of features can be equally effective in

generating the optimal homogeneous group in regions with different geographic and meteorological characteristics.

2.1 Introduction

Extreme rainfall events can result in devastating impacts. Under the influence of climate change, more frequent and intensive extreme rainfall events have been observed around the world (Seneviratne et al., 2012). In 1998, a series of extreme rainfall events on the Yangtze River affected the living situation for more than 223 million people, which caused the estimated economic damages at 166,600 million Yuans (Zong & Chen, 2000). In 2005, Canada experienced one of the most devastating floods in the province of Alberta and large areas in southern Ontario, and suffered CAD 800 million insured damages (Dan Sandink, 2013). In the same year, Hurricane Katrina, one of the deadliest and costliest storms in the history of the United States, resulted in nearly 1,500 fatalities (Boyd, 2010). In 2013, extreme floods in Central Europe, which were caused by long periods of heavy rains, resulted in €12 billion overall losses (Khazai et al., 2013). Under the current trend of climate change, the frequency of extreme rainfall events is expected to increase in some parts of world (Goswami et al, 2006; Cai et al., 2014). Considering the potentially increasing damages that can be caused by future extreme rainfall, accurate intensity-duration-frequency (IDF) curves are needed for updating the drainage infrastructure especially in urban areas.

Regional frequency analysis, which extends the data records from one site by gathering the observations from sites that share similar rainfall characteristics, is widely used for IDF estimation at specific sites. However, the estimates of design rainfall using traditional regional IDF curves are becoming increasingly inaccurate in the nonstationary environment. One of the major problems that causes this inaccuracy is the ineffective formation of the homogeneous region. Normally, the stations in the homogeneous group are selected based on their rainfall similarity to the target site. The similarity indicators can be divided into the categories of site characteristics and at-site statistics (Hosking & Wallis, 1997). Site characteristics, which are the physical representations of the weather stations (such as the geographic location, the orientation of the landscape, etc.), are used in the traditional geographic HGF process under the assumption that geographic proximity indicates rainfall similarity (Hosking & Wallis 1997; Ahmad et al.

2013; Burn, 2014; Haddad et al. 2015). However, extreme rainfall events that are assumed to remain stationary have been altered by climate change across Canada since 1950, and the patterns of their changes are not spatially uniform (Warren & Lemmen, 2014). At a large scale, a large increase in precipitation totals is observed in the Arctic, while decreasing trends have been detected in the Prairies region in southern Canada (Zhang et al, 2000). As for the rainfall patterns at a local scale, many papers have stated that daily rainfall is affected by the urban surface heat fluxes (Shepherd, 2005; Miao et al, 2011; Li et al., 2013). Since the constant site characteristics, specifically the geographic features, cannot reveal the possible changes in the rainfall patterns, their effectiveness as similarity indicators in the HGF process needs to be questioned.

Two potential solutions have been proposed to solve this problem. One is to ignore the regionalization step, and directly use remote sensing rainfall records at regional scale for the regional IDF estimation (Marra et al., 2017). However, this approach is used to provide areal information of the IDF curves, and the high uncertainty associated with remote sensing rainfall records can still be a huge problem. The other is to search for the more effective similarity indicators, which will be the focus in this study. To adjust HGF process under climate change, the at-site statistics, such as the mean annual precipitation, the mean number of wet days, the ratio of minimum average 2-month precipitation to maximum average 2-month precipitation, the parameters of hydrologic distributions etc. (Easterling, 1989; Hosking & Wallis, 1997; Gaál & Kyselý, 2009; Yang et al., 2010), have been used as the additional indicators in the process. However, the at-sites statistics are only obtainable at the gauged stations and can only achieve reliable conclusions when long rainfall records are available. To resolve this issue, atmospheric variables have been used as the new similarity indicators. Satyanarayana and Srinivas (2008) considered the region's hydro-meteorology through including monthly mean series of the large-scale atmospheric variables such as the specific humidity, temperature, precipitable water, wind velocity and wind direction, together with the geographical factors as the indicators for the HGF. Gabriele and Chiaravalloti (2013) used the Convective Available Potential Energy and the Q vector Divergence as indicators in the homogeneous macro region formation, and the Vertically Integrated Moisture Flux for the formation of homogeneous sub-regions. These methods are conducted under the assumption that the same set of features can serve as the optimal similarity indicators to detect rainfall patterns at different stations. Asong et al. (2015) advanced the former approach by using canonical correlation analysis to select atmospheric variables,

teleconnection indices and geographical site attributes that could explain the spatial patterns of precipitation of rainfall events in Canadian Prairie Provinces. However, the impacts of these selected indicators on the monthly precipitation were not consistent at different locations over the study area. Since there is a spatial difference among the extreme rainfall patterns at stations with different meteorological and geographic characteristics, should this difference be considered when selecting the similarity indicators for representing these different extreme rainfall patterns at various stations? This paper proposes an objective-oriented automatic feature selection and weighting algorithm to select the optimal feature combination for describing the rainfall patterns at the target site, which is used to form the homogeneous group that has the lowest uncertainty in the quantile estimates.

In the context of the above research gaps, this proposed methodology seeks to answer the following questions:

1) Do the potential similarity indicators respond to extreme rainfall events on the same temporal scale?
2) Is there a spatial difference in the similarity indicators that are used to distinguish rainfall patterns among the input stations at various locations?
3) Can urbanization alter the extreme rainfall regime, thus changing the frequency and magnitude of extreme rainfall events for an urban area?

These questions are explored through the proposed searching algorithm in a three-layer framework. The remainder of the paper is organized as follows. Section 2.2 presents the basic procedures for conducting regional IDF analysis. Section 2.3 describes the automatic feature selection and weighting algorithm in a three-layer design. Sections 2.4 and 2.5 present the data used in the study and the results from the analysis. The paper concludes in Section 2.6 with conclusions and a discussion of potential avenues for future work.

2.2 Regional IDF estimation

An Intensity-Duration-Frequency (IDF) curve is used to describe the occurrence probability of certain rainfall intensity and can provide essential information for civil infrastructure design to reduce the potential damage caused by heavy rainfall. To quantify the associated risk of these extreme rainfall events and capture the changes of the rainfall temporal distribution, extreme

events are commonly described by their return periods (T). In this research, IDF curves are estimated using the annual maximum rainfall series (AMS), which contains independent observations generated from different hydro-climatological events.

Regional frequency analysis is performed under the assumption that rainfall series from all of the stations in the homogeneous region share the same frequency distribution. The quantile estimate can be described by the following index-event equation (Dalrymple, 1960):

$$\hat{R}_{T,D} = \bar{R}_D \hat{x}_{T,D} \tag{2-1}$$

where D indicates the duration of the rainfall events under consideration, $\hat{R}_{T,D}$ is the regional estimate of T-year event for D duration at a specific site, $\bar{R}_D$ is the site's index event at D duration and $\hat{x}_{T,D}$ is the regional estimate of the dimensionless growth curve. The estimates can be obtained through the following steps (Donald H. Burn, 2014):

1) Data screening. To satisfy the stationary requirement in the traditional frequency analysis, the rainfall series with a change point or a trend identified through Pettitt test or Mann–Kendall nonparametric test are removed.
2) Attribute selection. Geographic distance is used to identify stations sharing similar rainfall patterns with the target site.
3) HGF. Homogeneous groups can be formed using the selected indicators through different pooling or clustering approaches.
4) Evaluation of the homogeneous region. A heterogeneity measure is evaluated for the initial pooling group and revisions of this group will be conducted if the heterogeneity measure does not meet the criteria of homogeneity.
5) Estimation of quantiles. The scaled time series from the group obtained from step (4) is used to identify the appropriate distribution and the corresponding quantile function. The rainfall quantile at a target site can be estimated by applying Equation (2-1) to the distribution obtained from step (5).
6) Uncertainty quantification. Parametric samplings through 1000 simulations are used to quantify the uncertainty in the rainfall quantile estimates (Hosking & Wallis, 1997)

2.3 Methodology

To address the research questions and consider the possible alterations in the rainfall regime caused by climate change and urbanization, this paper explores the possibility of including relevant meteorological factors, topographic features and urban impact indicators as the similarity measures for HGF through the feature selection and weighting process in a three-layer framework. The procedures in the proposed algorithm can be briefly summarized as: 1) Based on the rainfall mechanism, potential similarity indicators at all three layers are selected. 2) The relationships between the most responsive indicators to the extreme rainfall events and less responsive ones are analyzed to determine their appropriate temporal resolutions (TRs). 3) Homogeneous groups are formed using the selected similarity indicators and the uncertainty in quantile estimates at the target site is quantified. During the HGF process, feature weighting is used to reduce the impact of feature correlation on the formation process. 4) Step (3) is repeated multiple times using a search algorithm, and the optimal combination of similarity indicators is determined based on the quantified uncertainties.

The details of the proposed methodology are described in the following sections. Based on the mechanism of extreme rainfall events, a three-layer design is proposed in Section 2.3.1, and potential features that can affect extreme rainfall events at each layer are identified. To obtain the optimal TRs for the potential features at each layer, entropy differences and correlation coefficients between the most responsive indicators and less responsive ones at different TRs are used as the measures in Section 2.3.2. In Section 2.3.3 and 2.3.4, Fuzzy C-Means clustering is proposed as the tool for HGF, and Lagrange Multiplier is used to reduce the impact of feature correlation on the clustering results. Tabu Search, which is used as the search algorithm, is described in Section 2.3.5.

2.3.1 Three-layer design

Three stages can be identified in the process of extreme rainfall formation: cloud formation, rainfall generation and the change of rainfall intensity above the urban surface. Based on the possible heights of these stages, a three-layer design is proposed as the framework in which related features at each stage are selected as shown in Figure 2-1. The primary goal of this design is to distinguish the possible temporal differences among the features for the three layers.

The first layer of the proposed framework includes the atmosphere beyond the Planetary Boundary Layer (PBL), and is designed to consider the climate change impacts on the formation of the clouds that produce the extreme rainfall events at a large scale. Cloud is a suspended water or crystal mass in the air, and forms when the water vapor in the atmosphere condenses on the nuclei such as dust or ice.

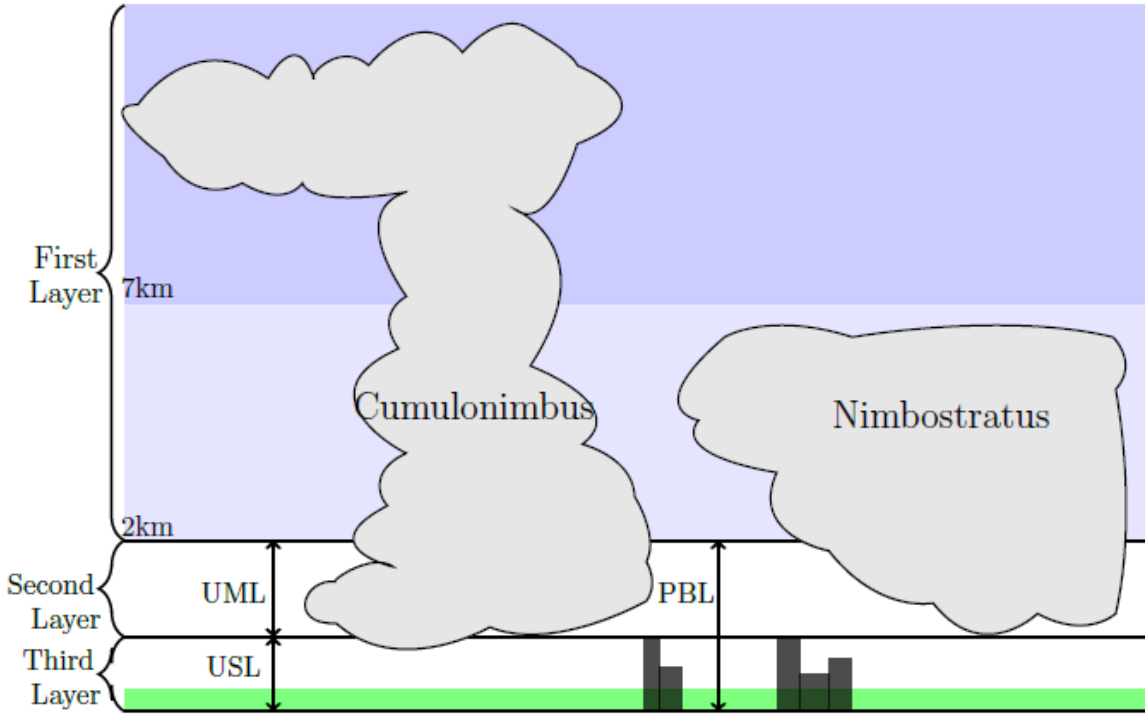


Figure 2- 1 The vertical structure of the three-layer design: Planetary Boundary Layer (PBL), Urban Mixing Layer (UML), Urban Surface Layer (USL). In the figure, the green area indicates the height of vegetation cover, the dark grey boxes indicate buildings (Adapted from Figure 1 in Shepherd, 2005)

Cloud formation is affected by several atmospheric variables including air temperature, geopotential height, specific humidity, U-component and V-component of wind velocity. Thus, these features at three pressure levels (300hPa, 500hPa and 700hPa) are included as potential similarity indicators at this layer. In contrast to the rainfall events that are caused by low level clouds, extreme rainfall events are normally caused by two types of clouds located at middle or high levels of the troposphere: cumulonimbus and nimbostratus, which are corresponding to the production of convective and stratiform rainfall (Houze, 1989; Houze Jr., 1997). Cumulonimbus is a multi-level cloud that starts as a low level cloud (cumulonimbus calvus) and expands to the middle (cumulonimbus capillatus) even high level (cumulonimbus incus) due to the unstable

atmosphere. This cloud type is capable of producing severe convective rainfall events in a short period. Nimbostratus forms in the middle level of the troposphere because of the rising warm air, and occurs along a warm or occluded front where the stratiform rainfall can be produced. To describe the atmosphere instability in the process of these two cloud formations, the Convective Available Potential Energy (CAPE) index, which represents the amount of energy that would be needed to lift a parcel of air through a certain distance in the atmosphere (Moncrieff & Miller, 1976), is selected and acts as the most responsive similarity indicator to the extreme rainfall events in the first layer (Gabriele & Chiaravalloti, 2013a)

After a cloud is formed, rainfall can be generated when the requirement of cooling is met, which happens within the PBL, i.e. in the Urban Mixing Layer (UML) which is the second layer of the framework. As the linkage between the first and third layer, the UML is under the influence from both climate change and urbanization at a regional scale (Collier, 2006). Based on the cloud types that can be used to generate the extreme rainfall events, the types of rainfall can be divided into the convective rainfall, and stratiform rainfall type which can be further divided into relief and frontal rainfall depending on the underlying geographic topography. Thus the geographic attributes including latitude and longitude of the potential stations, and the same category of atmospheric variables used in the first layer but at lower pressure levels (850hPa and 925hPa) are collected as potential indicators. What makes the extreme rainfall events different from the normal rainfall events is the abundance of water supply to feed the events (Gabriele & Chiaravalloti, 2013a). Thus the vertical integral of divergence of moisture flux, which describes the transport of the net atmosphere moisture flux per unit volume, is selected as the most responsive indicator in this layer.

The final layer is the Urban Surface Layer (USL), where rainfall intensity can be affected by the urban surface energy in sub-regional or even local scale. Many papers have stated that the local rainfall climatology in the urban environment has been altered due to the following mechanisms (Li et al., 2013): 1) Urban heat island effects. The presence of high buildings can increase surface roughness, and also affect the local energy fluxes or disrupt the thermal balance. This results in the locally enhanced convergence and increased surface temperature, which causes changes in the rainfall patterns in the urban area (Roth & Oke, 1993; Roth, 2000; Bornstein & Lin, 2000; Shepherd, 2005); 2) Urban canopy effects. The lack of the canopy covers and the

altered underlying surface characteristics can affect the surface temperatures or even divert precipitating systems (Shepherd, 2005; Loughner et al., 2012); 3) Urban aerosol effects. The increasing accumulation of aerosols in the atmosphere can both decrease and increase rainfall amount because of their radiative and cloud condensation nuclei (CCN) activities (Diem & Brown, 2003; Kaufman & Koren, 2006; Jin et al. 2010; Rosenfeld et al., 2014): a) radiation effects. On one hand, the aerosols impede the process of solar radiation reaching the land surface thus reducing the amount of available water for evaporation and cloud formation. On the other hand, aerosols can absorb solar radiation and cause the warming in the lower atmosphere, which can strengthen Asia monsoon circulation and increase local precipitation. b) CCNs activities. While the added CCNs can accelerate the process of cloud drops turning into raindrops, they can also cause the adverse effects if the cloud drops are too small. Thus relevant features are selected as rainfall similarity indicators in the last layer: The urban energy flux including the surface sensible heat flux and the surface latent heat flux (Miao et al., 2011), Photosynthetically Active Radiation index (PAR), Surface Net Solar Radiation (SNSR), Surface Net Thermal Radiation (SNTR) and the Surface Roughness (SR). Potential features at all three layers are listed in Table 2-1.

Table 2- 1 Potential feature dataset for each layer at the three layer design

Target layer	Potential features
First layer	Air temperature (Air), Geopotential height (Geo), Specific humidity (Sphu), U-component (Uwind) and V- component (Vwind) of the wind velocity (at the 300hPa, 500hPa and 700hPa pressure level), Convective Available Potential Energy (CAPE) and Q vector Divergence (QD)
Second layer	Air temperature (Air), Geopotential height (Geo), Specific humidity (Sphu), U-component (Uwind) and V- component (Vwind) of the wind velocity (at the 850hPa and 925hPa pressure level), Vertical Integral of Divergence of Moisture Flux (VIDWV), Latitude, Longitude, Elevation
Third layer	Urban Surface Sensible Heat Flux (SHTFL), Urban surface Latent Heat Flux (LHTFL), Photosynthetically Active Radiation index (PAR), Surface Net Solar Radiation (SNSR), Surface Net Thermal Radiation (SNTR), Surface Roughness (SR)

2.3.2 TR detection of the features series

Based on the description in Section 2.3.1, the potential features at each layer are under the influence of various phenomena and thus can be related to the rainfall events at different temporal scales. It is necessary to obtain representative values of those features at the optimal

TRs that can best describe the rainfall related patterns at the different layers. The correlations and entropy differences between the most responsive features and the less responsive ones at different TRs are used as indicators to measure their similarities and determine the optimal TRs at the three layers. While correlation coefficient works best among linear correlated features, the entropy difference may provide better explanations when the input features are nonlinearly correlated (Quiroz et al., 2011).

The discrete wavelet decomposition (DWT) is used to obtain the orthonormal bases of feature values at different TRs. In the discrete form, the wavelet transform of a function $f(t)$ is defined as the integral transform as (Kumar & Foufoula-Georgiou 1997):

$$Wf(m_w, n) = \lambda_0^{-m_w/2} \int_{-\infty}^{\infty} f(t)\psi\left(\lambda_0^{-m_w} t - nt_0\right)dt \tag{2-2}$$

$$\psi_{m_w,n}(t) = \frac{1}{\sqrt{\lambda_0^{m_w}}}\psi\left(\frac{t - nt_0\lambda_0^{m_w}}{\lambda_0^{m_w}}\right) = \lambda_0^{-m_w/2}\psi\left(\lambda_0^{-m_w} t - nt_0\right) \tag{2-3}$$

where $f(t)$ indicates the one-dimensional time series; λ is the scale parameter; t is a location parameter (indicates time in time series); $\overline{\psi}_{\lambda,t}(\mu)$ is the complex conjugate of $\psi_{\lambda,t}(\mu)$ (in this study, the Least Asymmetric wavelet); m_w and n are an integer and λ_0 is a fixed dilation step greater than 1. Details of the multiresolution analysis can referred to Mallat (1989) and Daubechies (1990).

To focus the methodology solely on the annual extreme rainfall events, the feature values should match the date of occurrence for the annual extreme rainfall events and be extracted at the optimal TRs at each layer.

2.3.3 Homogeneous group formation

The primary goal of this study is to obtain homogeneous groups that can generate the lowest uncertainty in the quantile estimates. Normally, HGF can be conducted using two approaches: pooling and clustering. The pooling method , such as the Region of Influence (ROI) approach, gathers the stations centred around the pre-set target site based on their similarities (Donald H.

Burn, 1990). However, the problem of determining the number of stations in the target group to achieve the optimal balance between the number of stations in the pooling group and its homogeneity is still unsolved (Donald H. Burn, 2014). Here, the clustering, specifically the Fuzzy C-Means clustering is used for the HGF (Bezdek et al, 1984).

Unlike non-overlapping clustering, Fuzzy clustering provides the possibility for one site having partial membership in more than one cluster, which has the advantage to form the overlapping clusters for undistinctive data with vague boundaries (Bezdek et al, 1984). Furthermore, in the situation when only limited number of input stations are available, the possibility of non-target site being the cluster centre allows the clustering method to form better groups to generate the lowest uncertainty in quantile estimates.

Fuzzy C-Means clustering is conducted through the minimization of the following objective function (Bezdek et al, 1984):

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m}\left\|x_i - c_j\right\|^2, \qquad 1 \le m < \infty \tag{2-4}$$

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left[\frac{\left\|x_i - c_j\right\|}{\left\|x_i - c_k\right\|}\right]^{\frac{2}{m-1}}} \tag{2-5}$$

where N is the number of stations; C is the number of clusters; u_{ij} is the membership value of station i in the cluster j, and can be calculated using Equation (2-5) ; x_i is the value of data points in station i; c_j is the center value of the cluster j; and m is the weighting exponent in the Fuzzy C-Means cluster.

2.3.4 Feature weighting

To deal with the possible issues caused by feature correlation during clustering, the Lagrange Multiplier is used (Borgelt, 2008). The objective function that is used in Lagrange Multiplier can be described as:

$$\varepsilon = \sum_{V=1}^{V} (w_V)^p \sum_{i=1}^{N} \sum_{k=1}^{M} \delta_{ik} \left\| \vec{x}_i^{\,V} - \vec{m}_k^{\,V} \right\|^2 \tag{2-6}$$

where p is the exponential parameter that is used to control the sparsity of the feature weightings. Research has shown that the clustering results remain stable when p increases to a certain value (Xu, et al. 2014), and is set to 6 in this study; V is the number of variables in the cluster; w_V is the weighting value for each variable; N is the number of stations involved in the cluster; M is the number of center groups obtained from clustering; δ_{ik} is the belonging value of each station to each cluster; $\vec{x}_i^{\,V}$ is the time series vector for a variable for station i; $\vec{m}_k^{\,V}$ is the time series vector of certain variable for each cluster center k .

Based on the objective function in Equation (2-6), the Lagrange formula that will be applied to obtain the weighting value for each variable is (Xu, et al., 2014):

$$L(\varepsilon, \lambda) = \varepsilon(w) + \lambda \left(\sum_{V=1}^{V} w_V - 1 \right) \qquad subject\ to: \sum_{V=1}^{V} w_V = 1 \tag{2-7}$$

Based on Equation (2-7), the corresponding weightings can be derived from the derivative and are listed as (Xu, et al., 2014):

$$w_V = \frac{1}{\sum_{V'=1}^{V} \left(\frac{D_V}{D_{V'}} \right)^{1/(p-1)}}, \qquad P > 1; \quad D_V = \sum_{i=1}^{N} \sum_{k=1}^{M} \delta_{ik} \left\| x_i^V - m_k^V \right\|^2 \tag{2-8}$$

2.3.5 Search algorithm

To address the first research question and select the optimal combination of similarity indicators for HGF, Tabu Search, which is regarded as the most effective tool to obtain the optimal subsets of relevant attributes in the feature selection domain, is used (Zhang & Sun, 2002). The Tabu Search is a combinatorial optimization algorithm that avoids the possibility of the searching process being stuck in a local optimum by crossing the boundaries of local optimality through using the inferior objective values in the process (Glover, 1986).

Three types of memory structures are involved in the search process: short, intermediate and long-term. The short-term memory or the regency-based memory, which is constantly modified in the search, is used to record the Tabu list and guide the search process (Glover, 1989; Zhang & Sun 2002). It is constructed by labeling the attributes that have been visited recently as the Tabu-active attributes, and these attributes will not be used again in the following search (Glover, 1989). While Tabu list is recorded to reduce the visiting of pre-visited attributes, the aspiration criteria is used to restrict the pre-visited moves (i.e. the Tabu move), and can be updated if the objective function value resulted from using the solution in consideration is better than the current aspiration value (Glover, 1989). In summary, the short-term Tabu Search can be carried out through the steps showing in Figure 2-2 (Sait & Youssef, 1999; Zhang & Sun, 2002). The intermediate memory and long term memory are used in the intensification and diversification process separately to either re-evaluate historically good solutions or incorporate features that have not been previously included as new solutions (Glover, 1989).

```
1. Set the initial values for a feasible solution, Tabu list and aspiration level.
2. For i in each iteration {
3.      Generate the neighbour solutions
4.      Find the next best solution through the comparison of their objective values
5.      # update the initial values.
6.      If (The best solution does not contain Tabu-active attributes {
7.          Best solution can be admissible
8           Update the Tabu list and aspiration level
9.      } else {
10.         If (The best solution generates better objective value than the current aspiration value)
11.     {
12.             Best solution can be admissible
13              Update the Tabu list and aspiration level
14          }
15      }
16.     Increment iteration number i
```

Figure 2- 2 Algorithmic description of Tabu Search (Adapted from Figure 3 in Tahir et al, 2007)

Considering the large number of potential features involved in Tabu Search, the process is established under the assumption that the best feature combinations that generates the optimal objective value at each layer can result in the best final outcome for the whole process. During the process, the searching for the optimal feature combination at each layer will be conducted separately, as shown in Figure 2-3. To avoid over-clustering caused by conducting the search at every layer, the comparison of the objective values between the groups obtained from the higher layer and the lower layer are used as the criteria to determine the necessity of conducting the procedure at the lower layer. If the lower layer generates a better objective function value compared to that from the higher layer, the search results from the lower layer is acceptable. If this condition is not met, the proposed procedure will not be conducted in the lower layer, and the algorithm will directly move to the next layer instead. The objective value being used in the searching process is the uncertainty of quantile estimates generated from the formed group, i.e., the average widths of confidence interval (CI) for the rainfall quantile estimates at different return periods.

2.3.6 Automatic feature selection and weighting algorithm

Based on the above sections, the automatic feature selection and weighting algorithm can be conducted through the following procedures. The flow chart of the proposed search algorithm is illustrated in Figure 2-3.

1) Original Feature Gathering. The time series of the potential rainfall-related features for the initial group of stations at all three layers are extracted at available spatial resolutions.
2) Temporal Scaling. Based on the method presented in Section 2.3.2, wavelet decomposition is used to extract the feature values at different TRs.
3) Feature Scaling. The method of scaling to unit length is used to standardize the range of each feature value, can be achieved by using equation $x' = \frac{x}{\|x\|}$, where x is the feature vector before scaling, $\|x\|$ is the Euclidean length of the vector, and x' is the vector after scaling.
4) Feature selection and weightings at the first layer or the higher layer. The hybrid searching algorithm can be conducted through the following procedures:
   (a) Initial selection of features. Randomly select certain number of features from the potential feature datasets in the higher layer, and record them in the Tabu list.

(b) Initial equal weightings for the selected features.

(c) Fuzzy C-Means clustering. Use the weightings from step (b) and the selected features from step (a) in the Fuzzy C-Means clustering to form the homogeneous group for the target site.

(d) Features weightings recalculation. Based on the clustering results from step (c), re-calculated the features weightings using Lagrange Multiplier.

(e) Repeat step (c) and (d). Use the feature weightings from step (d) as the new weightings, and repeat the steps (c) and (d). This process continues until the weighting obtained from step (d) is close to the weightings used in step (c). Then use these final weightings in step (d) to obtain the final homogeneous group for the target site.

(f) Objective value calculation. Use the formed group obtained from step (e) for the objective value calculation, i.e., the average CI widths for the quantile estimates at different return periods. Record this objective value in the searching process.

(g) Repeat steps (a) – (f) certain times. Based on the recorded memory of the objective values and Tabu list, select the next possible combination of features, and repeat the above process.

(h) Searching stop. After the pre-set number of iterations in the search process, select the optimal set of features based on the recorded memory of objective values.

5) Comparison. Compare the formed group from step 4) with the original input group for uncertainty in quantile estimates of the target site through the following steps: a) Uncertainty quantification. To quantify the uncertainty in the quantile estimates, parametric sampling is used to calculate the CI widths (Hosking & Wallis, 1997). b) Uncertainty Comparison. Calculate the ratio of the CI widths between the formed homogeneous groups from lower layer and the original input group from the higher layer. c) Boxplots. Repeat step b) 1000 times and box plots are created based on these calculated ratios. If the median values of the ratio in this boxplot was less than one at most of the return periods, the formed group at the lower layer would be considered as a better homogeneous group and accepted as the input group for repeating step 4) at the next layer. If not, the original group of stations would be used as the input for repeating step 4) at the next layer.

6) Repeat steps 4) and 5) at the second layer using the resulting group from previous layer as the input.

7) Repeat steps 4) and 5) at the third layer using the resulting group from previous layer as the input.
8) Final homogeneous group is obtained. The resulting group from step 7) will be used as the final homogeneous group for the quantile estimation at target site.

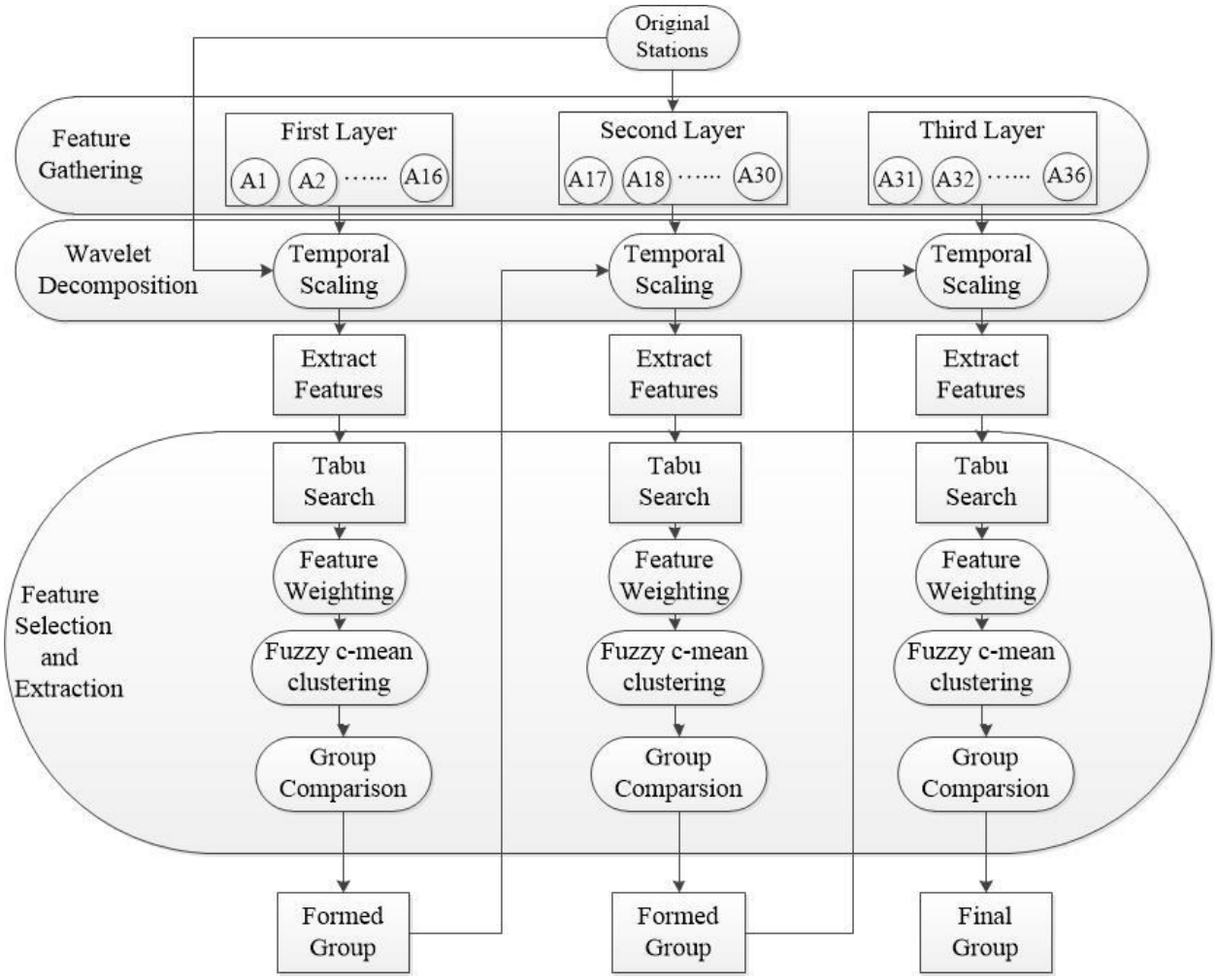


Figure 2- 3 The flowchart of Automatic Feature Selection and Weighting algorithm

2.4 Data and study area

2.4.1 Study area

The proposed methodology is tested in four regions in Canada. Across Canada, there are eight different climate regions: Arctic, Pacific Maritime, Cordilleran, Taiga, Boreal, Prairie, Southeastern and Atlantic Maritime. To verify the assumption of the proposed methodology that different optimal feature combinations can be used as the similarity indicators for HGF with

different input stations at various regions, four regions with a large number of rainfall stations were selected as the tested regions:

1) Region 1 is along the west coast and includes the IDF stations from Yukon Territory, Northwest Territories and intensive urbanized area in British Columbia. Due to the close proximity to the ocean and unique geographic characteristic in the mountain area, this region mainly experiences the Pacific Maritime and Cordilleran climate type; thus stratiform rainfall will be the main rainfall type throughout the year.
2) Region 2 is in the west and specifically includes weather stations from the Prairie and the adjacent Boreal climate region, where convective extreme rainfall is the main extreme rainfall type in the summer.
3) Region 3 is in the Boreal and Southeastern regions and includes stations from Ontario and Quebec, both of which have some intensive urbanized areas. While convective extreme rainfall events may be the main rainfall type in the Boreal region, stratiform rainfall is perceived to be more common in the Southeastern region due to its closeness to large water bodies.
4) Region 4 is in the Atlantic region and includes the stations from Nova Scotia, New Brunswick, Prince Edward Island and Newfoundland, where rainfall events are affected by cyclonic storms from the Atlantic Ocean. Due to the limited number of IDF stations in the Atlantic region, some nearby sites from Quebec are also included.

2.4.2 Extreme rainfall datasets

Based on the methodology presented in Section 2.3, two categories of information are required before conducting the searching process: the values of the extreme rainfall at different durations and their corresponding dates of occurrence. The first category of information is used for the rainfall quantile estimation during the uncertainty comparison and can be retrieved from the Engineering Climate Datasets provided by Environment and Climate Change Canada (ECCC) in the form of the annual maximum rainfall series (AMS) at nine durations ranging from 5 min to 24 h (5, 10, 15, and 30 minutes and 1, 2 6, 12, and 24 h). If the data records were not complete during the analysis period, the historical dataset also provided by ECCC could be used as an alternative source for providing missing AMS data points for the duration of 24h. For the second

category, the corresponding dates of occurrence for the annual extreme rainfall events were obtained based on the information provided from both the historical dataset on ECCC website and the total precipitation from ERA-Interim database provided by European Centre for Medium-Range Weather Forecasts (ECMWF) at the resolution of 0.125 degrees.

2.4.3 Potential features

Based on the methodology presented in Section 2.3.1, the potential features that on a yearly basis correspond to the date of occurrence for the AMS are extracted at different TRs from two sources: NOAA Global Ensemble Forecast System Reforecast (GEFS/R) and ERA-Interim Database from ECMWF (Berrisford et al., 2011;Hamill et al., 2013).

a) GEFS/R: The geopotential height, temperature, U-component and V-component of the wind and the specific humidity were extracted at five pressure levels (925hPa, 850hPa, 700hPa, 500hPa, 300hPa), plus the convective available potential energy, were collected at a resolution of 1 degree. All the data were extracted at the 3-hourly temporal interval basis.

b) ERA-Interim: The second layer indicator Vertical Integral of Divergence of Moisture Flux (VIMF) at a resolution of 1 degree, and the third layer indicators including Photosynthetically Active Radiation index (PAR), Surface Net Solar Radiation (SNSR), Surface Net Thermal Radiation (SNTR) and the Surface Roughness (SR) were extracted at a resolution of 0.125 degree. All of these indicators are extracted at the 6-hourly temporal interval basis.

2.4.4 Data Screening

Based on the availability of the observations from the above sources, the period for the search algorithm was set from 1985 to 2004. However, in the uncertainty comparison among different formed groups, all the available data points in the stationary rainfall period from the Engineering Climate Datasets were used considering the statistical accuracy that can be achieved with a larger sample size. Thus during the process of data screening, the selected stations in each region should have stationary extreme rainfall time series for two periods (the period from 1985 to 2004 and the whole available time period obtained from Engineering Climate Datasets), since the extreme rainfall series from shorter period could be regarded as the representatives for the longer period if they were both stationary. This objective was achieved through applying Pettitt test and Mann–Kendall nonparametric test with the block bootstrap resampling (Donald H. Burn, 2014).

The final number of stations with stationary rainfall series for each region is: 86 stations in Region 1 with VANCOUVER INTL A being the target site; 78 stations in Region 2 with the CALGARY INT'L being the target station; 162 stations in Region 3 with TORONTO CITY being the target station; and 72 stations in Region 4 with GANDER AIRPORT CS being the target station.

2.5 Application

To demonstrate the effectiveness of the proposed methodology and address all the research questions, the analysis in the following sections were conducted: 1) to answer the first research question, the methodology in Section 2.3.2 was applied in Section 2.5.1 to distinguish the optimal TRs for the features at different layers. 2) To answer the second question, all correlation coefficients and entropy differences between the most responsive indicators and less responsive ones at the first and second layers for all the stations were compared in Section 2.5.1 to demonstrate the spatial differences among their extreme rainfall pattern indicators. 3) To prove the necessity of conducting the algorithm at all three layers and answer the third question, the uncertainty in the quantile estimates from the homogeneous groups that were generated by using the algorithm at lower layer were compared to that in the formed groups from the higher layer in Section 2.5.2. 4) To illustrate the effectiveness of the proposed methodology, the algorithm was applied to form the homogeneous groups for rainfall events with different durations at four different regions in Section 2.5.2 and 2.5.3, and the uncertainty in their quantile estimates were compared to that from using at-site analysis and also that in the formed groups generated by using a geographic approach in Section 2.5.4.

2.5.1 Temporal scaling for the potential features

Both entropy differences and correlation coefficients between the most responsive feature and less responsive ones at different TRs are used to determine the optimal TRs at first and second layers. Since the relationship among the potential features at each layer may not be linear, entropy differences may provide more reliable conclusions than correlation coefficients. In the first layer, correlations between the CAPE and the rest of the first layer features at all stations reach the highest values at the 128- day TR (level 10 decomposition), and their corresponding entropy differences are reasonably small at this level, thus 128 days is selected as the optimal TR

for the features at the first layer. This 128-day TR is reasonable for the following reasons: 1) While precipitation process is a mesoscale atmospheric processes, it can be affected by the large scale phenomena including large clouds formation, warm and cold fronts collision etc. 2) In the three-layer design, the features in the first layer are used as the similarity indicators to gather stations that share similar extreme-rainfall related atmospheric characteristics at large scales. 3) Studies shows that extreme events-related atmospheric movements, specific CAPE-related features, are strong seasonal correlated with the seasonal migration of monsoon over West Africa and India (Murugavel et al., 2012; Meukaleuni et al., 2016).

The situation is more complicated at the second layer. For the majority of the stations, the correlation coefficients between the VIMF and the rest of features at the second layer increase steadily as the decomposition level increases until it reaches either level 10 or 9 (64-day TR) and the declines in their corresponding entropy differences become barely noticeable after level 9. Thus the 64-day TR is used as the optimal TR for extracting feature values at the second layer. In the third layer, the optimal TR is set to the duration of the extreme rainfall events of concern (i.e., 24-hour in this case) for two reasons: 1) No universal TR for all the stations can be obtained based on the values of entropy differences and correlation coefficients. 2) To meet the needs of considering extreme events at different durations for different construction purposes.

The spatial difference of the correlations between potential features for the stations in Region 1 is presented in Figure 2-4. Although at each station, the highest correlation coefficient values between the most responsive features and less responsive ones were all reached at level 10 or 9 decomposition for the first or second layers, major differences are still observed among these correlation values across different stations. At each station (which is described as each column in Figure 2-4), all the potential features at first (Figure 2-4(a)) and second (Figure 2-4(b)) layers have different levels of correlations to the extreme rainfall mechanisms. For each potential feature (which is described as each row in Figure 2-4), its correlations to extreme rainfall events vary largely among different stations at both layers. Thus the assumption of spatial difference for similarity indicators existing at different stations is valid. The same situation can be applied in Regions 2, 3 and 4.

However the approach of determining similarity indicators by choosing the ones that have higher correlation values at the target site is not suggested for the following reasons: 1) The correlation

or entropy differences obtained is not equivalent to the representation of the mutual information among different features. 2) Even if the mutual information were used as a potential measure, the approach of choosing the features that are highly relevant to the most responsive features might result in biased clustering as these features may not represent all aspects in the complex rainfall system. 3) In the stations where low correlations have been detected across all the features, such as the first column in Figure 2-4 (a), the approach of choosing comparatively higher correlated features as the similarity indicators may result in inferior HGF.

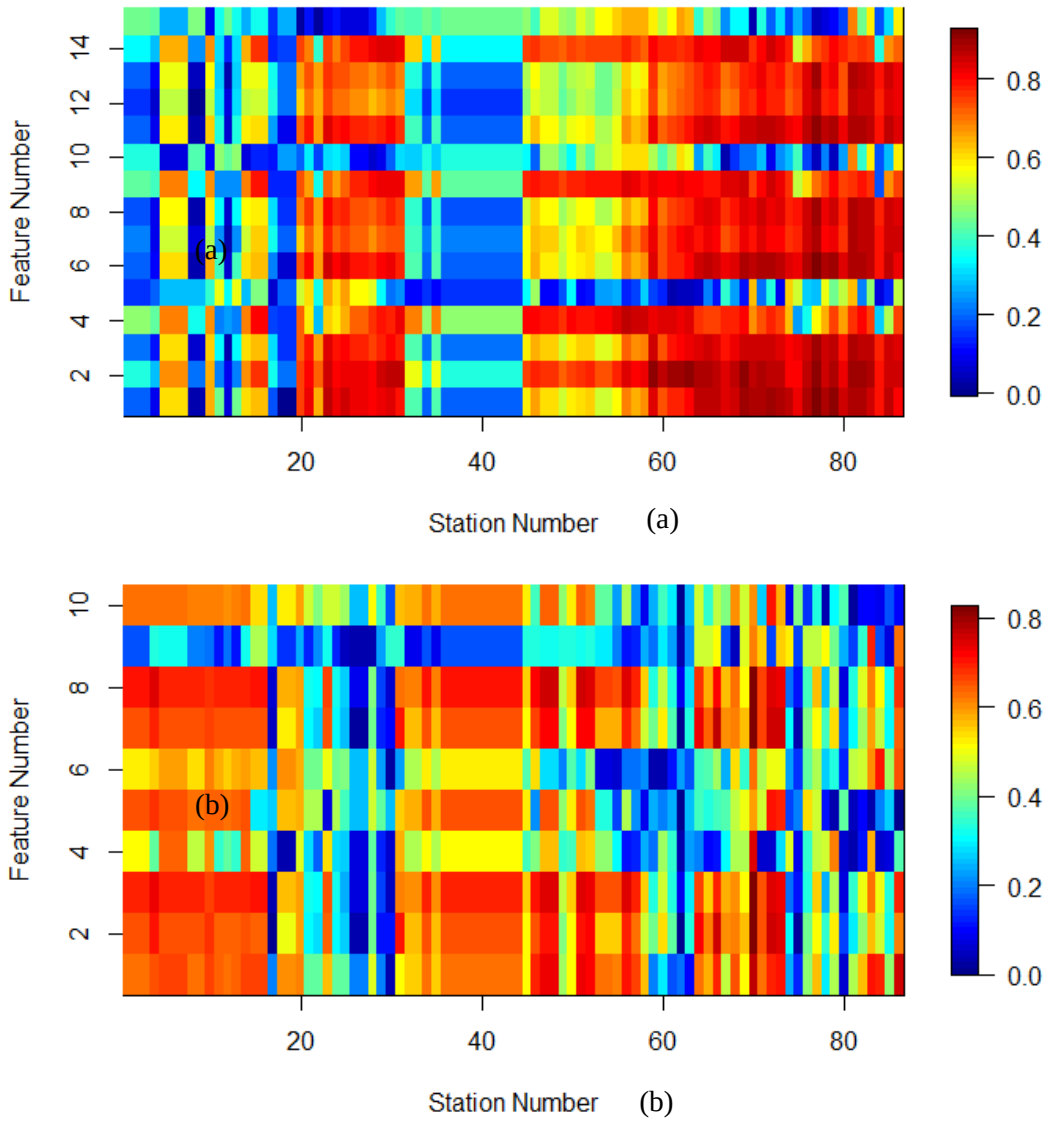


Figure 2- 4 Graphic display of the correlation coefficient values between the most responsive features and the less responsive ones in the first (a) and second (b) layers at level 10 and 9 decomposition from the stations in Region 1. The column number indicates the stations in Region 1, and the row number indicates the potential features in first and second layers. Each grid represents correlation value between the most responsive feature and one of the less responsive ones at one station.

2.5.2 HGF in Tabu Search

The cluster number was set to 2 in the Fuzzy C-Means clustering for the following reasons: 1) the larger the formed group the lower the uncertainty in the quantile estimates. The goal of this study is to form homogeneous groups with the lowest uncertainty in the quantile estimates not to find the optimal partition of the input stations in the clustering. In the formed group, both group heterogeneity and small group size can increase the uncertainty in the final quantile estimates. But the influence from group size normally outweighs the impact from heterogeneity as the geographic proximity of the original input group guarantees a certain level of group homogeneity. 2) The pre-knowledge of the cluster number for the first two layers. The clustering at the first layer is constructed to distinguish the stations whose extreme rainfall events are mainly caused by cumulonimbus or nimbostratus clouds. However, the extreme rainfall events at certain durations, especially the longer ones, are likely caused by the combination of different cumulonimbus or nimbostratus clouds. For those stations whose rainfall events are caused by the combination of cumulonimbus and nimbostratus, the second layer is used to further separate the input stations into the convective and stratiform rainfall type. For those stations whose rainfall events are caused by the nimbostratus, the clustering at the second layer is used to further separate rainfall type into relief and frontal rainfall. 3) The validity measures including partition coefficient (PC) index, partition entropy (PE) and the modified Xie–Beni index are not very effective in determining the optimal cluster number in this study (Xie & Beni, 1991; Wang & Zhang, 2007; Satyanarayana & Srinivas, 2011). As a conclusion, to gather a large number of stations to form the homogeneous group and reduce the computational cost without jeopardizing the search for the optimal feature combination, the search for the optimal feature combination was conducted with the fixed value of the weighting exponent ($m=2$) and the cluster number ($n=2$) for the clustering at all layers.

Based on the methodology presented in Section 2.3.4 and 2.3.5, the automatic feature selection and weighting algorithm is conducted consecutively in three layers. At each layer, the number of iterations is set to 60 with 3 repetitions in Tabu Search.

Two possible issues should be noted when selecting the optimal feature combination based on the memories recorded in Tabu Search: 1) Due to the stochastic characteristics in the estimation for the objective function values for each homogeneous group formed, one choice of feature

combination can repeat several times in the Tabu Search with different objective function values. Thus the frequency of the feature combinations appearing in the top 100 in the recorded memory is also taken into consideration during the determination of optimal feature selection. 2) The high correlations among the input features can cause many sets of feature combinations, with different weightings, to generate the same clusters. Thus, the combinations with the minimal number of features that can be used to form the homogeneous group is selected as the final choice. The selected feature combination with the corresponding weightings are presented in Table 2-2. The homogeneous groups at each layer in four regions can then be generated using the selected features; the groups are presented in Figure 2-5.

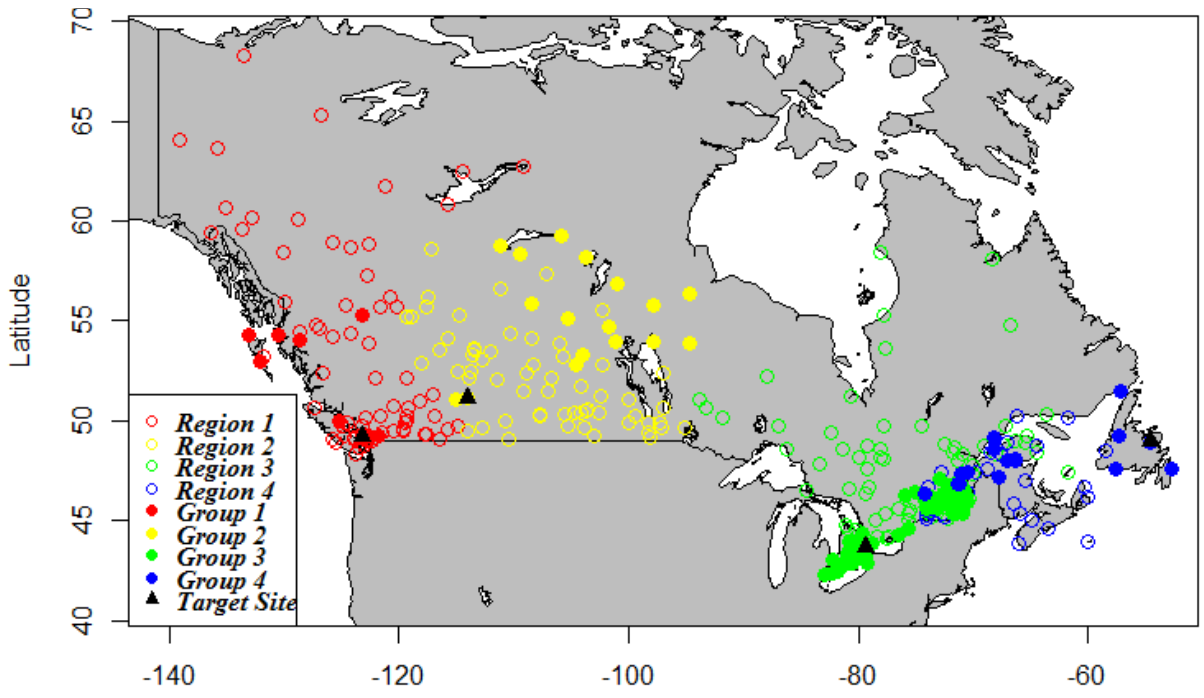


Figure 2- 5 Graphic display of homogeneous groups and original input groups for extreme rainfall events at 24h duration in Region 1, 2, 3 and 4. The circle dots in red, yellow, green and blue in the graph indicate the original input stations, while the solid dots at four colors represent the homogeneous groups in four regions. The black triangles represent the target sites in four different regions.

Table 2- 2 Feature selection and weighting results from Three-layer search algorithm

(a) results from the procedure at first layer

		air300	air500	air700	geo300	geo500	geo700	vw300	vw500	vw700	uw300	uw500	uw700	sphu300	sphu500	sphu700	CAPE
Region 1	Selection	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	0
	Weightings			0.243		0.239					0.167		0.166		0.184		
Region 2	Selection	1	0	0	1	0	0	0	0	0	0	0	1	1	1	0	0
	Weightings	0.242			0.236								0.166	0.177	0.178		
Region 3	Selection	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1
	Weightings						0.308					0.233				0.245	0.214
Region 4	Selection	0	0	0	1	1	1	0	0	0	0	1	0	0	1	1	0
	Weightings				0.190	0.188	0.185					0.140			0.148	0.149	

(b) results from the procedure at second layer

		air850	air925	geo850	geo925	vw850	vw925	uw850	uw925	sphu850	sphu925	VIMF	Latitude	Longitude	Elevation
Region 1	Selection	1	0	1	0	0	1	0	0	1	1	0	0	0	0
	Weightings	0.236		0.235			0.166			0.181	0.181				
Region 2	Selection	1	0	1	0	0	0	0	1	0	0	0	0	0	0
	Weightings	0.366		0.370					0.264						
Region 3	Selection														
	Weightings														
Region 4	Selection	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	Weightings									1					

(c) results from the procedure at third layer

		LHTFL	SHTFL	PRS	SSR	SNR	SRH
Region 1	Selection	0	0	0	1	0	0
	Weightings				1		
Region 2	Selection	1	1	1	1	0	0
	Weightings	0.239	0.248	0.259	0.254		
Region 3	Selection	1	0	0	0	0	1
	Weightings	0.510					0.490
Region 4	Selection	0	0	0	1	0	1
	Weightings				0.520		0.480

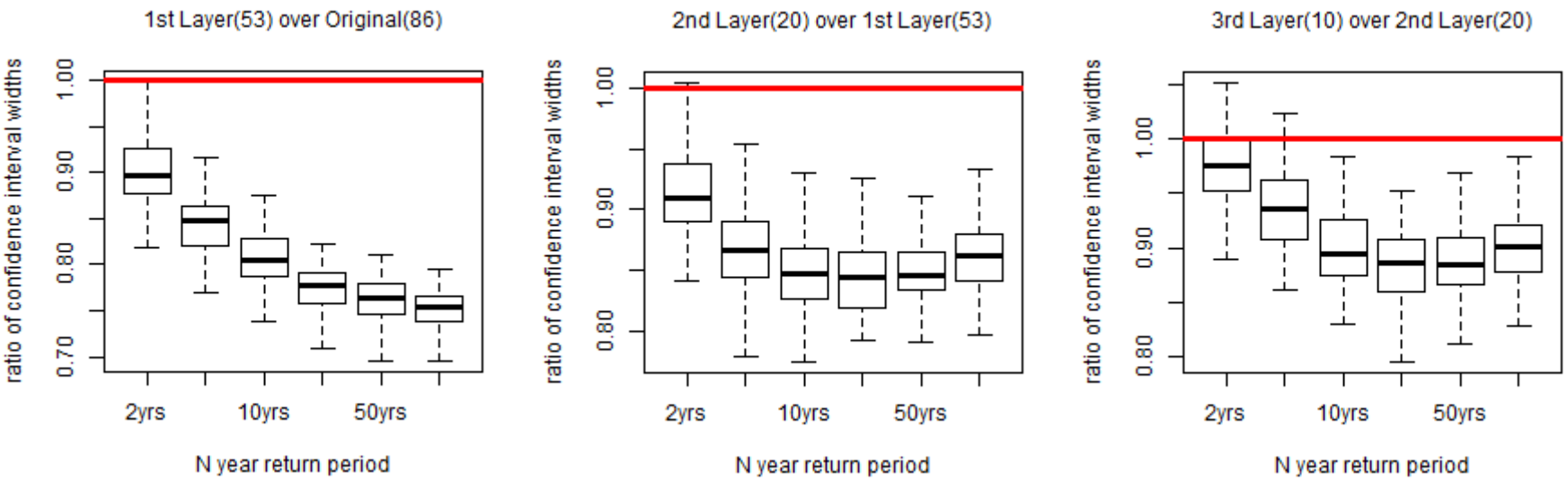
According to the graphic displays in Figure 2-5 and the boxplots for comparison of uncertainty in the quantile estimates in Figure 2-6, several findings can be observed: 1) The algorithm at first layer serves its purpose. As shown in Figure 2-5, all of the stations from each final homogeneous group (Groups 1, 2 and 3) in Regions 1, 2 and 3 lie within the same climate region and its adjacent area. In Region 4, the limited number of rainfall stations in the climate region are insufficient to produce accurate estimates and stations from other climate regions are also included in Group 4 to improve the accuracy of the estimates. 2) The implementation of judging criterion is effective. During the searching process in Region 3, the formed group from the second layer is not admissible as the uncertainty in its quantile estimates is higher than that from the formed group generated from the first layer as shown in the second plot in Figure 2-6 Region 3. This forces the search process to skip the procedure in the second layer and step directly to the third layer. 3) The three-layer framework is effective. The uncertainty in the quantile estimates from the formed groups is decreasing after each layer even when the number of stations in these group is also decreasing. 4) The third layer in the proposed algorithm can be equally effective in the less urbanized area. In Regions 2 and 4 where most of the stations are located in the less urbanized area, only a small number of stations are removed from the formed group generated from the procedure at the second layer, and the final homogeneous region resulted from the procedure at third layer can still generate the most accurate quantile estimates among all the formed groups. 5) Within each homogeneous group, the gathered stations may be located far away from the target site for the following reasons: a) the geographic location is just one potential similarity indicators used in HGF in the second layer. The final formed group does not entirely relied on the geographic proximity. b) The Fuzzy C-Means clustering allows for the possibility of non-target site being the centre of the formed group since the generated clusters are heavily depending on the information from the input group. In the case where some stations locating away from the target site, these stations may still have the feature values that are similar to the cluster centre. c) Even under the circumstance that the selected stations are comparatively far away from the target station, all the stations in the formed homogeneous group are still located in the same climate region.

2.5.3 Homogeneous group modification for rainfall series in short durations

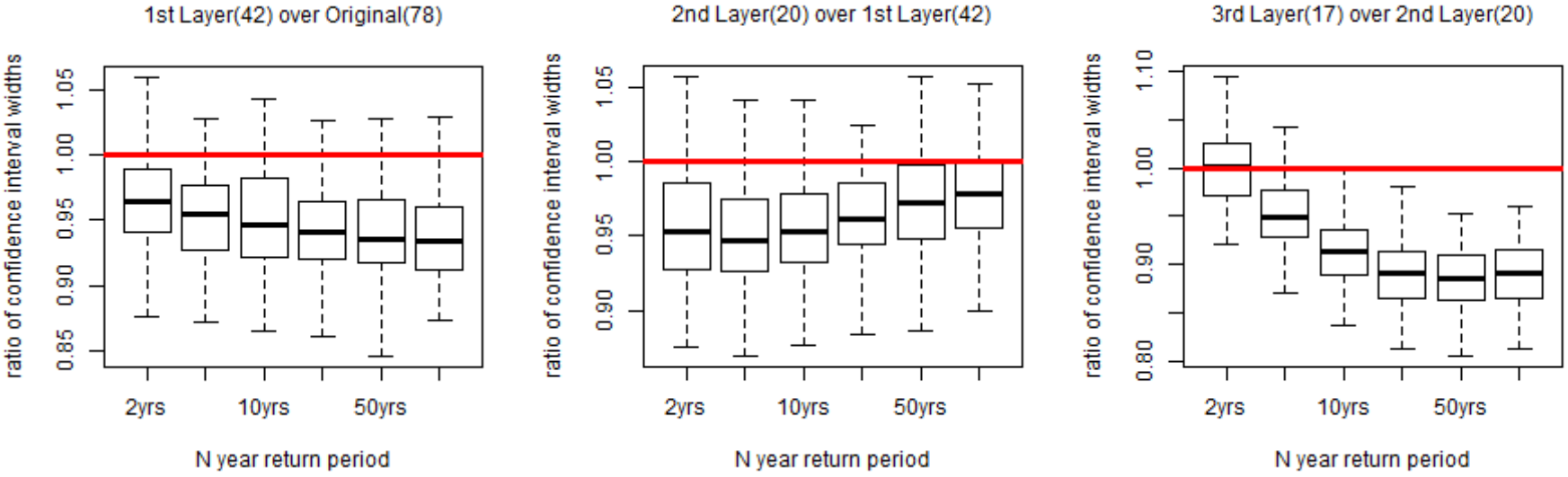
For the rainfall events that have durations shorter than 24h, the three-layer algorithm is simplified to the procedure at the first layer, since information about the date of occurrence for short duration events is not available. The reasons to only apply the first layer of proposed methodology in the HGF are: 1) the temporal step selected to extract feature values is 128-day, nearly a four-month period. Considering the extreme rainfall patterns before 2007, there is a great chance that the date of occurrence of the short and long extreme rainfall events happen in the same four-month period. 2) The first layer is used to gather stations that share similar extreme rainfall condition above the PBL. No matter the length of concerned duration, the pre-conditions for the occurrence of extreme rainfall events especially above the PBL are similar, even when the long- and short- duration events do not happen in the same time frame. Thus, the approach of using the same feature values as the similarity indicators can be justified. 3) To factor the difference in the rainfall amounts among different durations into the HGF for the extreme rainfall events at short durations, the annual maximum rainfall series at short durations will be used to calculate the objective values in the new searching process.

Based on the assumption that the shorter duration events occurred in the same four-month time window as the 24h duration events, this simplified procedure can be conducted. The homogeneous groups for the rainfall events at shorter durations are shown in Figure 2-7, and the following findings can be observed:

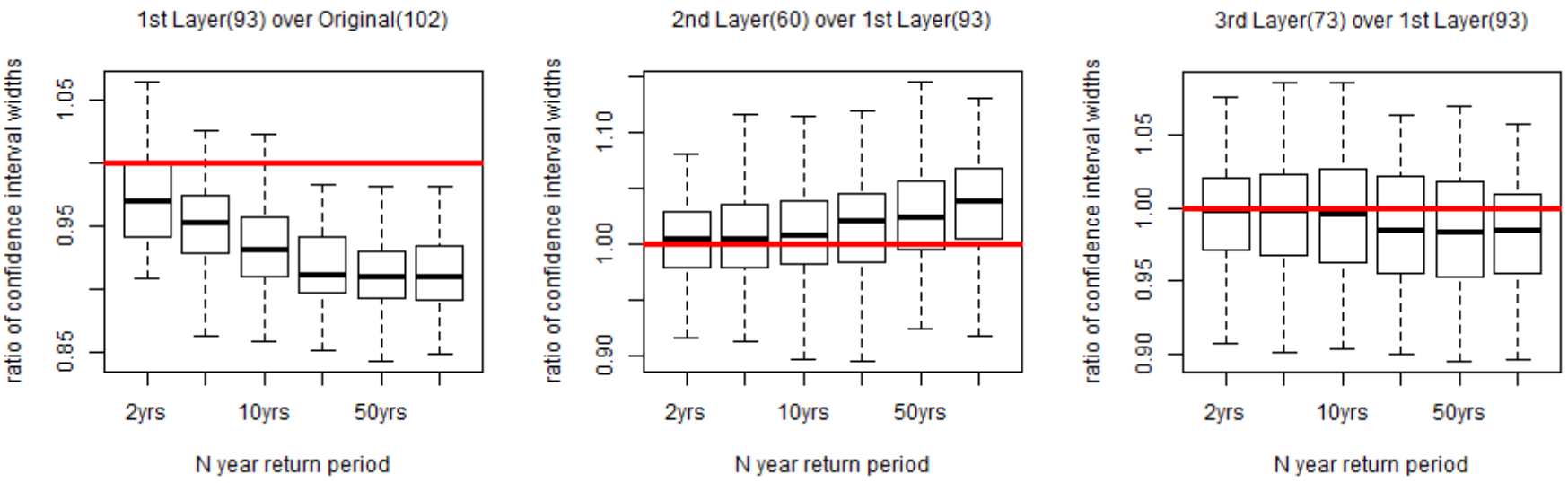
1) Region 1: The homogeneous group for the rainfall events in all of shorter durations is the same and consists of 38 stations in one climate region shown as the red dots (Group 1) in Figure 2-7(a) and 2-7(b). For the 24 hour rainfall series, the homogeneous group obtained from the procedure at the first layer consists of 53 stations, and has the same level of uncertainty in the quantile estimates as that which resulted from this 38-station group. Following the principle of gathering as many stations as possible without increasing the uncertainty in the quantile estimate, the 53- station group is regarded as the better homogeneous group for the quantile estimates of 24h duration rainfall events. Due to the proximity to the ocean, most of extreme rainfall events in this region are caused by stratiform rainfall, which can last for a long duration and increases the possibility that rainfall at all durations are caused by the same rainfall event and thus share a similar homogeneous region.



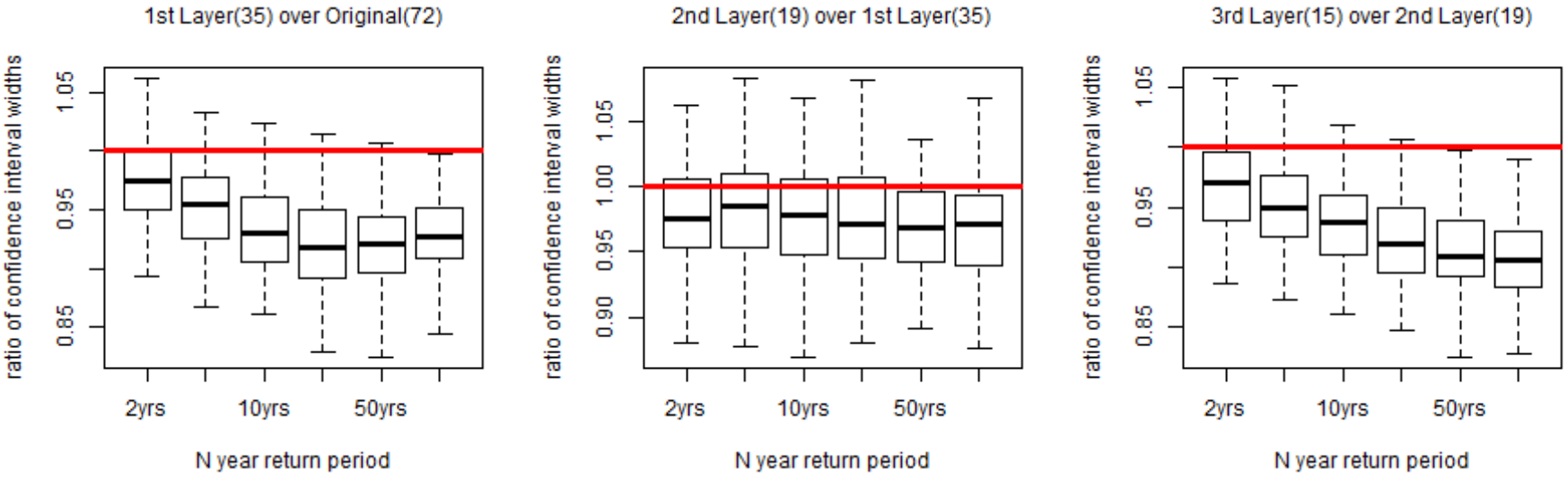
Region1



Region2



Region3



Region4

Figure 2- 6 Boxplots of the ratio of CI widths among different formed groups in Regions 1, 2, 3 and 4 using the AMS at 24-h duration. The boxplots compare the homogeneous group from procedures at higher layer to that from previous lower layer, and numbers in the brackets indicate the number of stations in the formed homogeneous groups. The red line in every boxplot indicates where the value equals to 1.

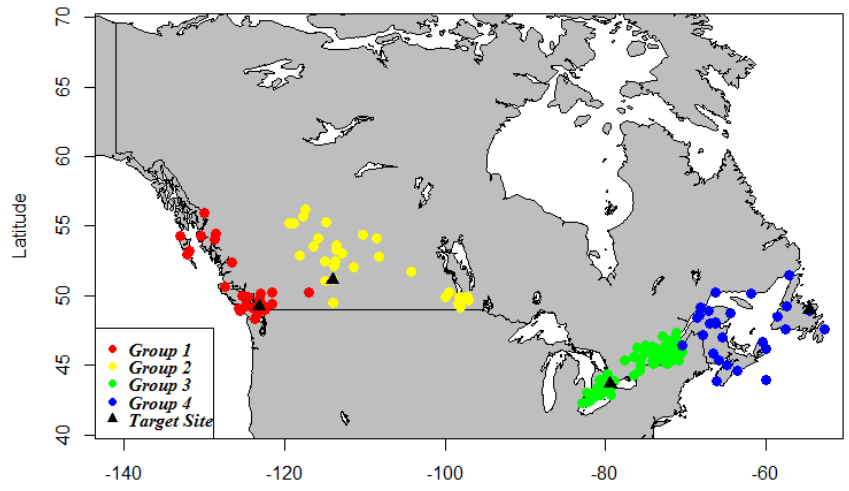
2) Region 2: The homogeneous groups for the rainfall events that have durations shorter than 1h share the same group which consists of 31 stations shown as the yellow dots (Group 2) in Figure 2-7(a). For the rainfall at all longer durations, the homogeneous region consists of 42 stations shown as the yellow dots (Group 2) in Figure 2-7(b). Region 2 is in the Prairies, where convective extreme rainfall events occur during shorter time windows. Extreme rainfall events at longer durations are likely caused by the stratiform rainfall events or the combination of convective and stratiform rainfall events.

3) Region 3: The same homogeneous group with 93 stations (shown as green dots of Group 3 in Figure 2-7(a) and 2-7(b)) for the rainfall events at all durations are identified and are located inside the Southeastern region. The main rainfall type in the Southeastern region is stratiform rainfall that can last for long durations.

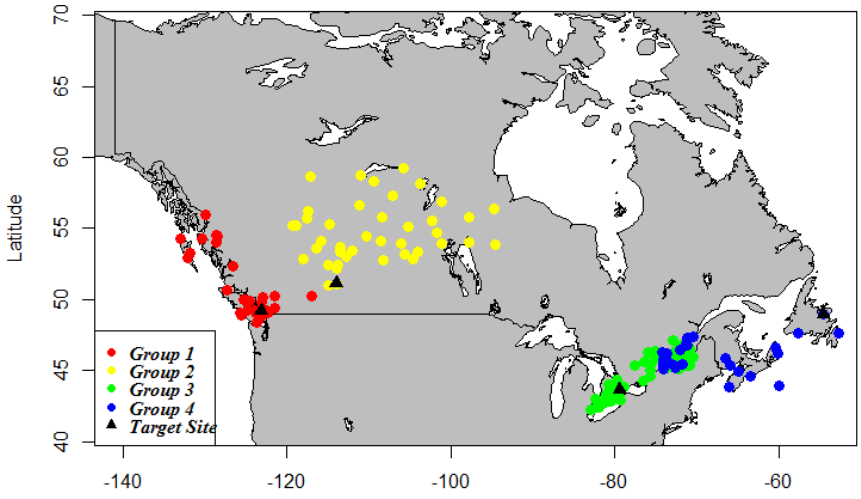
4) Region 4: The homogeneous group allocation in this region is similar to the situation in Region 2. In summary, two homogeneous groups for the rainfall events are identified: 27-station group for the rainfall events with the durations less than 1h shown as the blue dots in Group 4 from Figure 2-7(a), 29-station group for the events with durations between 1h to 12h shown as the blue dots in Group 4 from Figure 2-7(b). For the homogeneous region for the 24h duration rainfall events, the group of 35 stations obtained from the procedure at the first layer has the same level of uncertainty as the 29-station group. This situation can be explained by the same reasons noted above for Region 2.

2.5.4 Comparison between proposed approach and traditional geographic approach

To prove the homogeneity of the formed homogeneous groups in the selected regions, the homogeneity (H) test which can be constructed by using L-CV, L-skew or L-kurtosis was used (Hosking & Wallis, 1997). Details of the homogeneity test can be found in Appendix 1. If one of these three calculated H measures for the formed group is less than 2, the formed groups can be regarded as homogeneous. The HGFs for the AMS at 24-h duration in Region 1, 3 and 4 meet above requirement as their lowest positive H measures equal to 0.16, 1.80 and 1.18, respectively. For the homogeneous group in Region 2, all of the H measures are negative because of the positive correlation among the included sites (Castellarin, et al., 2008).



(a)



(b)

Figure 2- 7 Graphic display of homogeneous groups for extreme rainfall events at duration 5 min, 15 min, and 30 min in (a) and at duration 1h, 2h, 6h and 12h in (b) for Region 1, 2, 3 and 4 The solid dots in red, yellow, green and blue represent the homogeneous groups in four regions. The black triangles represent the target sites in four different regions.

To demonstrate the effectiveness of the proposed methodology, the obtained homogeneous regions are compared with the pooling groups generated from the traditional geographic approach in which the geographic distance from the target stations is used as the similarity indicator in the process (Burn, 2014). To reduce the possible impacts from sample size on the uncertainty in the quantile estimates, the formed groups from the proposed method and the groups generated using the geographic approach have the same number of stations.

The boxplots for this comparison of uncertainty in the quantile estimates for the rainfall series at 24-h duration are shown in Figure 2-8. The median values of the ratios in all boxplots at all return periods are less than one, which indicates that the selected features at all three layers are more effective than the geographic distance as similarity indicators to form the homogeneous regions. In terms of the applicability of the first layer algorithm of proposed methodology for the rainfall events at shorter durations, the same procedure is conducted to obtain the boxplots for comparison, and the same conclusions are reached.

To further validate the proposed methodology, two additional comparisons are conducted using the AMS at 24-h duration: 1) the comparison of CI widths in the quantile estimates from using the at-site approach and the proposed regional methodology. Based on the results shown in Figure 2-9, the formed homogeneous groups at four target sites generate narrower CI widths than that from using at-site approach for the longer return periods, and generally perform equally or inferior for the 2-year return period in terms of the uncertainty in the estimates. 2) The comparison of root mean square error (RMSE) in the growth curves of the HGFs from using proposed method and the traditional geographic approach. In this study, the RMSEs are the average differences between at-site and regional growth curves at different return periods, and can be calculated by using the procedures introduced by Mostofi Zadeh, et al. (2019). Based on the results shown in Table 2-3, the HGFs from using the proposed methodology generally generate lower RMSEs than that from using traditional geographic approach at all return periods.

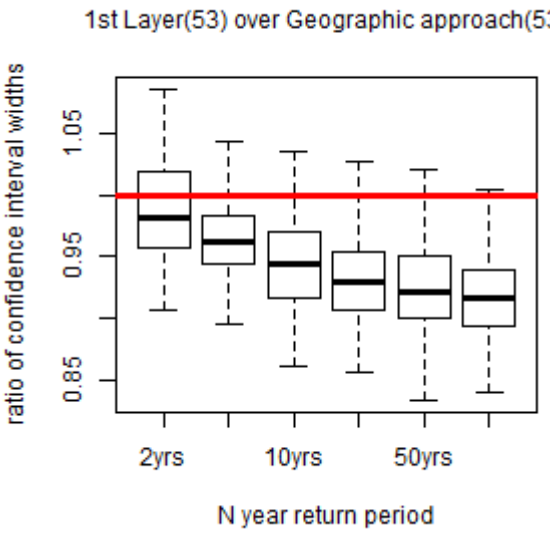
Table 2- 3 Summary of RMSEs from using different HGF approaches with the AMS at 24-h duration

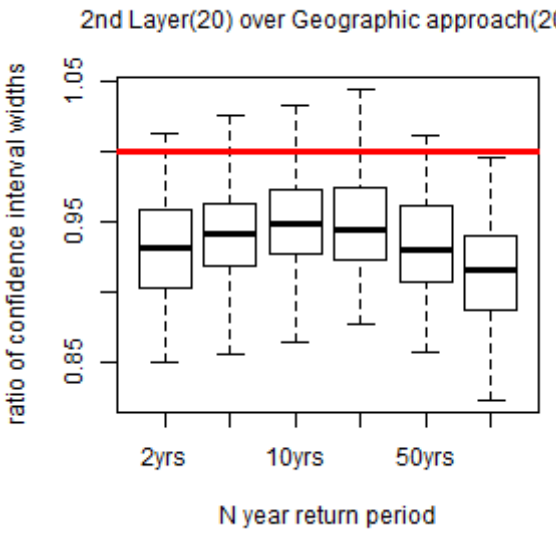
	Return Period						
Testing Regions	Approach	2-year	5-year	10-year	25-year	50-year	100-year
Region 1	Proposed method	0.016	0.012	0.013	0.030	0.047	0.065
	Traditional geographic method	0.014	0.018	0.036	0.063	0.084	0.107
Region 2	Proposed method	0.010	0.006	0.011	0.016	0.019	0.021
	Traditional geographic method	0.017	0.027	0.050	0.078	0.099	0.120
Region 3	Proposed method	0.009	0.020	0.030	0.042	0.051	0.059
	Traditional geographic method	0.016	0.019	0.039	0.067	0.089	0.113
Region 4	Proposed method	0.017	0.019	0.031	0.051	0.067	0.085
	Traditional geographic method	0.011	0.020	0.037	0.055	0.068	0.081

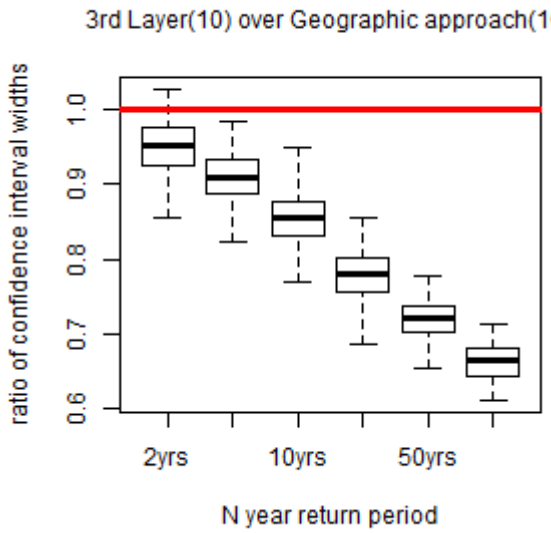
2.6 Conclusions

A HGF process based on the feature selection and weighting algorithm in a three-layer design is proposed to consider the spatial difference of similarity indicators at various regions and reduce the uncertainty for the quantile estimation in regional IDF estimation. Based on the stages of the extreme rainfall formation at different heights, the process extracted both atmospheric and geographic features to consider the possible alteration of rainfall patterns caused by climate change and urbanization at certain locations. During the process, the hybrid searching, which is the combination of Tabu Search, Fuzzy C-Means clustering and Lagrange multiplier is used to select the optimal feature combination and form the homogeneous group that generates the lowest uncertainty in the quantile estimates.

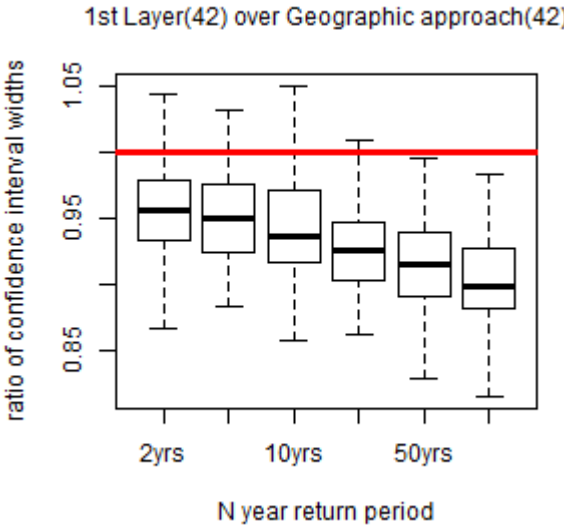
The advantages of the proposed methodology can be summarized into the following points: our method 1) considers the possible changes in the homogeneous region for the rainfall events causing by climate change and urbanization. 2) Considers spatial difference of the similarity indicators at different regions for conducting more effective HGF. 3) Allows the possibility of letting the data itself to drive the group centre and best utilize the most information provided from the available weather stations. The proposed method is applied in four regions across Canada with different rainfall types. The results are evidently superior in comparison with the traditional geographic approach.

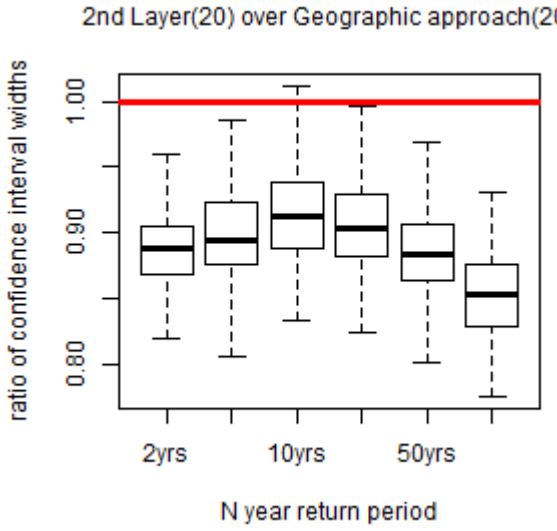


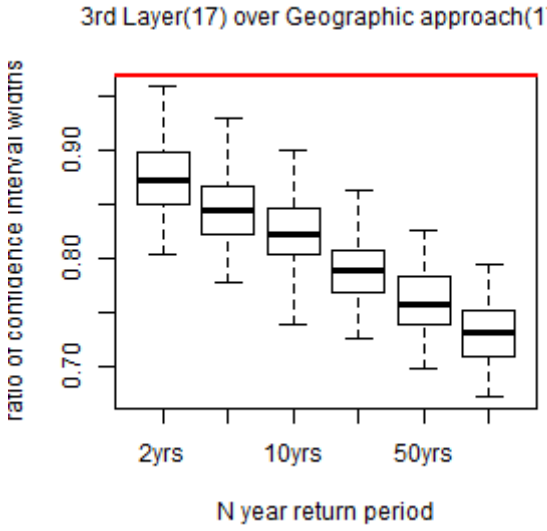




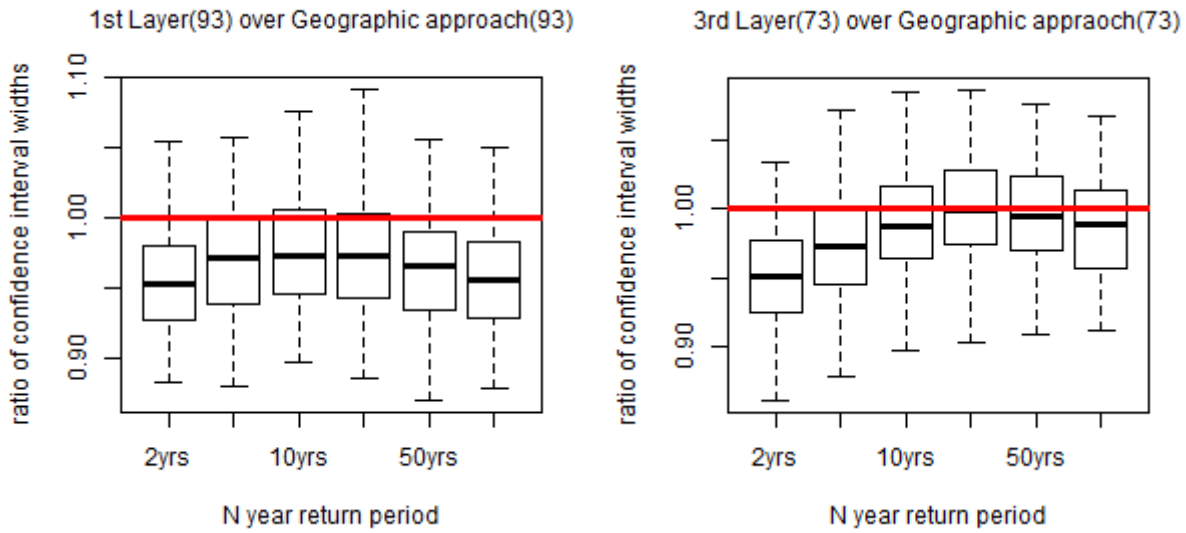
Region1



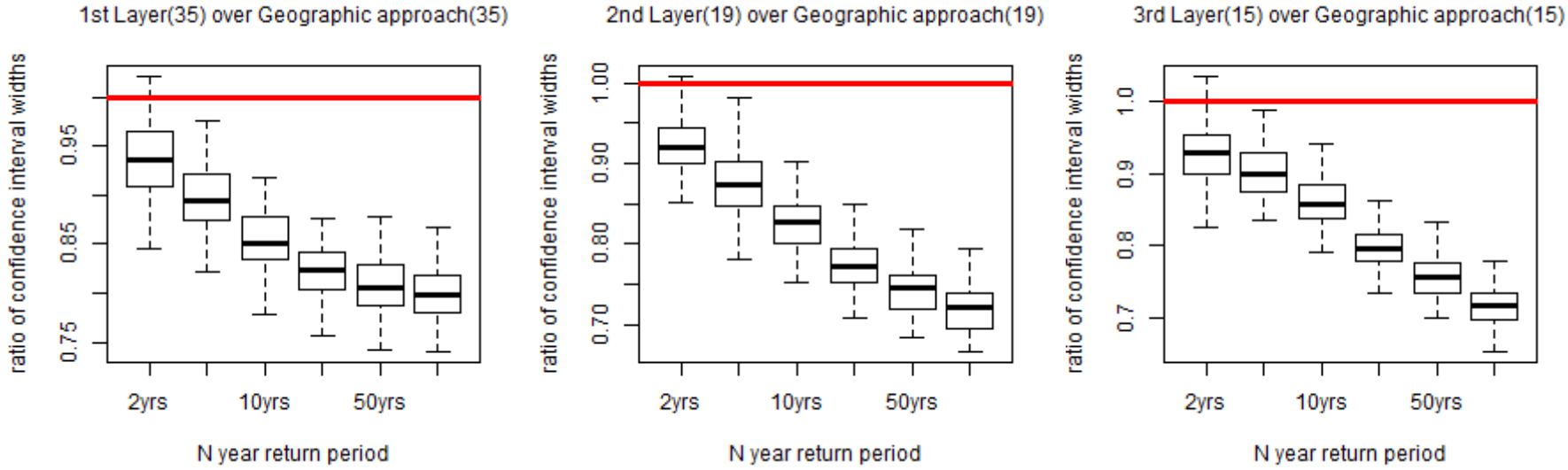




Region2

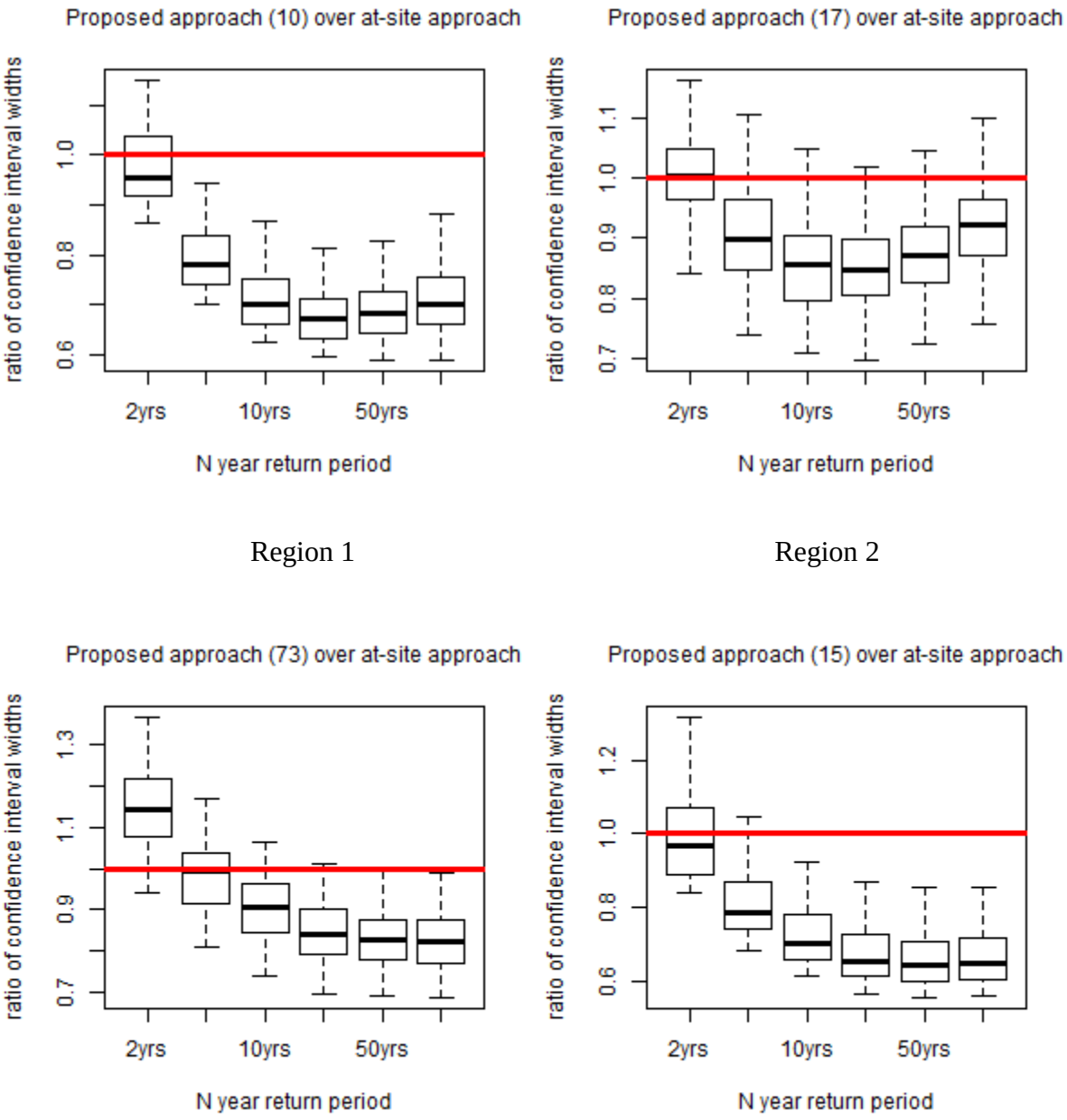


Region3



Region4

Figure 2- 8 Region 3 and 4 Boxplots of the ratio of CI widths between formed groups obtained from using the proposed approach and the geographic approach in Regions 1, 2, 3 and 4 using the AMS at 24-h duration. The boxplots are conducted at all three layers in which the proposed homogeneous groups are compared with the groups that have the same number of stations formed using geographic approach. The numbers in the brackets indicate the number of stations in the formed homogeneous groups. The red line in every boxplot indicates where the value equals to 1.



Region 1 Region 2

Region 3 Region 4

Figure 2- 9 Boxplots of the ratio of CI widths between proposed approach and at-site approach for four target sites in the selected regions using the AMS at 24-h duration. Each target site contains more than 50 data points.

Transition Paragraph A

The three-layer searching algorithm proposed in the previous chapter was successfully applied for homogeneous group formation in the historical stationary environment considering the impacts of climate change. However, climate change may result in spatial alteration of extreme rainfall patterns from changes in the near-surface moisture and increases in the global temperature. Hence, the previously optimal homogeneous groups may no longer be appropriate with respect to the accuracy and uncertainty in regional rainfall quantile estimates. To obtain optimal homogeneous groups for different target sites in the future context, an adjustment procedure, conditioned upon the previous optimal homogeneous groups, is proposed in this chapter. To enhance the adjusted homogeneous groups in the future context, improvements in the previously established three-layer searching algorithm have been made. The improvements include the replacement of the feature weighting method with feature extraction techniques to deal with the impacts of nonlinear correlation amongst the selected similarity indicators on the homogeneous group formation process and the addition of the Region of Influence approach as a method for homogeneous group formation. Under the assumption that the extreme rainfall series remain stationary within future 30-year periods, the previous optimal homogeneous groups are adjusted to the new optimality in a testing loop.

Chapter 3 An improved algorithm for homogeneous group formation for regional frequency analysis of extreme rainfall events under climate change

This chapter is built upon the submitted article with the same title in Global and Planetary Change. Minor differences between the submitted paper and the chapter have been made to facilitate consistency and coherence.

Yang, Z., & Burn, D. H. (2019). An improved algorithm for homogeneous group formation for regional frequency analysis of extreme rainfall events under climate change. Submitted to Global and Planetary Change.

Summary

An improved three-layer searching algorithm is proposed to differentiate the impacts of climate change and urbanization on the extreme rainfall patterns at target sites in regional frequency analysis. Previously, it has been shown that fuzzy c-mean clustering coupled with Lagrange multiplier is capable of forming effective homogeneous groups for the extreme rainfall events while reducing the correlation impacts. While feature weighting is relatively inferior to feature extraction on the effectiveness of extracting non-redundant information from the selected similarity indicators, the clustering approach in the homogeneous group formation is heavily dependent on the original group of input stations. Additionally, the homogeneous groups formed by using the historical observations may experience spatial alterations under climate change impacts. The new version of the three-layer searching algorithm proposed in this paper improves the effectiveness of homogeneous group formation in historical and future contexts in a number of ways. The approach improves the effectiveness of the non-redundant information extraction by using Isometric Feature Mapping, a non-linear dimensional reduction method. The effectiveness of the homogeneous group formation is improved by adding the Region of Influence (ROI) approach as an additional homogeneous group formation technique in addition to fuzzy c mean clustering (FCM). While the groups formed using the FCM approach can be characterized by the extreme rainfall patterns at or close to the group center, the extreme rainfall characteristics at groups formed using the ROI approach can be described by the patterns at pre-set target sites. An adjustment procedure is proposed to adjust the homogeneous group in the historical context for the consideration of the spatial alteration in the future context. The improved group formation process is tested in two regions in Canada with different atmospheric

characteristics. The improved version of the three-layer searching algorithm improves the performance of the searching algorithm in the historical context and fills the gap of providing a homogeneous group formation approach in the future context.

3.1 Introduction

Human activities have long been recognized as the main factors behind global climate change (Mann et al, 2017). With the increasing amount of human activities altering the natural environment around the world, unprecedented extreme weather events caused by climate change are expected in the future (Stott et al., 2010; Trenberth, 2011). Many papers have explored the patterns of extreme rainfall events around the world and detected temporal changes for many areas. Mondal and Mujumdar (2015) used ENSO index, global average surface air temperature and local mean temperature changes as the extreme rainfall indicators over India and detected non-stationarity at different geographic locations, on a grid scale, for the 100-year extreme rainfall event. Simonovic et al. (2017) analyzed the frequency and intensity of extreme rainfall events for nine durations across Canada and found extreme precipitation decreases in the west-central part of the country but increases in the east coast areas. Fernandes and Rodrigues (2017) compared the extreme rainfall events over Brazil between two periods, 1979–1999 and 2000–2015, and concluded that significant alterations had been detected in both the rainfall amount and rainfall mechanisms. Ma et al (2015) analyzed the daily rainfall records for 632 stations from 1969 to 2013 in China and concluded that the frequency of heavy rainfall events was increasing over southeast China while the frequency of moderate and small rainfall events was decreasing. All of above papers have used an at-site approach to quantify the frequency and intensity of the extreme rainfall events at a specific location. Both at-site and regional frequency analysis for the extreme rainfall events are widely applied for the characterization of extreme rainfall patterns. While the former utilizes only the records from target site for the quantile estimation, the latter utilizes data from stations sharing similar rainfall patterns to provide more accurate estimates for the extreme rainfall intensities.

Different levels of climate change impacts on heavy precipitation events have been observed or expected at different temporal periods (IPCC, 2013). Between 1950 and early 21st century (the historical context), low probability of changes have been observed in the frequency of heavy

precipitation events on a global scale (IPCC, 2013). Thus, in the historical context, the frequency analysis of the extreme rainfall events is normally conducted under the assumption that the stations exhibit stationary patterns, in which regional approach is used to improve the rainfall quantile estimates (Asong et al., 2015; Easterling, 1989; Gaál & Kyselý, 2009; Hosking & Wallis, 1997; Yang et al., 2010; Gabriele & Chiaravalloti, 2013; Satyanarayana & Srinivas, 2008).

To improve the regional frequency analysis of the extreme rainfall events in the historical context, many papers have explored methods for homogeneous group (HG) formation. Different similarity indicators have been used as the representation of the extreme rainfall patterns at various geographic locations. In the historical context, the at-site statistics and site characteristics, i.e. the mathematical and the physical representations of the extreme rainfall patterns for a certain period, can be used as similarity indictors (Hosking and Wallis, 1997). Common at-site statistics that can be used in the HG formation are the mean annual precipitation, the mean number of wet days etc. (Easterling, 1989; Hosking & Wallis, 1997; Gaál & Kyselý, 2009; Yang et al., 2010). At-site statistics require long records for calculation and also lack the capability to indicate the changes in the extreme rainfall under climate change. The site characteristics, including geographic or atmospheric features (Hosking and Wallis, 1997; Satyanarayana and Srinivas, 2008; Ahmad et al., 2013; Gabriele and Chiaravalloti, 2013; Burn, 2014; Asong et al., 2015; Haddad et al., 2015), can reveal the potential alterations in the HGs, and are preferable choices of similarity indicators for the HGs formation in the historical context. Based on the structure of the atmospheric layers, Yang and Burn (2019) proposed a three-layer searching algorithm to distinguish the spatial and temporal differences of similarity indicators used for HG formation on the various locations using site characteristics. To further improve the HG formation in the historical stationary context using the three-layer framework, the first goal of this paper is to propose an improved version through using a more efficient feature extraction method and adding an additional HG formation technique.

However, the probability of changes in the frequency of heavy precipitation events is expected to increase to the medium level at the early stage of 21st century, and accelerate to high level at the later 21st first century based on different climate simulations (IPCC, 2013). Thus, in the future

context (the majority of the 21st century), nonstationary patterns for the extreme rainfall events at different weather stations should be considered in regional frequency analysis.

Under the impacts of climate change, extreme rainfall systems can experience both spatial shifting and temporal changes that may result in changes in magnitudes for regional extreme intensities and spatial alterations in the current HGs. Many papers have stated the possibility of spatial changes being detected in the distribution of rain belts caused by global warming (Yang et al., 2015; Putnam and Broecker, 2017). The shifted rain belt can cause the redistribution of moisture available in the atmosphere and thus changes in extreme rainfall patterns. Mbengue and Schneider (2013) have detected shifted storm tracks, which are closely related to the formation of extreme cyclones that can result in extreme rainfall events for coastal regions. By ignoring the potential changes in the formation process of rainfall HGs, biased and inaccurate quantile estimates may be obtained when using a regional approach.

To conduct regional frequency analysis under climate change and provide more accurate estimates for the extreme rainfall intensity, the second goal of this paper is filling the gap of searching for HGs that can be used for the frequency analysis of target sites in the future context. The unavailability of atmospheric features in the future context means that the approach of using site characteristics for the HG formation is impossible, thus the at-site statistics are used instead. Conditioned upon the information from previous HGs, the new HGs are adjusted through identifying the non-compatible stations within the original HGs and the new compatible sites located near to the target stations in terms of similarity distance. To improve the optimality in the adjusted HGs, this procedure is conducted in a testing loop until the extreme rainfall patterns revealed in the final HGs are consistent with the patterns shown in the relevant rainfall cluster. This procedure is also included as a part of the improved version of the three-layer searching algorithm.

By utilizing the improved version of the three-layer searching framework for the HGs in the historical and future contexts, this paper seeks to achieve the following goals:

1) Generate improved HGs in comparison with the original version in the historical context.
2) Generate new HGs to achieve a new level of optimality in the future context.

The rest of this paper is structured as follows: The basic concepts of the original three-layer design are introduced in Section 3.2. The approach used to reduce the feature correlation impact and the additional homogeneous group formation approach in the improved version is introduced in Sections 3.3.1 and 3.3.2. The rationale and details of the adjustment procedures are described in Section 3.3.3. To demonstrate the superiority of the proposed methodology and achieve above two goals, study areas are described in Section 3.4 and application is conducted in Section 3.5.

3.2 Regional frequency analysis using the three-layer searching algorithm

Regional frequency analysis is conducted by using all the observations from the rainfall homogeneous group that is formed by gathering stations that are similar in terms of rainfall. Under the assumption that the stations in the homogeneous group share similar extreme rainfall patterns established in the index-event model, the regional rainfall quantile estimates can be described as (Dalrymple, 1960):

$$\hat{R}_T^P = \overline{R}\hat{x}_T^P \tag{3-1}$$

Where $\hat{R}_T^P$ is the quantile estimate of T-year event for a site, $\overline{R}$ is the index event; $\hat{x}_T^P$ represents the regional growth curve. To form the HGs in a stationary environment, the original version of three layer searching algorithm is applied with the annual maximum series (AMS).

A three-layer searching algorithm, which considers the climate change impacts on the spatial distribution of extreme rainfall, is used to form the optimal homogeneous region for a target station in a historical stationary environment (Yang and Burn, 2019). This framework is established based on the altitudes (heights) for the three possible stages in the formation of extreme rainfall processes: cloud formation, rainfall generation and alteration of rainfall intensity above the surface (Yang & Burn, 2019). As shown in Figure 2-1, each of these stages occurs sequentially at the corresponding layer: the first layer (atmosphere beyond the planetary boundary layer), the second layer (Urban Mixing Layer) and the third layer (Urban Surface Layer) (Collier, 2006; Shepherd, 2005; Yang & Burn, 2019). In the first layer, relevant features such as the Convective Available Potential Energy (CAPE) index and other atmospheric attributes between 300hPa height and 700hPa height can be selected as the potential similarity indicators to distinguish two types of extreme rainfall-induced clouds: cumulonimbus or nimbostratus (Houze, 1989; Houze Jr., 1997), and can be used as indicators to describe the

extreme rainfall patterns being affected by climate change (Gabriele & Chiaravalloti, 2013; Yang & Burn, 2019). The procedure at the first layer results in a macro homogeneous group that delineates the spatial extent of large-scale extreme rainfall events for the target site (Yang & Burn, 2019). The second and third layers are set to further distinguish the potential impacts on the HGs formation caused by atmospheric and geographic attributes at a regional scale and urbanization at a local scale, during which atmospheric attributes at lower heights, geographic attributes plus features associated with urban heat effect, urban canopy effect and urban aerosol effect, are selected to measure the rainfall similarity among the stations in the HG from the previous layer (Collier, 2006; Gabriele & Chiaravalloti, 2013; Li et al., 2013). The final group from the procedures at the third layer, which generates low uncertainty and small Root Mean Square Error (RMSE) in the quantile estimates, is regarded as the optimal HG for the regional frequency analysis at a target site. The potential features for all three layers are summarized in Table 2-1, and the detailed explanations for selecting these features can be found in the original paper (Yang & Burn, 2019).

The HG formation procedures in the original version of three-layer searching algorithm are briefly described as the following (Yang and Burn, 2019):

1) Feature gathering. Based on the structure set by the three-layer framework, the features that can be used as the potential similarity indicators in the HGs at each layer are gathered at different spatial resolutions.
2) Temporal resolution determination. The goal of stage separation in the three-layer framework is the consideration for the temporal difference for the features at different layers. In the original version of the three-layer searching algorithm, the correlations or entropy differences among the selected feature values at different temporal resolutions are used to determine the optimal temporal resolutions for the features at the first and second layer (Quiroz et al., 2011).
3) Feature values at optimal temporal scales. The feature values corresponding to the date of occurrences for the extreme rainfall events for the rainfall stations are extracted at the optimal temporal step, 128- and 64- day resolutions respectively for the first two-layer features. Explanations for selecting these two resolutions can be found in the original paper (Yang and Burn, 2019).

4) HG formation process at first layer. The extracted feature from previous steps are used in the FCM for the HG formation (Bezdek et al, 1984), during which the Lagrange Multiplier is used to improve the clustering results through reducing the correlation impacts (Borgelt, 2008; Xu, et al. 2014).
5) Feature combination searching. Tabu search, a feature selection technique, is applied to search for the optimal feature set to generate low uncertainty in the quantile estimates (Glover, 1986; Glover, 1989; Zhang and Sun 2002). The width of confidence intervals (CI) for the quantile estimates at target stations are used as the objective function in the search.

HG formation in the following layers. Repeat Step (4) and (5) at the second and third layers, during which the HGs resulting from the higher layer are the input group of the stations at the lower layers. The resulting HGs from the last layer are used as the optimal HGs for the target sites. The HGs at all the layers are verified by showing narrower CIs than the group resulting from the previous layer.

3.3 Improved Three-layer Searching Algorithm

This section is used to describe the methodology for the improved version of the three-layer algorithm. In Section 3.3.1, the application of the nonlinear dimensional reduction technique, Isometric Feature Mapping (ISOMAP), in the extraction of non-redundant feature information is introduced. In Section 3.3.2, the reasons and procedures of using both FCM and ROI approaches for the HGs formation in the historical context are outlined. In Section 3.3.3, the adjustment procedure that uses at-site statistics for the HGs formation in the future context is introduced. The verification details and procedure summary are described in Section 3.3.4 and 3.3.5.

3.3.1 Dimensionality reduction

To reduce the cross-correlation impacts on the HG formation process and extract the non-redundant information from possible feature combinations, the feature weighting method, Lagrange multiplier, was used in the original version of the three-layer searching algorithm. In this study, the nonlinear feature extraction technique Isometric Feature Mapping (ISOMAP) is used instead to improve the performance of non-redundant information extraction from the potential feature combinations.

Feature extraction, the method to extract non-redundant information from the input features, can be used to reduce feature correlation impacts on the similar station gathering results. Feature extraction can be divided into linear and nonlinear approaches based on the functions that are used in the process. Principal component analysis (PCA), as one of the most common feature extraction tools in regional rainfall frequency analysis (Baeriswyl and Rebetez, 1997; Chen and Hong, 2012), uses the orthogonal transformation to remove the linear correlation among the input features and creates a lower dimensional dataset by adopting the eigenvectors from its decomposition, and can work extremely efficiently for normally distributed time series (Satyanarayana and Srinivas 2008; Chen and Hong 2012). However, PCA cannot provide accurate feature extractions when the input features are highly non-linearly correlated. The common nonlinear feature extraction techniques, such as ISOMAP and Locally Linear Embedding (LLE), preserve the nonlinear feature structure through using the geometric relations between the target point and its neighborhood. LLE is conducted based on the assumption that each point can be locally represented by the weighted linear combinations of its neighbor points, through which the global nonlinear structure can be preserved by minimizing the reconstruction errors from using these weightings (Lee et al., 2004). The weightings can be calculated through solving the least-squares problem of the error functions (Roweis and Saul, 2000). However, LLE normally requires large input data sets in each dimension to achieve reasonable accuracy, thus is more time-consuming compared to ISOMAP. ISOMAP is proposed to be used as the main tool to remove information redundancy in this paper.

ISOMAP is a graph-based algorithm that retains the intrinsic geometry by re-constructing the relationships between all pairs of points on a weighted graph using their shortest geodesic distances (Tenenbaum et al, 2000). The distance matrix is then double centered before being used in eigenvalue decomposition to obtain the low dimension features (Lee et al., 2004). The procedures for the ISOMAP approach can be summarized into three steps (Tenenbaum et al, 2000):

1) Neighborhood definition. Based on the distance between all pair points on the target manifold, the neighbors for certain data point are selected by either choosing the ones whose distance are below the pre-set threshold or finding the pre-set number of nearest neighbor points. Then weighted graphs G are used as the representations for the relations among these selected neighbors.

2) Shortest path definition. Based on the weighted graphs on the target manifold obtained from the previous step, the shortest distances among all the pairs of points in each neighbourhood are computed, and this shortest distance is used as the geodesic distances to form the distance matrix $D_G = \{d_G(a,b)\}$ between pair of points a,b. To search for the shortest path, Floyd's algorithm or Dijkstra's algorithm is applied.

3) Construct the data on a new dimension. The method of multidimensional scaling is used on the distance matrix obtained from step (2). The intrinsic geometry on the target manifold can be guaranteed through embedding the data in d-dimensional Euclidean space Y (Wu, et al., 2004). The new data points in space Y should be minimized using the following equation:

$$E = \|\tau(D_G) - \tau(D_Y)\|_{L^2} \tag{3-2}$$

where D_Y represents the Euclidean distances matrix $\{d_Y(a,b) = \|y_a - y_b\|\}$; In Equation (3-2), the $\|A\|_{L^2}$ form represents the matrix norm $\sqrt{\sum_{a,b} A_{a,b}^2}$, and τ is the operator that computes inner products of the distances.

The feature correlation can be reduced by applying the ISOMAP procedure to the group of gathered feature values in two-dimension data frame at each year. The non-redundant information extracted from relevant features will be used as the similarity indicators in the HG formation.

3.3.2 Homogeneous group formation

The approaches for conducting the HG formation in regional frequency analysis can be divided into clustering and pooling methods. The former approach, as an unsupervised method, divides the input stations using the selected indicators and can reach the overall optimality among the sub-groups. While the latter method, as a supervised method, is usually preferred for the HG formation when enough stations are presented for the identification of similar stations for a target station.

In the original version of the three-layer searching algorithm, FCM, which has been proven to be an effective clustering approach in regional frequency analysis (Asong et al., 2015; Goyal &

Gupta, 2014;Satyanarayana & Srinivas, 2011), is applied for the HG formation. The objective function that is minimized in the FCM process can be described as (Bezdek et al, 1984; Yang & Burn, 2019):

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m}\left\|x_i - c_j\right\|^2, \quad 1 \le m < \infty \tag{3-3}$$

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left[\frac{\left\|x_i - c_j\right\|}{\left\|x_i - c_k\right\|}\right]^{\frac{2}{m-1}}} \tag{3-4}$$

where N and C are the number of stations and clusters; u_{ij}, which is calculated using Equation (3-4), represents the level of membership for station i in the cluster j; x_i represents each data point in station i; c_j represents the cluster j center; and m, which is set to 2 in this study, is the weighting exponent (Yang & Burn, 2019).

To improve the HG formation process, the region of influence approach (ROI) is added as an alternative pooling approach for the formation in the historical context. The ROI pooling approach, which is introduced by Burn (1990), is conducted under the assumption that multiple stations can experience similar hydro-meteorological patterns. During the HG formation process, the decision of whether particular stations can be involved as part of homogeneous region is determined by measuring their similarity to the target site and can be calculated by using Euclidean distance (Donald H. Burn, 1990):

$$dist_{pq} = \sqrt{\sum_{k=1}^{n}\left(X_{k,p} - X_{k,q}\right)^2} \tag{3-5}$$

where $X_{k,p}$ is the value of k_{th} similarity indicator at station p; and n is the total number of similarity indicators; p represents the target station; q represents one of the stations that could be a member of the pooling group. Stations with higher similarity with the target site, implying lower $dist_{pq}$ values, are first included in the pooling group (Shu & Burn, 2004). The formed groups are evaluated using the heterogeneity measure, which can calculated using the following equation (Hosking and Wallis, 1997; Castellarin et al., 2008):

$$H_k = \frac{V_k - \mu_{V_k}}{\sigma_{V_k}}; \qquad for\, k = 1, 2, 3. \tag{3-6}$$

where:

$$V_1 = \sum_{i=1}^{R} n_i \left(t_{2(i)} - \bar{t}_2 \right)^2 \Big/ \sum_{i=1}^{R} n_i \tag{3-7}$$

$$V_2 = \sum_{i=1}^{R} n_i \left[\left(t_{2(i)} - \bar{t}_2 \right)^2 + \left(t_{3(i)} - \bar{t}_3 \right)^2 \right]^{1/2} \Big/ \sum_{i=1}^{R} n_i \tag{3-8}$$

$$V_2 = \sum_{i=1}^{R} n_i \left[\left(t_{3(i)} - \bar{t}_3 \right)^2 + \left(t_{4(i)} - \bar{t}_4 \right)^2 \right]^{1/2} \Big/ \sum_{i=1}^{R} n_i \tag{3-9}$$

and the regional value of L-Cv, L-Cs, and L-kurtosis is represented by $\bar{t}_2$, $\bar{t}_3$ and $\bar{t}_4$; the L-moments at individual station are represented by $t_{2(i)}, t_{3(i)}$ and $t_{4(i)}$; n_i is the sample size at site i; and R represents the station number in the HG. The HG is considered as acceptably homogeneous when H_k is smaller than 1, and possibly homogeneous when H_k is greater than 1 but smaller than 2 (Hosking and Wallis, 1997).

3.3.3 Adjustment Procedure

The atmospheric conditions for the extreme rainfall event formations are expected to change in the future context and these spatial alterations can be categorized into two aspects: 1) potential spatial alterations in the extreme cloud formation caused by the increases of CAPE values at a large scale. As the warming trends across the tropical and subtropical areas being observed in the ensembles of global climate models (GCMs) simulations, the increases of CAPE extreme values in the corresponding areas are detected (Singh, et al., 2017). These CAPE value increases can affect the extreme cloud formation at the first layer of the proposed framework as it is the most influential feature at that layer. 2) Potential spatial changes in the extreme rainfall intensities is driven by the increased variability of the near-surface moisture in the future context at the regional or local scale based on climate model simulations (Pendergrass, et al., 2017). This variability can cause future heterogeneity in the current HGs. Since extreme rainfall events are the results of extreme weather condition happening at appropriate geographic location, which allows the large buyout of energy ascending into atmosphere or the occlusion between warm fronts and cold fronts, the changes in the extreme weather conditions can disrupt the original

mechanism of extreme rainfall event formation over the target sites and furthermore can cause heterogeneity or spatial alterations in the current HGs.

Although non-stationarity has been assumed for the extreme rainfall events under climate change, the World Meteorological Organization has suggested using 30-year timeframe for the characterization of weather events and the detection of changes in the climate normal under climate change (Charron, 2014). In this study, the HG formations in the future context are conducted in 30-year windows: 2011-2040, 2041-2070 and 2071-2100 separately for the following reasons: 1) Stationarity in the projected rainfall series can be observed within 30-year timeframes under different projected scenarios in Canada, especially for the 24-h rainfall series (Mailhot, et al., 2007; Mailhot, et al., 2012; Requena et al., 2019). 2) The 30-year length of data records allows reliable regional estimates for the fitted parameters for the regional growth curve.

In the future 30-year timeframe, different levels of spatial changes can be observed in the HGs for various stations within the same extreme rainfall system. In places where the impacts from geographic features outweigh those from the atmospheric conditions on the extreme formation process, their HGs tend to remain optimal under climate change. In reverse conditions, changed atmospheric conditions over the current HGs can cause some of the included stations to be non-compatible with patterns described at the target site in the future context. The proposed procedures in this section are used to remove these non-compatible stations and add potential new compatible ones. Compared to the traditional geographic approach for the HG formation that only relies on the geographic proximity to the target sites in the future context, the proposed methodology is conducted upon the information from the optimal HGs in the previous period and provides more reliable formations.

To achieve the goal of non-compatible station removal and compatible station addition, a four-stage procedure is proposed:

1) Stage 0: Identify similarity indicators. In the historical context, the site characteristics including the atmospheric features at the target site are used for the HG formation at the target site. However, this approach is not usable in the future context because the same future site characteristics are not available. Thus at-site statistics are used instead for the similar station gathering in the future context. Traditionally, the at-site statistics are calculated

through using the data records from the target site. This approach can be problematic since limited records of data are available for the future extreme rainfall characterization at the target site. Thus the rainfall statistics calculated from using all the rainfall records from the included stations in the HG for the target site at the previous period, i.e. at-group statistics, are used as the first similarity indicators for the characterization for the extreme rainfall patterns at the target site. Apart from the at-group statistics, the level of overlapping stations among the HGs for different target sites can be used as the second similarity indicator to increase the probability that HGs having close rainfall statistic values share similar rainfall patterns.

2) Stage 1: Identify the new extreme rainfall patterns at the target site in the future period. Based on the description in Stage 0, at-group statistics are used for HG formation at the target site in the future context. However, there is a concern about at-group statistics calculated using the previous HG being an inaccurate representation for future extreme rainfall patterns at target station, which is reasonable considering the existence of non-compatible stations within the previous HG in the future period. To reduce negative impacts from this concern, stations with close at-group statistics are gathered initially, during which fuzzy c mean clustering is used to reach the optimal division of input stations.
3) Stage 2: Station elimination and adding. The overlapping stations between the HG at the target site and the gathered HGs in the relevant cluster are selected as the compatible stations in the adjusted HG for the target site. Although the at-group statistics for similar station gathering may be better choice as a similarity indicator than the at-site statistics because of less uncertainty using larger samples, it may still cause the gathering of the stations having similar rainfall statistics but not sharing similar rainfall patterns because of limited data records. Thus the second similarity indicator, the level of overlapping stations between different HGs, is used. By using the level of overlapping stations between the HG at the target site and the remaining HGs within the same cluster, the non-compatible stations for potential removal are the ones showing least level of overlapping, and the compatible stations for potential addition are the ones showing high levels of overlapping among the remaining HGs.
4) Stage 3: Finalize HG at the target site in the future context. In Stage 2), different levels of overlapping can be identified for station removal and addition. The selection for the final HG

is the one showing the best objective function value with smallest number of station removals and additions.

To further reduce the impacts from the potential inaccurate representation for the future rainfall patterns resulting from using the optimal HGs in the previous period, the procedures from Stage 0 to 3 are used in a testing loop through the following procedures: (1) Identify the HGs that are likely to experience spatial alterations in the future context. (2) All the current HGs in the historical context are used as the initial groups for the calculation of rainfall statistics and extreme rainfall clustering at Stage 1. (3) Using the relevant clusters from step (2), conduct the adjustment procedures for the spatial altered HGs. (4) Using the adjusted HGs from step (3), re-conduct the clustering of extreme rainfall patterns in the future context. (5) Compare the stations in the new clusters from step (4) with the previous ones from step (2). If large differences of station selections are detected, the new adjusted HGs from step (4) would be used as new input HGs in step (2) and repeat step (2) – (5). The loop stops when small or no difference is observed between the new clusters and the previous ones.

3.3.4 Verification

The verification of the HGs are conducted separately in historical and future contexts. In the historical context, the verification process includes the following goals: 1) Goal 1: The superiority of the proposed HGs over their geographic counterparts, which can be obtained through using geographic proximity as the similarity indicator in the HG formation (Donald H. Burn, 2014). 2) Goal 2: The superiority of regional frequency analysis using the proposed HGs over at-site frequency analysis at target site. 3) Goal 3: The necessity of constructing three layers for the optimal feature selections for the HGs formation.

To achieve Goal 1, Test 1 is applied to conduct HG comparison using the Root Mean Square Error (RMSE), which are average differences between the regional growth curve and reliable at-site growth curves at stations with long data records at different return periods (Mostofi Zadeh et al., 2019). During the process, the average at-site growth curves are calculated using stations with no less than 40 years of data records, in the historical context, within the HG. The HG with lower average RMSE values at six return periods (2-year, 5-year, 10-year, 25-year, 50-year and 100-year) can provide more accurate quantile estimates.

To achieve Goal 2, Test 2 which calculates the ratio of CI widths between at-site and regional quantile estimates is used. In this study, the CI widths for the quantile estimates at six return periods for each HG is calculated using the methods introduced by Hosking and Wallis (1997), which involves the process of large sample simulation. In at-site frequency analysis, the procedures introduced by Burn (2014), which involves the repetitive random sample generation using a nonparametric bootstrap and the calculation of CI widths under the assumption that they are asymptotically normally distributed, is used (Stedinger, et al., 1993). This test can only be conducted at the target stations with data records longer than 40 years.

To achieve Goal 3 and identity the main feature influencing the extreme rainfall patterns in historical context, the division of extreme rainfall patterns in the testing region is conducted using the mutual information among the selected optimal feature combinations at different target HGs as similarity indicators in clustering, during which cluster validity indices such as Partition Entropy (PE), Partition Coefficient (PC) and Modified Partition Coefficient (MPC) are used to determine the optimal number of clusters at each layer (Wang and Zhang, 2007). The mutual information content (MIC) can be calculated using the following equations (Shannon, 1948):

$$H = -\frac{\sum_{i=1}^{N} p_i \log_2 p_i}{\log_2 L} \tag{3-10}$$

$$I(X,Y) = H(X) + H(Y) - H(X,Y) \tag{3-11}$$

$$H(X,Y) = \sum_i \sum_j p(x_i, y_j) \log_2 \frac{p(x_i \mid y_j)}{p(x_i)} \tag{3-12}$$

where H is the entropy of the data with a record length equal to L broken into N bins; p_i is the probability of an observation falling into the bin; $p(x_i, y_j)$ is the joint probability; $p(x_i \mid y_j)$ is the conditional probability between X and Y; X and Y are the non-redundant information extracted from one optimal feature combination being applied at these two target HGs. High MIC values indicate high dependence between the tested variables or a large reduction in uncertainty, i.e. high similarity between the feature combinations at two target HGs. Since the selected feature combinations can be used to represent the extreme rainfall patterns at target HGs, high similarity between the extreme rainfall patterns at two target HGs can be assumed.

The procedures for constructing the similarity distance matrix that can be used in the clustering are: 1) select one HG as the target HG, and record its optimal feature combination. 2) The MIC between this target HG and one of the remaining HGs in the test region is measured using the extracted non-redundant information from the recorded optimal feature combination from step 1) at these two HGs in Equation (3-11). 3) Repeat Step 2) at all the remaining stations in the test region. 4) Re-select different target HGs, and repeat Steps 1) to 3). This procedure stops when it has been applied at every station in the testing region.

In the future context, the verification of the HGs is conducted through comparing the proposed HGs with their geographic counterparts using the same Test 1 used in historical context.

3.3.5 Procedures for the improved three-layer searching algorithm

The improved version of three-layer searching algorithm for the formation and adjustment of HGs in historical and future contexts can be carried out through the following procedures in the two stages as shown in Figure 3-1:

Stage 1: HGs formation in the historical context.

1) Initialization. To initialize the searching process, the potential similarity indicators including atmospheric features and geographic features and the historical rainfall records at the included stations from the test region are gathered as input for the searching process.
2) Feature scaling. Based on the correlation and entropy analysis in the original version, the feature values are extracted at 128-day and 64-day resolution for the first and second layer respectively, while the feature values at the third layer are extracted at a daily scale (Yang and Burn, 2019). The date of occurrence for all of the extracted feature values corresponds to that of the extreme rainfall events at 24-h duration (Yang & Burn, 2019).
3) Feature extraction. To extract non-redundant information, ISOMAP is applied at the two-dimensional data frame that consists of the extracted feature values at optimal temporal resolutions for all input stations. During the extraction, the number of nearest neighbours is set to a large percent of the input stations at each layer for the following reasons: a) to guarantee the performance for final HG formation, large number of similar stations should be included after the procedure at each layer; b) to gather a sufficient number of similar rainfall

sites at each layer, the small scale differences among selected features can be smoothed or eliminated (Samko, et al., 2006).

4) HGs formation in the three-layer framework. The non-redundant information extracted from the selected feature combinations are used in both ROI and FCM for the HG formation at each layer, during which the HGs from the higher layer are the input groups for conducting the HG formation in the lower layers.
5) Feature selection. Tabu search, the same feature selection method used in the original version of the three-layer searching algorithm, is applied to find the optimal combination of similarity indicators that can be used in the similar station gathering process. To guarantee a sufficient number of stations in the HGs, the same objective values used in the original version, the average CI widths for the rainfall quantile that can be estimated by using the approach proposed by Hosking and Wallis (1997) , is used to guide the search (Yang & Burn, 2019).
6) Homogeneous group finalization. The HGs from using ROI and FCM are compared in terms of their performance, the one with better performance is selected as the final group.
7) Repeat step 1) to 6) at all the stations included in the test region.
8) Verification for the HGs in the historical context.

Stage 2: HGs adjustment in the future context

9) Identify the HGs that can experience spatial alterations in the next analysis period. Since the determination of optimal HGs in the historical context are selected based on their superiority over the geographic counterparts, these selected HGs retain their superiority in the future context if they experienced no spatial alterations. Thus the spatial altered HGs in the future context are identified by choosing the ones showing higher average RMSE values at six return periods (2-year, 5-year, 10-year, 25-year, 50-year and 100-year) compared to their geographic counterparts.
10) Initial identification for potential stations sharing similar extreme rainfall patterns as the target sites with spatial altered HGs. In this study, the initial HGs used for the similar rainfall station gathering at the next period are the optimal HGs at the current period.
11) Adjustments for the spatial altered HGs. The initial similar rainfall stations identified from step 10) are used in the four-stage procedures within the testing loop proposed in Section 3.3.3 for the compatible rainfall stations selection and non-compatible station removal.

During the process, the new HG, which consists of the compatible stations, is verified through the comparison with the corresponding geographic counterparts in terms of average RMSE values.

12) Verification for the HG adjustment in the future context

3.4 Study Area and dataset

3.4.1 Study area

To test the robustness of the improved three-layer searching algorithm, testing regions with different atmospheric and geographic characteristics are selected. Based on the geographic layout of climate regions and the available weather stations along the coastal area and inland in Canada, two test regions are selected:

1) Region 1 includes the stations in the Pacific Maritime and Cordilleran climate region (the areas in British Columbia and the adjacent province of Alberta), although the proposed procedures are only conducted on the 75 stations in British Columbia (BC). Because of the large geographic and atmospheric differences between British Columbia and Alberta, different extreme rainfall patterns are expected in these two regions. The inclusion of the stations in Alberta in the initial input stations is to test the effectiveness of the proposed methodology being able to distinguish different rainfall patterns.
2) Region 2 locates west-central part of Canada (includes provinces of Alberta, Saskatchewan and Manitoba), and includes the stations in the Prairie and the Boreal climate region.

3.4.2 Datasets

Based on the proposed methodology for the improved three-layer searching algorithm, the datasets are selected for different purposes:

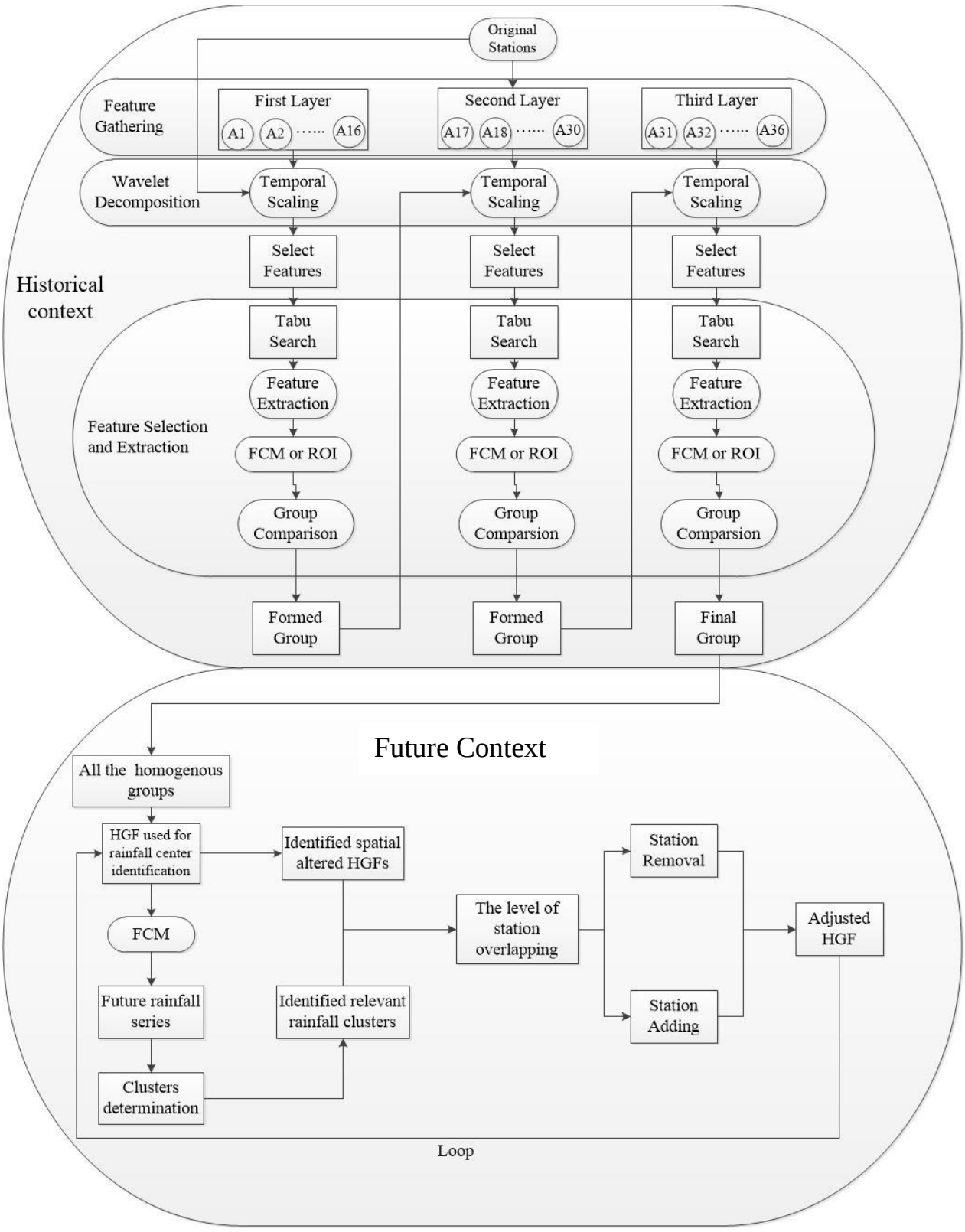


Figure 3- 1 The framework of the improved three-layer searching algorithm.

1) Extract the AMS for different stations in historical context. The same sets of historical AMS used in the original version of three-layer searching algorithm is applied here for the test of the improved version in historical context (Yang & Burn, 2019). Thus, all the AMS at durations 5, 10, 15, and 30 minutes and 1, 2 6, 12, and 24 h, which were retrieved from the historical dataset and the Engineering Climate Datasets from Environment and Climate Change Canada, and their corresponding dates of occurrence at 24-h duration from 1985 to 2004, which were extracted from the historical dataset from Environment and Climate Change Canada website (http://climate.weather.gc.ca/historical data/search_historic_data_e.html/), were applied in this study (Yang & Burn, 2019). The corresponding dates of occurrence at 24-h duration from 1985 to 2004 are used for extracting feature values.
2) Select stations with stationary AMS in historical context. To use the proposed methodology for the HGs formation in historical stationary context, the change-point detection test (Pettitt test) and the trend detection test (Mann–Kendall nonparametric test with the block bootstrap resampling) were applied for the purpose of removing stations with nonstationary trends in the historical context (Donald H. Burn, 2014). Thus, stations which exhibit stationary behavior for the available AMS were selected as inputs in the proposed methodology (Yang & Burn, 2019). Details of this step can be found in the original version of three-layer searching algorithm (Yang & Burn, 2019). The final numbers for stations with stationary rainfall series at each region are: 102 stations in Region 1; 78 stations in Region 2.
3) Extract the historical features values at the optimal temporal scales. To demonstrate the superiority of the improved version of the three-layer searching algorithm, the same sets of potential feature values used for the HG formation in the original version were re-used in this study. Using the NOAA Global Ensemble Forecast System Reforecast (GEFS/R:.esrl.noaa.gov/psd/forecasts/reforecast2/) and ERA-Interim Database (www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim), the feature values between 1985 and 2004 were extracted at 128-, 64- and 1- day temporal resolutions for all three layers. The details of this process can be found in the original approach (Yang & Burn, 2019). The reason for choosing time window 1985 to 2004 for extracting the feature values has been described in the original approach (Yang & Burn, 2019).

Extract the AMS for selected stations in future context. To adjust the HGs in the future context, the daily rainfall series from the NA-CORDEX archive with CanESM2/CanRCM4 model under the Representative Concentration Pathway (RCP) 4.5 scenario at 25-km (0.22°) spatial resolution were extracted (Bias corrected data records were not available at the time this study was conducted). The 24-h AMS rainfall projection series are extracted annually from March to November in British Columbia and from April to November in Region 2 (Requena et al., 2019). In this study, the adjustments are conducted in 2011-2040, 2041-2070 and 2071-2100

3.5 Applications

This section is used for the following tasks: 1) In Section 3.5.1, the improved three-layer searching algorithm was applied for HGs formation at every station in two testing regions in Canada in the historical context, and the results are compared with that from using the original version of the algorithm and a traditional geographic approach. 2) The identification of extreme rainfall clusters within each test region in the historical context are determined in Section 3.5.2 using the HGs from the previous section. 3) The adjustments for the spatially altered HGs in the future context are conducted in Section 3.5.3.

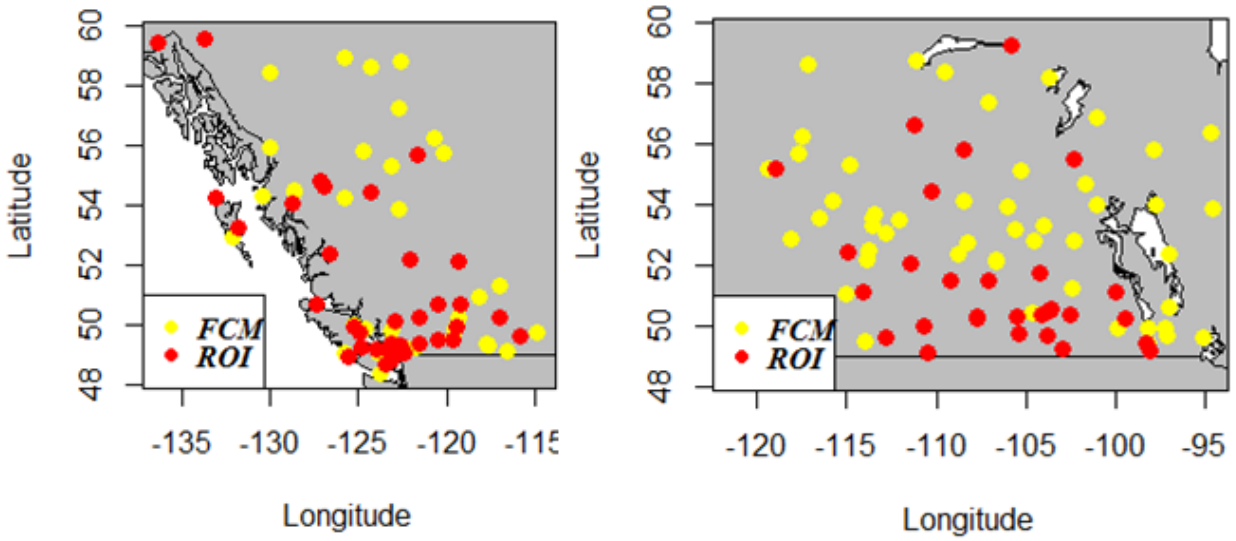
3.5.1. Homogeneous group comparison in the historical context

Based on the procedures described in Section 3.3.5, the improved three-layer searching algorithm is applied at every station in the testing regions. To test the effectiveness of the proposed methodology, Test 1 and 2 described in Section 3.3.4 were applied in the following analysis to verify the improvements in the new three-layer searching algorithm using 24-h AMS, and the following conclusion can be obtained:

1) The proposed procedure is effective for gathering stations sharing similar extreme rainfall patterns. Although stations from different climate regions are gathered as the initial inputs in Region 1, the proposed procedures can remove all the stations from the non-relevant climate region in the final HGs for different target stations.

2) The need for using both pooling and clustering methods as the potential HG formation approach. In the improved version searching algorithm, the ROI approach has been added as the additional HG formation technique to increase the station gathering efficiency. Based on

the procedures in Section 3.3.4, the final HG is the one generating lower RMSE values. As shown in Figure 3-2, FCM and ROI can be equally effective since the final HGs are the mix of groups generated from using FCM and ROI approach: the final HGs for 49 of 75 stations reach lower RMSE values when using ROI approach compared to the FCM approach in BC province from Region 1, while the final HGs for 29 of 78 stations achieve the same conclusion in Region 2.



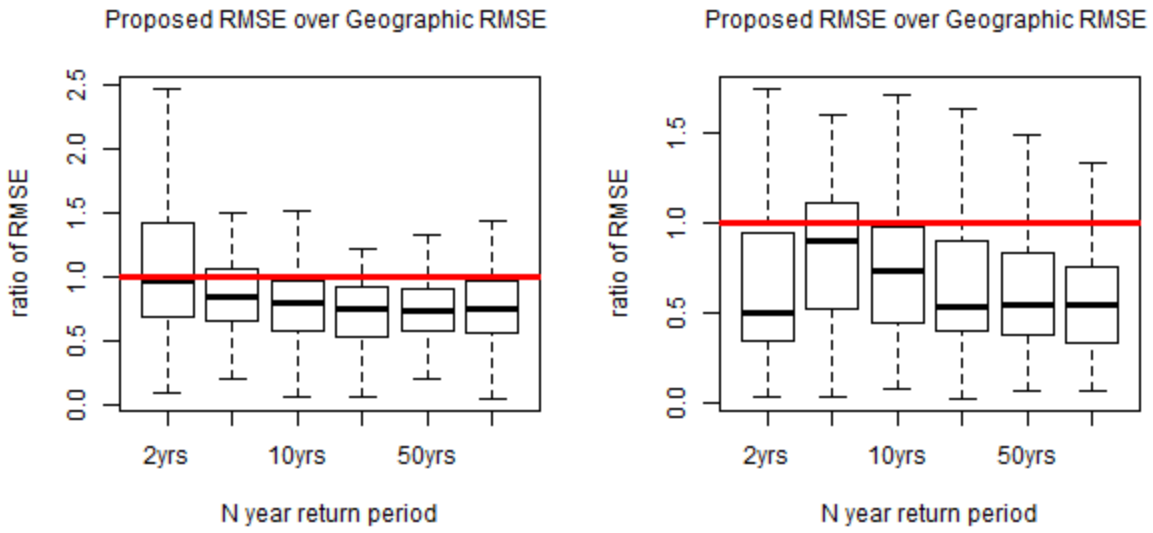
Region 1 Region 2

Figure 3- 2 The selection of the HG formation methods using the improved three-layer searching algorithm at different target sites in Region 1 and 2 in the historical context

3) The new version of three-layer searching algorithm can generate superior HGs for different target stations. The superiority of the proposed methodology is proven through applying Test 1 and 2 in the historical context in three aspects: a) The improved version can generate better HGs than that from using the original version in terms of the RMSE values. The improved algorithm performs better compared to the original version at 63 out of 75 stations in Region 1, and at 63 out of 78 stations in Region 2. b) The improved version performs better than the traditional geographic approach. Figure 3-3 displays the ratios of the RMSE values for all the HGs from using the proposed methodology over the values from using a traditional geographic approach in the historical context. In the proposed methodology, the average RMSE value at six return periods is used as the selection criterion for the final HGs at different target stations. That is the reason for the RMSE ratios at certain return periods for some stations staying above 1 in Figure 3-3. However, the results show lower RMSE values

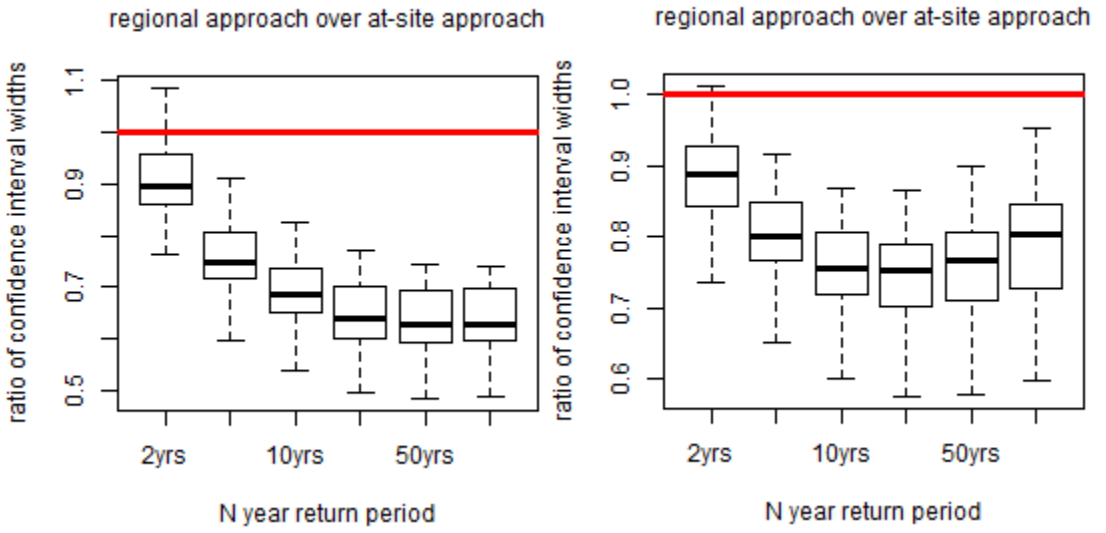
in the HGs at six return periods for the majority of stations can be achieved by using the improved version of the proposed methodology in both testing regions, as the median values of the RMSE ratios maintain below 1 for all return periods. c) The improved version of regional frequency analysis performs better than the at-site approach. To compare the CI widths generated from using the proposed regional frequency analysis and the at-site approach, the information of CI widths at stations with data records longer than 40 years are summarized in Figure 3-4. At each station, the median values for the ratios of the CI widths from using the proposed regional approach over the at-site approach using a large set of bootstrap resamples are calculated at different return periods in the historical context. These median values for all the selected stations are presented as boxplots in Figure 3-4. The results show the median values of the CI width ratios for all the stations at the return period longer than 2-years are below 1, thus the majority of HGs from proposed methodology generate lower uncertainty, i.e. narrower CI widths, at return periods longer than 2-years.

Due to the unavailability of the dates of occurrence information for the AMS at the shorter durations in historical context, the procedures described in Section 3.3.5 cannot be performed through all three layers for the AMS at the other eight durations. Adopting the assumption that the AMS at shorter durations happened in the same four-month period as the ones at 24-h duration, which is used in the original version (Yang and Burn, 2019), the obtained feature combination rankings that are generated by conducting the procedure for the first layer with 24-h AMS in Tabu search, can provide reliable information about the optimal feature combinations that can be used for the optimal HGs with AMS at shorter durations. The HGs through using the selected feature combinations generated lower average RMSE values compared to their geographic counterparts for the extreme rainfall events at shorter durations.



Region 1 Region 2

Figure 3- 3 Boxplots of the ratios of RMSEs among the HGs from the proposed methodology over geographic approach with the 24-h AMS in Region 1 and Region 2 in historical context. The line of value 1 is plotted in red in every boxplot.



Region 1 Region 2

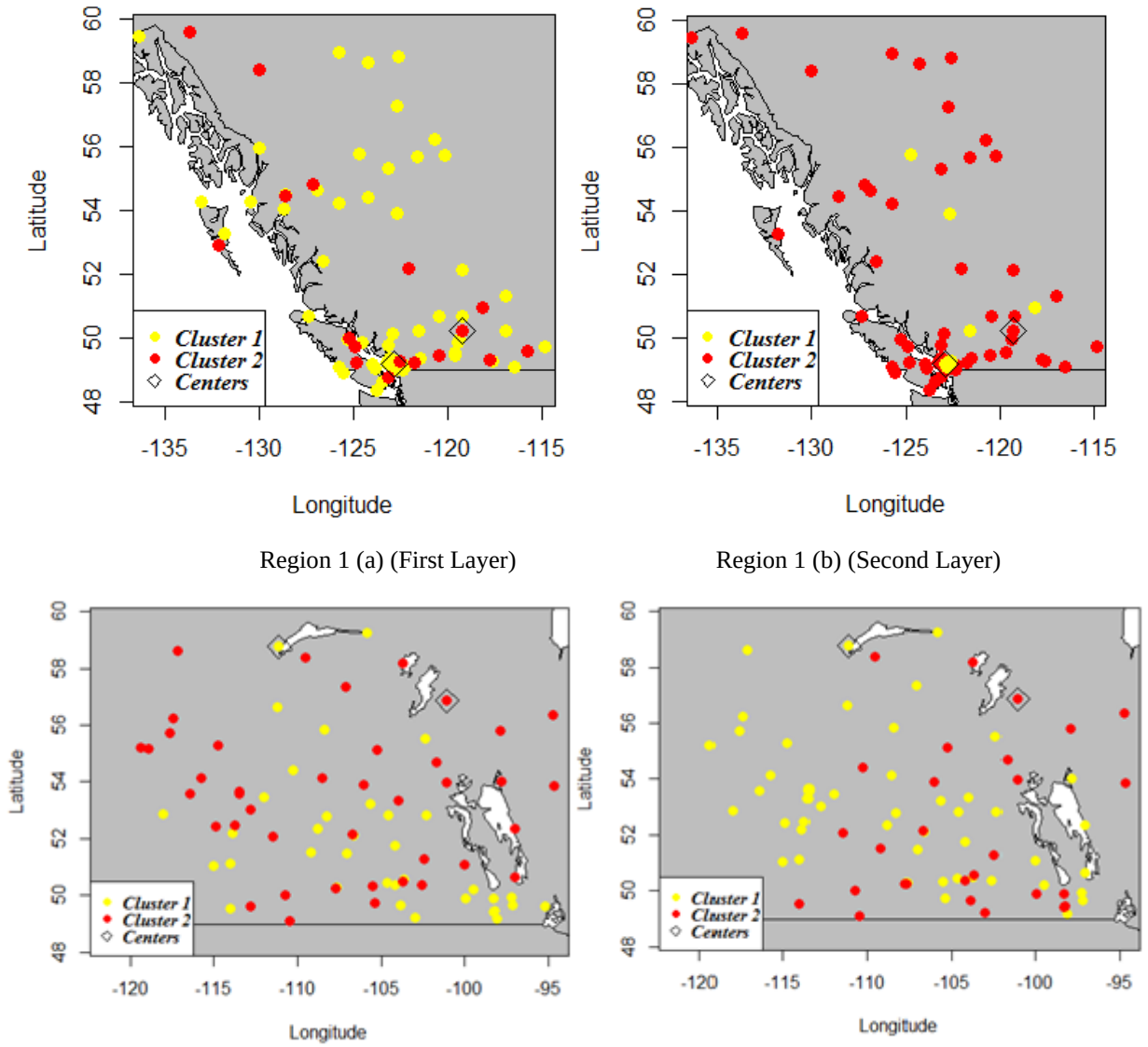
Figure 3- 4 Boxplots for the median values of the CI width ratios between the HGs from the proposed regional approach over the at-site approach at stations with data records longer than 40 years at six return periods in historical context with the 24-h AMS in Region 1 and Region 2. The line of value 1 is plotted in red in every boxplot.

3.5.2. The division of extreme rainfall pattern in the historical context

The division of the extreme rainfall events at 24-h duration are conducted through applying all the optimal feature combinations resulting from the proposed algorithm at the first two layers in the procedures described in Section 3.3.5. In this section, FCM is used as the choice of clustering method, and the division is conducted based on the feature combinations from all the HGs in the testing region. Several points can be observed from the divisions of extreme rainfall patterns at first and second layer in testing regions:

1) The division numbers for the first two layers are set to 2 based on the validity indices. This setting is consistent with the purposes of using the proposed procedures at the first two layers in the design: the intent to separate the cumulonimbus or nimbostratus caused extreme rainfall events at first layer, and the purpose of separating the convective rainfall events from the stratiform rainfall at second layer (Yang & Burn, 2019). The divisions resulting from conducting the procedures at the first and second layers are shown in Figure 3-5. Due to the fact that the HGs at some stations reach the optimality by conducting the proposed procedures only at the first layer, these stations are not included in the division process at the second layer thus the difference in the number of stations showing in the division graphs for the first and second layer.
2) Different stations are contained in the sub-division for the same centres at different layers. The division centres and their corresponding optimal feature combinations in Region 1 and 2 are shown in Table 3-1. Although the division centres at different layers are similar, differences among the constituent stations in each division between two layers are observed. This indicates that different information can be revealed when using the feature combinations at different layers.
3) Information overlapping is observed between the optimal feature combinations for the division centres at the first two layers. In terms of the cloud and rainfall mechanism at the first and second layer, there is a certain level of similarity between the divisions in the testing regions as the geographical proximity of included stations in the same climate regions indicating rainfall similarity at macro and regional scales. However, the urbanization impacts on the extreme rainfall patterns are limited to a local scale, thus completely different dominant feature combinations, or low mutual information, have been observed for the last layer.

4) There are different levels of station dispersions with regard to the divisions at different regions. Compared to the station division in Region 1, the gathered stations in each cluster in Region 2 are more dispersed since convective rainfall events at small spatial scales are common in this climate region.



Region 1 (a) (First Layer)

Region 1 (b) (Second Layer)

Region 2 (a) (First Layer)

Region 2 (b) (Second Layer)

Figure 3- 5 The rainfall pattern divisions through using the mutual information among the optimal feature combination at different HGFs as the similarity indicators at the first layer (Region1 (a) and Region2 (a)) and second layer (Region1 (b) and Region2 (b)).

3.5.3. Homogeneous groups adjustment in the future context

This section is used to adjust the spatially altered HGs to the new optimality in the future context.

Regardless of temporal patterns exhibited at different stations in the future context (from 2011 to 2100), all the selected stations from the historical context are used for the spatial adjustment of HGs at this stage for the following reasons: a) the temporal pattern of the target site is not an effective indicator for the spatial changes in the corresponding HG. Thus, the approach of determining the previous HG experiencing spatial changes based on the trend analysis of the target site is not effective. b) the adjustment procedure is conducted under the assumption that stationary patterns are exhibited at 30-year window in the future context (Mailhot, et al., 2007; Mailhot, et al., 2012; Requena et al., 2019).

In this study, the acceptable HGs in the future context are the ones that generated lower average RMSEs than their geographic counterparts as this criterion is used to determine if the previous optimal HGs experience the spatial alterations in the future context. Thus to search for the future optimal homogeneous groups for the extreme rainfall frequency analysis at a target site, the procedures introduced in Section 3.3.3 are conducted using all the optimal homogeneous groups from the historical context obtained from Section 3.5.1, during which the regional growth curve of different quantiles at upper tail are used as the similarity indicators for the initial pattern clustering in the future context. Then within the relevant cluster for different target sites, the level of overlapping stations is used as the additional indicator to reduce the possibility of non-pattern-similar stations gathering from only using the first indicator, and to conduct the non-compatible station removal and compatible station addition. In the calculation for the average RMSE values, the definition for stations with long data records are the ones having no less than 30 years records in future context, thus the projected rainfall records for all the included stations in the HGs are used for the RMSE value calculation.

However, one problem remains in the application of the adjustment procedure: the identification of true rainfall pattern clusters in the future context, i.e., the groups of stations that share similar rainfall patterns with different target sites. Since the cluster number selection obtained from the validity indices in the FCM results can be questionable and the goal is to form the optimal HGs for different target sites, the adjustment procedure is conducted using a range of cluster numbers and the final adjustments for the selected HG are the ones generating the lowest average RMSE values with a reasonable number of stations (no less than 10).

Based on the adjusted HGs and extreme rainfall clusters at different periods, the following conclusions can be obtained: 1) Different percentages of the stations experience spatial alterations in the HGs in the two regions. The number of stations that experience spatial alterations are 23, 11 and 18 for the different 30-year periods in Region 1, while 27 stations experience spatial alterations for each 30-year period in Region 2. The more volatile changes in Region 2 can be caused by the fact that small scale extreme rainfall events, i.e., convective rainfall, are more common in Region 2 thus more susceptible to alteration under climate change. 2) The adjustment procedure is effective for two reasons: a) the adjusted HGs can provide lower average RMSE values than that from the group using the traditional geographic approach since this rule is incorporated in the selection of adjusted HGs in the adjustment procedure; and b) The adjusted HGs resulting from the adjustment procedure generate lower RMSE values than their original counterparts before the procedure in the two selected regions. Figure 3-6 presents the boxplots of the RMSE ratios from adjusted HGs over the original HGs at six return periods in three consecutive analysis periods and shows that while the median values of the RMSE ratios are below 1 in Region 1 for all return periods, the same median values in Region 2 remain smaller than 1 at most return periods apart from the 5-year return period in the first 30-year period. 3) Several formed HGs in the historical context show consistent behavior across all future analysis periods. The historical HGs for 40 stations in Region 1 and 22 stations in Region 2 maintain optimality during all three future periods. As shown in Figure 3-7, which displays the RMSE ratios from these historical HGs over their geographic counterparts at six return periods in three consecutive analysis periods, most of the median values are below 1 apart from the 5-year return period in Region 2 at the first and third 30-year periods

3.6 Conclusions and Discussion

Considering the impacts from climate change and urbanization, the improved version of the three-layer searching algorithm selects atmospheric features as the optimal similarity indicators for the HG formation based on the previous established three-layer framework. In the original version of three-layer searching algorithm, the feature selection technique, Tabu search, is used to consider the spatial distinctions the rainfall similarity indicators at different target sites, while the establishment of the three-layer framework is used to distinguish temporal differences among the selected atmospheric similarity indicators at different atmospheric heights, during which

wavelet analysis is applied to extract feature values at different temporal resolutions (Yang & Burn, 2019). The new version is able to improve the HG formation process in the historical context and search for the optimal HGs in the future context.

Compared to the original version of the three-layer searching algorithm, new version is able to improve the HG formation process in the historical and future contexts through the following aspects: 1) The feature extraction method, ISOMAP, replaces the feature weighting techniques for the extraction of non-redundant information that is used as the rainfall similarity indicators for the target station as ISOMAP can account for the potential nonlinear correlation among different similarity indicators and thus improve the efficiency of HG formation process. 2) To improve the efficiency in the HG formation process in the historical context, ROI is added as an additional formation approach in this new version. 3) Considering the impacts from climate change and urbanization on the HG formation process, and the possibility of HGs experiencing spatial alternation in the future context, an adjustment procedure is introduced in a testing loop that is conditioned on the optimal HGs from the historical context to form new optimal HGs in the future context. While the HG adjustment of this improved three-layer searching algorithm can perform efficiently in the future context, the HG formation in the historical context is computationally-intensive because of the feature searching process but it is only conducted once for each target station.

The improved version has been tested in four climate regions in Canada that experience different extreme rainfall patterns using one rainfall projection scenario. The proposed methodology generates evidently superior results in comparison with the traditional geographic approach in both historical and future contexts. Future research could explore the following issues: 1) The use of an ensemble of different HGs resulting from using different rainfall projections under the same RCP; 2) The application of the proposed methodology using downscaled rainfall projections at different geographic locations.

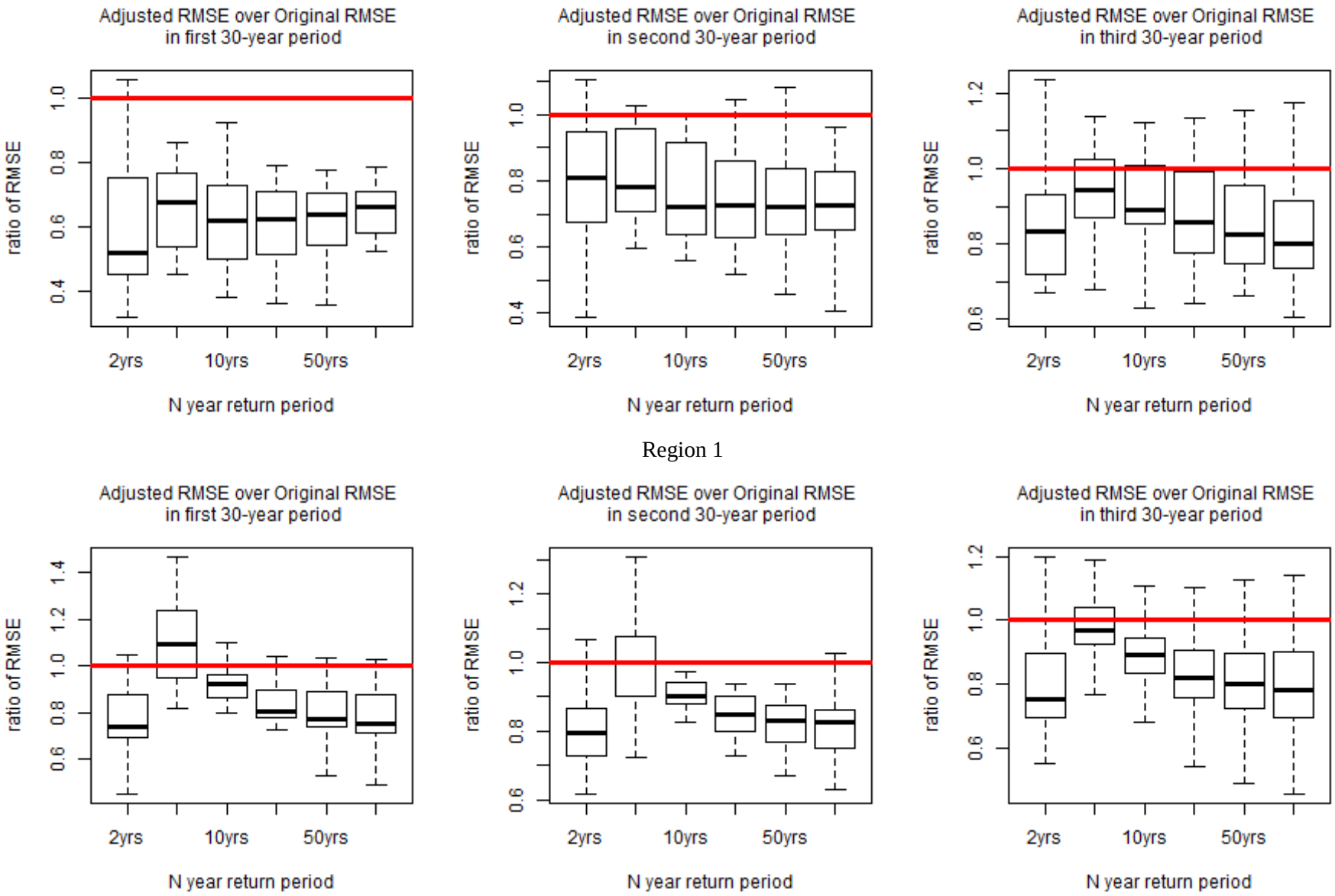


Figure 3- 6 Boxplots of the ratios of RMSE between the adjusted HGFs over the original ones using the future 24-h AMS in Region 1 and Region 2. The line of value 1 is plotted in red in every boxplot.

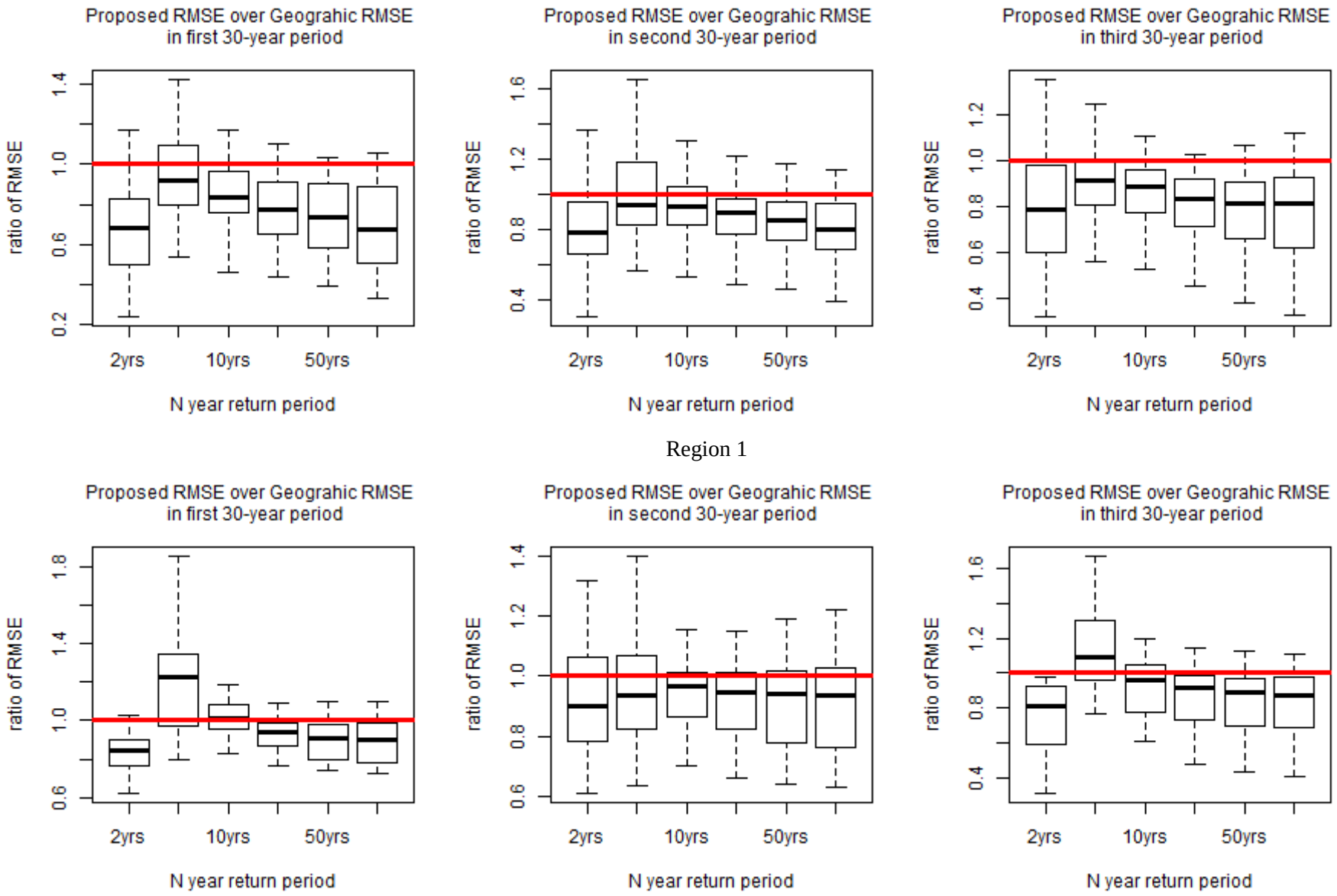


Figure 3- 7 Boxplots of the ratios of RMSE from the consistent HGFs generating from the proposed methodology over the geographic approach using the historical 24-h AMS in Region 1 and Region 2. The line of value 1 is plotted in red in every boxplot.

Table 3- 1 Optimal feature combinations from the division centres at first and second layers using the improved three-layer search algorithm

First Layer		Feature	air300	air500	air700	geo300	geo500	geo700	vw300	vw500	vw700	uw300	uw500	uw700	sphu300	sphu500	sphu700	CAPE
	Region 1	SURREY KWANTLEN PARK	0	1	0	1	0	1	0	0	0	1	1	0	0	0	1	1
		VERNON	0	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0
	Region 2	LYNN LAKE A	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
		FORT CHIPEWYAN A	1	1	1	0	1	0	0	0	0	0	0	1	0	0	1	0
Second Layer		Feature	air850	air925	geo850	geo925	vw850	vw925	uw850	uw925	sphu850	sphu925	VIMF	Elevation	Latitude	Longitude		
	Region 1	SURREY KWANTLEN PARK	0	1	0	0	0	0	0	0	0	1	0	0	0	1		
		VERNON	0	1	0	1	1	0	1	1	1	1	0	0	0	0		
	Region 2	LYNN LAKE A	0	0	0	0	0	0	0	0	0	0	0	0	0	1		
		FORT CHIPEWYAN A	1	1	0	1	0	0	0	0	0	0	0	0	1	1		

Transition Paragraph B

Under the assumption that extreme rainfall series remain stationary within future 30-year periods, the proposed methodology in the previous chapter can characterize extreme rainfall quantiles under climate change using regional frequency analysis in a stationary environment. However, when using a 90-year period as the time frame for the characterization of extreme rainfall events, the extreme rainfall records at certain stations can exhibit non-stationarity. Traditionally, regional frequency analysis of extreme rainfall events in a nonstationary environment is conducted based on the assumption that the parameters in the regional growth curve exhibit linear trends on a yearly basis. In this chapter, an alternative framework for the quantification of extreme rainfall quantiles in a nonstationary environment is proposed, which captures the dynamic changing behavior of extreme rainfall events used in regional frequency analysis. The proposed methodology can be used to described the non-monotonic change behavior of extreme rainfall events especially, for short return periods.

Chapter 4 The unscented Kalman filter-based algorithm for regional frequency analysis of extreme rainfall events in a nonstationary environment

This chapter is built upon a submitted article with the same title in Journal of Hydrology. Minor differences between the submitted paper and the chapter have been made to facilitate consistency and coherence.

Yang, Z. (2019). The application of unscented Kalman filter for regional frequency analysis of extreme rainfall events in a nonstationary environment. Submitted to Journal of Hydrology.

Summary

For regional frequency analysis of extreme rainfall events in a nonstationary environment, the location and scale parameters in the regional growth curves have been described as the weighted combinations of at-site estimates from using the time-covariate functions. However, under different climate change scenarios, the monotonic trend assumption about the parameters in regional growth curves can be questionable. To consider the non-monotonic change behavior of extreme rainfall events in a nonstationary environment, an unscented Kalman filter-based regional approach, is proposed for location, scale and shape parameters estimation using all the stations within a homogeneous group. To consider nonlinear state transition function (STF) and measurement function (MF) used for extreme value analysis, the unscented Kalman filter (UKF) is applied for the filtering of state estimates in the proposed methodology. To avoid UKF performance degradation caused by inaccurate state and measurement noise covariances, a Monte Carlo simulation-based procedure is proposed for the noise estimation. To reduce the uncertainty associated with the distribution fitting process and consider the spatial correlations amongst the stations in a homogeneous group, balancing resamples are used to improve state estimates in regional frequency analysis. This UKF-based approach addresses the non-monotonic change behavior of the extreme rainfall events in a nonstationary environment, and fills the gap of considering spatial correlation for dynamic state estimation in the regional frequency analysis. This proposed methodology was successfully applied to three homogeneous groups in Canada, and goodness-of-fit tests were used to validate the parameter estimates and prove the superiority over traditional covariate approach for providing effective rainfall quantile estimates for the upper tail of the distribution.

4.1 Introduction

Extreme rainfall-induced flood has become a life-threatening natural hazard worldwide (Zong & Chen, 2000; Boyd, 2010; Khazai et al., 2013; Sandink, 2013). In the last two decades, 3,148 flooding events (including rainfall-induced and snowmelt events) affected 1.5 billion people with direct economic losses of $656 billion (CRED and UNISDR, 2018). This is mainly caused by ineffective flood mitigation facilities in both rural and urban areas under climate change. Non-stationarity in the extreme rainfall events has been observed in India, Canada, Brazil, and China etc. (Ma et al 2015; Mondal and Mujumdar, 2015; Simonovic et al., 2016; Fernandes and Rodrigues, 2017). To balance the construction costs of flood mitigation system and the potential damages caused by future extreme rainfall events, accurate estimates of extreme rainfall quantiles are essential for the sustainable planning of the urban drainage system considering engineering, economic, administrative and cultural aspects (Zhou, 2014).

In a nonstationary environment, several methods have been proposed for the quantification of extreme rainfall frequency using at-site or regional frequency approach. at-site approach focuses on the estimation of time-dependent location and scale parameters through maximum likelihood or Bayesian inference (Katz, et al., 2002; Nadarajah, 2005; Cheng, et al., 2014). Using multiple records from stations in a homogeneous group, regional approach considers the time-dependent nonstationarity in two components: the parameters in the regional growth curve function; the index-event value which represents the mean extreme rainfalls at individual stations (Nam, et al., 2015). Assumed that non-stationarity is only detected in the first and second moments of the extreme rainfall statistics, Cunderlik and Burn (2003) estimated location and scale parameters using the de-trending records within a homogeneous group; O'Brien & Burn (2014) propoesed to use the weighted combination of location and scale parameters resulted from applying nontationary at-site approach at each station in a homoegenous group; Hanel, et al. (2009) considered the non-stationarity in both regional growth curve and index-events values through using unscaled records.

Several issues arise from applying above methods in a nonstationary environment: 1) Insufficient data records to quantify the changes in the extreme rainfall events in at-site approach. To reach low uncertainty in the quantile estimates at a return period (T), the record length should exceed

two times the T (Jacob, et al., 1999). The annual maximum records which satisfy this requirement (especially when T equals to 100) are rare. Thus, the regional approach is used in this study to reduce the uncertainty in the estimates through gathering records from several stations. 2) The linear assumption about the changing patterns in growth curve parameters may be impractical. The nonstationary extreme rainfall patterns are caused by changes in the interactions amongst synoptic processes, large-scale circulations and anthropogenic aerosols (Wang & Zhou, 2005; IPCC, 2013; Ohba, et al., 2015), and have been observed in historical records (Burn, et al., 2011). Gao et al. (2016) addressed this issue using Generalized Additive Models for Location, Scale, and Shape (GAMLSS) in at-site approach. Due to the increasing variability of near-surface moisture (Pendergrass, et al., 2017), an approach that considers the non-monotonic patterns of the parameters in regional growth curve is necessary for the regional frequency analysis in a nonstationary environment.

Dynamic state estimation (DSE) is widely used to estimate the non-monotonic change patterns of the extreme values. In at-site approach, Bayesian inference with linear transition functions (Gaetan & Grigoletto, 2004; Johnston & Djurić, 2019), or particle filtering (Toulemonde, et al., 2013) have been applied. However, considering the spatial correlation amongst the stations in a homogeneous group, these methods are difficult to apply in the regional frequency analysis.

To conduct the DSE in regional frequency analysis of extreme rainfall in a nonstationary environment, UKF-based approach is proposed, and aims to address the following issues that cannot be achieved in traditional Kalman filter (Wan & Van Der Merwe, 2000; Hadwin & Peterson, 2017): 1) Issue 1: non-Gaussian distributed state parameters and measurements. 2) Issue 2: the nonlinear STFs and MFs used for the quantification of extreme rainfall quantiles. 3) Issue 3: the unreliable state estimates caused by using insufficient samples at each state. The possibility of using large samples to improve Kalman filter is proven in Ensemble Kalman filter. But it is not applicable in this study as it is designed for the DSE of Gaussian variables with linear STFs and MFs. Particle filtering is another DSE method involves sampling, and is not recommended as enormous number of samples are needed to reduce uncertainty (Hadwin & Peterson, 2017). 4) Issue 4: the performance degradation caused by inaccurate noises used at different states.

To address the non-monotonic change patterns of extreme rainfall in a nonstationary environment, and fill the gap of considering the spatial correlation for DSE in the regional frequency analysis, a one-step forward procedure which is constructed using UKF with resamples is proposed in this paper. The rest of this paper is organized in following sequence: The traditional method for regional frequency analysis of the extreme rainfall events in a nonstationary environment is introduced in Section 4.2. The proposed methodology is described in Section 4.3. Simulations that are used to explore the changes of state noises under different scenarios and prove the effectiveness of the proposed methodology are described in Section 4.4. The application of the proposed methodology is presented in Section 4.5. The paper concludes in section 4.6 with conclusions and a discussion of potential avenues for future works.

4.2 Regional frequency analysis in a nonstationary environment

Regional frequency analysis assumes that rainfall series from stations in a homogeneous group share the similar extreme rainfall distributions, and are used for the characterization of extreme rainfall patterns. Regional approach is advantageous for the reduction of uncertainty in the quantile estimates, especially at long return periods, in comparison with at-site approach. In regional frequency analysis, quantile estimates are calculated using a index-event equation (Dalrymple, 1960):

$$Q_i(T) = \xi_i q(T) \tag{4-1}$$

where $Q_i(T)$ is the extreme rainfall quantile of site i at T-year return period, ξ_i is the index-event of site i, mean of the extreme rainfall records; and $q(T)$ is the regional dimensionless growth curve.

In Canada, the extreme rainfall records, the annual maximum rainfall series (AMS) in this study, are quantified using the Generalized Extreme Value (GEV) distribution which can be described as (Jenkinson, 1955; Alila, 1999):

$$f(y) = GEV(y) = \frac{1}{\sigma} t(y)^{\kappa+1} e^{-t(y)}, \quad t(y) = \begin{cases} \left(1 + \left(\frac{y-\mu}{\sigma}\right)\xi\right)^{-1/\kappa}, & \xi \neq 0 \\ e^{-(y-\mu)/\sigma}, & \xi = 0 \end{cases} \tag{4-2}$$

Where μ, σ and κ are the location, scale, and shape parameters respectively.

In a nonstationary environment, Equation (4-1) is modified to consider the changes in regional growth curves (O'Brien & Burn, 2014):

$$Q_i(T,t) = \xi_i q(T,t) \tag{4-3}$$

Where the parameters in $q(T)$ are weighted combination of the estimates from applying generalized linear models (GLM) at each station using following equation (O'Brien & Burn, 2014):

$$\begin{aligned} \mu(t) &= u_0 + u_1 t \\ \sigma(t) &= \alpha_0 + \alpha_1 t \\ \kappa(t) &= \kappa \end{aligned} \tag{4-4}$$

where $\mu(t)$, $\sigma(t)$ and $\kappa(t)$ are the estimates of the location, scale and shape parameters at time t ; u_0, u_1 and α_0, α_1 are the regression parameters used for the estimation. The parameters estimates in Equation (4-4) can be estimated using maximum likelihood (O'Brien & Burn, 2014). This traditional GLM-based regional parameter estimation method will be used as the comparison approach in application.

4.3 Methodology

To consider the non-monotonic change behavior of the extreme rainfall events and fill the gap of considering the spatial correlation for DSE in regional frequency analysis of extreme rainfall values, a one-step forward procedure is proposed in this section. In Section 4.3.1, the rationale and applicability of dynamic state estimation for the regional frequency analysis are explained, and the solutions for Issue 1 and 2 are introduced. In Section 4.3.2, UKF algorithm is described, and the solution for Issue 3 is proposed. In Section 4.3.3, the Monte Carlo simulation-based methods for state and measurement noise estimation are proposed to provide the solution for Issue 4. The one-step forward procedure is summarized in Section 4.3.4.

4.3.1 Dynamic State Estimation (DSE)

The DSE technique can be conducted through the following steps (Shivakumar & Jain, 2008) :

1) State parameters. In this study, three parameters: location, scale and shape parameters in the regional growth curve, are used to represent the statistical characteristics of the system at target states. To consider the non-stationarity in regional growth curves, additional regression parameters are added as state parameters to reflect their changes between successive states.
2) Identification of STFs and MFs. The STF which is used to describe the transition of state parameters is:

$$x_k = f\left(x_{k-1}\right) + w_k \tag{4-5}$$

where x_k is the state parameter vector at state k, and contains all the state parameters described in Step 1); $f(\cdot)$ represents the transition function; w_k is the state noise vector at state k.

The MF which is used to connect state parameters with measurements is:

$$z_k = h\left(x_k\right) + v_k \tag{4-6}$$

where z_k is the measurement vector at state k; $h\left(\cdot\right)$ represents the function used for measurement calculation using the state parameters; v_k is the measurement noise vector at state k.

3) State Prediction. State parameters at state k can be predicted using STFs (Shivakumar & Jain, 2008) .
4) State Filtering and updating. After the state parameters at state k having been predicted, they can be filtered when measurements at state k are available to be used in the MF. This filtering can be achieved through minimizing the sum of measurement residuals and predicted state vector residuals (Shivakumar & Jain, 2008).

However, because of non-Gaussian and nonlinear problems mentioned in Issue 1 and 2, the DSE algorithm, specifically the Kalman filter, may not generate accurate estimates in the regional frequency analysis of extreme rainfall events. To resolve these problems, the following procedures are proposed: 1) To address Issue 2, the UKF algorithm is adopted for state parameter estimation. UKF algorithm conducts the state propagation using a chosen set of sample points to

consider the nonlinearity in STFs and MF, and reduces the inaccurate state estimates caused by non-gaussian noises(Wan & Van Der Merwe, 2000). The state approximations using sampling points can be accurate to the third order of Taylor series expansion for Gaussian inputs, and second order for non-Gaussian inputs (Van Der Merwe & Wan, 2001). 2) To further improve the accuracy in state estimates and address Issue 1, state parameters are transformed before being used in UKF algorithm. In the GEV distribution, location parameters and log-transformed scale parameter can be consider as Gaussian distributed variables (Park, 2005; Yoon et al., 2010). The shape parameter follows a beta distribution and needs to be transformed into a Gaussian variable through quantile-to-quantile transformation before being used as a state parameter (Park, 2005; Yoon et al., 2010).

The potential STFs for the transformed distribution parameters and regression parameters are linear functions (LLL) in Equation (4-7) or quadratic functions (QQQ) in Equation (4-8):

$$PD_i^{k+1} = PR_{i,1}^{k+1} * PD_i^k + PR_{i,2}^{k+1} + w_k \quad PR_{i,1}^{k+1} = PR_{i,1}^k + w_k \qquad PR_{i,2}^{k+1} = PR_{i,2}^k + w_k \quad i = 1,2,3 \tag{4-7}$$

$$\begin{aligned} PD_i^{k+1} &= PR_{i,1}^{k+1} * PD_i^k + PR_{i,2}^{k+1} * \left(PD_i^k\right)^2 + PR_{i,3}^{k+1} + w_k \\ PR_{i,1}^{k+1} &= PR_{i,1}^k + w_k \qquad PR_{i,2}^{k+1} = PR_{i,2}^k + w_k \qquad PR_{i,3}^{k+1} = PR_{i,3}^k + w_k \quad i = 1,2,3 \end{aligned} \tag{4-8}$$

where PD_1^k, PD_2^k and PD_3^k represent the state estimates for location, log-transformed scale, and Gaussian-transformed shape parameters at state k; $PR_{i,1}^k, i = 1,2,3$, $PR_{i,2}^k, i = 1,2,3$ and $PR_{i,3}^k, i = 1,2,3$ represent the estimates of regression parameters for location, log-transformed scale, and Gaussian-transformed shape parameters at state k.

The quantile function of the GEV distribution in regional frequency analysis is used as the MF:

$$z_k = h\left(\mu_k, \sigma_k, \kappa_k\right) + v_k = \begin{cases} \mu_k + \left(\left(-\ln\left(y\right)\right)^{\left(-\kappa_k\right)} - 1\right) * \sigma_k / \kappa_k + v_k & \kappa_k \neq 0 \\ \mu_k + \left(-\ln\left(y\right)\right) * \sigma_k + v_k & \kappa_k = 0 \end{cases} \tag{4-9}$$

where $\mu_k, \sigma_k, \kappa_k$ are the regional estimates of location, scale and shape parameters at state k and can be obtained using the PD estimates in the state parameter vector; z_k contains the annual maximum observations from all the stations within a homogeneous group at state k. The extreme quantile observations are widely used as measurements in DSE process (Gaetan & Grigoletto,

2004; Johnston & Djurić, 2019). $y = j/(N+1), j = 1,2\cdots N$ is a set of space-even probabilities and used for measurement calculation, and N is the number of measurements at each state. This setting aims to provide a vector of quantile calculations that can be compared with measurements for the filtering.

4.3.2 Unscented Kalman Filter (UKF)

The standard UKF algorithm can be conducted through the following steps (Meng, et al., 2016; Zheng, et al., 2018):

1) Initialization. Set the values for initial state estimates and state error covariance matrix.

$$\hat{x}_0 = E[x_0] \qquad P_0 = E\left[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T\right] \tag{4-10}$$

where $\hat{x}_0$ is the initial state estimate; P_0 is the initial state error covariance matrix.

2) Sigma points selection.

The unscented transformation (UT), a procedure that can be used to estimate the statistics for a random variable experiencing nonlinear transformation (Wan & Van Der Merwe, 2000), is used to select sample points. They form a matrix $\phi_{k-1/k-1}$ consisting of $2L+1$ sigma vectors based on the following equations (Julier & Uhlmann, 1997):

$$\begin{aligned}
&\phi_{0,k-1/k-1} = \hat{x}_{k-1/k-1} \\
&\phi_{i,k-1/k-1} = \hat{x}_{k-1/k-1} + \left(\sqrt{(L+\lambda)\hat{P}^{xx}_{k-1/k-1}}\right)_i \qquad i = 1,\cdots,L \\
&\phi_{i,k-1/k-1} = \hat{x}_{k-1/k-1} - \left(\sqrt{(L+\lambda)\hat{P}^{xx}_{k-1/k-1}}\right)_{i-L} \qquad i = L+1,\cdots,2L
\end{aligned} \tag{4-11}$$

where $\hat{x}_{k-1/k-1}$ is the filtered state vector at state $k-1$; $\hat{P}^{xx}_{k-1/k-1}$ is the filtered state error covariance at state $k-1$; $\lambda = \alpha^2(L+\xi) - L$ is a scaling parameter; ξ is a secondary scaling parameter that is set to 0; α indicates the spread of the sigma points around $\hat{x}_{k-1/k-1}$ and is set to a very small positive value;.

3) Time update

the predictions for the state vector $\bar{x}_{k/k-1}$ and state error covariance $\bar{P}^{xx}_{k/k-1}$ at state k can be conducted (Meng et al., 2016):

$$\bar{x}_{k/k-1} = \sum_{i=0}^{2L} W_i^{(m)} f\left(\phi_{i,k-1/k-1}\right) \tag{4-12}$$

$$\bar{P}^{xx}_{k/k-1} = \sum_{i=0}^{2L} W_i^{(c)} \left(f\left(\phi_{i,k-1/k-1}\right) - \bar{x}_{k/k-1} \right)\left(f\left(\phi_{i,k-1/k-1}\right) - \bar{x}_{k/k-1} \right)^T + Q_k \tag{4-13}$$

where Q_k is the state noise covariance at state k. W_i represents the weights for selected points ($W_0^{(m)} = \lambda/(L+\lambda)$, $W_0^{(c)} = \lambda/(L+\lambda) + (1-\alpha^2+\beta)$, $W_i^{(m)} = W_i^{(c)} = 1/\{2(L+\lambda)\}, i = 1,\cdots,2L$, β is used to incorporate prior knowledge of the distribution for the variable and set to 2 for Gaussian distributions).

Predictions for the selected sigma points can be conducted:

$$\begin{aligned}
&\phi_{0,k/k-1} = \bar{x}_{k/k-1} \\
&\phi_{i,k/k-1} = \bar{x}_{k/k-1} + \left(\sqrt{(L+\lambda)\bar{P}^{xx}_{k/k-1}}\right)_i \qquad i = 1,\cdots,L \\
&\phi_{i,k/k-1} = \bar{x}_{k/k-1} - \left(\sqrt{(L+\lambda)\bar{P}^{xx}_{k/k-1}}\right)_{i-L} \qquad i = L+1,\cdots,2L
\end{aligned} \tag{4-14}$$

4) Measurement update

Predictions for the measurement $\bar{z}_{k/k-1}$, measurement innovation covariance $\bar{P}^{zz}_{k/k-1}$ and cross covariance matrix $\bar{P}^{xz}_{k/k-1}$ can be conducted:

$$\bar{z}_{k/k-1} = \sum_{i=0}^{2L} W_i^{(m)} h\left(\phi_{i,k/k-1}\right) \tag{4-15}$$

$$\bar{P}^{zz}_{k/k-1} = \sum_{i=0}^{2L} W_i^{(c)} \left[h\left(\phi_{i,k/k-1}\right) - \bar{z}_{k/k-1} \right]\left[h\left(\phi_{i,k/k-1}\right) - \bar{z}_{k/k-1} \right]^T + R_k \tag{4-16}$$

$$\bar{P}^{xz}_{k/k-1} = \sum_{i=0}^{2L} W_i^{(c)} \left[f\left(\phi_{i,k-1/k-1}\right) - \bar{x}_{k/k-1} \right]\left[h\left(\phi_{i,k/k-1}\right) - \bar{z}_{k/k-1} \right]^T \tag{4-17}$$

where R_k is the measurement noise covariance.

5) Kalman gain calculation

Kalman gain, K_k, can be calculated using $K_k = \bar{P}^{xz}_{k/k-1}\left(\bar{P}^{zz}_{k/k-1}\right)^{-1}$, the estimates for the filtered values of state vector $\hat{x}_{k/k}$ and state error covariance matrix $\hat{P}^{xx}_{k/k}$ at state k can be obtained:

$$\hat{x}_{k/k} = \bar{x}_{k/k-1} + K_k\left(z_k - \bar{z}_{k/k-1}\right) \tag{4-18}$$

$$\hat{P}^{xx}_{k/k} = \bar{P}^{xx}_{k/k-1} - K_k \bar{P}^{zz}_{k/k-1}\left(K_k\right)^T \tag{4-19}$$

However, high level of uncertainty in the state estimates can be generated from UKF algorithm, especially in the quantification of regional growth curves using limited number of measurements. In at-site approach, ensemble Kalman filter uses a Monte Carlo simulation to improve the state estimates, but is designed for the Gaussian variables propagating through linear transitions (Houtekamer & Mitchell, 1998).

To improve state estimates for regional frequency analysis and consider spatial correlation in DSE, a one-step forward procedure using resamples is proposed in this section. This procedure can be conducted through applying UKF algorithm through the successive states: the previous and target states, using the following steps: a) resamples at two states. 1000 sets of measurement samples, with the same sample size and average cross-correlation value, are generated at each state. b) PDs and PRs are estimated at the target state applying UKF algorithm through previous and target states using two random sets of samples. Thus, obtain one set of the PDs and PRs estimates at target state. c) 1000 sets of PDs estimates can be obtained by repeating step b) with 1000 sets of samples at each of the two states. d) the average PD values can be obtained from 1000 state estimates, and are used to quantify the final regional growth curve at target state. This one-step forward procedure is conducted sequentially at every target state for the state parameter estimation for regional frequency analysis in a nonstationary environment.

4.3.3 Estimation of the noise covariance

To avoid performance degradation in UKF algorithm and provide the solution for Issue 4 (Meng et al., 2016; Zheng et al., 2018), noise estimation procedures are proposed. Several methods have been proposed for the noise covariance estimation: 1) online methods. Noise covariances for state and measurements are estimated using the information at each state and can be further divided into four categories (Bavdekar, et al., 2011): a) correlation methods. Commonly applied in linear dynamic systems, noise covariance can be estimated using least-square method based on recurrent autocorrelations between error covariance and innovation covariance (Mehra, 1970; Bélanger, 1974; Odelson, et al., 2006). b) Bayesian methods. Numerical optimisations or expectation-maximization methods are used for noise covariance estimation in nonlinear systems through maximizing the likelihood function (Dempster, et al., 1977; Shumway & Stoffer, 2000). c) Monte Carlo simulations. Large state samples are generated conditioned upon priori knowledge and can be used to determine state noise covariance (Valappil & Georgakis, 2000). 2) adaptive methods. Based on the state error and innovation covariance, the adjusted noise covariances can be constructed using different methods. Cho and Choi (2006) proposed a sigma-point based receding-horizon Kalman filter (SPRHKF) to resist the deterioration in the state estimates caused by using inaccurate noises. Zheng et al. (2018) presented a covariance correction method to adjust the noise covariances through applying a weighting factor. However, the adaptive approach is not favourable in this study since high uncertainty and bias can be introduced in the noise covariance adjustments from using one limited measurement sample. Thus, online method, specifically the Monte Carlo simulations is applied for the noise covariance estimation.

The noise, which can be referred as random error, is used to describe the random differences between the estimates and the true values at the target state. To resolve the problem that true values are unknown at each state, the noise estimations are conducted with 1000 generated samples.

The state noises are estimated through the following steps:

1) Sample generation for the measurements at the successive states. Based on the original samples at the successive states, one set of measurement samples with the same sample size and average cross-correlation value is generated at each of the two states.

2) PD sample generation at successive states. To generate reliable samples for the distribution parameters and consider the uncertainty associated with using limited number of measurements at each state, a Simulation-Based sample generation approach which utilizes both sampling prior and fitting prior in the generation is applied (Wang and Gelfand, 2002). The sampling prior, which is an informative uniform distribution, can be constructed based on the prior knowledge of the distribution parameters and is added to constrain the samples to a reliable portion of the parameter space (Wang & Gelfand, 2002). The fitting prior, a noninformative vague distribution with large variance, is used to fit the distribution once the samples are generated from the sampling prior (Wang & Gelfand, 2002). In this study, the sampling prior is the confidence intervals for the distribution parameters and constructed using the measurements from Step 1), while the fitting priors for location and log-transformed scale parameters are the normal distribution with mean set to zero and variance equalling to 10^4, and for the shape parameter is the beta distribution with the hyper parameters (u, v) set to (6, 9) (Park, 2005; Yoon et al., 2010). Random-walk Metropolis-Hastings (MS) algorithm is applied to generate distribution parameter samples using these two priors. The transition kernels are normally distributed with means set to zero and variances set to 0.1, 0.1, and 0.04 respectively for location, scale and shape parameter (Yoon et al., 2010). As a result, 1000 sets of distribution parameter samples at each state can be generated using the measurement samples from Step 1), and are transformed into the correct forms to be considered as PD samples. Details of the MS algorithm are in Appendix 2

3) PR calculation between successive states. The PD samples generated in Step 2) are used as the inputs in Equations (4-7) or (4-8) for the estimation of PRs. One set of PR samples can be obtained from this step

4) Noise estimation for the PDs. Based on the regression analysis in Step 3), mean squares of the differences between the PD estimates from the regression analysis and the PD samples at the target states are used as one random set of state noises for the PDs.

5) Noise estimation for the PRs. For each set of PDs generated in the successive states from Step 2), one set of local PRs can be generated to reach the equilibrium state in Equation (4-9) while staying close to the true PR values obtained from regression analysis in Step 3). The

mean squares of differences between all the local optimal PRs for each set of PDs and true PRs for all the PDs are regarded as the state noises for the PRs

6) Finalization for the PD and PR noise estimates. Repeat steps 1) to 5) 100 times, the average values of these 100 sets of noise estimates are used as the diagonal values in the state noise covariance matrix (Nilsson, 2006).

Compared to state noise covariance, measurement noise covariance has less impact on the performance of UKF algorithm (Meng et al., 2016). Measurement noises are the random errors between the measurement estimates and their true and unbiased values, and can be estimated using the input measurements (Valappil & Georgakis, 2000). Measurement noise can be affected by the sample size and average cross-correlation value amongst the stations in the homogeneous group. Large sets of random samples are used to estimate the measurement noise through following steps: 1) Initialization. Using the measurements at target state, the fitted GEV distribution and average cross-correlation values can be determined. 2) Sample generation. 1000 sets of random samples with the same sample size and average cross-correlation value can be generated using the parametric sampling based on the fitted GEV distribution from step 1). Within each set, the generated samples are arranged in increasing order. 3) Designed quantile calculation. Using the fitted GEV distribution in Equation (4-9), a set of quantiles at design probabilities can be obtained and are arranged in increasing order. 4) Measurement noise calculation. Mean squared value of measurement differences can be obtained using one sample sets from step 2) and the empirical quantiles from step 3). 5) Repeat step 4) for all 1000 samples from step 2), and the final average value for all the mean squared differences can be used as the diagonal values in the measurement noise covariance (Bryson, et al., 2008; Akhlaghi, et al., 2018; Zheng et al., 2018). The fitted GEV distribution in Step 1) may be unreliable as samples used in the distribution fitting may be insufficient, and this has been reflected in the process.

4.3.4 UKF application in regional frequency analysis

The one-step forward procedure for regional frequency analysis of extreme rainfall events in a nonstationary environment can be summarized to the following steps as shown in Figure 4-1:

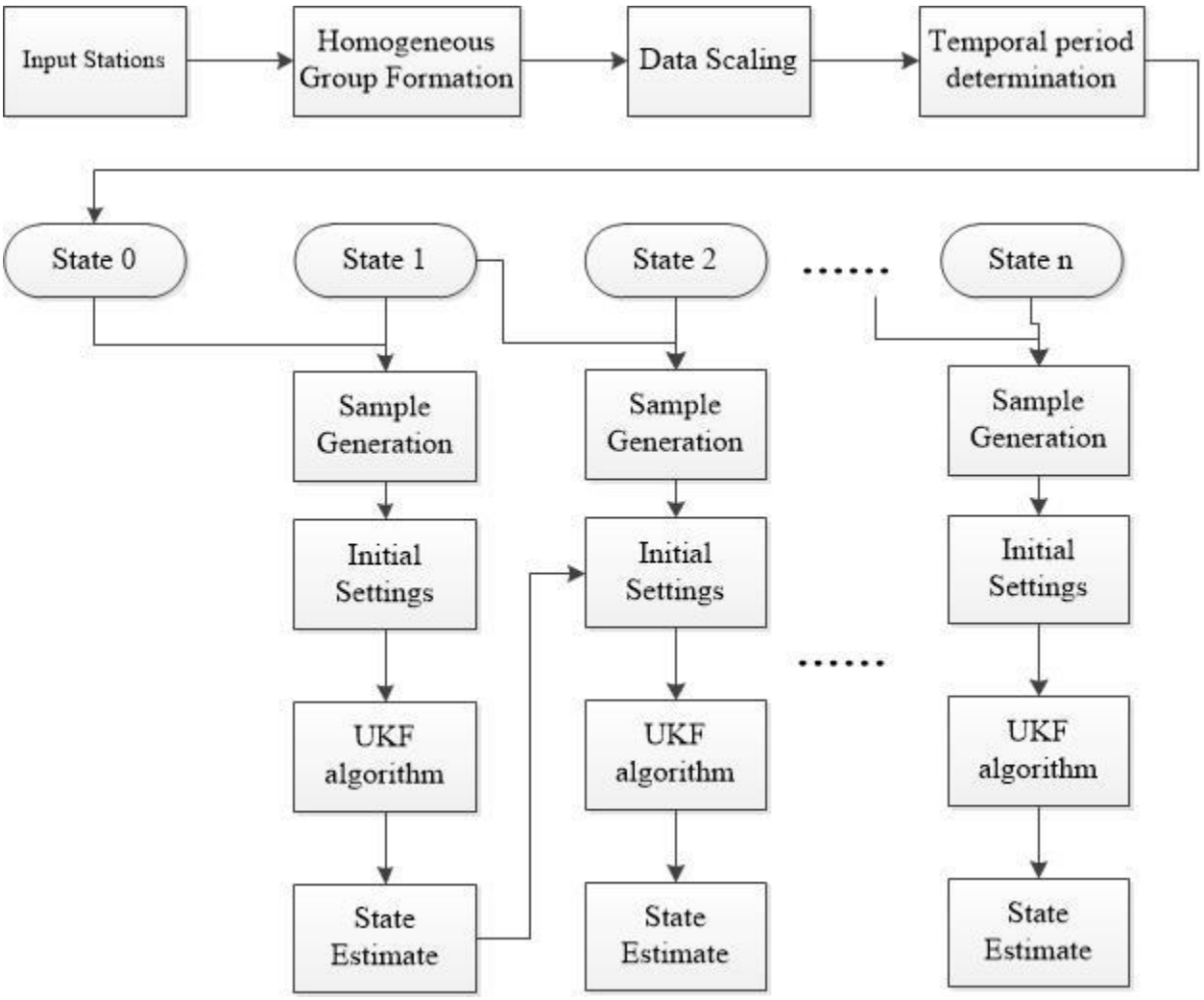


Figure 4- 1 The framework of the one-step forward procedure application for the regional frequency analysis in a nonstationary environment.

1) Homogeneous group formation. Trend centered pooling approach proposed by O'Brien and Burn (2014) is used to identify stations sharing similar rainfall patterns with the target site in a nonstationary environment. After applying the Mann–Kendall trend test at stations within the similar climate regions, select the stations showing similar trend (i.e. stationary, positive and negative trends) with the target site. Then, Region of Influence approach can be applied to further select the stations in the homogeneous group for the target site (Burn, 1990).

2) Observations processing. extreme rainfall series from the homogeneous group from Step 1) are scaled by their individual index-event values.

3) Temporal step selection. Since the proposed methodology is conducted under the assumption that the estimated reginal growth curve can be used to represent the extreme rainfall patterns

at each state (each temporal period), the number of rainfall observations within the selected temporal period should be enough to generate reliable rainfall statistics. In the traditional regional approach of frequency analysis, 5T station years should be included in the homogeneous group to provide reliable quantile estimates for the T-year return period (Robson & Reed, 1999). This rule is applied here for the determination of temporal steps for different return-periods.

4) The construction of State Zero. State Zero is used for the state estimates at State One. In this study, the measurements at State Zero are the replicates of the measurements at State One.

5) Sample generation. To generate reliable measurement samples that can be applied in the one-step forward procedure, a balanced resampling approach is used (Burn, 2003). Compared to the traditional resampling with replacements, this approach achieves a balance in the resamples by ensuring that each data point appears the same number of times through following steps: a) concatenate 999 copies of the original sample into an extended sample. b) randomly rearrange the original order in the extended sample. c) the final samples can be obtained through dividing the rearranged extended sample into 999 copies. Each one of the final 999 resamples plus the original sample (i.e. 1000 sets of samples in summary) will be used as the measurements for the estimation at each state. Furthermore, to preserve the spatial correlations among the stations in a homogeneous group, a vector resampling approach, in which balanced resampling procedure is conducted on the given years of the included stations, is implemented (Burn, 2003).
6) Application of the one-step forward procedure at State One. Randomly select the samples generated from step 5) at State Zero and One, the one-step forward procedure is applied for the PD estimation at State One. The noise covariances for all state parameters and measurements are estimated using the procedures described in Section 4.3.3. While the initial values for the PDs are the regional estimates through applying the method of L-moments in the original sample at State Zero, the initial values are set to 1 for the $PR_{i,1}$ and zero for the rest of PRs.

7) Application of the one-step forward procedure at latter states. With the samples generated from Step 5), the same one-step forward procedure with 1000 generated samples at each state

is applied for the PD estimates at latter states. The average state estimates, and the average error covariance for the state parameters and measurements obtained from previous states are used in the initial settings for the application of one-step forward procedure in the following states.

8) Uncertainty quantification for the estimated rainfall quantiles. 1000 rainfall quantiles at different return periods can be calculated from using 1000 samples of state estimates in the application of one-step forward procedure at Step 6) and 7). The 95% confidence intervals (CIs) at the target site are calculated based on these quantiles.

9) Verification and comparison. To verify the estimates of the distribution parameter, the goodness-of-fit test, specifically the Anderson-Darling (AD) Test, is used with the original samples at each temporal step. The parameter estimates are considered valid if the p-value from the AD test is larger than 0.05. To compare the parameter estimates from the proposed methodology and the traditional GLM-based method in the regional context, the AD test at the upper tail (ADup), which is conducted through assigning higher weights to the data at upper tail of the distribution, is used to compare their goodness-of-fit (Chernobai, et al., 2015). For details of the ADup test, see Appendix 3.

4.4 Simulations

This section is conducted for two purposes: 1) The exploration of state noises under different scenarios between successive states. In Section 4.4.1, different change scenarios for the distribution parameters between the successive states were constructed for this exploration. 2) The selection of optimal forms for STFs. In Section 4.4.2, the selection and testing for the ideal STF forms that could be used in the proposed methodology were conducted.

4.4.1. State Noise Estimation

Since measurement noises are affected by the function used in MF, the measurement sample size and average cross-correlation amongst the measurements from included stations, they can be considered as the same at different states when the three factors are fixed.

Th following simulations are used to explore the changes in state noises under different scenarios. In regional frequency analysis, four factors may affect the state noises in the proposed

methodology: (1) sample size of the extreme rainfall measurements from a homogeneous group, (2) magnitudes of changes for the distribution parameters between successive states, (3) average cross-correlation values amongst the stations included in the homogeneous group, and (4) the forms of STFs used in the UKF algorithm. Among these four factors, sample size has been a long-established factor affecting the uncertainty in parameter estimates (Adcock, 1997; Wang & Gelfand, 2002). This section explores the impacts from Factors (2), (3) and (4) through the following scenarios, in which the values for the three factors are set to:

1) Factor 2: In each scenario, two of the three parameters used in the regional growth curve remain the same at all states, while different magnitudes of changes for the other parameter are incorporated between successive states. The parameter values for all three scenarios are listed in Table 4-1 and the second state in the two successive states is regarded as the target state in the estimation.
2) Factor 3: For each scenario listed in Table 4-1, the state noises for the PDs and PRs are estimated using different average cross-correlation ranging from 0 to 0.6 with 0.1 being as the increment.
3) Factor 4: The two types of STFs described in Equations (4-7) and (4-8) are tested.

The details of the simulated group at each scenario are as follows: 1) Index-event values. The index events are set to 1 at all stations. 2) Station number and sample size. There are 20 stations, each with 20 data points at a state. 3) Sample generation. If the average cross-correlation value is zero, the 20 data points in each station are random values from a pre-set GEV distribution at each state and then multiplied by their index-events. For non-zero average cross-correlation values, samples at each state are generated through following steps: a) Normally-distributed sample matrix generation. Generate a group of data points using standard normal distribution, and divide them into a 20 by 20 matrix. b) Correlated normal-distributed samples. The sample matrix from step a) is multiplied by the Cholesky decomposition of a diagonal matrix with the correlation value being as the main diagonal entries. c) The cross-correlated sample that fits a GEV distribution. Transform correlated normally-distributed sample to the target GEV distribution.

To analyse the state noises generating from using the proposed methodology in Section 4.3.3 in above simulations, they are summarized and shown as Situation (a) and (b) in Figure 4-2. Situation (a): compare state noises with respect to the differences of distribution parameters

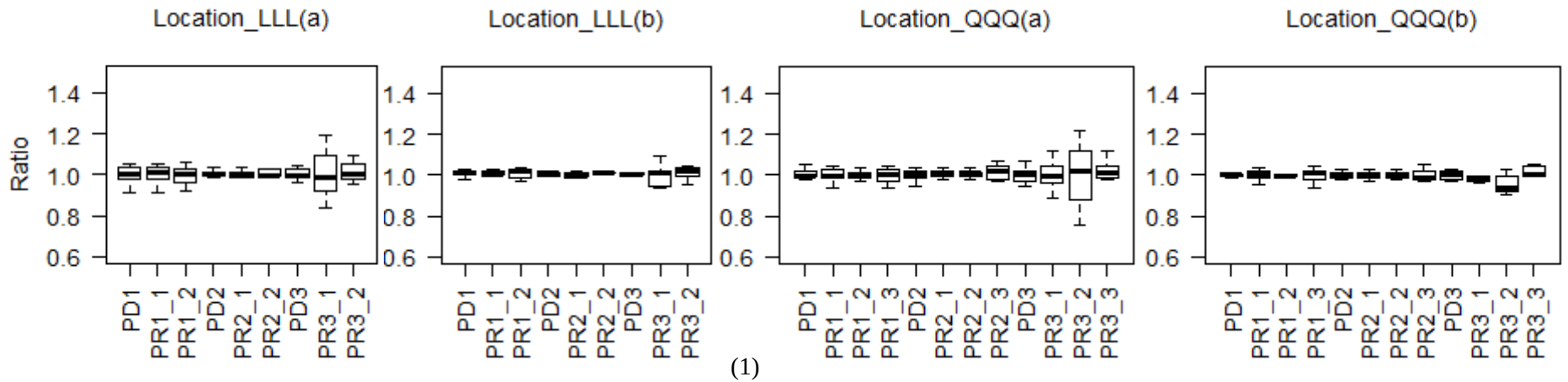
between successive states. In each scenario, the noise estimates for PDs and PRs using seven correlation values are first averaged at each state. Then for each parameter, a boxplot is created based on ratios between the individual average at each state over the overall average for all ten states. Situation (b): compare state noises with respect to different cross-correlation values. In each scenario, the noise estimates for PDs and PRs in all the ten states are first averaged for each cross-correlation value. Then for each parameter, a boxplot is created based on ratios between the individual average for each cross-correlation value over the overall average for all seven cross-correlation values. All the boxplots are created separately for STFs in LLL and QQQ forms.

The following conclusions can be drawn based on the boxplots in Figure 4-2: 1) The state noises are unlikely affected by Factor 2. In LLL (a) situations, for location and scale parameters, the interquartile ranges of the ratios for their PDs and PRs in all scenarios are close to zero; for shape parameters, slightly wider interquartile ranges are observed. This indicates no obvious change in the state noise is observed. Similar results are observed in QQQ (a) situation. 2) The state noises are unlikely affected by Factor 3. In LLL (b) situations, for almost all scenarios, the interquartile ranges of the ratios for almost all the state noises are close to zero, and a slightly wider interquartile range is observed for the PR3_2 in Scenario 3. Similar results are detected in QQQ (b) situation. 3) The state noise is not affected by Factor 4. In Situation (a) and (b), the two conclusions from previous analysis can be drawn using different choices of STFs in the proposed methodology.

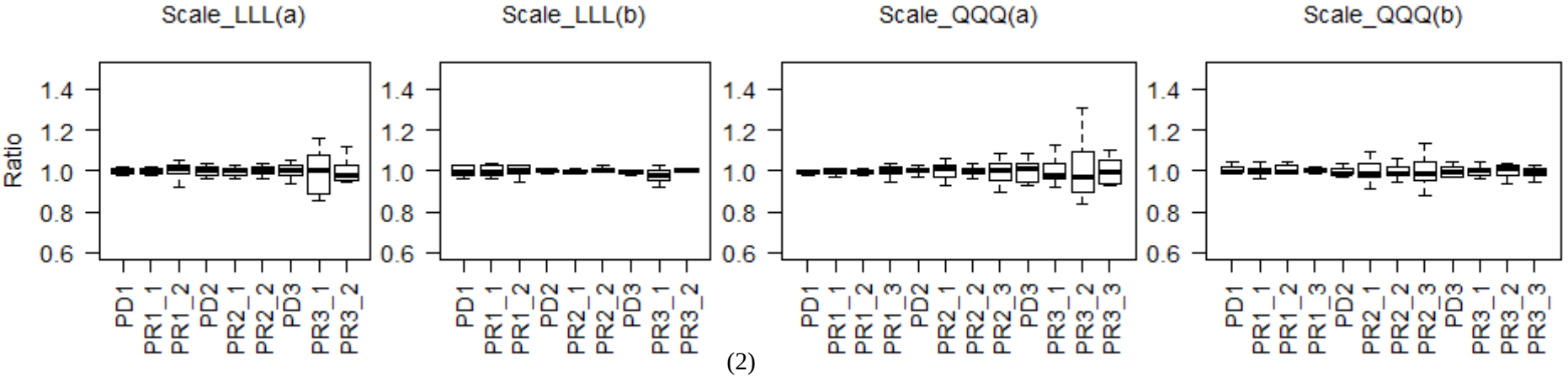
Based on the simulation results, when the homogeneous group for the target site are fixed across different states, the same state and measurement noises can be used in the one-step forward procedure for the regional frequency analysis in a nonstationary environment.

Table 4- 1 Three synthesized scenarios used for the state noise estimation in Section 4.4.1

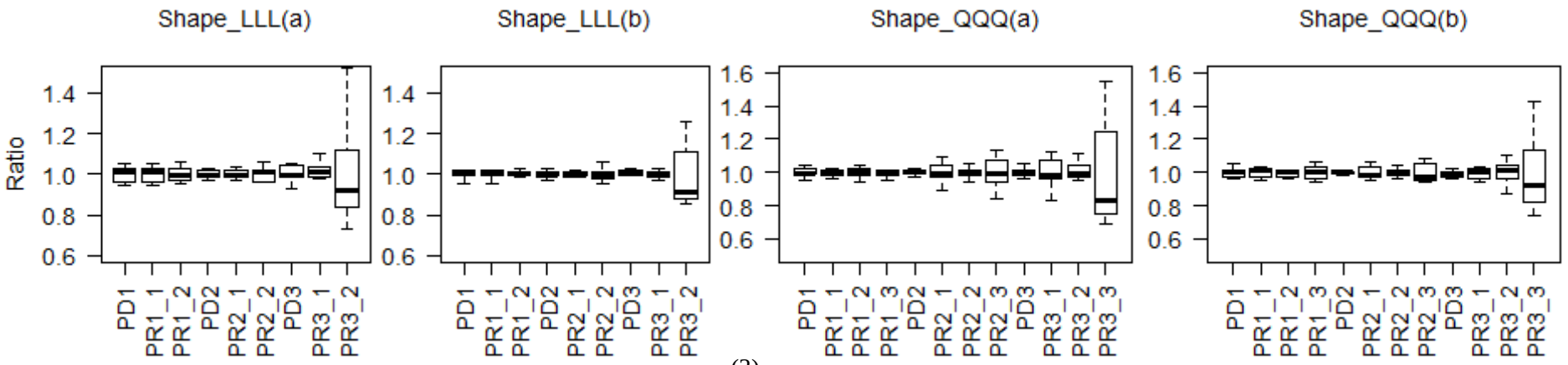
	Fixed Parameter		Changing Parameter											STF		Correlation
			State	State0	State1	State2	State3	State4	State5	State6	State7	State8	State9	State10			
Scenario1	Scale=0.27	Shape=-0.06	Location	0.670337	0.670337	0.67039	0.67089	0.676	0.686	0.706	0.756	0.856	1.056	1.456	LLL	QQQ	0-0.6
Scenario2	Location=0.87	Shape=-0.06	Scale	0.266221	0.266221	0.26623	0.26633	0.267	0.269	0.274	0.284	0.304	0.354	0.454	LLL	QQQ	0-0.6
Scenario3	Location=0.87	Scale=0.27	Shape	0.07656	0.07656	0.07655	0.07645	0.075	0.073	0.068	0.058	0.038	-0.012	-0.112	LLL	QQQ	0-0.6



(1)



(2)



(3)

Figure 4- 2 Boxplots for the ratio of estimated state noise over average state noise at two categories in three scenarios. Three scenarios have been described in Table 4-1: (1) Scenario 1; (2) Scenario 2; (3) Scenario 3. Two categories: (a) state noise summarized for the comparison among different magnitudes of the changes for the distribution parameters between successive states. (b) state noise summarized for the comparison among different cross-correlation values. LLL: STF in linear form. QQQ: STFs in quadratic form.

4.4.2. The selection of optimal STFs

To select the optimal STFs in the one-step forward procedure, the performance of STFs in LLL and QQQ forms are evaluated in a simulated homogeneous group. The details of this simulated group are as follows: 1) Station number and samples size at each state. There are 20 stations, and each with 10 data points at a state. 2) Number of states. 24 states are characterized by different GEV distributions. 3) Index-event values. The index events are set to 1 at all stations. 4) Cross-correlation values. The cross-correlation for this simulated group is set from 0 to 0.6 with 0.1 being used as the increment. 5) Sample generation. The same sample generation method used in Section 4.4.1 is applied here for the cross-correlated sample generation.

Table 4- 2 Trend description in three synthesized scenarios used in Section 4.4.2

		Period 1	Period 2	Period 3
Scenario 1	Location	Sinusoidal	Linear	Exponential
	Scale	Exponential	Linear	Sinusoidal
	Shape	Sinusoidal	Linear	Exponential
Scenario 2	Location	Sinusoidal	Linear	Exponential
	Scale	Linear	Exponential	Sinusoidal
	Shape	Sinusoidal	Linear	Exponential
Scenario 3	Location	Sinusoidal	Linear	Exponential
	Scale	Exponential in opposite direction	Linear	Sinusoidal in opposite direction
	Shape	Sinusoidal	Linear	Exponential

The three scenarios used to evaluate the performance of different STFs is presented in Table 4-2 and shown as the black lines in Figure 4-3 and 4-4. The reasons for choosing these scenarios are: 1) To evaluate the performance of the proposed methodology in different trends. Three trends: sinusoidal, linear and exponential are incorporated in the design of distribution parameters. The linear trends are designed to be in a constant state for evaluating the performance of proposed methodology in a stationary environment. 2) To evaluate the performance of proposed methodology when the changes in the distribution parameters are not synchronized. While location and shape parameter change at the same rate with different magnitudes in all three scenarios, the change directions of scale parameter vary: In Scenario 1, the scale parameter changes at the same rate and direction as the other distribution parameters but follows different trends; In Scenario 2, during the two thirds of the simulated states, the scale parameter changes

at different rates from the other parameters; In Scenario 3, the scale parameter changes at the same rates but follows opposite direction and different trends from the other parameters .

The parameters estimation using the proposed methodology in three scenarios can be conducted by following Step 4) to 7) introduced in Section 4.3.4. Based on the conclusions obtained in Section 4.4.2, the same state and measurement noises which are estimated using the procedure in Section 4.3.3 are applied at different states.

To select the optimal STFs for the proposed methodology, the estimates of the distribution parameters in all scenarios with average cross-correlation value being set to zero are conducted using STFs in LLL and QQQ forms. The parameter estimates from using STFs in LLL forms are shown in Figure 4-3, while those from using STFs in QQQ forms are shown in the first columns of each plot in Figure 4-4. Several conclusions can be obtained from comparing these estimates: 1) Trend depiction. The STFs in both LLL and QQQ forms can successfully describe the changing trends in location, scale and shape parameters in all three scenarios. 2) Noticeable differences are observed in the values of parameter estimates between using STFs in LLL and QQQ forms. The STFs in QQQ form generate better parameter estimates by showing values close to the design values for all distribution parameters especially for the shape parameter at all states, and scale and location parameters at stationary states. Thus, the STFs in QQQ form are selected as the optimal STFs in the proposed methodology.

To further demonstrate the effectiveness of the proposed methodology with STFs in QQQ form, it is tested in all three scenarios with different average cross-correlation values. The parameter estimates in three scenarios are depicted in Figure 4-4. The following conclusions are obtained: 1) The proposed methodology can accurately capture the trends for the location and scale parameters and reasonably simulate the trends in the shape parameter with different average cross-correlation values. 2) The accuracy for both location and scale parameter estimates reach the highest levels when average correlation is 0 and decreases when the average correlation value increases. The uncertainty for the shape parameter estimates reaches the lowest level when the average correlation value is 0.4 and decreases when the correlation value deviates from this value. 3) Since small differences are detected between the design values and estimated ones for the shape parameter, they have similar levels of goodness-of-fit based on AD test.

4.5 Application

This section is used to evaluate the proposed methodology, and compare it with the traditional GLM-based approach introduced in Section 4.2.

4.5.1. Study area and data processing

Following steps are conducted for the initial settings in the application of the proposed methodology for regional frequency analysis in a non-stationary environment:

1) The selection of testing regions. Two testing regions are selected in Canada: 1) Region 1: Includes most of the rainfall stations in British Columbia from Pacific Maritime and Cordilleran climate regions. 2) Region 2: Includes rainfall stations in Ontario and Quebec from Boreal and Southeastern climate regions

2) The extraction of 24-h AMS. The AMS of extreme rainfall records at 24-hr duration from 2011 to 2100 are extracted from the NA-CORDEX archive at 25-km (0.22°) spatial resolution using CanESM2/CanRCM4 model under the Representative Concentration Pathway (RCP) 4.5 scenario (Bias corrected data records were not available when this study was conducted). the 24-h AMS are extracted annually from March to November in British Columbia, and from April to November in Ontario and Quebec (Requena et al., 2019).

3) Trend analysis. Stations exhibiting trends need to first be identified. The Mann–Kendall nonparametric test is applied at every station using the extracted 24-h AMS for the trend detection (Mann, 1945; Kendall, 1948). Pettitt test is applied to remove stations with change-points. Thus, 15 stations with increasing trends are identified in Region 1 and 35 stations are identified in Region 2.

.

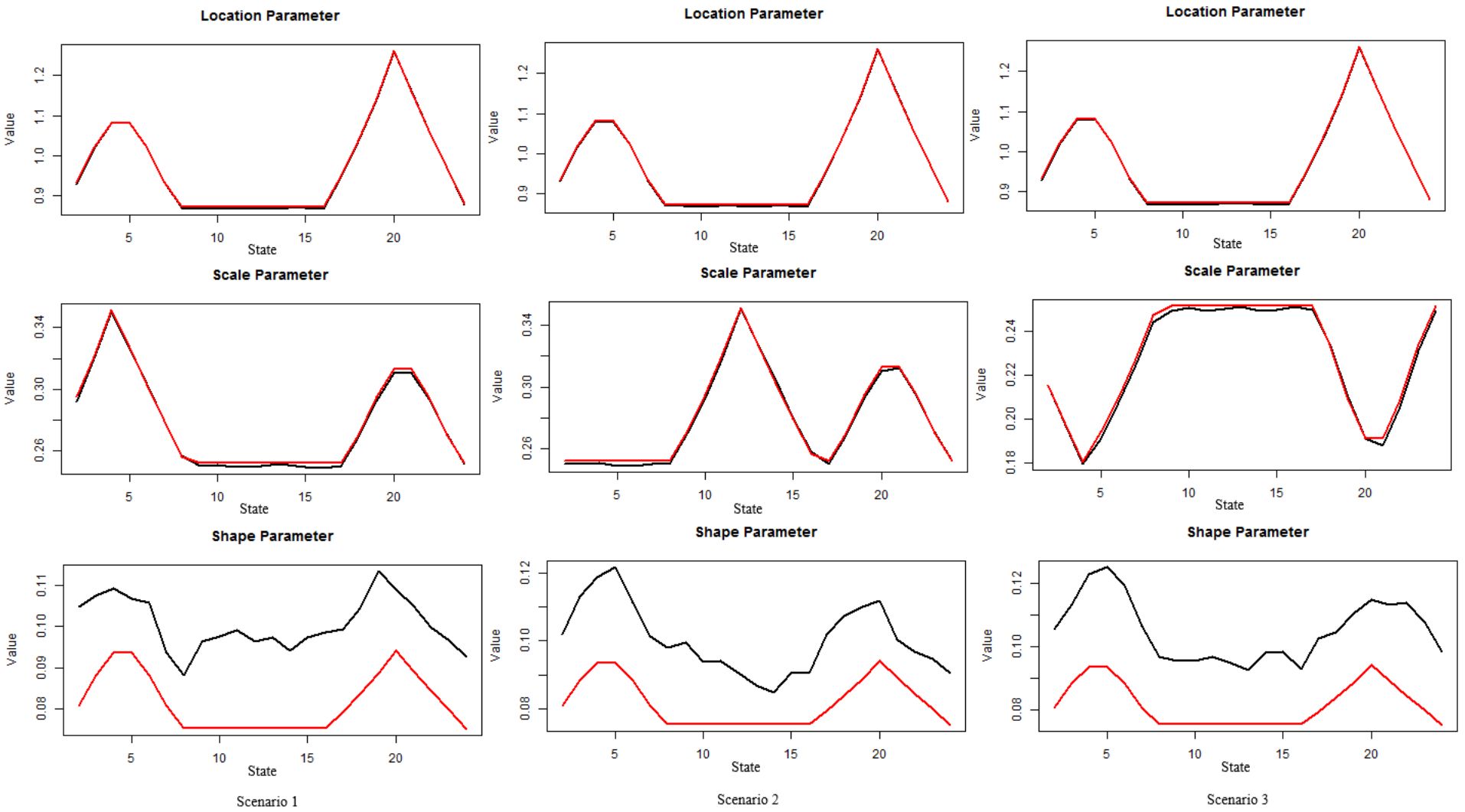
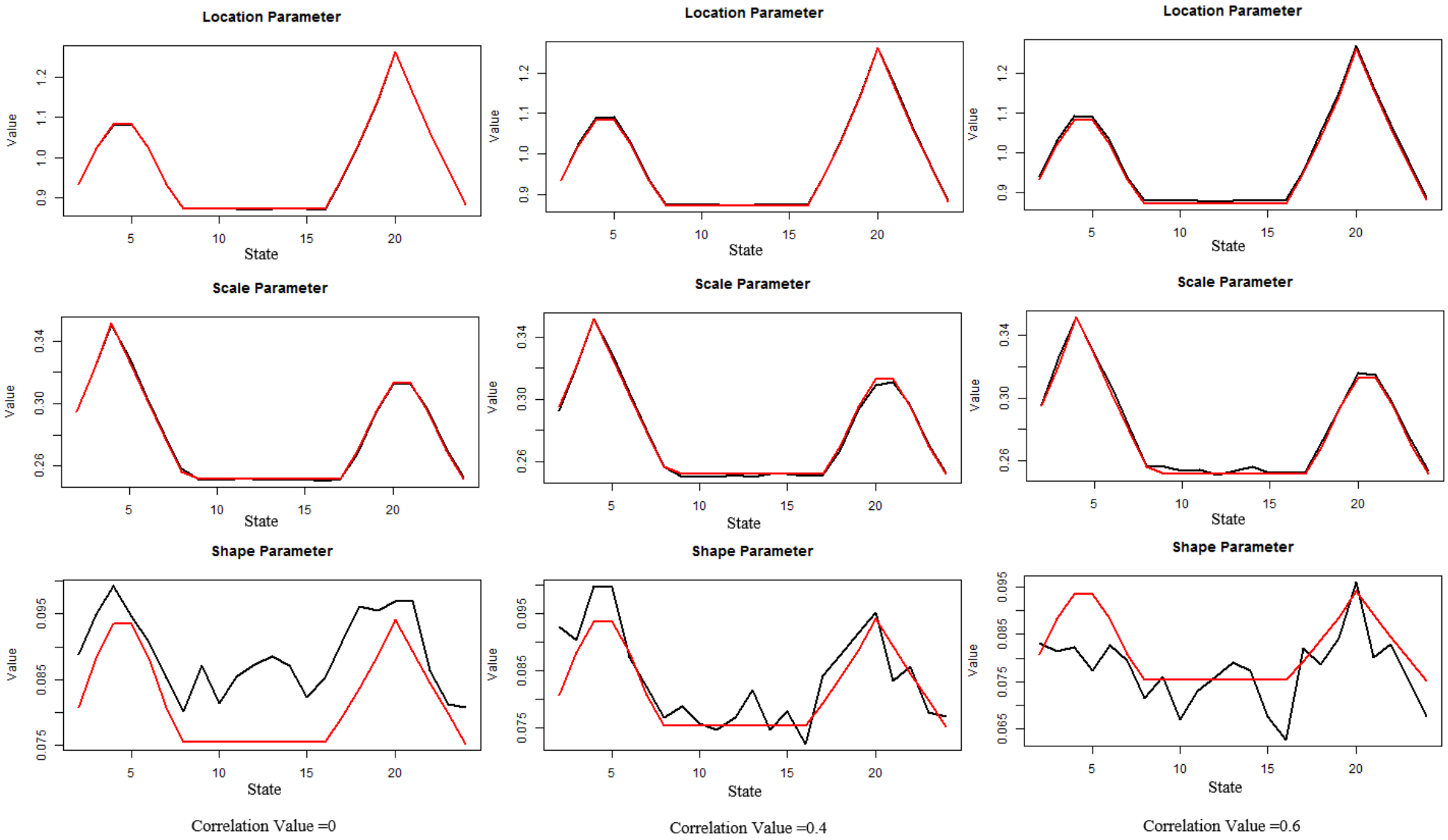
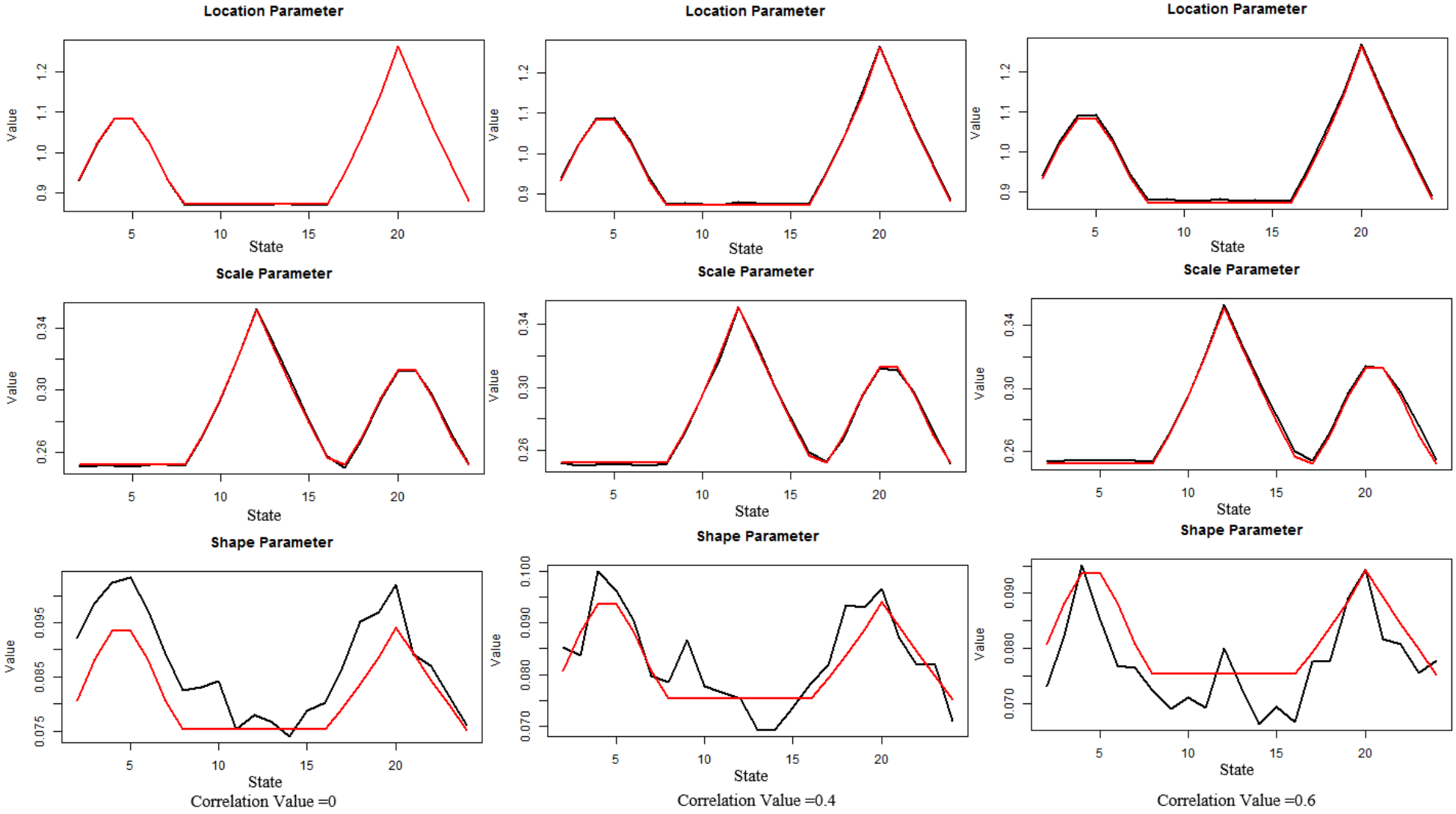


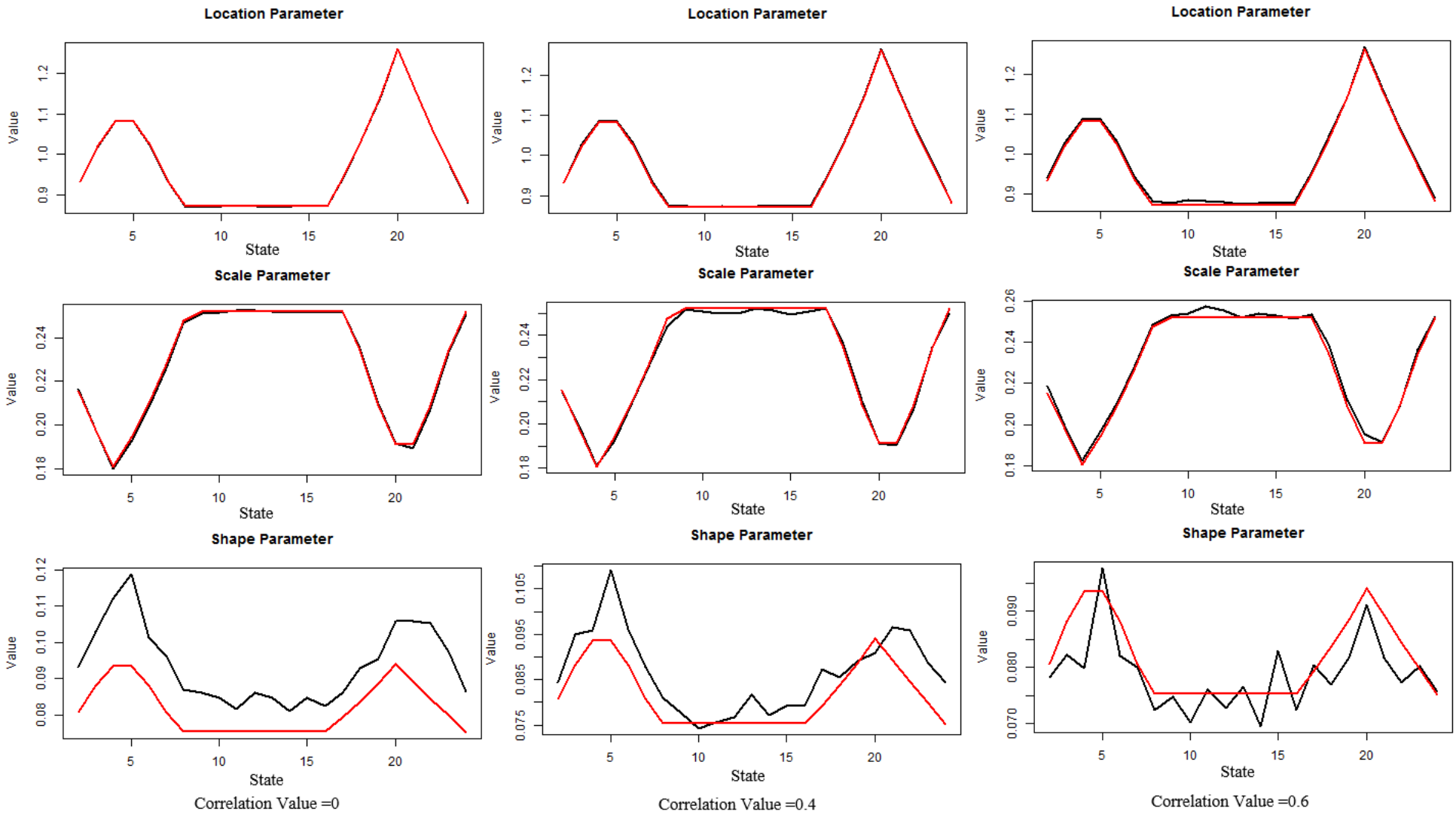
Figure 4- 3 The estimates of location, scale and shape parameters in three simulated scenarios (Columns 1, 2, and 3) using STFs in LLL form with average cross correlation equal to zero. The red line indicates the true values of the design parameters, while the black line shows the values for the estimates.


Location Parameter
Value
State
Location Parameter
Value
State
Location Parameter
Value
State
Scale Parameter
Value
State
Scale Parameter
Value
State
Scale Parameter
Value
State
Shape Parameter
Value
State
Shape Parameter
Value
State
Shape Parameter
Value
State
Correlation Value =0
Correlation Value =0.4
Correlation Value =0.6

Scenario 1


Location Parameter
Value
State
Location Parameter
Value
State
Location Parameter
Value
State
Scale Parameter
Value
State
Scale Parameter
Value
State
Scale Parameter
Value
State
Shape Parameter
Value
State
Correlation Value =0
Shape Parameter
Value
State
Correlation Value =0.4
Shape Parameter
Value
State
Correlation Value =0.6

Scenario 2



Scenario 3

Figure 4- 4 The estimates of location, scale and shape parameters in three simulated scenarios (Scenarios 1, 2 and 3) using STFs in QQQ form with average cross correlation equal to 0, 0.4 and 0.6 (Columns 1, 2, and 3). The three scenarios are different in the simulated trends for scale parameter; details are presented in Table 4-2. The red line indicates the true values of the design parameters, while the black line shows the values for the estimates.

4) Homogeneous group formation. Three homogeneous groups are formed: 1) Small homogeneous group for Station 1 (Climate ID: 1060841) in Region 1 (HG1). 6 stations are included with a heterogeneity measure of 0.92. 2) Small homogeneous group for Station 2 (Climate ID: 1160899) in Region 1 (HG2). 6 stations are included with a heterogeneity measure of 1.49. 3) Large homogeneous group for Station 3 (Climate ID: 7031375) in Region 2 (HG3). 30 stations are included with a heterogeneity measure of 1.2. The geographic display of these homogeneous groups is shown in Figure 4-5. In this step, the heterogeneity measure is calculated based on the difference between regional and at-site L-CVs (Hosking & Wallis, 1997).

5) the AMS Scaling. The extracted AMS at all selected stations are scaled by their index-event values.

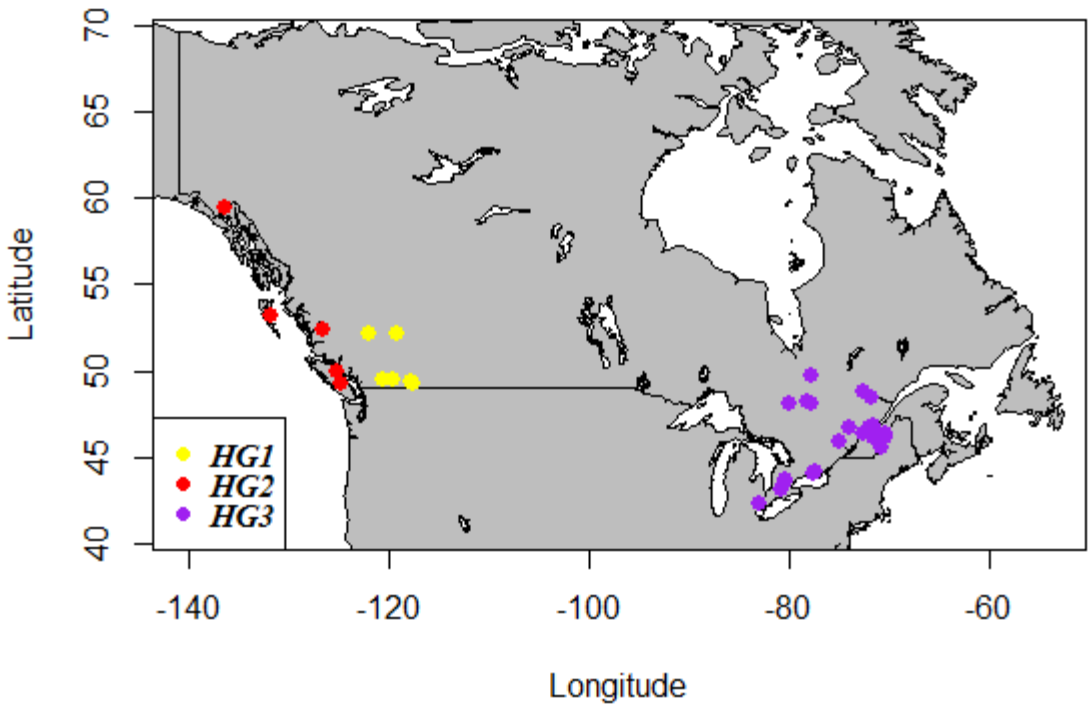


Figure 4- 5 Graphic displays of homogeneous groups HG1, HG2 and HG3 for Stations 1, 2 and 3.

4.5.2. Validation and comparison

This section is used for two purposes: the verification of the parameter estimates from the proposed methodology, and the comparison of goodness-of-fit between estimates from the proposed methodology and traditional GLM-based approach at the upper tail of the

measurements. Both procedures were conducted at different temporal steps for following purposes: 1) To prove the superiority of the proposed methodology over the traditional GLM-based method. Since the traditional GLM-based method is conducted on a yearly basis, the superiority of the proposed methodology can be confirmed when it is proven to provide better estimates on 1-year temporal step. 2) To capture the dynamic behavior of distribution parameters and reduce the uncertainty in the final quantile estimates. While small temporal steps should be used to meet the first requirement, larger uncertainty for the quantile estimates, especially at long return periods, can be generated. To balance these two goals, the 5T rule for gathering data records in the regional context is applied for the selection of temporal steps in the application of the proposed methodology for quantile estimation at different return periods (2-year, 5-year, 10-year, 25-year, 50-year and 100-year). Thus, the potential temporal steps used in this section are: 2-year, 5-year, 9-year, 22-year for the homogeneous groups formed at Station 1 and 2, 1-year, 2-year, 3-year, 5-year, 9-year and 15-year for the homogeneous groups formed at Station 3.

The parameter estimates from using the proposed methodology are verified through using AD test with the original data records at different temporal steps. The percentages of the valid estimates from using the proposed methodology and traditional GLM-based method at different temporal steps are shown in Table 4-3. The following conclusions can be obtained: 1) Both the proposed methodology and traditional approach can generate high percentages of valid estimates at different temporal steps. The proposed methodology is superior by providing higher percentages at different temporal steps. 2) The level of homogeneity can affect the validity of the parameter estimates in both approaches. The homogeneous group formed for Station 1 has the highest level of homogeneity and the highest percentage of valid estimates at different temporal steps.

To prove the superiority of the proposed methodology for providing better parameter estimates for the extreme rainfall quantile estimation at the upper tail, the ADup test is used as the goodness-of-fit test. The percentages of proposed methodology generating better parameter estimates than that from using the traditional GLM-based approach at different temporal steps are shown in Table 4-4. The following conclusions can be obtained: 1) Parameter estimates from the proposed methodology are better fits for the rainfall quantiles at upper tail than that from the traditional GLM-based approach. When using 1-year temporal step, the estimates from proposed

methodology are generally better fits at the upper tail as the percentages for being superior are higher than 50% for all three groups. For all groups, the level of homogeneity can affect the performance of the proposed methodology, since the percentage for parameter estimates being superior is the highest at the most homogeneous group for Station 1 while the lowest at the least homogeneous group for Station 2. With similar level of homogeneity, the performance of the proposed methodology is affected by sample size as the percentages for parameter estimates being superior are higher at the homogeneous group for Station 3 than that at Station 2. 2) The superiority of the proposed methodology is prominent when the temporal step is larger than 3-year since most of the percentages for parameter estimates being superior are higher than 80%.

Table 4- 3 The percentage of the parameter estimates from the proposed and traditional methodology that are valid using different temporal steps

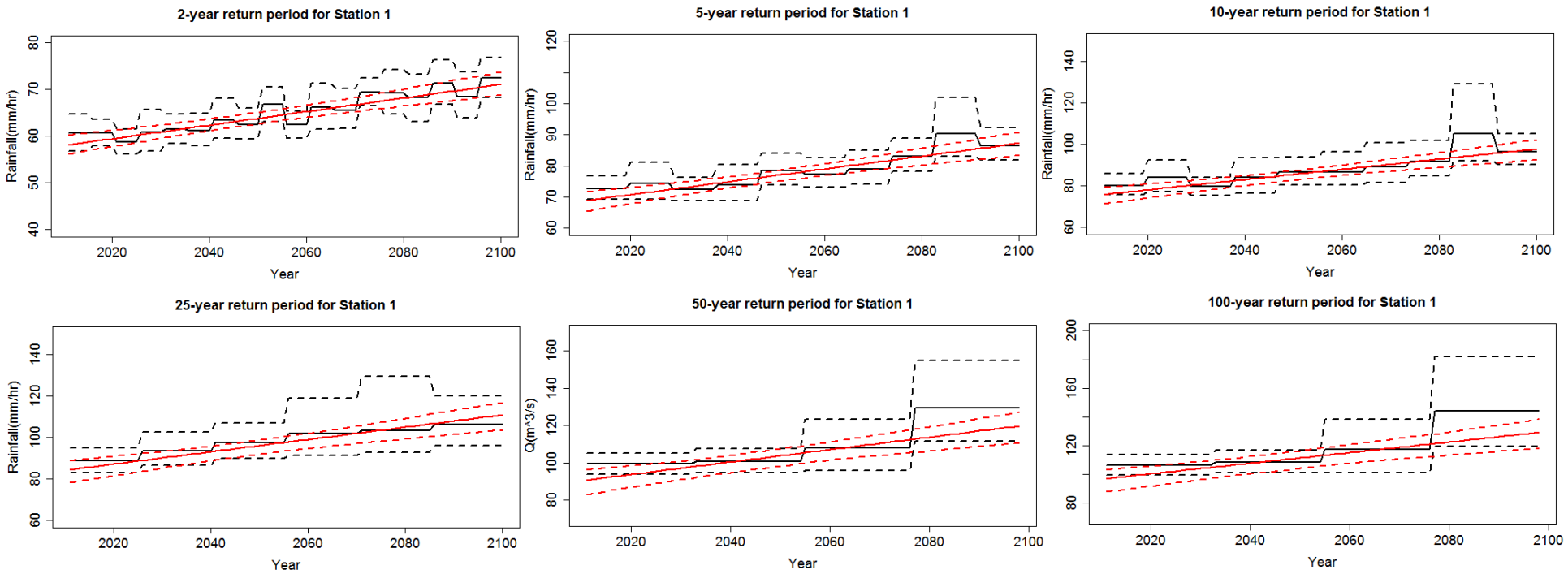
Temporal Step	1-year		2-year		3-year			
Method	Proposed	Traditional	Proposed	Traditional	Proposed	Traditional		
Station 1	98.89%	91.11%	100.00%	91.11%	100.00%	100.00%		
Station 2	92.22%	83.33%	91.11%	82.22%	100.00%	96.67%		
Station 3	70.00%	65.56%	80.00%	62.22%	96.67%	53.33%		
Temporal Step	5-year		9-year		15-year		22-year	
Method	Proposed	Traditional	Proposed	Traditional	Proposed	Traditional	Proposed	Traditional
Station 1	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Station 2	100.00%	88.89%	100.00%	90.00%	100.00%	100.00%	100.00%	100.00%
Station 3	94.44%	66.67%	100.00%	60.00%	100.00%	66.67%	100.00%	100.00%

Table 4- 4 The percentage of parameter estimates from the proposed methodology that are superior to that from using the traditional method when using different temporal steps

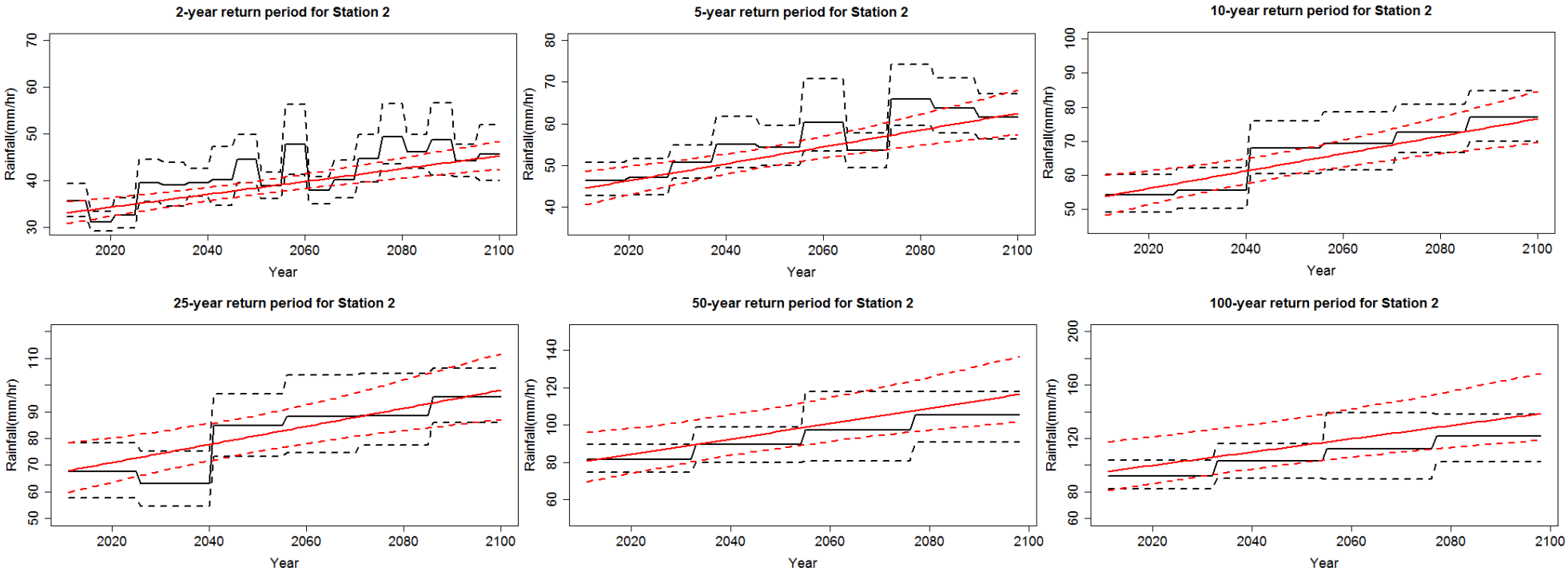
Temporal Step	1-year	2-year	3-year	5-year	9-year	15-year	22-year
Station 1	75.56%	51.11%	63.33%	88.89%	90.00%	50.00%	75.00%
Station 2	58.89%	77.78%	93.33%	100.00%	80.00%	83.33%	75.00%
Station 3	65.56%	53.33%	83.33%	83.33%	100.00%	83.33%	25.00%

Table 4- 5 The temporal steps used for the quantile estimation at different return periods for the three selected stations

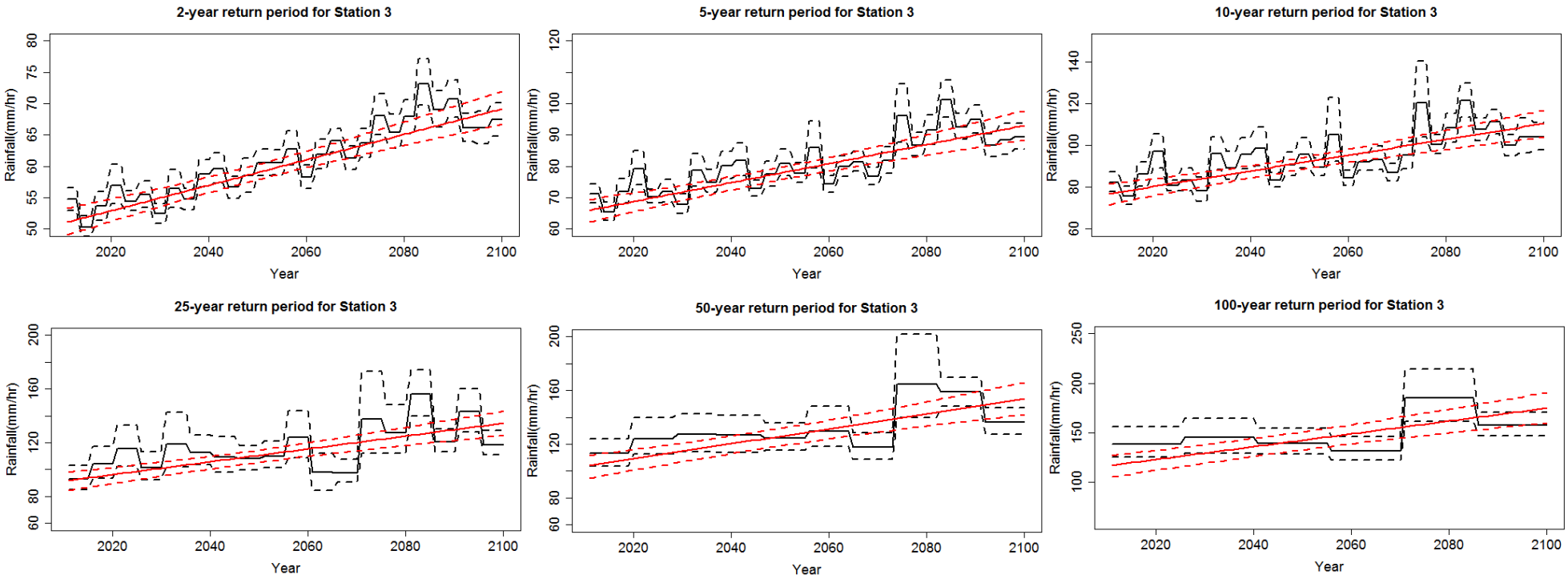
Return Period	2-year	5-year	10-year	25-year	50-year	100-year
Station 1	5-year	9-year	9-year	15-year	22-year	22-year
Station 2	5-year	9-year	15-year	15-year	22-year	22-year
Station 3	3-year	3-year	3-year	5-year	9-year	15-year


2-year return period for Station 1
Rainfall(mm/hr)
Year
5-year return period for Station 1
Rainfall(mm/hr)
Year
10-year return period for Station 1
Rainfall(mm/hr)
Year
25-year return period for Station 1
Rainfall(mm/hr)
Year
50-year return period for Station 1
Q(m^3/s)
Year
100-year return period for Station 1
Year

Station 1


2-year return period for Station 2
5-year return period for Station 2
10-year return period for Station 2
25-year return period for Station 2
50-year return period for Station 2
100-year return period for Station 2
Rainfall(mm/hr)
Year
2020
2040
2060
2080
2100

Station 2



Station 3

Figure 4- 6 Extreme rainfall quantiles and their associated 95% confidence intervals (CI) at Stations 1, 2 and 3 using the proposed methodology and traditional GLM-based method at different return-periods. The black solid line represents the quantiles estimated from using the proposed methodology; the 95% CI are described using the black dashed lines. The red solid line represents the quantiles estimated using the traditional method; the 95% CI are described using the red dashed lines.

4.5.3. Quantile estimation

Based on the comparison shown in Table 4-4 and following the 5T rule, the temporal steps used for the quantile estimation for the homogeneous groups at three selected stations are shown in Table 4-5. The quantile calculations with the 95% confidence intervals for Stations 1, 2 and 3 at different return periods are shown in Figure 4-6. Based on Figure 4-6, the following conclusions can be reached: 1) The basic trends of the extreme rainfall quantiles at different return periods from using the proposed methodology are consistent with that from using the traditional GLM-based method. All three homogeneous groups were formed by gathering stations showing increasing trends, thus the overall increasing trends of extreme rainfall quantiles at different return periods are consistent with the trend analysis. 2) Dynamic behavior of the extreme rainfall quantiles can be observed at different return periods. Compared to the linear increasing trend depicted in the extreme rainfall quantiles through using the traditional GLM-based method, more dynamic behavior for the extreme rainfall quantiles is observed, especially at short return periods for Station 1, and 2 and all return periods for Station 3.

4.6 Conclusions

Under climate change, the extreme rainfall patterns around the world are changing. The traditional GLM-based approach for the regional frequency analysis in a nonstationary environment is conducted under the assumption that the parameters in the regional growth curve are undergoing monotonic changes, and can be questionable since non-monotonic behavior of the extreme rainfall events having been observed.

To consider the dynamic behavior of the extreme rainfall in regional frequency analysis, a UKF-based procedure with spatial correlated resamples is established. To accurately capture the highly nonlinear relationships amongst the distribution parameters in the regional growth curve and consider their nonlinear transitions between successive states, the Unscented Kalman Filter is used. To reduce the uncertainty caused by limited number of rainfall records and consider the impacts form cross-correlation amongst the input stations, large sets of cross-correlated measurement resamples are generated for the quantification of the regional growth curve at each state. To reduce the possibility of performance degradation in the UKF application caused using

inaccurate noise covariance, Monte Carlo simulation-based method is used to explore the patterns of state and measurement noises.

Compared to the traditional approach, the proposed methodology is demonstrated to: 1) Fully utilize the advantage that the rainfall records from multiple stations are available for the parameter estimation in regional frequency analysis. 2) Consider the uncertainty associated with using limited rainfall records at each temporal period and the impacts from cross-correlation amongst the stations within a homogeneous group. 3) Consider the potential non-monotonic trends for the extreme rainfall quantiles at different return periods. The proposed methodology has provided reliable parameter estimates for the regional growth curve at stations exhibiting increasing trends and captured the non-monotonic change behavior of the extreme rainfall quantiles at different return periods

Chapter 5 General Conclusions

Accurate estimates for extreme rainfall intensity for various return periods at different geographic locations are essential for the effective planning and design of drainage systems in urban areas given the increasing occurrence of extreme rainfall events and the staggering costs from the associated damages. In a stationary environment, in which the statistics of the extreme rainfall records remain stable, frequency analysis of extreme rainfall events can be conducted through the exploration of historical observations. However, the approach of studying past-records may present unreliable rainfall quantile estimates under climate change since the basic assumption in the stationary approach may be violated. This thesis developed techniques to improve extreme rainfall quantile estimation in historical and future stationary environments through improving the homogeneous group formation process in regional frequency analysis, and to improve extreme rainfall quantile estimates in a future nonstationary environment through considering the potential non-monotonic change behavior of the parameters in the regional growth curve. The effectiveness of the developed approaches was examined using a large dataset of historical rainfall observations and projected rainfall series over Canada. The key findings of this research are outlined in the following sections.

5.1. Summary of results and conclusions

An approach to improve extreme rainfall quantile estimates in the historical stationary context through improving the homogeneous group formation using various geographic and atmospheric similarity indicators under climate change was developed in Chapter 2. To improve the homogeneous group formation process for target sites, the new approach, which is established using a three-layer framework, is used to select the optimal feature combination for representing extreme rainfall patterns at the target site. The proposed methodology has been demonstrated to provide promising homogeneous groups for four stations, from across Canada, with different geographic and atmospheric characteristics. Important findings from this work can be summarized as:

- A three-layer framework was established based on the stages of cloud formation, rainfall generation and the change of rainfall intensity over an urban surface, in the formation process of extreme rainfall events.

- The relationships between the atmospheric similarity indicators and annual extreme rainfall statistics were investigated for the first two layers of the three-layer framework. Differences were found in the selection of the optimal temporal resolutions used for the extraction of feature values at the first two layers, which verifies the necessity for the three-layer design.
- An automatic feature selection and weighting consisting of the hybrid searching algorithm of Tabu Search, Lagrange Multiplier and Fuzzy C-means clustering, was introduced for the searching for the optimal feature combination as the extreme rainfall similarity indicators for a target site.
- Optimal feature combinations used for the rainfall homogeneous group formation were selected at each of the four target stations. The various features selected for each station and the different correlation values with the annual maximum rainfall series demonstrated the need to identify similarity indicators for each target station.
- Less uncertainty and lower RMSE values were found in the homogeneous group generated from using the proposed methodology, thus confirming the improved precision of the results.

Chapter 3 provides a framework for regional rainfall quantile estimates under climate change based on the assumption that the extreme rainfall pattern remains stationary during 30-year timeframes: 2011 to 2040, 2041 to 2070 and 2071 to 2100. The improved homogeneous group formation procedures based on the three-layer searching framework in the historical stationary context, from Chapter 2, and the adjustment procedures for the previously formed homogeneous groups reaching new optimality in the future stationary context, were introduced in this chapter. To improve the three-layer searching algorithm for the homogeneous group formation in the historical context, a nonlinear feature extraction technique replaced the feature weighting method for the reduction of feature correlation impacts on the similar rainfall station gathering process and the Region of Influence approach was added as an additional homogeneous group formation technique for the gathering of similar rainfall stations. To adjust the previous homogeneous groups to new optimality in the future stationary context, an adjustment procedure was used conditional upon the homogeneous groups from the previous temporal period. During this process, the at-group rainfall statistics were used as the new similarity indicators to re-evaluate the rainfall similarity between different stations. The conclusions of this chapter are as follows:

- The application of ISOMAP as the feature extraction tool and the adding of Region of Influence as an additional homogeneous group formation method improved the performance

of the three-layer searching algorithm based on the resulting RMSE values from the formed homogeneous groups.

- The Fuzzy C-Mean clustering and the Region of Influence approach can be equally effective acting as the homogeneous group formation method in the three-layer searching algorithm.
- Under climate change, the changes in the spatial distribution of available moisture can cause spatial alteration of the extreme rainfall system. Thus, the optimal homogeneous groups that generate lower RMSE values compared to their geographic counterparts in the historical context are no longer optimal in the future context.
- The adjustment procedure can successfully revise the historical homogeneous groups to the new optimality in the future context in a hierarchical process. During the process, the rainfall statistics from applying the previous homogeneous groups in the future stationary context and the level of overlapping stations among the previous homogeneous groups are used as the new similarity indictors.

Chapter 4 provides an alternative framework for rainfall quantile estimation for stations showing significant trends in the future context. An unscented Kalman filter-based method was applied to challenge the traditional assumption about monotonic changing behavior for the regional growth curve in a future nonstationary environment. By utilizing multiple data points from the homogeneous group at different temporal steps, the proposed methodology was able to estimate non-monotonic changing parameters in the regional growth curve at different return periods with a homogeneous group formed using a trend centered pooling approach. The conclusions from Chapter 4 are presented as follows:

- The adoption of quadratic functions for the state transitions of the GEV distribution parameters can successfully depict the non-monotonic trends in values for the location and scale parameter and reasonably estimate the values of the changing shape parameter in simulations.
- The use of large bootstrap resamples in the UKF algorithm can successfully reduce the uncertainty caused by quantifying three parameter values using one nonlinear measurement function and increase the accuracy in the parameter estimates.
- A Markov Chain Monte Carlo-based simulation approach was proposed for the estimation of the state noise covariance used in UKF. The results indicated stationary state noise

covariance for the GEV parameters were observed and were successfully applied for the parameter estimation in the proposed methodology.

- The application of the proposed one-step forward procedure can successfully reduce the possibility of performance degradation caused by insufficient samples and inaccurate noise covariance in the UKF algorithm.
- The proposed methodology generates improved parameters in the regional growth curve compared with parameter estimates obtained from a traditional method.

5.2. Future research

The system developed in this thesis provides a basic framework for regional frequency analysis of extreme rainfall events in a nonstationary environment. Future potential improvements to the methodology can be conducted in the following aspects:

1) The improvement in the identification of similar rainfall stations for the adjustment procedure in the short temporal period. In Chapter 3, the identification of similar rainfall stations was conducted using Fuzzy C-Mean clustering with the at-group statistics being used as similarity indicators. This procedure can achieve reasonable accuracy when a 30-year record length is available for the adjusted period in a stationary environment. However, the accuracy of the adjustment process may decrease for shorter record lengths. Furthermore, when using the Fuzzy C-Means clustering, the gathering of similar rainfall stations depends heavily on the overall information provided from the inputs. Future research will first use the dynamic Bayesian network for the identification of stations that can be used to represent the extreme rainfall patterns within the adjusted short period, then the Region of Influence approach will be applied for the gathering of similar rainfall stations for the identified representative stations (i.e., the center stations). Finally, conditional upon the optimal homogeneous group in the previous temporal periods and the collected stations for the relevant center station at the current adjusted period, the adjustment procedure proposed in Chapter 3 can be applied for the homogeneous group formation for the target site for the adjusted period. The improvements in the homogeneous group adjustment in the short temporal periods facilitate the application of the proposed methodology in Chapter 4 for the estimation of non-monotonic behavior of the extreme rainfall quantiles.

2) The ensemble estimates of regional rainfall quantiles from using the different projected rainfall series under the same climate change scenario. In both Chapters 3 and 4, the methodology was tested by using one projected rainfall series under one climate change scenario. Future work will endeavour to obtain ensemble estimates of extreme rainfall quantiles from using different projected rainfall series under the same climate change scenario. Under the impacts of climate change, climate precipitation projections simulated using different Global Climate Models (GCM) and downscaled by different Regional Climate Models (RCM) are available for the extreme rainfall quantile estimation. To account for the different performances of GCMs and RCMs, the weighted average of the extreme rainfall quantiles estimated from applying the proposed methodology in Chapter 3 in a stationary environment and the proposed methodology in Chapter 4 in a nonstationary environment with different rainfall projections, can be calculated.

Reference

Adcock, C. T. (1997). Sample size determination : A Review. *The Statistician*, *46*(2), 261–283.

Ahmad, N. H., Othman, I. R., & Deni, S. M. (2013). Hierarchical Cluster Approach for Regionalization of Peninsular Malaysia based on the Precipitation Amount. *Journal of Physics: Conference Series*, *423*, 012018. https://doi.org/10.1088/1742-6596/423/1/012018

Akhlaghi, S., Zhou, N., & Huang, Z. (2018). Adaptive adjustment of noise covariance in Kalman filter for dynamic state estimation. *IEEE Power and Energy Society General Meeting*, *2018-Janua*, 1–5. https://doi.org/10.1109/PESGM.2017.8273755

Alila, Y. (1999). A hierarchical approach for the regionalization of precipitation annual maxima in Canada. *Journal of Geophysical Research*, *104*(24), 31645–31655. https://doi.org/10.1029/1999JD900764

Allan, R. P., & Soden, B. J. (2008). Atmospheric warming and the amplification of precipitation extremes. *Science*, *321*(5895), 1481–1484. https://doi.org/10.1126/science.1160787

Asong, Z. E., Khaliq, M. N., & Wheater, H. S. (2015). Regionalization of precipitation characteristics in the Canadian Prairie Provinces using large-scale atmospheric covariates and geophysical attributes. *Stochastic Environmental Research and Risk Assessment*, *29*(3), 875–892. https://doi.org/10.1007/s00477-014-0918-z

Baeriswyl, P. A., & Rebetez, M. (1997). Regionalization of precipitation in Switzerland by means of principal component analysis. *Theoretical and Applied Climatology*, *58*(1–2), 31–41. https://doi.org/10.1007/BF00867430

Bavdekar, V. A., Deshpande, A. P., & Patwardhan, S. C. (2011). Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter. *Journal of Process Control*, *21*(4), 585–601. https://doi.org/10.1016/j.jprocont.2011.01.001

Bélanger, P. R. (1974). Estimation of noise covariance matrices for a linear time-varying stochastic process. *Automatica*, *10*(3), 267–275. https://doi.org/10.1016/0005-1098(74)90037-5

Berrisford, P., Dee, D. P., Poli, P., Brugge, R., Fielding, M., Fuentes, M., … Simmons, A. (2011). *The ERA-Interim archive Version 2.0*. (1), 23. Retrieved from https://www.ecmwf.int/node/8174

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, *10*(2–3), 191–203. https://doi.org/10.1016/0098-3004(84)90020-7

Borgelt, C. (2008). Feature weighting and feature selection in fuzzy clustering. *Fuzzy Systems, 2008. FUZZ-IEEE 2008.(IEEE World Congress on Computational Intelligence). IEEE International Conference On*, 838–844. https://doi.org/10.1109/FUZZY.2008.4630468

Bornstein, R., & Lin, Q. (2000). Urban heat islands and summertime convective thunderstorms in Atlanta: Three case studies. *Atmospheric Environment*, *34*(3), 507–516. https://doi.org/10.1016/S1352-2310(99)00374-X

Boyd, E. C. (2010). Estimating and Mapping the Direct Flood Fatality Rate for Flooding in Greater New Orleans Due To Hurricane Katrina. *Risk, Hazards & Crisis in Public Policy*, *1*(3), 87–110. https://doi.org/10.2202/1944-4079.1017

Bryson, A. E., & Ho, Y.-C. (1975). *Applied Optimal Control*. Washington, DC: Hemisphere, 481 pp.

Burn, D H. (2003). The use of resampling for estimating confidence intervals for single site and pooled frequency analysis. *Hydrological Sciences Journal*, *48*(1), 25–38. https://doi.org/DOI 10.1623/hysj.48.1.25.43485

Burn, Donald H. (1990). Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resources Research*, *26*(10), 2257–2265. https://doi.org/10.1029/WR026i010p02257

Burn, Donald H. (2014). A framework for regional estimation of intensity-duration-frequency (IDF) curves. *Hydrological Processes*, *28*(14), 4209–4218. https://doi.org/10.1002/hyp.10231

Burn, Donald H., Mansour, R., Zhang, K., & Whitfield, P. H. (2011). Trends and Variability in Extreme Rainfall Events in British Columbia. *Canadian Water Resources Journal*, *36*(1), 67–82. https://doi.org/10.4296/cwrj3601067

Bush, E., D.S. Lemmen, & Editors. (2019). *Canada's Changing Climate Report*. Government of Canada,Ottawa,ON. 444 p.

Cai, W., Borlace, S., Lengaigne, M., Van Rensch, P., Collins, M., Vecchi, G., … Jin, F. F. (2014). Increasing frequency of extreme El Niño events due to greenhouse warming. *Nature Climate Change*, *4*(2), 111–116. https://doi.org/10.1038/nclimate2100

Calenda, G., Mancini, C. P., & Volpi, E. (2009). Selection of the probabilistic model of extreme floods: The case of the River Tiber in Rome. *Journal of Hydrology*, *371*(1–4), 1–11. https://doi.org/10.1016/j.jhydrol.2009.03.010

Castellarin, A., Burn, D. H., & Brath, A. (2008). Homogeneity testing: How homogeneous do heterogeneous cross-correlated regions seem? *Journal of Hydrology*, *360*(1–4), 67–76. https://doi.org/10.1016/j.jhydrol.2008.07.014

Charron, I. (2014). *A Guidebook on Climate Scenarios: Using Climate Information to Guide Adaptation Research and Decisions*. Montreal, Quebec, Canada.

Chen, L.-H., & Hong, Y.-T. (2012). Regional Taiwan rainfall frequency analysis using principal component analysis, self-organizing maps and L-moments. *Hydrology Research*, *43*(3), 275. https://doi.org/10.2166/nh.2012.032

Cheng, L., AghaKouchak, A., Gilleland, E., & Katz, R. W. (2014). Non-stationary extreme value analysis in a changing climate. *Climatic Change*, *127*(2), 353–369. https://doi.org/10.1007/s10584-014-1254-5

Chernobai, A, Rachev, S. T., & Fabozzi, F. . (2015). Composite Goodness-of-Fit Tests for Left-Truncated Loss Samples. *Handbook of Financial Econometrics and Statistics*, 1–2903. https://doi.org/10.1007/978-1-4614-7750-1

Chernobai, Anna, Rachev, S. T., & Fabozzi, F. J. (2005). Composite goodness-of-fit tests for left-truncated loss samples. In *Technical Report*. https://doi.org/10.1007/978-1-4614-7750-1_20

Cho, S. Y., & Choi, W. S. (2006). *Robust Positioning Technique in Low-Cost*. *55*(4), 1132–1142.

Collier, C. G. (2006). The impact of urban areas on weather. *Quarterly Journal of the Royal Meteorological Society*, *132*(614), 1–25. https://doi.org/10.1256/qj.05.199

CRED and UNISDR. (2018). *New UNISDR and CRED report: Economic Losses, Poverty & Disasters (1998 - 2017)*. Geneva, Switzerlandand Brussels, Belgium.

Cunderlik, J. M., & Burn, D. H. (2003). Non-stationary pooled flood frequency analysis. *Journal of Hydrology*, *276*(1–4), 210–223. https://doi.org/10.1016/S0022-1694(03)00062-3

Dalrymple, T. (1960). Flood Frequency Analyses. *Geological Survey Water-Supply Paper*, 1543-A.

Daubechies, I. (1990). The Wavelet Transform , Time-Frequency Localization and Signal Analysis. *IEEE Transactions on Information Theory*, *36*(5), 961–1005.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data Via the EM Algorithm . In *Journal of the Royal Statistical Society: Series B (Methodological)* (Vol. 39). https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

Diem, J. E., & Brown, D. P. (2003). Anthropogenic Impacts on Summer Precipitation in Central Arizona, U.S.A. *The Professional Geographer*, *55*(3), 343–355. https://doi.org/10.1111/0033-0124.5503011

Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A., & Maher, N. (2016). More extreme precipitation in the worldâ €TM s dry and wet regions. *Nature Climate Change*, *6*(5), 508–513. https://doi.org/10.1038/nclimate2941

Easterling, D. R. (1989). Regionalization of thunderstorm rainfall in the contiguous United States. *International Journal of Climatology*, *9*(6), 567–579. https://doi.org/10.1002/joc.3370090603

Environment and Climate Change Canada. (2017). 8. The Year of the Urban Flood. Retrieved from Environment and Climate Change Canada website: https://www.ec.gc.ca/meteo-weather/default.asp?lang=En&n=D05E090A-1

Fernandes, L. G., & Rodrigues, R. R. (2017). Changes in the patterns of extreme rainfall events in southern Brazil. *International Journal of Climatology*, *1352*(August 2017), 1337–1352. https://doi.org/10.1002/joc.5248

Fletcher, T. D., Shuster, W., Hunt, W. F., Ashley, R., Butler, D., Arthur, S., … Viklander, M. (2015). SUDS, LID, BMPs, WSUD and more – The evolution and application of terminology surrounding urban drainage. *Urban Water Journal*, *12*(7), 525–542. https://doi.org/10.1080/1573062X.2014.916314

FloodNet. (2015). FloodNet - NSERC Network - Enhanced flood forecasting and management capacity in Canada. Retrieved from http://www.nsercfloodnet.ca/

Forestieri, A., Lo Conti, F., Blenkinsop, S., Cannarozzo, M., Fowler, H. J., & Noto, L. V. (2018). Regional frequency analysis of extreme rainfall in Sicily (Italy). *International Journal of Climatology*, *38*(January), e698–e716. https://doi.org/10.1002/joc.5400

Gaál, L., & Kyselý, J. (2009). Comparison of region-of-influence methods for estimating high quantiles of precipitation in a dense dataset in the Czech Republic. *Hydrology and Earth System Sciences*, *13*(11), 2203–2219. https://doi.org/10.5194/hess-13-2203-2009

Gabriele, S., & Chiaravalloti, F. (2013a). Searching regional rainfall homogeneity using atmospheric fields. *Advances in Water Resources*, *53*, 163–174. https://doi.org/10.1016/j.advwatres.2012.11.002

Gabriele, S., & Chiaravalloti, F. (2013b). Using the Meteorological Information for the Regional Rainfall Frequency Analysis: An Application to Sicily. *Water Resources Management*, *27*(6), 1721–1735. https://doi.org/10.1007/s11269-012-0235-6

Gaetan, C., & Grigoletto, M. (2004). Smoothing Sample Extremes with Dynamic Models. *Extremes*, *7*, 221–236.

Gao, L., Huang, J., Chen, X., Chen, Y., & Liu, M. (2016). Risk of Extreme Precipitation under Nonstationarity Conditions during the Second Flood Season in the Southeastern Coastal Region of China. *Journal of Hydrometeorology*, *18*(3), 669–681. https://doi.org/10.1175/jhm-d-16-0119.1

Garc´ıa-Mar´ın, A. P., Ayuso-Mu ˜noz, J. L., Taguas-Ruiz, E. V., & Estevez, J. (2011). Regional analysis of the annual maximum daily rainfall in the province of Malaga (southern Spain) using the principal component analysis. *Water and Environment Journal*, *25*(4), 522–531. https://doi.org/10.1111/j.1747-6593.2011.00251.x

Glover, F. (1986). Future Paths for Integer Programming and Links to Artificial Intelligence. *Computers and Operations Research*, *13*(5), 533–549.

Glover, F. (1989). Tabu Search-Part I. *ORSA Journal on Computing*, *1*(3), 190–206.

Goswami, B. N., Venugopal, V., Sengupta, D., Madhusoodanan, M. S., & Xavier, P. K. (2006). Increasing Trend of Extreme Rain Events Over India in a Warming Environment. *Science*, *314*(5804), 1442–1445. https://doi.org/10.1126/science.1132027

Goyal, M. K., & Gupta, V. (2014). Identification of Homogeneous Rainfall Regimes in Northeast Region of India using Fuzzy Cluster Analysis. *Water Resources Management*, *28*, 4491–4511. https://doi.org/10.1007/s11269-014-0699-7

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., & Wallis, J. R. (1979). Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, *15*(5), 1049. https://doi.org/10.1029/WR015i005p01049

Haddad, K., Johnson, F., Rahman, A., Green, J., & Kuczera, G. (2015). Comparing three methods to form regions for design rainfall statistics: Two case studies in Australia. *Journal of Hydrology*, *527*, 62–76. https://doi.org/10.1016/j.jhydrol.2015.04.043

Hadwin, P. J., & Peterson, S. D. (2017). An extended Kalman filter approach to non-stationary

Bayesian estimation of reduced-order vocal fold model parameters. *The Journal of the Acoustical Society of America*, *141*(4), 2909–2920. https://doi.org/10.1121/1.4981240

Hamill, T., Bates, G., Whitaker, J., Murray, D., Fiorino, M., & Galarneau, T. (2013). A Description of the 2nd Generation NOAA Global Ensemble Reforecast Data Set. *NOAA Earth System Research Lab, Physical Sciences Division Bouder, Colorado, USA*, 10. Retrieved from https://www.esrl.noaa.gov/psd/forecasts/reforecast2/README.GEFS_Reforecast2.pdf

Hanel, M., Buishand, T. A., & Ferro, C. A. T. (2009). A nonstationary index flood model for precipitation extremes in transient regional climate model simulations. *Journal of Geophysical Research Atmospheres*, *114*(15), 1–16. https://doi.org/10.1029/2009JD011712

Hosking, J. R. M., & Wallis, J. R. (1997). *Regional Frequency Analysis: An approach based on L-moments, Cambridge University Press, New York.* https://doi.org/doi:10.1017/CBO9780511529443

Houtekamer, P., & Mitchell, H. L. (1998). Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review*, *126*(3), 796–811.

Houze Jr., R. a. (1997). Stratiform precipitation in the tropics: A meteorological paradox? *Bull. Amer. Meterol. Soc.*, *78*(10), 2179–2196. https://doi.org/10.1175/1520-0477(1997)078<2179:SPIROC>2.0.CO;2

Houze, R. A. (1989). Observed structure of mesoscale convective systems and implications for large-scale heating. *Quarterly Journal of the Royal Meteorological Society*, *115*(487), 425–461. https://doi.org/10.1002/qj.49711548702

Insurance Bureau of Canada. (2017). Spring flooding in Ontario and Quebec caused more than $223 million in insured damage. Retrieved September 19, 2019, from http://www.ibc.ca/on/resources/media-centre/media-releases/spring-flooding-in-ontario-and-quebec-caused-more-than-223-million-in-insured-damage

Insurance Bureau of Canada. (2018). Toronto Flood Causes over $80 Million in Insured Damage. Retrieved September 19, 2019, from http://www.ibc.ca/on/resources/media-centre/media-releases/toronto-flood-causes-over-80-million-in-insured-damage

IPCC. (2013). *Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J.* Retrieved from doi:10.1017/CBO9781107415324

Jacob, D., Reed, D. W., & Robson, A. J. (1999). *Choosing a Pooling Group: Flood Estimation Handbook*. Wallingford, UK.: Institute of Hydrology.

Jin, M., Shepherd, J. M., & Zheng, W. (2010). Urban Surface Temperature Reduction via the Urban Aerosol Direct Effect: A Remote Sensing and WRF Model Sensitivity Study. *Advances in Meteorology*, *2010*, 1–14. https://doi.org/10.1155/2010/681587

Johnston, D. E., & Djurić, P. M. (2019). A Recursive Bayesian Model for Extreme Values. *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal*

Processing (ICASSP), 5062–5066. https://doi.org/10.1109/ICASSP.2019.8682828

Julier, S. J., & Uhlmann, J. K. (1997). A New Extension of the Kalman Filter to Nonlinear Systems. *In Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls*.

Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, *25*(8–12), 1287–1304. https://doi.org/10.1016/S0309-1708(02)00056-8

Kaufman, Y. J., & Koren, I. (2006). Smoke and pollution aerosol effect on cloud cover. *Science*, *313*(5787), 655–658.

Kendall, M. G. (1948). Rank correlation methods. In *Rank correlation methods.* Oxford, England: Griffin.

Khazai, B., Bessel, T., Möhrle, S., Dittrich, A., Schröter, K., Mühr, B., … Trieselmann, W. (2013). *June 2013 Flood in Central Europe - Focus Germany Report 1 – Update 2: Preconditions, Meteorology, Hydrology*. *1*(June), 1–13.

Kumar, P., & Foufoula-georgiou, E. (1997). Wavelet analysis for geophysical applications. *Reviews of Geophysics*, *35*(4), 385–412.

Lee, F., Scherer, R., Leeb, R., Schlögl, A., Bischof, H., & Pfurtscheller, G. (2004). Feature mapping using PCA, Locally Linear Embedding and Isometric Feature Mapping for EEG-based brain computer interface. *Digital Imaging in Media and Education, Proceedings of the 28th AAPR Workshop*.

Lehmann, J., Coumou, D., & Frieler, K. (2015). Increased record-breaking precipitation events under global warming. *Climatic Change*, *132*(4), 501–515. https://doi.org/10.1007/s10584-015-1434-y

Li, D., Bou-Zeid, E., Baeck, M. L., Jessup, S., & Smith, J. a. (2013). Modeling Land Surface Processes and Heavy Rainfall in Urban Environments: Sensitivity to Urban Surface Representations. *Journal of Hydrometeorology*, *14*(4), 1098–1118. https://doi.org/10.1175/JHM-D-12-0154.1

Loughner, C. P., Allen, D. J., Zhang, D. L., Pickering, K. E., Dickerson, R. R., & Landry, L. (2012). Roles of urban tree canopy and buildings in urban heat island effects: Parameterization and preliminary results. *Journal of Applied Meteorology and Climatology*, *51*(10), 1775–1793. https://doi.org/10.1175/JAMC-D-11-0228.1

Ma, S., Zhou, T., Dai, A., & Han, Z. (2015). Observed changes in the distributions of daily precipitation frequency and amount over China from 1960 to 2013. *Journal of Climate*, *28*(17), 6960–6978. https://doi.org/10.1175/JCLI-D-15-0011.1

Mailhot, A., Beauregard, I., Talbot, G., Caya, D., & Biner, S. (2012). Future changes in intense precipitation over Canada assessed from multi-model NARCCAP ensemble simulations. *International Journal of Climatology*, *32*(8), 1151–1163. https://doi.org/10.1002/joc.2343

Mailhot, A., Duchesne, S., Caya, D., & Talbot, G. (2007). Assessment of future change in intensity-duration-frequency (IDF) curves for Southern Quebec using the Canadian

Regional Climate Model (CRCM). *Journal of Hydrology*, *347*(1–2), 197–210. https://doi.org/10.1016/j.jhydrol.2007.09.019

Mailhot, A., Lachance-Cloutier, S., Talbot, G., & Favre, A. C. (2013). Regional estimates of intense rainfall based on the Peak-Over-Threshold (POT) approach. *Journal of Hydrology*, *476*, 188–199. https://doi.org/10.1016/j.jhydrol.2012.10.036

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*(7), 674–693. https://doi.org/10.1109/34.192463

Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, *13*(3), 245–259.

Mann, M. E., Lloyd, E. A., & Oreskes, N. (2017). Assessing climate change impacts on extreme weather events: the case for an alternative (Bayesian) approach. *Climatic Change*, *144*(2), 131–142. https://doi.org/10.1007/s10584-017-2048-3

Marra, F., Morin, E., Peleg, N., Mei, Y., & Anagnostou, E. N. (2017). Intensity–duration–frequency curves from remote sensing rainfall estimates: comparing satellite and weather radar over the eastern Mediterranean. *Hydrology and Earth System Sciences*, *21*(5), 2389–2404. https://doi.org/10.5194/hess-21-2389-2017

Mbengue, C., & Schneider, T. (2013). Storm track shifts under climate change: What can be learned from large-scale dry dynamics. *Journal of Climate*, *26*(24), 9923–9930. https://doi.org/10.1175/JCLI-D-13-00404.1

Mehra, R. K. (1970). On the Identification of Variances and Adaptive Kalman Filtering. *IEEE Trans. Automat. Control*, *15*(12), 175–184. https://doi.org/10.1063/1.3057931

Meng, Y., Gao, S., Zhong, Y., Hu, G., & Subic, A. (2016). Covariance matching based adaptive unscented Kalman filter for direct filtering in INS/GNSS integration. *Acta Astronautica*, *120*, 171–181. https://doi.org/10.1016/j.actaastro.2015.12.014

Meukaleuni, C., Lenouo, A., & Monkam, D. (2016). Climatology of convective available potential energy (CAPE) in ERA-Interim reanalysis over West Africa. *Atmospheric Science Letters*, *17*(1), 65–70. https://doi.org/10.1002/asl.601

Miao, S., Chen, F., Li, Q., & Fan, S. (2011). Impacts of urban processes and urbanization on summer precipitation: A case study of heavy rainfall in Beijing on 1 August 2006. *Journal of Applied Meteorology and Climatology*, *50*(4), 806–825. https://doi.org/10.1175/2010JAMC2513.1

Moncrieff, M. W., & Miller, M. J. (1976). The dynamics and simulation of tropical cumulonimbus and squall lines. *Quarterly Journal of the Royal Meteorological Society*, *102*(432), 373–394. https://doi.org/10.1002/qj.49710243208

Mondal, A., & Mujumdar, P. P. (2015). Modeling non-stationarity in intensity, duration and frequency of extreme rainfall over India. *Journal of Hydrology*, *521*, 217–231. https://doi.org/10.1016/j.jhydrol.2014.11.071

Mostofi Zadeh, S., Durocher, M., Burn, D. H., & Ashkar, F. (2019). Pooled flood frequency analysis: a comparison based on peaks-over-threshold and annual maximum series.

Hydrological Sciences Journal, *64*(2), 121–136. https://doi.org/10.1080/02626667.2019.1577556

Murugavel, P., Pawar, S. D., & Gopalakrishnan, V. (2012). Trends of Convective Available Potential Energy over the Indian region and its effect on rainfall. *International Journal of Climatology*, *32*(9), 1362–1372. https://doi.org/10.1002/joc.2359

Nadarajah, S. (2005). Extremes of daily rainfall in West Central Florida. *Climatic Change*, *69*(2–3), 325–342. https://doi.org/10.1007/s10584-005-1812-y

Nam, W., Kim, S., Kim, H., Joo, K., & Heo, J. H. (2015). The evaluation of regional frequency analyses methods for nonstationary data. *IAHS-AISH Proceedings and Reports*, *371*(1), 95–98. https://doi.org/10.5194/piahs-371-95-2015

Nilsson, M. (2006). Kalman filtering with unknown noise covariances. *Proc. Reglermote*, (6), 1–4. Retrieved from http://eprints.sics.se/423

O'Brien, N. L., & Burn, D. H. (2014). A nonstationary index-flood technique for estimating extreme quantiles for annual maximum streamflow. *Journal of Hydrology*, *519*, 2040–2048. https://doi.org/10.1016/j.jhydrol.2014.09.041

Odelson, B. J., Lutz, A., & Rawlings, J. B. (2006). The autocovariance least-squares method for estimating covariances: Application to model-based control of chemical reactors. *IEEE Transactions on Control Systems Technology*, *14*(3), 532–540. https://doi.org/10.1109/TCST.2005.860519

Ohba, M., Kadokura, S., Yoshida, Y., Nohara, D., & Toyoda, Y. (2015). Anomalous Weather Patterns in Relation to Heavy Precipitation Events in Japan during the Baiu Season. *Journal of Hydrometeorology*, *16*(2), 688–701. https://doi.org/10.1175/jhm-d-14-0124.1

Park, J. S. (2005). A simulation-based hyperparameter selection for quantile estimation of the generalized extreme value distribution. *Mathematics and Computers in Simulation*, *70*(4), 227–234. https://doi.org/10.1016/j.matcom.2005.09.003

Pendergrass, A. G., Knutti, R., Lehner, F., Deser, C., & Sanderson, B. M. (2017). Precipitation variability increases in a warmer climate. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-17966-y

Putnam, A. E., & Broecker, W. S. (2017). Human-induced changes in the distribution of rainfall. *Science Advances*, *3*(5), e1600871. https://doi.org/http://dx.doi.org/10.1126/sciadv.1600871

Quiroz, R., Yarlequé, C., Posadas, A., Mares, V., & Immerzeel, W. W. (2011). Improving daily rainfall estimation from NDVI using a wavelet transform. *Environmental Modelling & Software*, *26*(2), 201–209. https://doi.org/10.1016/j.envsoft.2010.07.006

Requena, A. I., Burn, D. H., & Coulibaly, P. (2019). Estimates of gridded relative changes in 24-h extreme rainfall intensities based on pooled frequency analysis. *Journal of Hydrology*, *577*(June), 123940. https://doi.org/10.1016/j.jhydrol.2019.123940

Robson, A., & Reed, D. (1999). Statistical Procedures for Flood Frequency Estimation. In *Flood Estimation Handbook, Volume 3*. Wallingford, UK.: Institute of Hydrology.

Rosenfeld, D., Rosenfeld, D., Lohmann, U., Raga, G. B., Dowd, C. D. O., Kulmala, M., … Andreae, M. O. (2014). Flood or Drought : How Do Aerosols Affect Precipitation ? *Science*, *1309*(2008), 1309–1314. https://doi.org/10.1126/science.1160606

Roth, M. (2000). Review of atmospheric turbulence over cities. *Quarterly Journal of the Royal Meteorological Society*, *126*(564), 941–990. https://doi.org/10.1002/qj.49712656409

Roth, M., & Oke, T. R. (1993). Turbulent transfer relationships over an urban surface. I. Spectral characteristics. *Quarterly Journal of the Royal Meteorological Society*, *119*(513), 1071–1104. https://doi.org/10.1002/qj.49711951311

Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science (New York, N.Y.)*, *290*(5500), 2323–2326. https://doi.org/10.1126/science.290.5500.2323

Sait, S. M., & Youssef, H. (1999). *Iterative Computer Algorithms with Applications in Engineering: Solving Combinatorial Optimization Problems* (1st ed.). Los Alamitos, CA, USA: IEEE Computer Society Press.

Samko, O., Marshall, A. D., & Rosin, P. L. (2006). Selection of the optimal parameter value for the Isomap algorithm. *Pattern Recognition Letters*, *27*(9), 968–979. https://doi.org/10.1016/j.patrec.2005.11.017

Sandink, D. (2016). Urban flooding and ground-related homes in Canada: an overview. *Journal of Flood Risk Management*, *9*(3), 208–223. https://doi.org/10.1111/jfr3.12168

Sandink, Dan. (2013). *Urban flooding in Canada: Lot-side risk reduction through voluntary retrofit programs , code interpretation and by-laws*.

Satyanarayana, P., & Srinivas, V. V. (2008). Regional frequency analysis of precipitation using large-scale atmospheric variables. *Journal of Geophysical Research*, *113*(D24), 1–16. https://doi.org/10.1029/2008JD010412

Satyanarayana, P., & Srinivas, V. V. (2011). Regionalization of precipitation in data sparse areas using large scale atmospheric variables - A fuzzy clustering approach. *Journal of Hydrology*, *405*(3–4), 462–473. https://doi.org/10.1016/j.jhydrol.2011.05.044

Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C. M., Kanae, S., Kossin, J., … Zhang, X. (2012). *Changes in climate extremes and their impacts on the natural physical environment. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastr*. Cambridge, UK, and New York, NY, USA.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, *Vol.27*(1948), 379–423.

Shepherd, J. M. (2005). A review of current investigations of urban-induced rainfall and recommendations for the future. *Earth Interactions*, *9*(12). https://doi.org/10.1175/EI156.1

Shivakumar, N. R., & Jain, A. (2008). A Review of Power System Dynamic State Estimation Techniques. *2008 Joint International Conference on Power System Technology and IEEE Power India Conference*, 1–6. https://doi.org/10.1109/ICPST.2008.4745312

Shu, C., & Burn, D. H. (2004). Homogeneous pooling group delineation for flood frequency analysis using a fuzzy expert system with genetic enhancement. *Journal of Hydrology*, *291*(1–2), 132–149. https://doi.org/10.1016/j.jhydrol.2003.12.011

Shumway, R. H., & Stoffer, D. S. (2000). *Time Series Analysis and its Applications*. https://doi.org/10.1007/978-3-319-52452-8

Simonovic, S. P., Schardong, A., & Sandink, D. (2016). Mapping Extreme Rainfall Statistics for Canada under Climate Change Using Updated Intensity-Duration-Frequency Curves. *Journal of Water Resources Planning and Management*, 04016078. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000725

Simonovic, S. P., Schardong, A., & Sandink, D. (2017). Mapping Extreme Rainfall Statistics for Canada under Climate Change Using Updated Intensity-Duration-Frequency Curves. *Journal of Water Resources Planning and Management*, *143*(3), 04016078. https://doi.org/10.1061/(ASCE)WR.1943-5452

Singh, M. S., Kuang, Z., Maloney, E. D., Hannah, W. M., & Wolding, B. O. (2017). Increasing potential for intense tropical and subtropical thunderstorms under global warming. *Proceedings of the National Academy of Sciences*, *114*(44), 11657–11662. https://doi.org/10.1073/pnas.1707603114

Stedinger, J., Vogel, R., & Foufoula-Georgiou, E. (1993). Frequency Analysis of Extreme Events. *Handbook of Hydrology*, *18*.

Stott, P. A., Gillett, N. P., Hegerl, G. C., Karoly, D. J., Stone, D. A., Zhang, X., & Zwiers, F. (2010). Detection and attribution of climate change: A regional perspective. *Wiley Interdisciplinary Reviews: Climate Change*, *1*(2), 192–211. https://doi.org/10.1002/wcc.34

Tahir, M. A., Bouridane, A., & Kurugollu, F. (2007). Simultaneous feature selection and feature weighting using Hybrid Tabu Search/K-nearest neighbor classifier. *Pattern Recognition Letters*, *28*(4), 438–446. https://doi.org/10.1016/j.patrec.2006.08.016

Tenenbaum, J. B., Silva, V. De, & Langford, J. C. (2000). *A Global Geometric Framework for Nonlinear Dimensionality Reduction*. *290*(December), 2319–2323. https://doi.org/10.1126/science.290.5500.2319

Toulemonde, G., Guillou, A., & Naveau, P. (2013). *Particle filtering for Gumbel-distributed daily maxima of methane and nitrous oxide*. (November 2012), 51–62. https://doi.org/10.1002/env.2192

Trenberth, K. E. (2011). Attribution of climate variations and trends to human influences and natural variability. *Wiley Interdisciplinary Reviews: Climate Change*, *2*(6), 925–930. https://doi.org/10.1002/wcc.142

Valappil, J., & Georgakis, C. (2000). Systematic estimation of state noise statistics for extended Kalman filters. *AIChE Journal*, *46*(2), 292–308. https://doi.org/10.1002/aic.690460209

Van Der Merwe, R., & Wan, E. A. (2001). The square-root unscented Kalman filter for state and parameter-estimation. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *6*, 3461–3464.

Wan, E. a. A., & Van Der Merwe, R. (2000). The unscented Kalman filter for nonlinear estimation. *Technology*, *v*, 153–158. https://doi.org/10.1109/ASSPCC.2000.882463

Wang, F., & Gelfand, A. E. (2002). A simulation-based approach to Bayesian sample size determination for performance under a given model and for separating models. *Statistical Science*, *17*(2), 193–208. https://doi.org/10.1214/ss/1030550861

Wang, W., & Zhang, Y. (2007). On fuzzy cluster validity indices. *Fuzzy Sets and Systems*, *158*(19), 2095–2117. https://doi.org/10.1016/j.fss.2007.03.004

Wang, Y., & Zhou, L. (2005). Observed trends in extreme precipitation events in China during 1961-2001 and the associated changes in large-scale circulation. *Geophysical Research Letters*, *32*(9), 1–4. https://doi.org/10.1029/2005GL022574

Warren, F. J., & Lemmen, D. S. (2014). Synthesis. *Canada in a Changing Climate: Sector Perspectives on Impacts and Adaptation*, 1–18. Retrieved from http://www.nrcan.gc.ca/environment/resources/publications/impacts-adaptation/reports/assessments/2014/16309

Wu, Y., Chan, K. L., & Wang, L. (2004). Face recognition based on discriminative manifold learning. *Proceedings - International Conference on Pattern Recognition*, *4*, 171–174. https://doi.org/10.1109/ICPR.2004.1333731

Xie, X. L., & Beni, G. (1991). A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, pp. 841–847. https://doi.org/10.1109/34.85677

Xu, Y. M., Wang, C. D., & Lai, J. H. (2014). Weighted Multi-view Clustering with Feature Selection. *Pattern Recognition*, *53*, 25–35. https://doi.org/10.1016/j.patcog.2015.12.007

Yang, S., Ding, Z., Li, Y., Wang, X., Jiang, W., & Huang, X. (2015). Warming-induced northwestward migration of the East Asian monsoon rain belt from the Last Glacial Maximum to the mid-Holocene. *Proceedings of the National Academy of Sciences*, *112*(43), 13178–13183. https://doi.org/10.1073/pnas.1504688112

Yang, T., Shao, Q., Hao, Z.-C., Chen, X., Zhang, Z., Xu, C.-Y., & Sun, L. (2010). Regional frequency analysis and spatio-temporal pattern characterization of rainfall extremes in the Pearl River Basin, China. *Journal of Hydrology*, *380*(3–4), 386–405. https://doi.org/10.1016/j.jhydrol.2009.11.013

Yang, Z., & Burn, D. H. (2019). Automatic Feature Selection and Weighting for the Formation of Homogeneous Groups for Regional IDF Estimation. *Journal of Hydrology*, *575*(May), 292–307. https://doi.org/10.1016/j.jhydrol.2019.05.015

Yoon, S., Cho, W., Heo, J.-H., & Kim, C. E. (2010). A full Bayesian approach to generalized maximum likelihood estimation of generalized extreme value distribution. *Stochastic Environmental Research and Risk Assessment*, *24*(5), 761–770. https://doi.org/10.1007/s00477-009-0362-7

Zhang, H., & Sun, G. (2002). Feature selection using tabu search method. *Pattern Recognition*, *35*(3), 701–711. https://doi.org/10.1016/S0031-3203(01)00046-2

Zhang, X., Vincent, L. A., Hogg, W. D., & Niitsoo, A. (2000). Temperature and precipitation trends in Canada during the 20th century. *Atmosphere - Ocean*, *38*(3), 395–429. https://doi.org/10.1080/07055900.2000.9649654

Zheng, B., Fu, P., Li, B., & Yuan, X. (2018). A robust adaptive unscented kalman filter for nonlinear estimation with uncertain noise covariance. *Sensors (Switzerland)*, *18*(3). https://doi.org/10.3390/s18030808

Zhou, Q. (2014). A review of sustainable urban drainage systems considering the climate change and urbanization impacts. *Water (Switzerland)*, *6*(4), 976–992. https://doi.org/10.3390/w6040976

Zong, Y., & Chen, X. (2000). The 1998 flood on the Yangtze, China. *Natural Hazards*, *22*(2), 165–184. https://doi.org/10.1023/A:1008119805106

Appendix 1 The Procedure of Heterogeneity Testing

The probability weighted moments (PWMs) of the random variable X with the cumulative distribution function $F(X)$ can be defined as (Greenwood et al., 1979):

$$M_{p,r,s} = E\left[X^p \{F(X)\}^r \{1-F(X)\}^s\right] = \int_0^1 \left[x(F)\right]^p F^r (1-F)^s dF \tag{A1.1}$$

$$\alpha_r = \int_0^1 x(u)(1-u)^r du, \qquad \beta_r = \int_0^1 x(u)u^r du \tag{A1.2}$$

Based on the above equation, the calculation of the L-moments is conducted by implementing the shifted Legendre polynomials as the followings

$$P_r^*(u) = \sum_{k=0}^{r} p_{r,k}^* u^k \tag{A1.3}$$

Where u^k are orthogonal between -1 and 1, and

$$p_{r,k}^* = \frac{(-1)^{r-k}(r+k)!}{(k!)^2(r-k)!} \tag{A1.4}$$

So the L-moments of the X can be described as the following equation:

$$\lambda_{r+1} = (-1)^r \sum_{k=0}^{r} p_{r,k}^* \alpha_k = \sum_{k=0}^{r} p_{r,k}^* \beta_k \tag{A1.5}$$

The method of L-moments can be considered as the improved technique based on the probability weighted moments (PWMs) (Hosking & Wallis, 1997). In the parameter estimation of the GEV model, four L-moments are frequently used: λ_1 is the L-location or mean of the distribution; λ_2 is the L-scale; τ is the L-skewness; τ_3 is the L-skewness; τ_4 is the L-Kurtosis, which can be calculated based on the following equations (Hosking & Wallis, 1997):

$$\lambda_1 = \xi + \alpha\{1-\Gamma(1+k)\}/k \tag{A1.6}$$

$$\lambda_2 = \alpha(1-2^{-k})\Gamma(1+k)/k \tag{A1.7}$$

$$\tau = \lambda_2 / \lambda_1 \tag{A1.8}$$

$$\tau_3 = 2\left(1-3^{-k}\right)/\left(1-2^{-k}\right)-3 \tag{A1.9}$$

$$\tau_4 = \left\{5\left(1-4^{-k}\right)-10\left(1-3^{-k}\right)+6\left(1-2^{-k}\right)\right\}/\left(1-2^{-k}\right) \tag{A1.10}$$

After the L-moments being calculated, the following procedures can be used for the Heterogeneity Testing:

1) Calculate the regional coefficient of L-CV, L-CS or L-kurtosis using:

$$t^R = \sum_{i=1}^{N} n_i t^{(i)} \Big/ \sum_{i=1}^{N} n_i \tag{A1.11}$$

where N is the total number of stations in the region; n_i is the record length of station i; $t^{(i)}$ is the at-site L-CV, L-CS or L-kurtosis; t^R is the regional L-CV, L-CS or L-kurtosis.

2) Determine the weighted measures of dispersion (Castellarin et al., 2008):

$$V_1 = \left\{\sum_{i=1}^{N} n_i \left(t_{2(i)} - \bar{t}_2\right)^2 \Big/ \sum_{i=1}^{N} n_i\right\}^{1/2} \tag{A1.12}$$

$$V_2 = \sum_{i=1}^{N} n_i \left[\left(t_{2(i)} - \bar{t}_2\right)^2 + \left(t_{3(i)} - \bar{t}_3\right)^2\right]^{1/2} \Big/ \sum_{i=1}^{N} n_i \tag{A1.13}$$

$$V_3 = \sum_{i=1}^{N} n_i \left[\left(t_{3(i)} - \bar{t}_3\right)^2 + \left(t_{4(i)} - \bar{t}_4\right)^2\right]^{1/2} \Big/ \sum_{i=1}^{N} n_i \tag{A1.14}$$

where $\bar{t}_2, \bar{t}_3$ and $\bar{t}_4$ are the regional mean of L-CV, L-CS and L-kurtosis; $t_{2(i)}, t_{3(i)}, t_{4(i)}$ and n_i are the values of L-CV, L-Cs, L-kurtosis and sample size for site i; N is the number of sites in the homogeneous group.

3) Estimate the corresponding kappa distribution based on the regional average L-moments.

4) Determine the theoretical mean μ_V and variance σ_V through the repeated simulations based on the estimated kappa distribution.

5) Calculate the heterogeneity measure based on the following equation:

$$H = \frac{(V - \mu_V)}{\sigma_V} \tag{A1.15}$$

The pooling group can be considered as "acceptably homogeneous" if $H < 1$, "possibly heterogeneous" if $1 \leq H < 2$, and "definitely heterogeneous" if $H \geq 2$.

Appendix 2 Overview of random-walk Metropolis-Hastings (MS) algorithm

Generally, the posterior density distribution $f\left(\theta|y^{(n)}\right)$ of the parameter can be described as:

$$f\left(\theta|y^{(n)}\right)=\frac{f\left(y^{(n)}|\theta\right)f(\theta)}{f\left(y^{(n)}\right)} \tag{A2.1}$$

where $y^{(n)}$ is the observations with sample size equalling to n; $f\left(y^{(n)}\right)$ is the marginal distribution of the observations; θ represents the parameters used in the target distribution; $f(\theta)$ is the prior distribution of the parameters (F. Wang & Gelfand, 2002) and can be calculated using:

$$f\left(y^{(n)}\right)=\int f\left(y^{(n)}|\theta\right)f(\theta)d\theta \tag{A2.2}$$

Thus Equation (A2.1) can be translated into the following:

$$f\left(\theta|y^{(n)}\right)=\frac{f\left(y^{(n)}|\theta\right)f(\theta)}{\int f\left(y^{(n)}|\theta\right)f(\theta)d\theta}=\frac{L\left(\theta|y^{(n)}\right)f(\theta)}{\int L\left(\theta|y^{(n)}\right)f(\theta)d\theta} \tag{A2.3}$$

where $f\left(y^{(n)}|\theta\right)$ equals the likelihood function $L\left(\theta|y^{(n)}\right)$, indicating the probability of the value $y^{(n)}$ given the current parameters θ. Applying the above equation in the prior environment, the above equation becomes:

$$f\left(\theta|y^{(n)}\right)=\frac{L\left(\theta|y^{(n)}\right)f(\theta)}{\int L\left(\theta|y^{(n)}\right)f(\theta)d\theta} \tag{A2.4}$$

Due to the difficulty of calculating the integral in the above posterior distribution, the Markov Chain Monte Carlo (MCMC) simulation, specifically the random-walk Metropolis-Hastings (MS) algorithm, is used here to generate the samples from this posterior distribution to obtain the relevant sample statistics (Yoon et al., 2010).

The MS algorithm is developed to solve the difficulty of direct sampling in the original MCMC algorithm. In the process, an acceptance probability function was introduced to speed the time for the MCMC algorithm to converge to the stationary distribution within an acceptable error.

$$\alpha_t = \min\left\{1, \frac{q\left(\theta^{(j)}\middle|\theta^*\right)}{q\left(\theta^*\middle|\theta^{(j)}\right)} \frac{P\left(\theta^*\middle|y^{(n)*}\right)}{P\left(\theta^{(j)}\middle|y^{(n)*}\right)}\right\} \tag{A2.5}$$

where $\theta^{(*)}$ is the optimal state of parameters in the previous steps, and $\theta^{(j)}$ is the proposed parameters at the current state. The transition kernel for the parameters $q\left(\theta^{(*)}\middle|\theta^{(j)}\right)$ is used to compare the previous optimal state and the current state.

Appendix 3 Overview of the Anderson-Darling Statistic

For a given sample of size n, $x_1, x_2, \cdots, x_n$, let $x_1 \le x_2 \le \cdots \le x_n$, the empirical distribution function (EDF) $F_n(x)$ can be calculated through the following equations (Chernobai et al., 2005):

$$F_n(x) = \frac{number\ of\ observations \le x}{n} = \begin{cases} 0 & x \le x_1 \\ i/n & x_i \le x \le x_{i+1} \quad i = 1, 2, \cdots, n-1 \\ 1 & x_n \le x \end{cases} \tag{A3.1}$$

To compare the goodness-of-fit of the given sample to a given distribution, the vertical differences between the EDF and the cumulative distribution function (CDF) $F_\theta(x)$ associated with the given distribution can be used as the measure. The quadratic class of the vertical differences, which is grouped as the Cramer-von Mises family, can be calculated through the following equation (Anna Chernobai et al., 2005):

$$Q = n \int_{-\infty}^{\infty} \left(F_n(x) - F_\theta(x)\right)^2 \psi\left(F_\theta(x)\right) dF_\theta(x) \tag{A3.2}$$

where $\psi\left(F_\theta(x)\right)$ is a weighted function that assigns weights to sample differences. In the case for Anderson-Darling statistics, $\psi\left(F_\theta(x)\right) = \left\{F_\theta(x)\left(1 - F_\theta(x)\right)\right\}^{-1}$.

Normally, the AD test outperforms its counterparts through placing higher weights on the upper and lower tails and can be calculated through the following equation (Chernobai et al., 2005; Calenda et al., 2009):

$$AD^2 = -n - \frac{1}{n}\sum_{j=1}^{n}\left(2n - 2j + 1\right)\log\left(1 - F_\theta\left(x_j\right)\right) + \frac{1}{n}\sum_{j=1}^{n}\left(1 - 2j\right)\log\left(F_\theta\left(x_j\right)\right) \tag{A3.3}$$

This AD test can be modified to place more weight on the upper tail through replacing the weight function in Equation (A3.2) with $\psi\left(F_\theta(x)\right) = \left\{\left(1 - F_\theta(x)\right)\right\}^{-2}$, and the AD statistic at the upper tail can be calculated through the following equation (Anna Chernobai et al., 2005):

$$AD_{up}^2 = 2\sum_{j=1}^{n}\log\left(1 - F_\theta\left(x_j\right)\right) + \frac{1}{n}\sum_{j=1}^{n}\left(2\left(n - j\right) + 1\right)\frac{1}{1 - F_\theta\left(x_j\right)} \tag{A3.4}$$

(u, v) are set to (6, 9) in the beta prior distribution (Park, 2005).